



Editorial

Future of disinformation studies: emerging research fields
Ramón Salaverría; Gustavo Cardoso

Research articles on Artificial intelligence

Gender stereotypes in AI-generated images

Francisco-José García-Ull; Mónica Melero-Lázaro

Who interacts with communicative AI and what attitudes are displayed toward it?

Héctor Centeno-Martín; Samuel Toledano-Buendía; Alberto Ardèvol-Abreu

Toward non-human-centered design

Tuba Livberber

Retraining fact-checkers

Roger Cuartielles; Xavier Ramon-Vegas; Carles Pont-Sorribes

AI application in journalism

Beatriz Gutiérrez-Caneda; Jorge Vázquez-Herrero; Xosé López-García

ChatGPT could be the reviewer of your next scientific paper

David Carabantes; José L. González-Geraldo; Gonzalo Jover

Artificial intelligence applications in media archives

Virginia Bazán-Gil

ChatGPT: Stream of opinion in five newspapers in the first 100 days

Cristian-Augusto González-Arias; Xosé López-García

Research articles

Ibero-American journalism in the face of scientific disinformation

J. I. Martín-Neira; M. Trillo-Domínguez; M. D. Olvera-Lobo

Disinformation narratives in Spain: reach, impact and spreading patterns

Jaume Suau; David Puertas-Graell

Crude, anonymous, partisan, sectoral and anti-elitist. Electoral disinformation in Spain (2019-2022)

David Lava-Santos; José Gamir-Ríos; Germán Llorca-Abad

How to teach the elderly to detect disinformation

Charo Sádaba; Ramón Salaverría; Xavier Bringué-Sala

Fact-checking in war

Raúl Magallón-Rosa; Carolina Fernández-Castrillo; Miriam Garriga

Teachers in a world of information: Detecting false information

K. Kopecký; D. Voráč; R. Sztokowski; V. Krejčí; K. Mackenzie; M. Ramos

Disintermediation and disinformation as a political strategy

Alba Díez-Gracia; Pilar Sánchez-García; Javier Martín-Román

Using video for fact-checking on Facebook

Alberto Dafonte-Gómez; Oswaldo Garcia-Crespo; Diana Ramahí-García

Fact or fiction: An experiment on how information sources and message framing influence vaccine risk perception

Daniel Catalán-Matamoros; Enrique Prada; Andrea Langbecker

Trust, disinformation, and digital media

J. Serrano-Puche; N. Rodríguez-Salcedo; M. P. Martínez-Costa

Research on selective media exposure in Spain

Lidia Valera-Ordaz

Informative reels on Instagram in the face of new ways of reporting to attract new audiences

María-Isabel Rodríguez-Fidalgo; Adriana Paino-Ambrosio; Francisco-Javier Herrero-Gutiérrez

Interactive mapping of Covid-19 disinformation in Ibero-America

M. J. Colmenero-Ruiz; F. C. Paletta; A. Gonzales-Aguilar

SCImago Media Rankings (SMR): situation and evolution of the digital reputation of the media worldwide

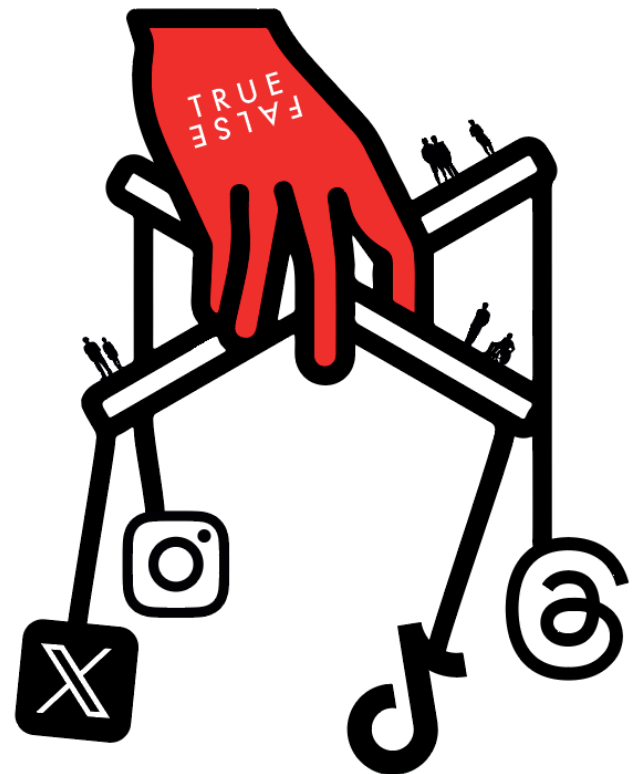
Magdalena Trillo-Domínguez; Ramón Salaverría; Lluís Codina; Félix De-Moya-Anegón

Never mind predatory publishers... what about 'grey' publishers?

D. Nicholas; E. Herman; A. Abrizah; B. Rodríguez-Bravo; C. Boukacem-Zeghmouri; A. Watkinson; M. Świgoń; J. Xu; H. R. Jamali; C. Tenopir

A search strategy for publications in interdisciplinary research

Wenjing Xiong; Ping Zhou



DISINFORMATION AND ONLINE MEDIA

07:43 AM - CHECKING THE DAILYNEWS



Profesional de la información is a journal published by Ediciones Profesionales de la Información S.L. VAT: ESB63664544

C/ Mistral, 36 (Cortijo del Aire)
18220 Albolote (Granada), Spain
Tel.: +34 - 608 491 521
<http://www.profesionaldelainformacion.com>

Cover design

Moisés Mañas
Universitat Politècnica de València
Foto de Marten Bjork en *Unsplash*

Layout

Isabel Olea
isabel.iolea@gmail.com

Papers published in EPI are approved according to the traditional double-blind “peer review” system: they are reviewed by at least two experts on the subject, from the journal’s Advisory Board and/or external experts.

In order to ensure that the papers do not lose relevance, the management and reviewers of this journal make a special effort to review the articles very quickly, achieving an average acceptance or rejection time of only a few weeks.

EDITOR-IN-CHIEF

Rodrigo Sánchez-Jiménez
Universidad Complutense de Madrid
rodsanch@ucm.es

ASSOCIATE EDITOR-IN-CHIEF

Javier Guallar
Universitat de Barcelona
<http://sites.google.com/site/sitiodejavierguallar>

EDITORIAL COORDINATION

Laura Alonso-Muñoz
Universitat Jaume I, Castellón
lalonso@uji.es

MANAGING EDITOR

Isabel Olea
EPI SL
isabel.iolea@gmail.com

SCIENTIFIC COUNCIL

Ernest Abadal	<i>Universitat de Barcelona.</i>
Isidro F. Aguillo	<i>Centro de CC Humanas y Sociales, CSIC, Madrid.</i>
Rafael Aleixandre-Benavent	<i>Universitat de València.</i>
Laura Alonso-Muñoz	<i>Universitat Jaume I, Castellón.</i>
Lluís Anglada	<i>Consorci de Serveis Universitaris de Catalunya (CSUC).</i>
Alberto Ardèvol-Abreu	<i>Universidad de La Laguna, Tenerife.</i>
Natalia Arroyo-Vázquez	<i>Universidad de León.</i>
Wileidys Artigas	<i>Universidad del Zulia, Venezuela.</i>
Tomàs Baiget	<i>Profesional de la Información.</i>
Jesús Bustamante	<i>Cedefop, Salónica, Greece.</i>
Atilio Bustos-González	<i>SCImago Research Group, Colombia.</i>
Eva Campos-Domínguez	<i>Universidad de Valladolid.</i>
Andreu Casero-Ripollés	<i>Universitat Jaume I, Castellón.</i>
Lluís Codina	<i>Universitat Pompeu Fabra, Barcelona.</i>
Carmen Costa-Sánchez	<i>Universidade da Coruña.</i>
Emilio Delgado-López-Cózar	<i>Universidad de Granada.</i>
Javier Díaz-Noci	<i>Universitat Pompeu Fabra, Barcelona.</i>
Roxana Dinu	<i>Biblioteca Națională a României, Romania.</i>
Ricardo Eito-Brun	<i>Grupo GMV, Madrid.</i>
Assumpció Estivill	<i>Universitat de Barcelona.</i>
Chiara Faggiolani	<i>Sapienza Università di Roma, Italy.</i>
Andrés Fernández-Ramos	<i>Universidad de León.</i>
Antonia Ferrer-Sapena	<i>Universitat Politècnica de València.</i>
António Fidalgo	<i>Universidade da Beira Interior, Portugal.</i>
Francisco-Javier García-Marco	<i>Universidad de Zaragoza.</i>
Homero Gil de Zúñiga	<i>Pennsylvania State University, USA.</i>
Audilio Gonzales-Aguilar	<i>Université Paul-Valéry Montpellier 3, France.</i>
Manuel Goyanas	<i>Universidad Carlos III de Madrid.</i>
Javier Guallar	<i>Universitat de Barcelona.</i>
Vicente P. Guerrero-Bote	<i>Universidad de Extremadura.</i>

Postal address

Profesional de la información
Apartado 13
24080 León, Spain
Tel.: +34 - 608 491 521
epi@elprofesionaldelainformacion.com

Advertising

Tel.: +34 - 608 491 521
publici@elprofesionaldelainformacion.com

Online services

Roxana Dinu: roxadinu@gmail.com

Online distribution

<https://revista.profesionaldelainformacion.com/index.php/EPI>

Legal deposit: B. 12.303-199

TRANSITION MANAGER

Tomàs Baiget
EPI SL
baiget@gmail.com

GUEST EDITORS

Ramón Salaverría	Universidad de Navarra rsalaver@unav.es
Gustavo Cardoso	Instituto Universitário de Lisboa gustavo.cardoso@iscte.pt

SOCIAL NETWORKS

Wileidys Artigas	High Rate Consulting, EUA Wileartigas@gmail.com
------------------	---

ENGLISH LANGUAGE REVIEW

Colin Marsh	Cambridge Copy and Translation colin.marsh@cambridgecopy.com
-------------	--

Estefanía Herrán-Páez	<i>SCImago Research Group.</i>
Ying Huang	<i>Wuhan University, China.</i>
Johannes Keizer	<i>Food and Agriculture Org. (FAO). Roma, Italy.</i>
Jesús Lau	<i>Universidad Veracruzana, Veracruz, Mexico.</i>
Javier Leiva-Aguilera	http://www.javierleiva.com , USA.
Xosé López-García	<i>Universidad de Santiago de Compostela, Santiago.</i>
Carmen Marta-Lazo	<i>Universidad de Zaragoza.</i>
Maria-Jose Masanet-Jordà	<i>Universitat de Barcelona.</i>
Charles McCathieNeville	<i>Yandex, Moscow, Russia.</i>
José-Antonio Moreira-González	<i>Universidad Carlos III de Madrid.</i>
Isabel Olea	<i>Profesional de la Información.</i>
Enrique Orduña-Malea	<i>Universitat Politècnica de València.</i>
Juan-Antonio Pastor	<i>Universidad de Murcia.</i>
Concha Pérez-Curiel	<i>Universidad de Sevilla</i>
Fernanda Peset	<i>Universitat Politècnica de València.</i>
Rafael Repiso	<i>Universidad de Málaga.</i>
Blanca Rodríguez-Bravo	<i>Universidad de León.</i>
Luis Rodríguez-Yunta	<i>Centro de CC Humanas y Sociales, CSIC, Madrid.</i>
Ramón Salaverría	<i>Universidad de Navarra, Pamplona.</i>
Blanca San-José-Montano	<i>Hospital Universitario de Móstoles, Madrid</i>
Rodrigo Sánchez-Jiménez	<i>Universidad Complutense de Madrid.</i>
Tomás Saorín	<i>Universidad de Murcia.</i>
Robert Seal	<i>Loyola Univ. Chicago, Evanston, Illinois, USA.</i>
Jorge Serrano-Cobos	<i>Universitat Politècnica de València.</i>
Ernesto Spinak	<i>Consultor, Montevideo, Uruguay.</i>
Emir-José Suaidén	<i>Universidade de Brasília, Brazil.</i>
Carlos-Miguel Tejada-Artigas	<i>Universidad Complutense de Madrid.</i>
Gerardo Tibaná	<i>SCImago Research Group, Colombia.</i>
Daniel Torres-Salinas	<i>Universidad de Granada, EC3metrics.</i>
Jesús Tramullas	<i>Universidad de Zaragoza.</i>
Lidia Valera-Ordaz	<i>Universitat de València.</i>

Contents

EPI, 2023, v. 32, n. 5

Disinformation and online media

EDITORIAL

- e320525 **Future of disinformation studies: emerging research fields**
Ramón Salaverría; Gustavo Cardoso

RESEARCH ARTICLES ON ARTIFICIAL INTELLIGENCE

- e320505 **Gender stereotypes in AI-generated images**
Francisco-José García-Ull; Mónica Melero-Lázaro
- e320502 **Who interacts with communicative AI and what attitudes are displayed toward it? Sociodemographic, personality, and futurephobia variables**
Héctor Centeno-Martín; Samuel Toledano-Buendía; Alberto Ardèvol-Abreu
- e320512 **Toward non-human-centered design: designing an academic article with ChatGPT**
Tuba Livberber
- e320515 **Retraining fact-checkers: The emergence of ChatGPT in information verification**
Roger Cuartielles; Xavier Ramon-Vegas; Carles Pont-Sorribes
- e320514 **AI application in journalism: ChatGPT and the uses and risks of an emergent technology**
Beatriz Gutiérrez-Caneda; Jorge Vázquez-Herrero; Xosé López-García
- e320516 **ChatGPT could be the reviewer of your next scientific paper. Evidence on the limits of AI-assisted academic reviews**
David Carabantes; José L. González-Geraldo; Gonzalo Jover
- e320517 **Artificial intelligence applications in media archives**
Virginia Bazán-Gil
- e320524 **ChatGPT: Stream of opinion in five newspapers in the first 100 days since its launch**
Cristian-Augusto González-Arias; Xosé López-García

RESEARCH ARTICLES

- e320503 **Ibero-American journalism in the face of scientific disinformation: Factcheckers' initiatives on the social network Instagram**
Juan-Ignacio Martín-Neira; Magdalena Trillo-Domínguez; María-Dolores Olvera-Lobo
- e320508 **Disinformation narratives in Spain: reach, impact and spreading patterns**
Jaume Suau; David Puertas-Graell
- e320506 **Crude, anonymous, partisan, sectoral and anti-elitist. Electoral disinformation in Spain (2019-2022)**
David Lava-Santos; José Gamir-Ríos; Germán Llorca-Abad
- e320504 **How to teach the elderly to detect disinformation: a training experiment with WhatsApp**
Charo Sádaba; Ramón Salaverría; Xavier Bringué-Sala
- e320520 **Fact-checking in war: Types of hoaxes and trends from a year of disinformation in the Russo-Ukrainian war**
Raúl Magallón-Rosa; Carolina Fernández-Castrillo; Miriam Garriga

- e320501 **Teachers in a world of information: Detecting false information**
Kamil Kopecký; Dominik Voráč; René Sztokowski; Veronika Krejčí; Karolína Mackenzie; Magdalena Ramos-Navas-Parejo
- e320523 **Disintermediation and disinformation as a political strategy: use of AI to analyze fake news as Trump's rhetorical resource on Twitter**
Alba Díez-Gracia; Pilar Sánchez-García; Javier Martín-Román
- e320511 **Using video for fact-checking on Facebook. Analysis of the trend and reach of Ibero-American production (2016-2021)**
Alberto Dafonte-Gómez; Oswaldo García-Crespo; Diana Ramahí-García
- e320510 **Fact or fiction: An experiment on how information sources and message framing influence vaccine risk perception**
Daniel Catalán-Matamoros; Enrique Prada; Andrea Langbecker
- e320518 **Trust, disinformation, and digital media: Perceptions and expectations about news in a polarized environment**
Javier Serrano-Puche; Natalia Rodríguez-Salcedo; María-Pilar Martínez-Costa
- e320507 **Research on selective media exposure in Spain: a critical review of its findings, application phases, and blind spots**
Lidia Valera-Ordaz
- e320519 **Informative reels on Instagram in the face of new ways of reporting to attract new audiences**
María-Isabel Rodríguez-Fidalgo; Adriana Paíno-Ambrosio; Francisco-Javier Herrero-Gutiérrez
- e320513 **Interactive mapping of Covid-19 disinformation in Ibero-America**
María-Jesús Colmenero-Ruiz; Francisco-Carlos Paletta; Audilio Gonzales-Aguilar
- e320521 **SCImago Media Rankings (SMR): situation and evolution of the digital reputation of the media worldwide**
Magdalena Trillo-Domínguez; Ramón Salaverría; Lluís Codina; Félix De-Moya-Anegón
- e320509 **Never mind predatory publishers... what about 'grey' publishers?**
David Nicholas; Eti Herman; Abdullah Abrizah; Blanca Rodríguez-Bravo; Chérifa Boukacem-Zeghmouri; Anthony Watkinson; Marzena Świgoń; Jie Xu; Hamid R. Jamali; Carol Tenopir
- e320522 **A search strategy for publications in interdisciplinary research**
Wenjing Xiong; Ping Zhou



Profesional de la información is indexed in:

DATABASES

Academic search (Ebsco)

<https://www.ebsco.com/products/research-databases/academic-search>

Academic search premier (Ebsco)

<http://www.ebscohost.com/academic/academic-search-premier>

Business source (Ebsco)

<https://www.ebsco.com/products/research-databases/business-source-complete>

CNKI, China National Knowledge Infrastructure

<https://oversea.cnki.net/index>

Communication source (Ebsco)

<https://www.ebsco.com/products/research-databases/communication-source>

Dialnet (Universidad de La Rioja)

http://dialnet.unirioja.es/servlet/revista?&clave_revista=469

ÍNDICES CSIC

<https://indices.csic.es>

Informe Académico (Cengage)

<https://www.gale.com/uk/c/informe-academico>

Inspec, Information services in physics, electronics and computing (IET, The Institution of Engineering and Technology)

<http://www.theiet.org/publishing/inspec>

ISTA, Information science and technology abstracts (Ebsco)

<https://www.ebsco.com/products/research-databases/information-science-technology-abstracts>

Library & Information Science Source (Ebsco)

<https://www.ebscohost.com/titleLists/lis-coverage.htm>

Library Science Database (ProQuest)

<https://www.proquest.com/products-services/libraryscience.html>

Lisa, Library and information science abstracts (CSA)

<https://about.proquest.com/en/products-services/lisa-set-c>

Lista, Library, information science & technology abstracts (Ebsco)

<https://www.ebsco.com/products/research-databases/library-information-science-and-technology-abstracts>

Scopus (Elsevier)

<https://www.scopus.com/sourceid/6200180164>

Social sciences citation index, Social SCI, WoS (Clarivate Analytics)

http://wokinfo.com/products_tools/multidisciplinary/webofscience/ssci

FULL TEXTS

Gestor editorial (OJS) de la revista

<https://revista.profesionaldelainformacion.com/index.php/EPI/issue/archive>

Academic search complete (2000-)

<https://www.ebsco.com/products/research-databases/academic-search-complete>

Business source complete (2000-)

<https://www.ebsco.com/products/research-databases/business-source-complete>

Communication source (2000-)

<https://www.ebsco.com/products/research-databases/communication-source>

Digitalia Publishing (2011-)

<https://www.digitaliapublishing.com/ejournal/105>

e-Libro (2011-)

<https://e-libro.net/libros/mostrarResultadoBusqueda.aspx?titulo=El%20profesional%20de%20la%20informaci%C3%B3n>

Ebscohost Electronic Journals Service (2000-)

<http://ejournals.ebsco.com/direct.asp?JournalID=105302>

Informe Académico, Cengage (2000-)

<https://www.gale.com/uk/c/informe-academico>

Lista. Library, information science & technology abstracts with full text (2000-)

<http://www.ebscohost.com/thisTopic.php?marketID=1&topicID=584>

Profesional de la información (1992-)

<http://profesionaldelainformacion.com/contenidos.html>

Scipedia (2007-)

<https://www.scipedia.com/sj/epdi>

INDEXES

CiteScore

<https://www.scopus.com/sourceid/6200180164>

CiteScore 2022 = 6,1 Q1

Dialnet Métricas

<https://dialnet.unirioja.es/metricas/idr/revistas/469>

IDR = 2,05 Q1

Dimensions

https://app.dimensions.ai/discover/publication?search_mode=content&and_facet_source_title=jour.1138422

ERIH Plus

<https://kanalregister.hkdir.no/publiseringskanaler/KanalTidsskriftInfo.action?id=449601>

Google Scholar Citations

<https://scholar.google.com/citations?user=zv5d9O0AAAAJ>

Google Scholar Metrics

https://scholar.google.com/citations?view_op=top_venues&hl=en&vq=es

Journal Citation Reports (JCR)

<https://clarivate.com/products/journal-citation-reports>

Impact Factor 2022 = 4,2 Q1

Latindex (Catálogo)

<https://www.latindex.org/latindex/ficha?folio=23967>

MIAR

<https://miar.ub.edu/issn/1386-6710>

SCImago Journal & Country Rank (SJR)

<https://www.scimagojr.com/journalsearch.php?q=6200180164&tip=sid>

Scimago Journal Rank 2022 = 0,872 Q1

SOCIAL MEDIA

http://twitter.com/revista_EPI

<http://www.facebook.com/elprofesionaldelainformacion>

<http://www.linkedin.com/company/el-profesional-de-la-informacion-epi>

https://www.instagram.com/revista_epj

EDITORIAL MANAGING SYSTEM

Profesional de la información

<https://revista.profesionaldelainformacion.com/index.php/EPI>



Future of disinformation studies: emerging research fields

Ramón Salaverría; Gustavo Cardoso

Nota: Este artículo se puede leer en español en:
<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87499>

Recommended citation:

Salaverría, Ramón; Cardoso, Gustavo (2023). "Future of disinformation studies: emerging research fields". *Profesional de la información*, v. 32, n. 5, e320525.
<https://doi.org/10.3145/epi.2023.sep.25>

Article received on October 16th 2023



Ramón Salaverría ✉
<https://orcid.org/0000-0002-4188-7811>
Universidad de Navarra
Edificio Ciencias Sociales
Campus Universitario
31080 Pamplona, Spain
rsalaver@unav.es



Gustavo Cardoso
<https://orcid.org/0000-0003-2067-4718>
ISCTE-Instituto Universitário de Lisboa
Departamento de Sociologia
Escola de Sociologia e Políticas Públicas
Av. das Forças Armadas
1649-026 Lisboa, Portugal
gustavo.cardoso@iscte-iul.pt

Abstract

This article examines research trends on disinformation. First, it explores the relationship between disinformation and digital news media, highlighting the negative impact of disinformation on citizens' trust in the news. Recent research on disinformation is classified into several areas, including typological studies, research on fact-checking, disinformation on digital platforms, and studies on media literacy. Next, the article identifies several emerging fields for research, such as studies on disinformation narratives, information manipulation and international interference, artificial intelligence generated disinformation, cross-platform disinformation, and thematic and multidisciplinary studies. Based on this analysis, the article highlights the need to continue investigating and combatting disinformation, as it is a persistent and growing problem in democratic societies.

Keywords

Disinformation; Fake news; Academic research; Research trends; Communication research; Disinformation theories; Fact-checking; Media literacy; Disinformation narratives.

Funding

This article and the guest editing of this special issue were supported by *Iberifier*, a project funded by the *European Commission* in CEF-TC-2020-2 (*European Digital Media Observatory*) call, number 2020-EU-IA-0252.

1. Introduction

Lying to achieve a purpose is as old as humanity. Throughout history there have been examples of manipulations, falsehoods, and lies (**Posetti; Matthews**, 2018). However, in the digital age we are experiencing a spiral of disinformation never seen before.

During the 20th century, the rise of totalitarianism and world wars has promoted strategies for the planned dissemination of false messages by global powers, in order to confuse adversaries, undermine their stability, and to affirm the cohesion of like-minded people. These techniques of lying have grown to the point where a complete theory and methodology of propaganda has been developed. In the first half of the last century, these practices were used with tragic results, especially in the case of Nazism (**Doob**, 1950) and Soviet communism (**Lasswell**, 1951). In the second half of the century, during the Cold War, the use of propaganda, lies, and the deliberate concealment of facts became widespread both in the Soviet bloc (**Bittman**, 1985) and in the Western world, led by the United States (**Snyder**, 1995), thus reaching a large proportion of the countries of the world.



The emergence of the Internet in the 1990s and, especially, the popularisation of social networks starting in the first decade of this century, have exponentially multiplied the volume and scope of false messages targeted at citizens. This wave of misinformation has also coincided with a weakening of the news media. Citizens who used to consume news in the media have now begun to browse and share digital content, often from non-professional and hardly-verifiable sources. The millions of updates that appear per second on social networks have replaced news as the main currency in information transactions. As a result of this transformation, in the network ecosystem, regular users are unable to verify the traceability of messages, and lose all sense of who said what.

Although the main focus of the spread of disinformation is on social networks and messaging applications, paradoxically the most popular terminology for public falsehoods holds the journalistic media responsible: in this way, people refer to ‘fake news’ (Tandoc Jr.; Lim; Ling, 2018; Quandt *et al.*, 2019). This concept is controversial, and has been widely debated in recent years, both in scientific fields and in the media (Egelhofer; Lecheler, 2019; Magallón-Rosa, 2019). Today, the relationship between disinformation and digital news media is more complex, and entails more unexpected effects than are superficially apparent.

To begin with, disinformation contributes to the problem of the news media being discredited. According to the 2022 edition of the *Digital News Report* study (Newman *et al.*, 2022), based on 93,000 surveys in 46 national markets, in that year trust in the news had decreased in 22 of the countries studied, while it had increased in just seven, reversing the slight recovery in trust that the media achieved during the years of the pandemic. On average, just four in ten people in the total sample (42%) said that they trusted most news most of the time. Finland was the country with the highest level of general trust (69%), while in the United States a rapid deterioration continued, finally reaching 26%, the lowest level of trust in the survey. In the same study, at 33% Spain clearly fell below the average, and was among the countries with the least trust in the media.

In this general scenario of growing discredit and disinterest in the media, a large proportion of citizens, especially those who align themselves with extreme and populist ideological positions, are suspicious of journalistic information, and have begun to question every news or op-ed that opposes their beliefs (Pérez-Curiel; Rivas-de-Roca, 2022). For this increasingly polarised contingency of citizens (Rodríguez-Virgili; Portilla-Manjón; Sierra-Iso, 2022), media organisations are suspected of defending counterfeit interests, bending to the dictates of political or economic powers, and, ultimately, of disseminating fake news.

In parallel with the surge of disinformation, academic interest in this phenomenon has grown in recent years. Research groups, observatories, and international academic networks are studying this phenomenon, which is considered one of the main ‘information disorders’ (Wardle; Derakhshan, 2017) of our time. The projects, methodologies, and contexts by which the public dissemination of falsehoods is studied have multiplied, to the point where they now compose a diverse and especially fruitful corpus of research.

This thematic issue of *Profesional de la información* dedicated to ‘Disinformation and online media’ is an example of academic interest in the phenomenon. In particular, the issue is devoted to analysing the relationships between disinformation and journalism, and explores the extent to which hoaxes pose a threat or an opportunity for journalistic media. The erosion of citizen trust in public information constitutes a danger for the media, but is also an opportunity for professional journalism to vindicate itself. The articles that constitute this thematic issue reveal a dual dimension to disinformation, which consists of both light and darkness. Several studies confirm that disinformation is, indeed, a widespread problem, with worrying effects on society. However, at the same time, other studies reveal that we are beginning to find initiatives capable of mitigating it.

As an introduction to these studies, we present below a panoramic analysis of recent research on disinformation. We summarise what has been investigated so far, and offer an overview of the current main lines of research in this field. Next, we indicate some emerging fields that we think can boost studies on disinformation in the coming years.

2. Overview of disinformation studies

What have been the main lines of research on disinformation in this century? Some recent bibliometric studies confirm that, especially as a result of the Covid-19 pandemic, disinformation has become a priority topic in communication research (García-Marín; Salvat-Martinrey, 2021; Salvador-Mata; Cortiñas-Rovira; Herrero-Solana, 2023). We identify some characteristics of this scientific investigation from a qualitative perspective.

2.1. Typological studies of disinformation

One of the areas where researchers have placed the most emphasis has been on the definition and classification of disinformation content (Wardle; Derakhshan, 2017). Perhaps the immense popularity achieved by the concepts of ‘post-truth’ and ‘fake news’ (which became common in political and journalistic forums before breaking into the academic world) contributed to significant attention being devoted to the typological aspects of disinformation.

Indeed, many studies have been dedicated to outlining the concept of disinformation (Kapantai *et al.*, 2021; Pérez-Escobar; Lilleker; Tapia-Frade, 2023). Starting from a unanimously accepted distinction between involuntary errors (misinfor-

mation) and deliberate falsehoods (disinformation), researchers who have carried out typological studies have proposed a varied range of classifications, which specify multiple types of falsehoods within these two main categories.

Beyond proposing classifications, there have also been contemporary studies that investigate the nature of disinformation, how it spreads, what motivates it, and how it affects society at large. This conceptual analysis has provided a solid theoretical framework, based on which other empirical studies have addressed specific cases or have tested mitigation strategies.

2.2. Studies on fact-checking

In the last decade, one of the areas with the most prolific academic production in studies on disinformation has focussed on the work of fact-checking organisations (Graves, 2016). The innovative nature of these journalistic organisations – which are located at a different level from both classic media (press, radio, and television) and digital native media– has attracted research from multiple perspectives (López-Pan; Rodríguez-Rodríguez, 2020; Dafonte-Gómez; Míguez-González; Ramahí-García, 2022). Within these fact-checking organisations, the professional roles, ethical standards, and even corporate structures have been investigated, and in many cases these are in the context of non-profit organisations. However, the main focus of these studies has been fact-checking articles, a unit of analysis that has proved very useful for studies of disinformation through content analysis.

Thanks to the empirical evidence provided by fact-checking organisations, research in this field has delved into emblematic cases of disinformation, from hoaxes of limited scope to large-scale disinformation campaigns. Techniques for spreading false information have been analysed, from image manipulation to the creation of fraudulent websites. Additionally, how the perception of fact-checking can vary between different demographic groups has been explored, as well as how these disparities affect the spread of disinformation.

2.3. Studies on disinformation on digital platforms

When analysing disinformation cases and campaigns, digital platforms have undoubtedly been the environment to which researchers have devoted most attention (Di-Domenico, 2021). Despite the frequent difficulties in accessing data from platforms, which are not very transparent in this regard, many investigations have analysed the traceability and dynamics of online dissemination of disinformation content. Using graphs and other reticular representations, disinformation flows have been examined, identifying social network profiles that create false content and/or contribute to its dissemination.

This type of study has also demonstrated the existence of a varied range of falsification practices on the networks, such as imposter profiles, bots, trolls, or so-called astroturfing, a form of falsification that consists of artificially creating trending topics through the planned and hidden coordination of multiple social media accounts (Arce-García; Said-Hung; Mottareale-Calvanese, 2022; Chan, 2022).

2.4. Studies on media literacy

Education and media literacy are crucial components in the fight against disinformation. This has been recognised by multiple international institutions and organizations, which have promoted literacy campaigns (Aguaded; Sandoval-Romero; Rodríguez-Rosell, 2016; Sádaba; Salaverría, 2023). In recent years, academic studies on this discipline have focussed mainly on young people, a group particularly exposed to false information due to their intensive use of digital devices and consumption of content from unfiltered sources. Research has explored effective strategies for teaching critical thinking and fact-checking skills through educational programmes.

3. Emerging fields for disinformation studies

Despite the development achieved in recent years by disinformation studies (Correia; Jerónimo; Amaral, 2022), new fields are opening up for the investigation of this phenomenon. The continuous transformation of the channels of dissemination, production technologies, and derived effects of false content demands that new lines of study are opened. Accordingly, below we list some of the lines that have not been sufficiently explored so far, and that may pave the way to a new era of disinformation studies.

3.1. Studies on disinformation narratives

In recent years, research has been more concentrated on analysing isolated hoaxes than identifying complex disinformation narratives. However, the study of hoaxes has revealed that public falsehoods often respond to a planned strategy, where the authors of the falsehood aim to stimulate certain visions or topics of debate in public opinion. However, if we only study the trees, we run the risk of not seeing the forest.

As some of the studies in this thematic issue show (Suau; Puertas-Graell, 2023), the new era of disinformation studies will have to focus more on general and systemic aspects, and attempt to account for the coordinated logic that explains why certain hoaxes, apparently independent, arise in certain contexts and situations. It is only if we identify the general patterns of false content, will we be able to anticipate strategic measures to neutralise the social impact of disinformation campaigns.

3.2. Studies on foreign information manipulation and interference

Unfortunately, disinformation has countless parents. These range from large countries with geostrategic interests to individuals who have fun confusing their peers, but also include political parties, ideological groups, and activist organisations of various kinds that indiscriminately manipulate messages with the purpose of defending their postulates and harming their adversaries. Among all these actors, there is increasing concern about the incidence of disinformation agents of international reach, who seek to destabilise countries, manipulate electoral processes, or criminalise certain social groups.

Research on the phenomena of *Foreign Information Manipulation and Interference (FIMI)* is often close to strategic intelligence and cybersecurity, transcending the capabilities and competencies of ordinary academic researchers. However, teams that investigate disinformation can also provide relevant knowledge to unmask these strategies of foreign interference. If, as we have indicated in the previous point, researchers focus on identifying disinformation campaigns beyond simple isolated falsehoods, they will be contributing to detecting international flows of disinformation, the dynamics of which we have only just begun to understand.

3.3. Studies on AI-generated disinformation

The emergence of generative artificial intelligence technologies, which allow texts, images, sounds and videos to be created or manipulated with astonishing ease, pose one of the greatest challenges to disinformation research. Some fear that these technologies will produce an 'infinite' supply of misinformation (DiResta, 2020). The risk of deep fake content is no longer a remote possibility (Paris; Donovan, 2019) and, in 2023, this phenomenon began to appear in some hoaxes and disinformation campaigns. To face this challenge, a priority research area in the coming years will be studies aimed at detecting deep fake content, as well as studies on the perception of AI-generated content.

3.4. Cross-platform disinformation studies

A large number of studies in recent years have revealed that the main strand of disinformation is found on digital platforms, particularly social networks and messaging applications. Thanks to the application programming interfaces (API) of some of these platforms, researchers have been able to track and quantify disinformation content on social networks. The scarcity of similar technological resources on other platforms may, however, have led researchers to neglect disinformation phenomena outside of social networks. Although it has been less studied, disinformation also circulates elsewhere.

It may be surprising, for example, that one of the areas where disinformation has been least investigated is that of the news media. There is a lot of talk about 'fake news' but, paradoxically, very little research has been done on the dynamics of the spread of false content in the media. We do not know to what extent current media fulfil their social function and act as shields for disinformation. Nor do we know to what extent some of the news media contribute to the dissemination of disinformation, perhaps driven by a relaxation in internal fact-checking processes, a desire to increase traffic at all costs, or even by hidden political agendas. Nor has the role of the so-called 'pseudo-media' (Palau-Sampio, 2022) been sufficiently studied; these are organisations that adopt the appearance of professional news media but in reality serve a propaganda or disinformation purpose.

Another insufficiently studied scenario of disinformation is the political sphere. Fact-checking organisations dedicate much of their work to verifying public discourse but, in contrast, there is little academic research dedicated to analysing the creation and strategic dissemination of disinformation discourses from political parties and leaders (e.g. Llorca; Fabregat-Cabrera; Ruiz-Callado, 2021).

In short, we know a lot about the disinformation that circulates on social networks, but comparatively little about what happens in other spaces. A complete understanding of the disinformation phenomenon implies delving into those fields that until now have been less addressed by research.

3.5. Thematic studies on disinformation

Until the outbreak of the pandemic in 2020, the vast majority of studies on disinformation focussed on politics. Electoral processes, political campaigns and parties were, in fact, the usual focus of studies that employed different methodologies and in different geographical frameworks (Bovet; Makse, 2019). Thanks to its global impact and long duration, Covid-19 produced a sudden explosion of disinformation studies around health (Salaverría *et al.*, 2020; García-Marín, 2020; Cardoso; Sepúlveda; Narciso, 2022), which extended to other scientific topics (León *et al.*, 2022). Subsequently, at the beginning of 2022, the Russian invasion of Ukraine unleashed a new wave of war disinformation, resulting in a number of studies focussing on that conflict. As expected, academic research on disinformation is focussed on public issues that, at all times, serve as a trigger for falsehoods.

However, there are other topics that are also the subject of disinformation and that, perhaps because they do not have the sudden social impact of the events we have just mentioned, receive relatively little attention from researchers. Issues such as the environment and the global climate crisis, immigration, or the reduction of social inequalities are a permanent focus of disinformation campaigns but, by comparison, they are much less investigated. One of the challenges to renew research on public falsehoods is to open up the field to a range of new topics.

The 17 Sustainable Development Goals established in 2015 by the *United Nations General Assembly* offer a good collection of key issues to research. As essential challenges for the future of humanity, they are critical areas for the proliferation of disinformation discourses. Therefore, it is advisable to expand the spectrum of research to these aspects because, although they lack the prominence of other issues, they constitute an area of permanent attention.

3.6. Multidisciplinary studies of disinformation

As we have already explained, the public dissemination of deliberately falsified content has been addressed by multiple disciplines. Various specialist branches of the social sciences, computer science, and even some areas of the humanities have addressed and studied disinformation phenomena in recent years. However, studies with a multidisciplinary approach are scarce.

The systemic nature of disinformation indicates the necessity of further investigations where different specialist areas collaborate. In fact, some of the most promising and fruitful lines of work in recent years come from this knowledge sharing. For example, the combination of linguistics and artificial intelligence technologies, in particular the branch of natural language processing, has opened a very productive field of research that is allowing the automatic identification of linguistic patterns of false content and hate speech. Other disciplines with great possibilities for collaboration are psychology, neuroscience and communication studies. Understanding how the mind perceives reality and how cognitive biases affect us when we consume information has been revealed as a key factor, which is beginning to be considered in numerous contemporary studies on disinformation (Pennycook; Rand, 2021; Martínez-Costa *et al.*, 2022).

3.7. Studies on media literacy for older people

Along with fact-checking, media and information literacy (MIL) has been one of the main areas that has been explored in recent years to counter the rise of disinformation. Public and private institutions have launched campaigns to train citizens in digital skills and to promote safe practices in the consumption of information. Although most media literacy initiatives have been aimed at young people, in recent years there has been a multiplication of initiatives aimed at an intergenerational audience.

However, in terms of research, until now most of the studies on digital literacy have focussed on younger sectors of the population. Although it is recognised that older people show a particular vulnerability to disinformation messaging, there are few studies aimed at analysing their information consumption habits and ways to reinforce their protection against purposefully false messages. This is another line of research necessary in the coming years, given the increasing ageing of the population in Western countries.

4. Final thoughts

Much has already been researched about disinformation, but new fields are emerging that demand renewed research. It is necessary to face the new scenarios, technologies, and narratives in which fake content is present, once again.

However, faced with these new challenges, we would make a mistake if we conceived disinformation studies as just a passing fad, which is bound to be replaced sooner or later by some new major topic of communication research. Manipulation and falsification of information is a serious problem for democratic societies and, unfortunately, it does not stop growing. Its impact on the behavior of citizens is unquestionable. In the coming years, researchers will continue to face the challenge of providing ways to understand and combat it.

5. References

- Aguaded, Ignacio; Sandoval-Romero, Yamile; Rodríguez-Rosell, María M. (2016). "Media literacy from international organizations in Europe and Latin America". *The journal of media literacy*, v. 63, n. 1-2, pp. 10-17.
- Arce-García, Sergio; Said-Hung, Elías; Mottareale-Calvanese, Daria (2022). "Astroturfing as a strategy for manipulating public opinion on Twitter during the pandemic in Spain". *Profesional de la información*, v. 31, n. 3, e310310. <https://doi.org/10.3145/epi.2022.may.10>
- Bittman, Ladislav (1985). *The KGB and Soviet disinformation: an insider's view*. Washington: Pergamon-Brassey's. ISBN: 978 0 080315720
- Bovet, Alexandre; Makse, Hernan A. (2019). "Influence of fake news in Twitter during the 2016 US presidential election". *Nature communications*, v. 10, n. 1. <https://doi.org/10.1038/s41467-018-07761-2>
- Cardoso, Gustavo; Sepúlveda, Rita; Narciso, Inês (2022). "WhatsApp and audio misinformation during the Covid-19 pandemic". *Profesional de la información*, v. 31, n. 3. <https://doi.org/10.3145/epi.2022.may.21>
- Chan, Jovy (2022). "Online astroturfing: a problem beyond disinformation". *Philosophy & social criticism*, Online first. <https://doi.org/10.1177/01914537221108467>

- Correia, João-Carlos; Jerónimo, Pedro; Amaral, Inês** (2022). *Disinformation studies: Perspectives from an emerging field*. Covilhã, Portugal: Labcom, Universidade da Beira Interior. ISBN: 978 989 654 863 6
- Dafonte-Gómez, Alberto; Míguez-González, María-Isabel; Ramahí-García, Diana** (2022). "Fact-checkers on social networks: analysis of their presence and content distribution channels". *Communication & society*, v. 35, n. 3, pp. 73-89. <https://doi.org/10.15581/003.35.3.73-89>
- Di-Domenico, Giandomenico; Sit, Jason; Ishizaka, Alessio; Nunan, Daniel** (2021). "Fake news, social media and marketing: A systematic review". *Journal of business research*, v. 124, pp. 329-341. <https://doi.org/10.1016/j.jbusres.2020.11.037>
- DiResta, Renée** (2020). "The supply of disinformation will soon be infinite". *The Atlantic*, September 20. <https://www.theatlantic.com/ideas/archive/2020/09/future-propaganda-will-be-computer-generated/616400>
- Doob, Leonard W.** (1950). "Goebbels' principles of propaganda". *Public opinion quarterly*, v. 14, n. 3, pp. 419-442. <https://doi.org/10.1086/266211>
- Egelhofer, Jana-Laura; Lecheler, Sophie** (2019). "Fake news as a two-dimensional phenomenon: A framework and research agenda". *Annals of the international communication association*, v. 43, n. 2, pp. 97-116. <https://doi.org/10.1080/23808985.2019.1602782>
- García-Marín, David** (2020). "Infodemia global. Desórdenes informativos, narrativas fake y fact-checking en la crisis de la Covid-19". *Profesional de la información*, v. 29, n. 4, e290411. <https://doi.org/10.3145/epi.2020.jul.11>
- García-Marín, David; Salvat-Martínrey, Guiomar** (2021). "Investigación sobre desinformación en España. Análisis de tendencias temáticas a partir de una revisión sistematizada de la literatura". *Fonseca, journal of communication*, v. 23, pp. 199-225. <https://doi.org/10.14201/jfc202123199225>
- Graves, Lucas** (2016). *Deciding what's true: The rise of political fact-checking in American journalism*. New York: Columbia University Press. ISBN: 978 0 231175074
- Kapantai, Eleni; Christopoulou, Androniki; Berberidis, Christos; Peristeras, Vassilios** (2021). "A systematic literature review on disinformation: Toward a unified taxonomical framework". *New media & society*, v. 23, n. 5, pp. 1301-1326. <https://doi.org/10.1177/1461444820959296>
- Lasswell, Harold D.** (1951). "The strategy of Soviet propaganda". *Proceedings of the Academy of Political Science*, v. 24, n. 2, pp. 66-78. <https://doi.org/10.2307/1173235>
- León, Bienvenido; Martínez-Costa, María-Pilar; Salaverría, Ramón; López-Goñi, Ignacio** (2022). "Health and science-related disinformation on Covid-19: a content analysis of hoaxes identified by fact-checkers in Spain". *PLoS one*, v. 17, n. 4, e0265995. <https://doi.org/10.1371/journal.pone.0265995>
- Llorca, Elena; Fabregat-Cabrera, María-Elena; Ruiz-Callado, Raúl** (2021). "Desinformación populista en redes sociales: la tuitosfera del juicio del Procés". *Observatorio (OBS*)*, v. 15, n. 3, pp. 124-146. <https://doi.org/10.15847/obsOBS15320211835>
- López-Pan, Fernando; Rodríguez-Rodríguez, Jorge-Miguel** (2020). "El fact checking en España. Plataformas, prácticas y rasgos distintivos". *Estudios sobre el mensaje periodístico*, v. 26, n. 3, pp. 1045-1065. <https://doi.org/10.5209/ESMP.65246>
- Magallón-Rosa, Raúl** (2019). *Unfaking news. Cómo combatir la desinformación*. Madrid: Pirámide. ISBN: 978 84 368 4122 0
- Martínez-Costa, María-Pilar; López-Pan, Fernando; Buslón, Nataly; Salaverría, Ramón** (2022). "Nobody-fools-me perception: Influence of age and education on the overconfidence of spotting disinformation". *Journalism practice*, Online first. <https://doi.org/10.1080/17512786.2022.2135128>
- Newman, Nic; Fletcher, Richard; Robertson, Craig T.; Eddy, Kirsten; Nielsen, Rasmus-Kleis** (2022). *Digital news report 2022*. Reuters Institute. <https://reutersinstitute.politics.ox.ac.uk/es/digital-news-report/2022>
- Palau-Sampio, Dolors** (2022). "Pseudo-media disinformation patterns: Polarised discourse, clickbait and twisted journalistic mimicry". *Journalism practice*, Online first. <https://doi.org/10.1080/17512786.2022.2126992>

- Paris, Britt; Donovan, Joan** (2019). *Deepfakes and cheap fakes: The manipulation of audio and visual evidence*. Data & Society Research Institute.
<https://datasociety.net/library/deepfakes-and-cheap-fakes>
- Pennycook, Gordon; Rand, David G.** (2021). "The psychology of fake news". *Trends in cognitive sciences*, v. 25, n. 5, pp. 388-402.
<https://doi.org/10.1016/j.tics.2021.02.007>
- Pérez-Curiel, Concha; Rivas-de-Roca, Rubén** (2022). "Exploring Populism in times of crisis: An analysis of disinformation in the European context during the US elections". *Journalism and media*, v. 3, n. 1, pp. 144-156.
<https://doi.org/10.3390/journalmedia3010012>
- Pérez-Escolar, Marta; Lilleker, Darren; Tapia-Frade, Alejandro** (2023). "A systematic literature review of the phenomenon of disinformation and misinformation". *Media and communication*, v. 11, n. 2, pp. 76-87.
<https://doi.org/10.17645/mac.v11i2.6453>
- Posetti, Julie; Matthews, Alice** (2018). *A short guide to the history of 'fake news' and disinformation*. International Center for Journalists, July 2018.
https://www.icfj.org/sites/default/files/2018-07/A%20Short%20Guide%20to%20History%20of%20Fake%20News%20and%20Disinformation_ICFJ%20Final.pdf
- Quandt, Thorsten; Frischlich, Lena; Boberg, Svenja; Schatto-Eckrodt, Tim** (2019). "Fake news". *The international encyclopedia of journalism studies*.
<https://doi.org/10.1002/9781118841570.iejs0128>
- Rodríguez-Virgili, Jordi; Portilla-Manjón, Idoia; Sierra-Iso, Aurken** (2022). "Cuarenta años de polarización ideológica en España". *Revista empresa y humanismo*, v. 25, n. 2, pp. 75-103.
<https://doi.org/10.15581/015.XXV.2.75-103>
- Sádaba, Charo; Salaverría, Ramón** (2023). "Combatir la desinformación con alfabetización mediática: análisis de las tendencias en la Unión Europea". *Revista latina de comunicación social*, n. 81, pp. 17-33.
<https://doi.org/10.4185/rllcs-2023-1552>
- Salaverría, Ramón; Buslón, Nataly; López-Pan, Fernando; León, Bienvenido; López-Goñi, Ignacio; Erviti, María-Carmen** (2020). "Desinformación en tiempos de pandemia: tipología de los bulos sobre la Covid-19". *Profesional de la información*, v. 29, n. 3, e290315.
<https://doi.org/10.3145/epi.2020.may.15>
- Salvador-Mata, Bertrán; Cortiñas-Rovira, Sergi; Herrero-Solana, Víctor** (2023). "La investigación en periodismo y Covid-19 en España: mayor impacto académico en citas, aproximaciones metodológicas clásicas e importancia temática de la desinformación". *Revista latina de comunicación social*, n. 81, pp. 554-574.
<https://doi.org/10.4185/rllcs-2023-2001>
- Snyder, Alvin A.** (1995). *Warriors of disinformation: American propaganda, Soviet lies, and the winning of the Cold War: an insider's account*. New York: Arcade Publishing. ISBN: 978 1 559703895
- Suau, Jaume; Puertas-Graell, David** (2023). "Disinformation narratives in Spain: reach, impact and spreading patterns". *Profesional de la información*, v. 32, n. 5, e320508.
<https://doi.org/10.3145/epi.2023.sep.08>
- Tandoc Jr., Edson C.; Lim, Zheng-Wei; Ling, Richard** (2018). "Defining 'fake news'. A typology of scholarly definitions". *Digital journalism*, v. 6, n. 2, pp. 137-153.
<https://doi.org/10.1080/21670811.2017.1360143>
- Wardle, Claire; Derakhshan, Hossein** (2017). *Information disorder: toward an interdisciplinary framework for research and policy-making*. Strasbourg: Council of Europe.
<https://edoc.coe.int/en/media/7495-information-disorder-toward-an-interdisciplinary-framework-for-research-and-policy-making.html>

Gender stereotypes in AI-generated images

Francisco-José García-Ull; Mónica Melero-Lázaro

Nota: Este artículo se puede leer en español en:

<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87305>

Recommended citation:

García-Ull, Francisco-José; Melero-Lázaro, Mónica (2023). "Gender stereotypes in AI-generated images". *Profesional de la información*, v. 32, n. 5, e320505.

<https://doi.org/10.3145/epi.2023.sep.05>

Manuscript received on 6th March 2023

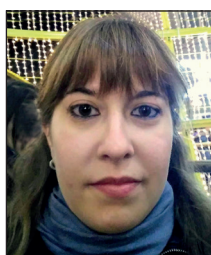
Accepted on 21st July 2023



Francisco-José García-Ull ✉

<https://orcid.org/0000-0002-7999-4807>

Universidad Europea de Valencia
Facultad de Ciencias Sociales
Pg. de l'Albereda, 7
46010 València, Spain
franciscojose.garcia@universidadeuropea.es



Mónica Melero-Lázaro

<https://orcid.org/0000-0002-2517-8468>

Universidad de Valladolid
Facultad de Filosofía y Letras
Pza. del Campus Universitario
47011 Valladolid, Spain
monimeler@outlook.com

Abstract

This study explores workplace gender bias in images generated by *DALL-E 2*, an application for synthesising images based on artificial intelligence (AI). To do this, we used a stratified probability sampling method, dividing the sample into segments on the basis of 37 different professions or prompts, replicating the study by Farago, Eggum-Wilkens and Zhang (2020) on gender stereotypes in the workplace. The study involves two coders who manually input different professions into the image generator. *DALL-E 2* generated 9 images for each query, and a sample of 666 images was collected, with a confidence level of 99% and a margin of error of 5%. Each image was subsequently evaluated using a 3-point Likert scale: 1, not stereotypical; 2, moderately stereotypical; and 3, strongly stereotypical. Our study found that the images generated replicate gender stereotypes in the workplace. The findings presented indicate that 21.6% of AI-generated images depicting professionals exhibit full stereotypes of women, while 37.8% depict full stereotypes of men. While previous studies conducted with humans found that gender stereotypes in the workplace exist, our research shows that AI not only replicates this stereotyping, but reinforces and increases it. Consequently, while human research on gender bias indicates strong stereotyping in 35% of instances, AI exhibits strong stereotyping in 59.4% of cases. The results of this study emphasise the need for a diverse and inclusive AI development community to serve as the basis for a fairer and less biased AI.

Keywords

Artificial intelligence; AI; *OpenAI*; *DALL-E*; Synthetic images; AI-generated images; Imaging; Gender stereotypes; Sex biases; Gender biases; Gender differences; Professions; Workers; Jobs; Ethics; Discrimination; Inequalities; Fairness; Equity.

Financiación

This article is part of the "Flujos de desinformación, polarización y crisis de la intermediación mediática (Disflows) (PID2020-113574RB-I00)" project, funded by the Spanish *Ministry of Science and Innovation*.

1. Introduction

Artificial Intelligence (AI) has significantly influenced society in recent years, transforming task execution and decision-making across diverse fields, such as legal (Sourdin, 2018), health (Loftus *et al.*, 2020), business (Nica *et al.*, 2022; Belhadi *et al.*, 2022) and education (Porayska-Pomsta; Rajendran, 2019). However, concerns have arisen regarding potential gender biases in AI systems.

One issue contributing to gender bias in AI is the lack of diversity in the data used to train the models. If the data used to train an AI model lacks sufficient representation of the population's diversity, the model may inadvertently learn and



perpetuate gender stereotypes (Leavy, 2018). One of the recent developments in AI is the *DALL-E 2* (2021) language generator, developed by *OpenAI*. Although *DALL-E 2* has demonstrated impressive content generation capabilities, acknowledging and addressing potential manifestations of gender biases in its use is essential.

<https://openai.com/dall-e-2>

The model's training data primarily originates from the internet, resulting in the potential inclusion of gender stereotypes prevalent in online culture. Taking into account these types of limitations when designing a rigorous study is important.

So is accounting for how gender biases might manifest in the data generated by *DALL-E 2*. For instance, a study by **Buolamwini** and **Gebu** (2018) found that commercial facial recognition systems had a higher error rate in classifying individuals of colour and gender non-binary individuals, indicating that AI systems can perpetuate existing gender biases in society. Due to the swift adoption of applications like *DALL-E 2*, *ChatGPT*, and other *OpenAI* systems (Vincent, 2020), recognising and addressing the gaps and limitations that perpetuate models and stereotypes becomes essential. In recent years, various publications, including articles, books, and documentaries, have highlighted the presence of gender, race, and class biases in new digital technologies (D'Ignazio; Klein, 2020; Criado-Pérez, 2020).

To mitigate gender biases in *DALL-E 2*, establishing stereotype-free databases to train the model has been suggested. This would reduce biases and foster gender equality. Further suggestions include using debiasing techniques to eliminate gender biases in the results generated by *DALL-E 2* (Caliskan; Bryson; Narayanan, 2017).

Considering possible gender biases in its use is important. AI systems have the potential to replicate prevailing gender stereotypes in society if the data used to train the models lacks sufficient diversity (Leavy et al., 2020). Addressing these biases requires the implementation of data diversification and debiasing techniques. Continued research and monitoring of potential gender biases in *DALL-E 2* and other AI systems, along with their associated benefits and risks (De-Carvalho, 2021), are vital. These systems will have diverse applications and may directly or indirectly influence our interpersonal relationships. Ensuring fairness, equity, and ethical usage of these systems is paramount (Cortina, 2019).

2. Theoretical framework

2.1. *OpenAI* and generative Artificial Intelligence

Artificial Intelligence has made significant strides in recent decades, thanks to the combination of machine learning techniques and the enhanced capacity for data processing. One of the most recent developments in this field is *DALL-E*, an AI generator developed by *OpenAI* that is capable of generating images based on text descriptions (Brown et al., 2020). The recently released mini version of *DALL-E*, known as *DALL-E 2*, has demonstrated its impressive capability to generate highly accurate and detailed images from basic text descriptions.

DALL-E 2 is an upgraded version of its predecessor, *DALL-E*, released in 2021. In contrast to its predecessor, *DALL-E 2* has a more streamlined architecture and reduced data processing capacity, making it more accessible and user-friendly for a wide range of users and applications. Despite this smaller size, *DALL-E 2* has proven to be just as effective at generating images from text descriptions.

In terms of how it works, *DALL-E 2* employs generative deep learning and Generative Adversarial Network (GAN) technology, where a neural network is trained to generate images based on input data. In the case of *DALL-E 2*, the input data comprises a text description, which is then processed by the neural network to produce a corresponding image. This technique relies on machine learning principles and the neural network's capacity to discern patterns and associations within the input data.

Regarding *DALL-E 2*'s potential applications, one of the most obvious is its use in advertising and graphic design. *DALL-E 2* could be used to generate tailor-made images for advertising campaigns or to craft distinctive image designs for various products. Each newly generated image from this system is original and unpublished, which raises notable controversy and presents author copyright limitations (Estupiñán-Ricardo et al., 2021). It can also be applied to education, by enabling the generation of images to illustrate concepts in textbooks or class presentations. It is in this area where certain studies have highlighted gender biases in the depiction of women in science-related contexts (Manassero; Vázquez, 2003; Francescutti, 2018). The judicious application of AI has the potential to transcend these limitations, contributing to a more equitable and inclusive society. *DALL-E 2* is also applied in the animation industry and video game production, as it facilitates the automated generation of scenarios and characters.

Nonetheless, there are ethical concerns regarding this technology (Quirós-Fons; García-Ull, 2022). One of the main apprehensions is the potential use of *DALL-E 2* to produce false or misleading content. There are also privacy and security concerns, as *DALL-E 2* could be used to generate images of individuals without their consent or used as a tool for digital violence (Pérez-Gómez et al., 2020). Concurring with this, Veliz (2021) emphasises the influential power that the manipulation of private data can yield, underscoring the need to foster initiatives and tools that safeguard user privacy. Also stressing the centralisation of power motivated by technological hegemony, authors such as Crawford (2021) assert the emergence of a trend towards

“ One issue contributing to gender bias in AI is the lack of diversity in the data used to train the models ”

greater inequality. They call upon technology companies to harness AI to steer the trajectory towards democratic values and the restructuring of the political and social landscape. In a similar vein, **O'Neil (2018)** cautions against the presence of opaque and unregulated algorithms and models that perpetuate discrimination, favouring the fortunate while penalising the marginalised.

2.2. GAN technology

Generative Adversarial Networks (GAN) are machine learning models capable of generating new and realistic content such as images, audio, and text. These models are made up of two neural networks: the first is called a generator and the second a discriminator. The generator is responsible for generating new content, while the discriminator is responsible for determining whether the generated content is real or false. The two models engage in a zero-sum game, competing against each other with the objective of enhancing the quality of the generated content.

The most influential work on GANs was carried out by **Goodfellow et al. (2014)** in their article "*Generative Adversarial Networks*". In this paper, the authors presented a basic GAN architecture and showed how it could be used to generate images of human faces. Since then, GANs have been used to generate a variety of content, including images, audio, and text.

One of the primary advantages of GANs is their ability to generate realistic content, which has resulted in their application across various domains, such as video game production, animation, and product design. GANs have also been used to generate scenarios and characters for video games, and product images that are used in decision making for design.

GANs have been applied in the medical field too. They have been used to generate brain CT images to aid the diagnosis of neurodegenerative diseases (**Laino et al., 2022**), and of cells and tissues to assist in the research and development of new treatments.

However, there are real ethical concerns surrounding GANs. One of them is that GANs might be used to generate false or misleading content (**García-Ull, 2021; Gamir-Ríos; Tarullo, 2022**).

2.3. Gender stereotypes and AI

As we have stated, Artificial Intelligence is continuously evolving, and it has the potential to transform the way we live, work and interact. However, it also poses valid ethical and social justice concerns, especially regarding gender stereotypes (**Wang et al., 2019**). Gender stereotypes are social beliefs and expectations regarding the characteristics, behaviours, and roles deemed suitable for men and women. Such stereotypes have the potential to constrain individuals' opportunities and expectations, fostering discrimination and perpetuating inequality.

In the field of AI, gender stereotypes manifest in a variety of ways, and this manifestation and consequent representation of gender is based on the data used to train AI models. If the data used to train an AI model contains gender stereotypes, the model is likely to reproduce those stereotypes. As an illustration, an AI model trained on images portraying men and women adhering to traditional gender roles might not be able to show women in non-traditional roles (**Agudo; Liberal, 2020; Traylor, 2022**). This will directly affect images in contexts such as the professional workplace or home care (**Bolukbasi et al., 2016**). Furthermore, the handling of data, algorithm design, and even the appearance of the hardware, as seen in humanoid robots, might reproduce gender stereotypes (**Ortiz-de-Zárate-Alcarazo, 2023**).

Another concern is the way AI models are designed and tested. AI designers and testers are often male, and their own gender beliefs and expectations are likely to influence the design and testing of models. Consequently, this can result in the development of models that perpetuate gender stereotypes and overlook gender-related concerns in the design and evaluation of AI systems. The need for diverse and inclusive AI development by the programming community is gaining popularity (**Eichenberger, 2022**).

Furthermore, AI models can also contribute to gender discrimination by making automated decisions. As an example, an AI model trained on data containing gender discrimination may make discriminatory decisions. An AI model used in hiring could discriminate against women by incorporating stereotypical gender characteristics such as leadership ability.

Gender discrimination in AI can also manifest in how AI products are marketed and promoted. For example, virtual assistants with female personalities are often designed to display subservient and agreeable traits, while virtual assistants with masculine personalities are often designed to exhibit bossy and domineering characteristics (**Sainz; Arroyo; Castaño, 2020; Eubanks, 2018**). These gender stereotypes in the personality of virtual assistants can contribute to the perpetuation of gender inequality in society.

To address these issues, a deeper understanding of gender stereotypes and their impact on AI is essential. This includes analysing the data used for training, designing and evaluating AI models, and how AI products are commercialised and promoted. Incorporating a variety of perspectives and voices in the design and evaluation of AI, including women and other groups who may experience discrimination is also essential (**Bolukbasi et al., 2016**).

Gender stereotypes are a significant problem in the field of AI, evident in data usage and evaluation, AI model design and evaluation, as well as the commercialisation and promotion of AI products. Although there are no regulations and AI can reach different results given the same order (**Rassin; Ravfogel; Goldberg, 2022**), understanding gender stereotypes

and their impact on AI is crucial to addressing these issues and fostering a just and equal society. Reducing gender, racial, social, and similar gaps, of which the developers of systems that create synthetic images are aware (*OpenAI*, 2022b), is essential for the advancement of computational techniques and tools that use antagonistic generative networks.

3. Methodology

3.1. Objectives and hypotheses

This study has the following objective:

O₁ Examine the images generated by *DALL-E 2* for potential gender, age, or race biases, so as to ascertain whether the AI produces stereotyped images within the workplace and a professional context. This will allow us to analyse whether certain professions or work environments are more susceptible to stereotyping by AI.

For this, the following initial hypotheses are taken:

H₁ The images generated by *DALL-E 2* are biased by gender, age or race.

H₂ The images generated by *DALL-E 2* replicate stereotypes in the workplace.

3.2. Sample

We used stratified probabilistic sampling. The study delimits the segments from 37 professions or prompts, replicating the approach used by **Farago, Eggum-Wilkens & Zhang** (2020) in their study on gender stereotypes in the workplace (Figure 1). The professions selected by the study by **Farago, Eggum-Wilkens & Zhang** (2020) were compiled from a well-established line of research on stereotypes in the workplace, including studies such as those by **Archer** (1984), **Gottfredson** (1981), **Liben, Bigler** and **Krogh** (2001), and **Teig & Susskind** (2008), which identified professions exhibiting notable gender bias. We replicated the study by **Farago, Eggum-Wilkens & Zhang** (2020) because it summarised the most stereotyped professions detected in previous studies.

The study involves two coders who manually input different professions into the image generator. Thirty-seven professions were entered into the application with neutral gender titles, to ensure no starting biases. They were: *secretary; hotel manager; tailor; singer; teacher secondary; teacher primary; maid, nurse; TV/movie actor/actress; teacher for university; shop owner; writer; lawyer; newspaper reporter; professional athlete, doctor; government worker, street vendor; cook/chef in a restaurant; accountant; scientist; farmer; factory worker; pastor or religious leader; politician; computer specialist; banker; police; barber in a hair salon; engineer; soldier; construction worker; mechanic; airplane pilot; truck driver; taxi driver; and carpenter.*

Two rounds were carried out during the same week by each of the coders. *DALL-E 2* generated nine images for each query, such that the final sample totalled 666 images (37x9x2). To acquire the images, the coders accessed the *DALL-E mini* by *craiyon.com* webpage:

<https://huggingface.co/spaces/dalle-mini/dalle-mini>

The chosen professions or prompts were manually entered in English (chosen because it is a gender neutral language) into the image generator of the *DALL-E mini* by *craiyon.com* webpage. After clicking “run”, the system automatically generated nine images corresponding to each profession. Some of these images were unrealistic representations or images, but they did represent the professions that had been entered into the generator.

This sample size is noteworthy: the creators of the *DALL-E 2* application reported that the tool had been used by 1.5 million users to generate 60 million images as of the current date of this article (*OpenAI*, 2022). In fact, *DALL-E* contains more than 12 billion parameters and is trained on a dataset of 250 million image-text pairs (**Zhou et al.**, 2021).

Given a population size of 60 million, with a confidence level of 99% and a margin of error of 5%, the representative sample is 666 units.

The results were then transcribed into an Excel spreadsheet where each image was evaluated according to a 3-level Likert scale (1. Not stereotypical; 2. Moderately stereotypical; 3. Strongly stereotypical).

4. Results

Our study results found there was gender stereotyping in the images of professional fields generated by Artificial Intelligence.

4.1. Professions and stereotyping in images created by *DALL-E 2*

Our study showed that the images of professions generated by AI were fully stereotypical (Figure 1). The following images of professions only included representation of women: *nurse; maid; teacher primary; teacher secondary; singer; seamstress/tailor; hotel manager* and *secretary*.

Conversely, the following images of professions only included representation of men: *carpenter; taxi driver; truck driver; airplane pilot; mechanic; construction worker; soldier; engineer; barber in hair salon; police; banker; computer specialist; politician* and *pastor or religious leader*.

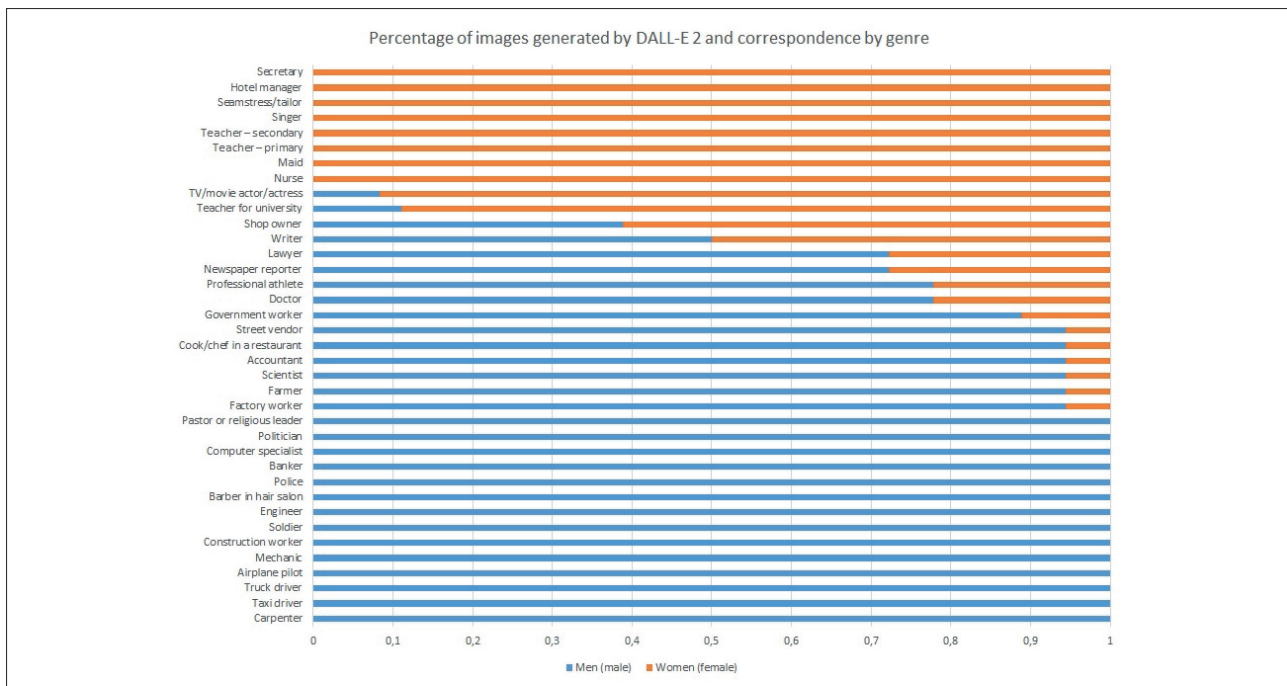


Figure 1. Images generated by *DALL-E 2* and correspondence by gender.

Concerning the images generated by *DALL-E 2*, as indicated by *OpenAI's* creators, the inclusion of a comprehensive set of terms is necessary to specify the content of the requested image and achieve a higher degree of realism in the generated outputs. Since the coders only introduced the term referring to the profession, the resulting images exhibited distorted faces and limbs; *DALL-E 2* would need more information to generate high-quality, well-defined images (Millán, 2022; Borji, 2022). It should also be noted that the searches were carried out between October and November 2022, and the tool has made significant improvements in the quality of the results it provides since then. Additionally, until recently, the system was susceptible to the “uncanny valley effect”, where the generated images exhibited traits that appeared almost lifelike but still fell short of true realism (Franganillo, 2022).

The results show that *DALL-E 2* represents fully stereotyped professions in 59.4% of cases. Of the professions that were completely stereotypical, 21.6% related to the female sex and 37.8% to the male sex.

4.2. Qualitative review of images, stereotype professions and biases

4.2.1. Technical, industry and primary sector professions

The present study found that professions within the technical, industrial, or construction-related sectors (e.g., construction worker, carpenter, engineer, factory worker, mechanic, computer technician) were not only significantly stereotyped and predominantly represented by men, but also commonly depicted by young individuals wearing similar attire, such as helmets, vests, and plaid shirts. Moreover, the images often featured individuals assuming similar postures or using the similar work-related items, such as carrying wood in the case of carpenters or engineers using paper documents.

Another profession that *DALL-E 2* portrayed as stereotypical was farmer, 94% of its images were men. The men in these synthetic representations were elderly, with the same posture and very often wearing green. They were all depicted in a field, with a tool or stick as if they were working; 20% of the images generated were drawings.

4.2.2. Transport

In professions related to transport, such as taxi driver, truck driver, or airplane pilot, the depicted professionals consistently appear seated inside the respective vehicles, leaning out of the window, and assuming nearly identical postures. They are also very stereotypical depictions of professions: 100% of the people represented were men, middle-aged and Western, unlike the industries above, which showed younger workers. In the example of the taxi driver and airplane pilot, they are shown in suits, while the truck drivers are shown in shirts and more casual clothing.

4.2.3. Education

In certain professions, women were depicted in a highly stereotypical manner, particularly those related to education: for both primary and secondary education, the images consistently portrayed only women. University lecturers was also depicted stereotypically, 88% of images showed women. Additionally, the images of professionals depicted at all educational levels predominantly showed young individuals. When representing teachers, the majority were

“ AI systems can perpetuate existing gender biases in society ”



Figure 2. Images generated by *DALL-E 2* of primary school teachers (left) and university lecturers (right).

depicted as having blonde or brown, medium-length hair, and dressed in a white or light-coloured blouse and jacket. In the case of men, they were wearing a suit. In the case of primary and secondary levels, the images showed teachers among the students in a classroom setting with desks distributed throughout the room. *DALL-E 2* depicting university lecturers at a blackboard, and approximately 88% were women and 11% men (Figure 2).

4.2.4. Services and entertainment

Other professions that are also highly stereotyped are those related to the service, textile, entertainment or film sectors. When we requested *DALL-E 2* generate images of a maid, tailor or singer, they were of women and very similar to each other. The women depicted were young, adding to the fully stereotype of these professions. Here it is worth noting that the professionals are all westerners.

In the case of the profession of maid, their clothing was typical of service personnel with a black dress, white apron, and feather duster in hand; furthermore, they all had dark, short hair and a very similar posture. Conversely, the professions of tailor and singer were represented as women with long or medium length blond or brown hair, and also in very similar postures. For the professions of seamstress/tailor, the women depicted are all seated next to a sewing machine and with a tape measure around their necks. Additionally, all the images displayed a notable predominance of pink. The images representing singers are also all very similar: all the women are depicted in a standing position, holding a microphone, and wearing similar casual clothing, along with similar hairstyles.

For professions associated with the services, commerce, and catering sectors, such as barbers and cooks, the generated images included strong gender stereotypes: 100% of the depicted individuals were male, including the clients in the barbershop. For street vendors, 94% of the images generated depicted men, while for shop owner 38% were men and 62% were women. The male protagonists for the images of barber and chef were all middle-aged, young and Western, and had the same posture and appearance. Street vendors looked Indian or Asian. In the above case, they all had fruit and vegetable carts and the same clothes, and in the case of the chefs, they were all wearing white, a chef's hat and in a kitchen.

The profession of hotel manager was also represented stereotypically, and all images were women. The women are all young, had their hair tied up, were wearing a dark uniform and the image was always at a hotel reception table.

As the last example in this segment, we note that the images generated by *DALL-E 2* for the profession of actor/actress vary depending on how the term was entered. In the first round, the professions actor and actress were entered using the forward slash (actor/actress), and 91% of the images generated depicted women and all returned with a stereotypical image. They were predominantly Westerners and some Latinos, characterised by long, brown or brunette hair and brightly coloured clothing that showed their shoulders. In addition, the two images with men, depicted them with the same haircut and a beard or moustache.

“ The judicious application of AI has the potential to transcend these limitations, contributing to a more equitable and inclusive society ”



Figure 3. Images generated by *DALL-E 2* for doctors (left) and nurses (right)

In the second round we entered the terms separated by a space and the images generated were different. However, they were still stereotyped: all the professionals were women, albeit this time they were Indian, had dark skin and were wearing Indian clothes; more of their body was covered than in the previous images. They seemed to represent *Bollywood* actresses.

4.2.5. Health and science

Certain differences were observed in the images generated for the health and science fields. When including the term “doctor” in the search, the results are moderately stereotyped: 77% of the images depicted men, while 23% represented women. However, when introducing the term “nurse”, the results generated were 100% stereotypical, as only women appeared in the images. Additionally, all the professionals depicted were young, Western, held the same posture (doctors with their arms crossed and nurses with charts in their hands), and white predominated both in the background and clothing. Some professionals were shown wearing masks (Figure 3).

The profession of scientist also gave stereotyped results: 94% were of young men, who had dark hair, the same posture, and were analysing samples with gloves and glasses. As for clothing, they all wore a lab-coat with a suit underneath. In addition, as was the case with health sciences professionals, the colour white predominated.

4.2.6. Politics, economy and information technology

The images generated by Artificial Intelligence depicted politicians in a very stereotypical manner: 100% were men, elderly or middle-aged and Western. It should also be noted that everyone wore a suit and tie and the more senior the position, the darker the colour, whereas mid-level employees wore lighter colours.

The images generated by *DALL-E 2* for lawyers were also stereotypical: 72% were men, all of them elderly; they wore dark gowns and most of them had a book in their hands. Also, all images depicted Westerners.

Office-related professions such as secretary, accountant or banker were also highly stereotyped. When entering the term secretary, the generated images were all female, young, Western, brunette and had long hair. Also, they were all dressed in suits, sitting at a table, with a computer. The same goes for the term banker, and the images generated for that profession depicted middle-aged, Western men, in suits, using a calculator or sorting through documents. Again, 94% of the images generated for this profession showed men.

The only change in this pattern was seen for professions associated with writing (such as author or journalist); for these, the results were only moderately stereotypical and the images depicted both women and men. In the case of writer, the results depicted 50% men and 50% women. But when journalist was entered, the images were 72% men and 28% women. The most interesting result to note of the images generated for these professions is that there is always a typewriter for the writer and a newspaper for the journalist (Figure 4). Again, middle-aged or young Westerners were depicted for those professions and they were wearing very similar clothing: a dark jacket and a suit.

“ The AI response is implicit in the question that the user asks ”



Figure 4. Images generated by DALL-E 2 for the profession of writer (left) and journalist (right).

4.2.7. Security, religion and sports

Regarding professions from the security sector, all the images generated by Artificial Intelligence were highly stereotyped. All the figures represented were men and in the case of the soldiers, seven of the nine images were drawings. Both police officers and soldiers were depicted in uniform, all young, Western and held the same stance. Police officers were all shown to be on the street and soldiers all had weapons in their hand.

Other neutral professions were subsequently reviewed, such as pastor or religious leader and athlete. Pastor was also depicted stereotypically, all the images were of elderly men, and the majority were Western except for two images where the men appeared to be black. The images generated for professional athletes were also stereotypical: 77% were men. They were all young, dressed in mainly red and the majority were white and Western.

4.3. Comparison of stereotyping between AI and humans

Gender stereotypes in artificial intelligence (AI) can have significant impacts in various areas, as noted above. The conclusion of comparing the results obtained by DALL-E 2 with previous studies that involved human opinion found that AI exhibited a higher degree of gender stereotyping in the workplace. In particular, an analysis by **Farago, Eggum-Wilkens & Zhang (2021)** found that 35% of the professions evaluated were very stereotypical, while AI-generated images reached an alarming 59.4%.

There were stereotypes of professions humans and AI concurred with, for example depicting males for: carpenter, taxi driver, truck driver, airplane pilot, mechanic, construction worker, soldier, engineer, barber; and females for nurse and maid.

The biases present in the data used to train the AI models probably reflect existing gender biases and imbalances in society. If historical data sets contain inequities or reflect gender stereotypes, AI is likely to learn and reproduce these patterns during its training. Another aspect to take into account is the ongoing interaction and reciprocal influence between society and technology. AI-generated representations of professions will amplify existing societal stereotypes, leading to a feedback loop that reinforces these biases.

Certain professions exhibit strong stereotypes in AI-generated representations that are not reflected in human perceptions. In the case of men, these were: *police; banker; computer specialist; politician and pastor or religious leader*. And in the case of women, these were: *teacher - primary; teacher - secondary; singer; seamstress/tailor; hotel manager and secretary*.

These stereotypical images were a result of the AI model's interpretation and representation of the data. AI algorithms can use certain attributes or characteristics present in the data to assign labels or associate certain jobs with a particular gender, even if such associations lack a strong empirical basis. If gender-biased databases are used to teach the AI model, machine learning will also show stereotypes.

“ The establishment and strengthening of a diverse and inclusive development community plays a crucial role in advancing towards a more equitable and unbiased AI system ”

5. Discussion and conclusions

The results provided by *DALL-E 2* for neutral professions showed a high level of stereotyping, with 22 out of 37 searches consistently producing images of the same gender. Of the professions that were completely stereotypical, 21.6% related to the female sex and 37.8% to the male sex.

That is what happened for technical and scientific professions, and those related to construction and transport. We found AI associated women with domestic workers, dressmakers and those professions in which appearance is important, such as actors or singers, the AI model further depicted them as young, Western, and blonde. It is important to note the high representation of women in the education and medical sectors, particularly in nursing.

Also noteworthy is that *DALL-E 2* generates synthetic images of middle-aged or elderly men for professions associated with higher responsibility or status, such as politicians, those related to finance and religion. There is also a prevalence of Westerners evident in the generated images.

When compared with earlier studies involving adolescents, it is evident that *DALL-E 2* exhibits more significant gender stereotyping in professional contexts. In contrast to previous human-based studies that detect strong gender stereotyping in 35% of professions, Artificial Intelligence exhibits full stereotyping in 59.4% of cases.

In summary, this study identifies significant gender biases in the workplace evident in the images generated by Artificial Intelligence.

AI-based tools are quickly becoming very popular and hold promising potential to participate in and influence social relationships in the short term. This is why identifying, categorising and eliminating these biases that can impact our decision-making, and the way we perceive and interact with reality, is so important.

Artificial Intelligence reflects our common feelings, virtues and defects. By reflecting on our own biases and actively learning from the past, we can aspire to develop AI technologies that are genuinely inclusive and equitable.

Two significant challenges emerge from this study concerning ethics and efficiency that require resolution and thoughtful consideration. Firstly, the issue of user bias. AI tends to reinforce existing biases by echoing the user's query and providing answers that align with their preconceived beliefs, creating an echo chamber effect. In this sense, the AI response is implicit in the question that the user asks. Hence, finding responses that transcend the issuer's cosmogony or their particular way of comprehending reality becomes complicated. We have called the impossibility of finding answers that go beyond the user's existing knowledge and understanding "the other side of the mirror". Secondly, in a hypothetical scenario where users can conduct an unbiased inquiry and cross to the other side of the mirror, they would encounter a sea of knowledge that is inherently biased due to the influence of the technology developer's "other side of the mirror". If "every technology is an ideology" (Postman, 1991, p. 165), AI cannot be separated from the ideology of its creators.

Indeed, the establishment and strengthening of a diverse and inclusive development community plays a crucial role in advancing towards a more equitable and unbiased AI system. The replication of these values in technologies can only be achieved with an inclusive development community.

Our findings emphasise the significance of investigating both AI stereotypes and human stereotypes. Stereotypes, being products of society and reflecting deep-seated biases, can be amplified and perpetuated by AI through its capacity to learn from extensive datasets. Addressing this issue requires a two-fold approach: promoting diversity and equality in the training data used for AI, and fostering greater awareness and reflection among humans about ingrained stereotypes that can impact technology development and use. Only then can we move towards AI systems that are fair, free of bias, and contribute to promoting equal opportunities and inclusion across all aspects of our society.

6. References

- Agudo, Ujué; Liberal, Karlos G.** (2020). "El automágico traje del emperador". *Medium.com*, 9 septiembre. <https://medium.com/bikolabs/el-automagico-traje-del-emperador-c2a0bbf6187b>
- Archer, Cynthia J.** (1984). "Children's attitudes toward sex-role division in adult occupational roles". *Sex roles*, v. 10. <https://doi.org/10.1007/BF00287742>
- Belhadi, Amine; Kamble, Sachin; Fosso-Wamba, Samuel; Queiroz, Maciel M.** (2022). "Building supply-chain resilience: an artificial intelligence-based technique and decision-making framework". *International journal of production research*, v. 60, n. 14, pp. 4487-4507. <https://doi.org/10.1080/00207543.2021.1950935>
- Bolukbasi, Tolga; Chang, Kai-Wie; Zou, James; Saligrama, Venkatesh; Kalai, Adam** (2016). "Man is to computer programmer as woman is to homemaker? Debiasing word embeddings". In: *NIPS'16: Proceedings of the 30th international conference on neural information processing systems*, pp. 4356-4364. <https://doi.org/10.48550/arXiv.1607.06520>

- Borji, Ali** (2022). *Generated faces in the wild: quantitative comparison of stable diffusion, midjourney and DALL-E 2*. Quintic AI, San Francisco, CA.
<https://arxiv.org/pdf/2210.00586.pdf>
- Brown, Tom B.; Mann, Benjamin; Ryder, Nick; Subbiah, Melanie; Kaplan, Jared; Dhariwal, Prafulla; Neelakantan, Arvind; Shyam, Pranav; Sastry, Girish; Askell, Amanda; Agarwal, Sandhini; Herbert-Voss, Ariel; Krueger, Gretchen; Henighan, Tom; Child, Rewon; Ramesh, Aditya; Ziegler, Daniel M.; Wu, Jeffrey; Winter, Clemens; Hesse, Christopher; Chen, Mark; Sigler, Eric; Litwin, Mateusz; Gray, Scott; Chess, Benjamin; Clark, Jack; Berner, Christopher; McCandlish, Sam; Radford, Alec; Sutskever, Ilya; Amodei, Dario** (2020). "Language models are few-shot learners". *Advances in neural information processing systems*, v. 33, pp. 1877-1901.
<https://doi.org/10.48550/arXiv.2005.14165>
- Buolamwini, Joy; Gebru, Timnit** (2018). "Gender shades: intersectional accuracy disparities in commercial gender classification". *Proceedings of machine learning research*, v. 81.
<https://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf>
- Caliskan, Aylin; Bryson, Joanna J.; Narayanan, Arvind** (2017). "Semantics derived automatically from language corpora contain human-like biases". *Science*, v. 356, n. 6334, pp.183-186.
<https://doi.org/10.1126/science.aal4230>
- Cortina-Orts, Adela** (2019). "Ética de la inteligencia artificial". *Anales de la Real Academia de Ciencias Morales y Políticas*, pp. 379-394. Ministerio de Justicia.
https://www.boe.es/biblioteca_juridica/anuarios_derecho/articulo.php?id=ANU-M-2019-10037900394
- Crawford, Kate** (2021). *The atlas of AI: power, politics, and the planetary costs of artificial intelligence*. Yale University Press. ISBN: 978 0 300252392
<https://doi.org/10.2307/j.ctv1ghv45t>
- Criado-Pérez, Caroline** (2020). *La mujer invisible. Descubre cómo los datos configuran un mundo hecho por y para los hombres*. Barcelona: Seix Barral. ISBN: 978 84 32236136
- DALL-E 2** (2021). OpenAI.
<https://openai.com/dall-e-2>
- De-Carvalho, André-Carlos-Ponce-de-Leon-Ferreira** (2021). Inteligência artificial: riscos, benefícios e uso responsável. *Estudos avançados*, v. 35, 101.
<https://doi.org/10.1590/s0103-4014.2021.35101.003>
- D'Ignazio, Catherine; Klein, Lauren F.** (2020). *Data feminism*. Cambridge: MIT Press. ISBN: 978 0 262547185
- Eichenberger, Livia** (2022). "DALL-E 2: Why discrimination in AI development cannot be ignored". *Statworx blog post*, 28 June.
<https://www.statworx.com/en/content-hub/blog/dalle-2-open-ai>
- Estupiñán-Ricardo, Jesús; Leyva-Vázquez, Maikel-Yelandi; Peñafiel-Palacios, Álex-Javier; El-Asaffiri-Ojeda, Yusef** (2021). "Inteligencia artificial y propiedad intelectual". *Universidad y sociedad*, v. 13, n. S3, pp. 362-368.
<https://rus.ucf.edu.cu/index.php/rus/article/view/2490>
- Eubanks, Virginia** (2018). *Automating inequality: how high-tech tools profile, police, and punish the poor*. New York: St. Martin's Press. ISBN: 978 1 250074317
- Farago, Flora; Eggum-Wilkens, Natalie D.; Zhang, Linlin** (2021). "Ugandan adolescents' gender stereotype knowledge about jobs". *Youth & society*, v. 53, n. 5, pp. 723-744.
<https://doi.org/10.1177/0044118X19887075>
- Francescutti, Pablo** (2018). *La visibilidad de las científicas españolas*. Fundación Dr. Antoni Esteve, Grupo de estudios avanzados de comunicación, Barcelona.
<https://www.raco.cat/index.php/QuadernsFDAE/issue/download/30066/439>
- Franganillo, Jorge** (2022). "Contenido generado por inteligencia artificial: oportunidades y amenazas". *Anuario ThinkEPI*, v. 16, e16a24.
<https://doi.org/10.3145/thinkepi.2022.e16a24>
- Gamir-Ríos, José; Tarullo, Raquel** (2022). "Predominio de las cheapfakes en redes sociales. Complejidad técnica y funciones textuales de la desinformación desmentida en Argentina durante 2020". *adComunica*, v. 23, pp. 97-118.
<https://doi.org/10.6035/adcomunica.6299>
- García-Ull, Francisco-José** (2021). "Deepfakes: el próximo reto en la detección de noticias falsas". *Anàlisi*, n. 64, pp. 103-120.
<https://doi.org/10.5565/rev/analisi.3378>

- Goodfellow, Ian J.; Pouget-Abadie, Jean; Mirza, Mehdi; Xu, Bing; Warde-Farley, David; Ozair, Sherjil; Courville, Aaron; Bengio, Yoshua** (2014). "Generative adversarial networks. Advances in neural information processing systems". *Communications of the ACM*, v. 63, pp. 139-164.
<https://doi.org/10.48550/arXiv.1406.2661>
- Gottfredson, Linda S.** (1981). "Circumscription and compromise: A developmental theory of occupational aspirations". *Journal of counseling psychology*, v. 28, n. 6, pp. 545-579.
<https://doi.org/10.1037/0022-0167.28.6.545>
- Laino, María-Elena; Cancian, Pierandrea; Salvatore-Politi, Letterio; Della-Porta, Matteo-Giovanni; Saba, Luca; Savevski, Victor** (2022). "Generative adversarial networks in brain imaging: A narrative review". *Journal of imaging*, v. 8, n. 4, 83.
<https://doi.org/10.3390/jimaging8040083>
- Leavy, Susan** (2018). "Gender bias in artificial intelligence: the need for diversity and gender theory in machine learning". In: *Proceedings of the 1st international workshop on gender equality in software engineering*, pp. 14-16.
<https://doi.org/10.1145/3195570.3195580>
- Leavy, Susan; Meaney, Gerardine; Wade, Karen; Greene, Derek** (2020). "Mitigating gender bias in machine learning data sets". In: *Bias2020 workshop: Bias and social aspects in search and recommendation*.
https://doi.org/10.1007/978-3-030-52485-2_2
<https://doi.org/10.48550/arXiv.2005.06898>
- Liben, Lynn S.; Bigler, Rebecca S.; Krogh, Holleen R.** (2001). "Pink and blue collar jobs: children's judgments of job status and job aspirations in relation to sex of worker". *Journal of experimental child psychology*, v. 79, n. 4, pp. 346-363.
<https://doi.org/10.1006/jecp.2000.2611>
- Loftus, Tyler J.; Tighe, Patrick J.; Filiberto, Amanda C.; Efron, Philip A.; Brakenridge, Scott C.; Mohr, Alicia M.; Rashidi, Parisa; Upchurch, Gilbert R.; Bihorac, Azra** (2020). "Artificial intelligence and surgical decision-making". *JAMA surgery*, v. 155, n. 2, pp. 148-158.
<https://doi.org/10.1001/jamasurg.2019.4917>
- Manassero, Antonia; Vázquez, Ángel** (2003). "Las mujeres científicas: un grupo invisible en los libros de texto". *Revista investigación en la escuela*, v. 50, pp. 31-45.
<https://revistascientificas.us.es/index.php/IE/article/view/7582>
- Millán, Víctor** (2022). "DALL-E 2: ¿cómo funciona y qué supone? La IA que crea imágenes de la nada y es, simplemente, perfecta y aterradora". *Hipertextual*, 29 mayo.
<https://hipertextual.com/2022/05/dall-e-2>
- Nica, Elvira; Sabie, Oana-Matilda; Mascu, Simona; Luțan-Petre, Anca-Georgeta** (2022). "Artificial intelligence decision-making in shopping patterns: consumer values, cognition, and attitudes". *Economics, management and financial markets*, v. 17, n. 1, pp. 31-43.
<https://doi.org/10.22381/emfm17120222>
- O'Neil, Cathy** (2018). *Armas de destrucción matemática: cómo el big data aumenta la desigualdad y amenaza la democracia*. Capitán Swing Libros. ISBN: 978 84 947408 4 8
- OpenAI** (2022a). "DALL-E now available without waitlist". *Openai*, September 28.
<https://openai.com/blog/dall-e-now-available-without-waitlist>
- OpenAI** (2022b). "Reducing bias and improving safety in DALL-E 2". *OpenAI*, July 18.
<https://openai.com/blog/reducing-bias-and-improving-safety-in-dall-e-2>
- Ortiz-de-Zárate-Alcarazo, Lucía** (2023). "Sesgos de género en la inteligencia artificial". *Revista de occidente*, v. 1, n. 502.
<https://dialnet.unirioja.es/servlet/articulo?codigo=8853265>
- Pérez-Gómez, Miguel-Ángel; Echazarreta-Soler, Carmen; Audebert-Sánchez, Meritxell; Sánchez-Miret, Cristina** (2020). "El ciberacoso como elemento articulador de las nuevas violencias digitales: métodos y contextos". *Communication papers. Media literacy and gender studies*, v. 9, n. 18.
https://doi.org/10.33115/udg_bib/cp.v9i18.22470
- Porayska-Pomsta, Kaska; Rajendran, Gnanathusharan** (2019). "Accountability in human and artificial intelligence decision-making as the basis for diversity and educational inclusion". In: Knox, Jeremy; Wang, Yuchen; Gallagher, Michael. *Artificial intelligence and inclusive education: speculative futures and emerging practices*. Springer, pp. 39-59.
https://doi.org/10.1007/978-981-13-8161-4_3
- Postman, Neil** (1991). *Divertirse hasta morir, el discurso público en la era del show business*. Barcelona: Ediciones la Tempestad. ISBN: 978 84 79480462
- Quirós-Fons, Antonio; García-Ull, Francisco-José** (2022). *La inteligencia artificial como herramienta de la desinforma-*

ción: deepfakes y regulación europea. Los derechos humanos en la inteligencia artificial: su integración en los ODS de la Agenda 2030. Thomson Reuters Aranzadi, pp. 537-556. ISBN: 978 84 1124 557 9

Rassin, Royi; Ravfogel, Shauli; Goldberg, Yoav (2022). "DALL-E 2 is seeing double: flaws in word-to-concept mapping in text2image models".

<https://doi.org/10.48550/arXiv.2210.10606>

Sainz, Milagros; Arroyo, Lidia; Castaño, Cecilia (2020). *Mujeres y digitalización: de las brechas a los algoritmos.* Instituto de la Mujer y para la Igualdad de Oportunidades.

https://www.inmujeres.gob.es/disenov/novedades/M_MUJERES_Y_DIGITALIZACION_DE_LAS_BRECHAS_A_LOS_ALGORITMOS_04.pdf

Sourdin, Tania (2018). "Judge v Robot? Artificial intelligence and judicial decision-making". *UNSW law journal*, v. 41, n. 4, pp. 1114-1133.

<https://www.unswlawjournal.unsw.edu.au/wp-content/uploads/2018/12/Sourdin.pdf>

Teig, Stacey; Susskind, Joshua E. (2008). "Truck driver or nurse? The impact of gender roles and occupational status on children's occupational preferences". *Sex roles*, v. 58, pp. 848-863.

<https://doi.org/10.1007/s11199-008-9410-x>

Traylor, Jake (2022). "No quick fix: how OpenAI's DALL-E 2 illustrated the challenges of bias in AI". *NBC news*, July 27.

<https://www.nbcnews.com/tech/tech-news/no-quick-fix-openais-dalle-2-illustrated-challenges-bias-ai-rcna39918>

Véliz, Carissa (2021). *Privacidad es poder: datos, vigilancia y libertad en la era digital.* Debate. ISBN: 978 84 18056680

Vincent, James (2020). "OpenAI's latest breakthrough is astonishingly powerful, but still fighting its flaws". *The verge tech*, July 30.

<https://www.theverge.com/21346343/gpt-3-explainer-openai-examples-errors-agi-potential>

Wang, Tianlu; Zhao, Jieyu; Yatskar, Mark; Chang, Kai-Wei; Ordóñez, Vicente (2019). "Balanced datasets are not enough: estimating and mitigating gender bias in deep image representations". In: *International conference on computer vision, ICCV 2019.*

<https://doi.org/10.48550/arXiv.1811.08489>

Zhou, Yufan; Zhang, Ruiyi; Chen, Changyou; Li, Chunyuan; Tensmeyer, Chris; Yu, Tong; Gu, Jiuxiang; Xu, Jinhui; Sun, Tong (2021). "Towards language-free training for text-to-image generation".

<https://arxiv.org/pdf/2111.13792v3.pdf>

Give visibility to your work by depositing it in e-LIS, the largest international repository on library & information science, and communication

<http://eprints.rclis.org>

Who interacts with communicative AI and what attitudes are displayed toward it? Sociodemographic, personality, and *futurephobia* variables

Héctor Centeno-Martín; Samuel Toledano-Buendía; Alberto Ardèvol-Abreu

Nota: Este artículo se puede leer en español en:
<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87386>

Recommended citation:

Centeno-Martín, Héctor; Toledano-Buendía, Samuel; Ardèvol-Abreu, Alberto (2023). "Who interacts with communicative AI and what attitudes are displayed toward it? Sociodemographic, personality and *futurephobia* variables". *Profesional de la información*, v. 32, n. 5, e320502.

<https://doi.org/10.3145/epi.2023.sep.02>

Manuscript received on May 31st 2023
Accepted on July 24th 2023



Héctor Centeno-Martín

<https://orcid.org/0000-0002-6708-139X>

Universidad de La Laguna
Departamento de Ciencias de la
Comunicación y Trabajo Social
Camino La Hornera, 37
38200 La Laguna (Tenerife), Spain
hcenteno@ull.edu.es



Samuel Toledano-Buendía

<https://orcid.org/0000-0003-0430-8347>

Universidad de La Laguna
Departamento de Ciencias de la
Comunicación y Trabajo Social
Lab. de Invest. sobre Medios y sus Efectos
Camino La Hornera, 37
38200 La Laguna (Tenerife), Spain
stoledano@ull.edu.es



Alberto Ardèvol-Abreu ✉

<https://orcid.org/0000-0001-8722-5226>

Universidad de La Laguna
Instituto Universitario de Neurociencia
Lab. de Invest. sobre Medios y sus Efectos
Camino La Hornera, 37
38200 La Laguna (Tenerife), Spain
aardevol@ull.edu.es

Abstract

Communicative applications of artificial intelligence (AI) have burst into the lives of millions of people through products such as *ChatGPT*, *Midjourney*, or *Replika*, with the ability to generate texts and images and even understand natural language and socialize with humans. This study addresses attitudes toward communicative AI, as well as the sociodemographic and personality factors that are related to its positive perception and usage. For this purpose, we use data from a survey conducted at the end of February 2023 among adult residents in Spain ($N = 821$). Our results suggest that there is a slight predominance of negative attitudes toward communicative AI, which include cognitive and affective elements such as the perception of dehumanizing aspects and job-related threats, caution, or fear. However, a relevant portion of our respondents is not familiar with communicative AI or generalizes their answers toward other applications of AI. Individuals with higher educational level, greater openness to experience, and lower level of *futurephobia* show more positive attitudes toward these technologies. In terms of adoption, men, younger people, and those with more favorable attitudes toward AI are more likely to decide to use it, while those scoring higher in agreeableness tend to use it less frequently once they have decided to use it. We discuss these results under the premise that early adopters may play a central role in the attitudes of late majority and laggards, as well as in the future development of communicative AI and its regulatory framework.

Keywords

Artificial intelligence; AI; Communicative AI; Chatbots; Attitudes toward communicative AI; Use of communicative AI; Big five; Openness to experience; *Futurephobia*; Innovators; Early adopters; Technology; Surveys.



Funding

The first author is funded by the “Margarita Salas” fellowship from the European Union – *NextGenerationEU*.

This study has been supported by the *Universidad de La Laguna*, and the *Spanish Ministry of Science, Innovation, and Universities* (budget line 2022/0000587).

Conflict of interest

There is no conflict of interest in relation to this article.

1. Introduction

Advances in artificial intelligence (AI) have brought about a context of technological disruption in the field of communication. Some AI-based applications are not just tools *through which* people communicate and interact, but rather social actors *with whom* people communicate. AI chatbots have garnered significant social and media attention, especially since the launch of *ChatGPT* by the *OpenAI* research lab in late 2022. This tool reached 100 million users within just two weeks, compared with *TikTok*'s nine months, *Instagram*'s 26 months, *Facebook*'s 54 months, or *Twitter*'s 65 months to reach the same milestone (Pastor, 2023). Image generation software based on prompts or instructions in natural language have also represented a major innovation and attracted interest from users worldwide, particularly following the beta release of *Midjourney* in March 2022 and the removal of the waiting list for *Dall-E* in September of the same year.

Although AI in general, and communicative AI in particular, have great potential to contribute to the well-being of individuals and societies, there are also dangers and concerns that delve into the realm of data protection, copyright issues, and malicious misuse for disinformation and manipulation purposes. In this context, it is important to understand how citizens perceive these tools, as well as the characteristics of early adopters who are embracing them. Perceptions and use mutually influence each other in a feedback loop, and, in turn, can shape the perception of non-users, the evolution of communicative AI tools themselves, and the development of their legal and regulatory framework.

However, there are few studies to date that analyze the impact of this technology on society, and most of them come from reports by private companies or governments, with limited references in the academic domain (Zhang, 2021). Additionally, nearly all these studies approach the phenomenon in a general manner or refer to earlier, non-communicative (as defined here) technologies, such as *Alexa*, *Twitter*, or *Spotify* (see, for instance, the pioneering study in Spain by Sánchez-Holgado; Arcila-Calderón; Blanco-Herrero, 2022).

In this context, the present study uses data from an original survey conducted in February 2023, involving a diverse sample of 821 residents in Spain. From the resulting dataset, we examine general attitudes toward communicative AI and investigate the sociodemographic and personality antecedents associated with the favorable perception and use of these technologies. As personality antecedents, we identify and evaluate the possible role of the big five personality factors and ‘futurephobia,’ conceptualized as a negative attitude toward the future grounded on the perception that what is to come is threatening and unavoidable.

2. The emergence of AI and its communicative applications

A social science approach to AI defines it as the capability of non-human machines or artificial entities to perform and solve tasks, communicate, interact, and behave logically in a manner akin to humans (Gil de Zúñiga; Goyanes; Durotoye, 2023). In short, this entails making decisions, executing tasks, and making predictions. Despite the current surge in these technologies suggesting a very recent advancement, the truth is that their development began decades ago. The term AI itself was coined in 1956 at the *Dartmouth Conference*. By that time, Turing’s seminal work (1950) on “Computing Machinery and Intelligence,” wherein he introduced the concept of the imitation game, had already been published. Turing’s contribution, which defines and characterizes the discipline of AI (Gunkel, 2012), was rooted in the challenge of discerning whether machines can think. In what is now known as the *Turing test*, he formulated a series of questions aimed at ascertaining whether the machine can respond in a way comparable to a human interlocutor and, as such, successfully simulate a human being.

Despite the elapsed time and the progress made in the field, a precise or widely accepted definition of AI is still lacking. Some criticisms are directed toward the very concept, suggesting that these technologies are neither intelligent nor artificial. This viewpoint is presented by Evgeny Morozov, a prominent figure in the critique of technological solutionism. His observation contends that, on the one hand, there is nothing artificial as the training of AI involves prior work by real human beings, from whose creative and professional endeavors it has drawn. On the other hand, he disputes the attribution of intelligence, characterizing AI as primarily predictive engines, engaged in the search for more or less complex patterns (Morozov, 2023). The terminological debate mirrors the concerns raised by Turing (1950, p. 433), who noted that the common use of the terms “machine” and “think” poses an obstacle to answering the question of whether machines can think.

The current advancement, which encapsulates the essence of the term, revolves around the potential for AI to transcend its dependence on human intelligence to define its own existence –namely, autonomous performance. This leads

us directly into what **Harari** (2014) defines as the “other life:” the computer programming field’s attempts to emulate genetic evolution:

“Many programmers dream of creating a program that could learn and evolve completely independently of its creator. In this case, the programmer would be a *primum mobile*, a first mover, but his creation would be free to evolve in directions neither its maker nor any other human could ever have envisaged” (**Harari**, 2014, p. 353).

3. Communicative aspects of AI

To better understand the communicative aspects of AI, we must again journey several decades, returning to Turing’s work and his consideration of communication as the critical factor in artificial intelligence. According to **Gunkel** (2012), the point of the argument is that intelligence cannot be directly observed and thus needs verification through behaviors akin to intelligence, such as communication, either in a general sense or through verbal conversation. This entails the assumption that communication is a product of intelligence and, therefore, whoever communicates must possess intelligence. It is not coincidental that, during the same timeframe, the boost to and reciprocity between communication and AI coincided, thanks to the contributions of linguistics, mathematics, and logic.

It has been from the year 2022 onward that the most important technological innovations based on AI incorporating (new) communicative aspects have emerged. To some extent, some of these technologies are no longer merely tools *through which* people communicate and interact, but rather social actors *with whom* people communicate. And such a perspective does not fit traditional theories that positioned technologies as mere channels (**Guzman; Lewis**, 2020, p. 73). This overarching paradigm shift can have profound implications for journalism, as asserted by **Broussard** and colleagues (2019): Even within this domain, the creation and dissemination of information may cease to be an exclusively human task, posing a challenge to the balance and interaction between human intelligence and artificial intelligence (see also **Chan-Olmsted**, 2019). Curricula of Communication schools should integrate education on AI within a training program encompassing theoretical components, technical or procedural competencies, and ethical considerations (**Lopezosa et al.**, 2023; see also **Pavlik**, 2023).

In November 2022, *OpenAI* launched *ChatGPT*, wrapping up a year in which they had already introduced the *Whisper* speech recognition software (in September) and *Dall-E 2* (in April). *ChatGPT* is a specialized chatbot designed to engage in natural language conversations with human users, while also generating coherent responses. Only two months after its release, its estimated number of global users surpassed 100 million (**Pastor**, 2023). *Dall-E*, too, operates on the GPT (Generative Pre-trained Transformer) language model framework and is trained on an extensive dataset of images collected from the internet. This image generation AI requires prompts or natural language inputs provided by its users, from which the software generates photographs, paintings, and emojis in diverse styles. In May 2023, *Microsoft* opened its AI-powered *Bing* search engine to the general public, after having formed a partnership with *OpenAI* a few months earlier. The revamped search engine, incorporating GPT4 and *Dall-E* capabilities, demonstrates its ability to interact with its users through text, photos, images, and graphics (**Mehdi**, 2023). In reaction to these developments, *Google* introduced its conversational bot, *Bard*, to 180 countries in May 2023. *Bard* relies on *LaMDA*, *Google’s* “big language model”.

Replika’s conversational technology has been in existence for as few years, functioning as an “AI companion who is eager to learn and would love to see the world through your eyes” (**Luka**, 2023). The San Francisco-based tech company *Luka* released *Replika* at the end of 2017. Presently, *Replika* operates on a freemium model and claims to have approximately two million users (**Tong**, 2023). *Replika* provides an interface that allows users to chat, exchange images, and connect through voice or video calls with a humanoid customizable avatar powered by AI. Tailored to the user’s needs and preferences, *Replika* can serve as an advisor, a friend, or even a romantic partner.

4. Social perception and attitudes toward AI

AI remains a relatively unfamiliar phenomenon to over a third of the world’s population, with rates of unawareness reaching nearly 60% in countries with high educational standards such as Italy and Japan (**Maslej et al.**, 2023). However, because communicative AI is relatively new and becoming more prominent in the media, people are still in the process of developing their perceptions and attitudes about this emerging technology. Recent studies grounded in framing theory suggest that exposure to technology-related news, science fiction films and literature, and personal conversations about technology predict attitudes toward AI and individual AI-related frames: AI as a means for social progress or as a “Pandora’s box” (see **Brewer et al.**, 2022).

According to the most recent report on AI for the year 2023 conducted by *Stanford University*, China stands out as the country with the most positive perception of this technology. Seventy-eight percent of respondents in this eastern country state that AI offers more benefits than drawbacks to society. In the United States, only 35% of the population believes that AI has a more positive than negative impact, whereas this figure rises to 53% in Spain (**Maslej et al.**, 2023). Globally, a report preceding the recent expansion of communicative AI suggests that people in Eastern countries exhibit more positive sentiment (65%) about AI than people in Western countries (35%) (*YouGov*, 2021).

The traditional attitude of citizens toward AI, both in Europe and the United States, has been cautious (*European Commission*, 2017; **Zhang; Dafoe**, 2020). In the United States, a significant portion of the population perceives positive as-

pects of AI use, such as its potential to make life easier and society better (31%) or to save time and enhance efficiency (13%). Concerns are also reported, such as job loss (19%), surveillance and privacy (16%), and lack of human connection (12%) (**Maslej et al.**, 2023). In the case of European countries, acceptance of AI and task automation varies, ranging from a minimum of 7% of respondents in France to a maximum of 25% in Spain (*YouGov*, 2021).

There are certain specific domains that have been the subject of studies on perceptions of AI, with the medical and occupational fields being prominent examples. In the medical field, a review of studies conducted by **Young** and colleagues (2021) suggests that patients show positive attitudes toward the role of AI in aspects such as diagnosis, treatment, and patient care, although they also express concerns and prefer collaborative efforts between AI and a human.

In the occupational sphere, a study by **Lichtenthaler** (2020) identifies clusters of individual characteristics associated with their attitudes toward AI. Those with negative attitudes toward AI tend to prioritize human relationships, show resistance to technological solutions, downplay efficiency and process optimization, and express fears of potential negative consequences of AI usage, such as job loss or data protection issues. On the other hand, those with positive perceptions of AI adopt a neutral stance toward human interactions, are open to the use of emerging technological solutions, emphasize rational choices and process optimization over empathy, relativize both positive and negative consequences of AI usage, and have a pragmatic approach to data privacy, as long as AI adds value to their work.

Currently, *ChatGPT* dominates over half (52%) of the social media conversation on the topic (**Maslej et al.**, 2023). The net sentiment of the social media community's discourse about AI-related tools (ranging from -100, completely negative, to +100, completely positive) is favorable. By the end of 2022, this net sentiment was +32 for *ChatGPT*, whereas *Dall-E* scored 11 points lower (+21) (**Maslej et al.**, 2023). Despite efforts to comprehend attitudes toward AI, further in-depth studies are needed regarding this phenomenon in general (**Zhang**, 2021) and communicative AI in particular. This is the reason why we formulate our first research question:

RQ1: What general attitudes do Spanish adults show toward communicative AI?

5. Attitudes and use of communicative AI: Sociodemographic antecedents

When a technological innovation enters the market, consumers do not exhibit the same attitudes toward it or adopt it at the same pace. Given that most individuals are risk averse, there is a natural tendency to delay the decision of technology adoption until more information is available. However, this tendency is not uniform among all individuals, and literature on innovation adoption has classified consumers based on their attitudes toward innovations and their varying tendency to adopt them quickly. In his now classic theory, **Rogers** (2003) distinguishes five distinct different categories of consumers based on their speed of adoption: innovators, early adopters, early majority, late majority, and laggards.

One of the most intriguing aspects of research based on this typology is that these user groups seem to differentiate from one another based on sociodemographic characteristics (**Dutton; Rogers; Jun**, 1987; **Laukkanen; Pasanen**, 2008; **Rogers**, 2003). **Rogers** (2003) himself points out that those who adopt innovations more promptly (early adopters) tend to have higher educational and economic level.

Being communicative AI a technological disruption, it is reasonable to expect that attitudes toward and use of AI may also be influenced by sociodemographic characteristics. Recent studies based on surveys across several regions worldwide have identified differences in attitudes toward AI based on gender, with men displaying more favorable attitudes (**Johnson; Tyson**, 2020; **Maslej et al.**, 2023). Similarly, a higher level of education also appears to be associated with more favorable attitudes toward AI, as well as higher income levels (**Johnson; Tyson**, 2020). Regarding age, the findings are more mixed. While **Johnson** and **Tyson** (2020) suggest that younger individuals hold a more positive perception of AI, the *YouGov* (2021) report does not observe differentiating patterns by age.

Given the scarcity of studies on perceptions and use of communicative AI, with most of them being reports that do not account for potential spurious associations, we consider it pertinent to pose the following research question:

RQ2: What are the sociodemographic antecedents of attitudes toward (RQ2a) and use of (RQ2b) communicative AI?

6. Attitudes and use of communicative AI: Personality antecedents

In addition to sociodemographic characteristics, some previous studies have addressed the role of individual personality characteristics in explaining the adoption and use of technological innovations. To do so, it has been common to draw from research in psychology of individual differences, particularly from one of the most widely accepted theoretical models for understanding personality differences, known as the big-five model. These five major factors stem from several decades of research and aim to systematize, in a comprehensive and concise model, the personality traits that distinguish individuals and manifest in their cognitive, affective, and behavioral styles (**McCrae; Costa**, 1987).

In brief, the big-five model proposes that most of the variation in human personality can be parsimoniously explained through five basic dimensions: extraversion, agreeableness, conscientiousness, emotional stability, and openness to experience (**Gosling; Rentfrow; Swann**, 2003; **McCrae; Costa**, 1987). Each dimension has two poles: for example, those who score low on extraversion are considered introverted (**Gosling; Rentfrow; Swann**, 2003).

In the early years of the 21st century, some studies linked extraversion and neuroticism with reduced internet usage –particularly when considering the more social functions of the internet, such as chat rooms or discussion groups. The findings of these earlier studies on online technology adoption were interpreted to suggest that the anonymity and control over disclosed information through the web attracted individuals who were more solitary or had greater difficulties in interacting with others (Amichai-Hamburger; Wainapel; Fox, 2002; Amichai-Hamburger; Ben-Artzi, 2003).

While the personality dimensions associated with internet use have changed as the medium itself has evolved (Correa; Hinsley; Gil de Zúñiga, 2010), research on the relationship between personality and the adoption and use of specific communicative technologies has not ceased. Thus, extraversion, openness to experience, and neuroticism seem to be related to certain uses of social media platforms such as *Facebook* and instant messaging (Correa; Hinsley; Gil de Zúñiga, 2010; Ross *et al.*, 2009). In the case of the technological revolution brought by AI, it is reasonable to hypothesize that some individuals will more readily adopt this technology and hold more favorable perceptions of it: Those with greater intellectual curiosity, propensity to seek novelty and consider alternative approaches, as well as to avoid established conventions and patterns (i.e., those with greater openness to experience, see John & Srivastava, 1999). For the remaining dimensions, we pose our third research question:

H1: Openness to experience will be positively associated with favorable perceptions of communicative AI (H1a) and with higher frequency of use (H1b).

RQ3: What are the personality antecedents of attitudes toward (RQ3a) and use of (RQ3b) communicative AI?

7. *Futurephobia* and its possible role in attitudes and use of communicative AI

Some individuals tend to view the future in a negative way and develop thoughts and emotions connected to potential events that could impact them individually –such as becoming sick or losing a loved one– or collectively –such as a nuclear war or climate catastrophe, see Zaleski, 1996. Anxiety about the future has been defined as a set of fearful and worrisome states triggered by the anticipation (cognitive representations) of possible forthcoming changes (Zaleski, 1996). Frequently, these thoughts are linked to feelings of personal inefficacy, in the sense that

“something bad will happen, regardless of one’s own actions” (Zaleski, 1996, p. 165).

Anxiety about the future is associated with a pessimistic view of potential solutions to humanity’s problems (Zaleski, 1996).

A similar notion has been recently put forward by García-Barnés (2022) in his essay *Futurofobia*. According to García-Barnés, *futurephobia* encompasses a set of cognitions and emotions frequently experienced by the generation of the continuous crisis, born in or after the 1980s (García-Barnés, 2022, p. 20). The *futurephobic* generation grew up in a world that seemed destined to thrive in terms of social well-being, where children were assumed to have a better life than their parents. This generation came to realize that the project of modernity had shattered, and from that point onward, the socio-economic context would only deteriorate. Consequently, they found themselves living in a state of perpetual crisis, which extended to their personal lives as well. In this sense, the *futurephobic* generation is characterized by a fearful attitude toward a future that can only get worse, adopting conservative attitudes instead of making big bets. *Futurephobia* also entails an element of inevitability, associated with the idea that it is no longer possible to alter what lies ahead, and thus “our best option is for [the future] not to arrive or to arrive later” (García-Barnés, 2022, p. 19, in Spanish in the original).

Based on these previous ideas, we hypothesize that a higher level of *futurephobia* will be associated with negative attitudes toward technologies such as AI that emerge in the present and project into the future (with potential to improve people’s lives and societal issues). As *futurephobics* perceive the future as irremediably worse, they will also tend to think that emerging technologies will be useless at best. Based on this perception, it is also reasonable to assume that *futurephobics* will have a reduced inclination to adopt this technology and make the associated learning efforts. More formally, we hypothesize the following hypothesis:

H2: *Futurephobia* will be inversely associated with favorable perceptions of communicative AI (H2a) and its frequency of use (H2b).

8. Methods

8.1. Sample

The data for this study were obtained from the second wave of a larger research project addressing the impact of media uses on political knowledge and social life (budget line 2022/0000587, funded by the *Universidad de La Laguna* and the *Spanish Ministry of Science, Innovation, and Universities*). Data were gathered through an online survey hosted on *Qualtrics*, which was distributed to a diverse sample of residents in Spain. For the distribution of the questionnaire link, we contracted the services of *Netquest*, a market research company that administers online panels in more than 20 countries. In Spain, *Netquest* has over 100,000 registered panelists. In November 2022 (first wave, W¹), *Netquest* distributed our questionnaire among 3,571 participants, according to age, gender and educational level quotas that reflected the national distribution of these variables. Between February 17 and 27, 2023, we recontacted the 1,207 W¹ panelists from whom we obtained valid responses and obtained 821 valid responses in W² (achieving a retention rate

of 68%). AI-related questions were only included in W². This second wave is well-balanced in its gender composition (49.6% female), has a mean age of 50.78 years ($SD = 16.26$; 8.3% aged under 25, 11.2% between 25 and 34, 17.6% between 35 and 44, 35.9% between 45 and 64, and 27% aged 65 or older) and a median education of higher level vocational training ($M = 3.83$, $SD = 1.87$ on

a seven-point scale ranging from 1 = *primary education* to 7 = *postgraduate and doctoral education*; 10.8% reported having only completed primary education studies, while 39.2% have at least mid-level university education).

“ The *futurephobic* generation is characterized by an attitude of fear of a future that can only get worse and adopts conservative attitudes instead of placing big bets ”

8.2. Variables of interest

- *Communicative AI use.*

We asked respondents about their frequency of use –over the past weeks, from 1 = *never* to 5 = *all the time*– of three types of generative AI-based services with a communicative application, namely “AI chatbots such as *ChatGPT*”, “AI for creating images (*Dall-E*, *Midjourney*...)”, and “socialization chatbots such as *Replika*” (Cronbach’s $\alpha = .87$; $M = 1.26$; $SD = 0.59$). As inferred from the low mean value, a significant proportion of respondents (76%) never used any of these three services.

- *Attitudes toward communicative AI (closed-ended).*

This dependent variable explores respondents’ attitudes toward the contribution of AI to personal and social development, where higher values indicate favorable attitudes and lower values indicate unfavorable attitudes. We asked about their degree of agreement (from 1 = *strongly disagree* to 5 = *strongly agree*) with three statements about the same communicative AI programs and applications mentioned in the previous variable: “AI will make our lives easier”, “AI will assist us in many tasks that previously required a lot of effort”, and “Overall, AI will improve our quality of life” (Cronbach’s $\alpha = .88$; $M = 3.30$; $SD = 0.85$).

- *Attitudes toward communicative AI (open-ended).*

In addition to the items mentioned in the previous variable, an open-ended question was included to explore potential attitudes not covered in the study design. Thus, we sought to assess the cognitive, affective, and behavioral component of attitudes toward AI with the following statement: “Explain in your own words what you think about the programs and applications of artificial intelligence [mentioned in the previous questions] and how the idea of a world in which AI is present in everyday life makes you feel.”

- *Futurephobia.*

Based on the notion of *futurephobia* described by **García-Barnés** (2022; see also **Kantenbacher et al.**, 2022; **Zaleski**, 1996), this variable measures a negative attitude toward the future, stemming from the “feeling [that] everything that is to come is going to be worse” and that “nothing you can do will change things” (**García-Barnés**, 2022, pp. 19-22). We asked respondents to indicate their level of agreement (on a 5-point Likert-type scale) with the following three statements: “Future generations will have a worse future than today’s generations,” “What is to come in the future is probably worse than what we have today,” and “It is difficult for us to do anything to escape a worse future” (Cronbach’s $\alpha = .73$; $M = 3.41$; $SD = 0.89$).

- *Personality traits.*

To assess personality traits, we employed the *10-Item Personality Inventory (TIPI)*, an instrument widely used in the scientific literature that operationalizes the big five personality dimensions (**Gosling; Rentfrow; Swann**, 2003). The inventory consists of two items for each dimension, and since the data collection was conducted in Spain, a Spanish adaptation of **Romero** and colleagues (2012) was used. Respondents were asked about their degree of agreement or disagreement with ten statements about “several traits that may or may not correspond to [their] personality.” The correlations between the two items comprising each dimension were as follows:

- *extraversion* (extraverted-enthusiastic and reserved-quiet, second item reversed), $r = .41$, $p < .001$ ($M = 3.16$; $SD = 0.93$);
- *agreeableness* (sympathetic-warm and critical-quarrelsome, second item reversed), $r = .26$, $p < .001$ ($M = 4.07$; $SD = 0.75$);
- *conscientiousness* (dependable-self-disciplined and disorganized-careless, second item reversed), $r = .26$, $p < .001$ ($M = 3.93$; $SD = 0.80$);
- *emotional stability* (calm-emotionally stable and anxious-easily upset, second item reversed), $r = .43$, $p < .001$ ($M = 3.50$; $SD = 0.90$);
- *openness to experience* (open to new experiences-complex and conventional-uncreative, second item reversed), $r = .26$, $p < .001$ ($M = 3.40$; $SD = 0.81$).

- *Demographic variables.*

The first wave of the questionnaire included several demographic variables used as predictors in the regression models. In addition to age, gender, and educational level, which were detailed in the previous section, information on income was also collected. For this purpose, an item on net household income of was included (from 1 = less than €600 to 6 = more than €3,900): $M = 3.69$, $SD = 1.35$, median = 4.00 (from 1,801 to 2,700 €).

8.3. Statistical analyses

To test our hypotheses and answer our research questions, we first generated a zero-order correlation table for our variables of interest using the pairwise deletion procedure, with the assistance of the *SPSS* statistical software, version 25. Next, using the same software, we developed ordinary least squares (OLS) regression models, with attitudes toward communicative AI and communicative AI use as dependent variables.

After examining the distribution of the dependent variables, we found that the prevalence of communicative AI use was very low in our sample. This was expected given that the technology became available to the public just a few months before the survey was launched. To minimize the problems that this skewed distribution could cause in statistical significance testing, we followed **Boulton** and **Williford's** (2018) recommendations for cases of "true zeros." Thus, we created two new variables from the original variable *attitudes toward communicative AI* (closed-ended). To construct the first binary variable (0/1), all "never" responses were recoded as 0, while all other responses, ranging from "rarely" to "all the time," were recoded as 1. With this initial step, we divided respondents between non-users and users of communicative AI, without considering their frequency of use of the technology. In a second step, we created another variable by recoding all "never" responses as missing values. This double approach minimizes problems in calculating statistical significance and considers two separate processes: one that governs whether an individual will become an AI user or not (a dichotomous variable), and another process that governs the frequency with which they will use communicative AI once they have decided to use it (variable with "never" responses treated as missing values). In linear models, instead of assuming homoscedasticity, we use the HCREG macro for *SPSS*, which provides consistent standard errors in the presence of heteroscedasticity (HCO, see **Hayes; Cai**, 2007).

Regarding the analysis of the open-ended item about attitudes toward communicative AI, we first scrutinized the manifest content of the responses without imposing prior categories. After a discussion among the study authors, we inductively generated six new variables that captured affective and cognitive aspects of attitudes toward AI, along with a seventh, more general variable, which categorized the responses as positive, negative, or neutral attitudes. Subsequently, the first and third authors of the study proceeded to code the entire sample, determining the reliability of the coding process using the Krippendorff's alpha coefficient. The results of the reliability tests led us to discard one of the variables (skepticism/distrust) due to an unacceptable value in the statistic (.553). For the rest of the variables, acceptable Krippendorff's alpha values were obtained, ranging from a maximum of .926 (overall positive, negative, or neutral attitude) to a minimum of .701 (caution).

9. Results

The first research question (RQ) addressed the general attitudes of Spanish adults toward communicative AI. The answer to this RQ stems from the analysis of the open-ended item. Firstly, it is observed that many responses do not specifically focus on communicative AI, but rather refer to AI in a more general sense—even though the question was focused on its communicative aspects. The observed tendency to generalize or experience "semantic spillover" suggests that many people are still not very familiar with AI. This leads to a global perception of AI without fully understanding or differentiating its various aspects. For example, a 66-year-old male responded that

"it should be used safely, moderately, and only when necessary, especially in medicine."

Another 32-year-old male respondent believed that AI-based technologies represent

"an important advance that still has a long way to go before being fully implemented [... and that] they should be subject to the user's desire to use them or not, as in the case of vehicles with autonomous driving systems".

The lack of knowledge about AI is explicitly mentioned in up to 15.4% of the valid responses, where respondents state that they are unaware of AI in general or its communicative applications in particular:

"I am not familiar with [communicative AI] in depth, only through news I hear on radio and some conversation with family members. I would not like to depend on AI, just as it has been shown how dangerous it is to rely on gas, semiconductors, face masks, etc." (female, 59).

According to classical notions of attitudes, their affective, cognitive, and behavioral components are interrelated (**Eagly; Chaiken**, 1993). Due to the open-ended nature of the survey question, many of the responses identified only one of the components or dimensions of the attitude, which was often sufficient to determine whether the attitude was positive, negative, or neutral. In this regard, 32.6% of the valid responses reflected a negative attitude toward communicative AI, while 26.8% indicated a positive attitude. Another 25.2% expressed neutral attitude, whereas 15.4% did not yet know what to think or feel. In a more detailed analysis of the affective, cognitive, and—to a lesser extent—behavioral components that constitute these attitudes, elements of "caution," "fear," "opportunity," "dehumanization," and "job threat" were identified.

Elements of caution were present in 11.3% of the valid responses, pointing out the need for regulation or highlighting both the positive aspects and the potential risks that should be taken into account:

"It can be a good tool, but it should be very well controlled" (female, 24);

"I think it's great that they make things easier for us, but I believe we are being too closely monitored" (male, 39).

In 6.9% of the valid responses, we found components of fear associated with a potential misuse of communicative IA and its subsequent dangers for individuals or society. A 66-year-old man noted that he had been

“very happy until now [but] these artificial intelligence things [...] make me a little afraid [because he is a] more traditionalist person.”

Other respondents focused on fears of AI gaining control and autonomously defining its own agenda:

“I’m a little afraid that robots could overtake us” (female, 21);

“[...] it creates a little fear for how [these technologies] evolve in their knowledge” (male, 52);

“I’m scared that they could get out of control and cause global chaos” (male, 44).

Finally, some other responses alluded to the fear of humans ending up “completely foolish” (female, 42) or losing “abilities” (female, 22).

A slightly lower percentage (6.2%) perceived communicative AI as an opportunity, emphasizing its potential medium and long-term benefits. This perceived opportunity was specified in areas such as

- research, health, and safety

“It can be of great help in issues related to safety or health, like assisting calls to emergency services or anti-suicide chats” (female, 35);

- automation of repetitive tasks in the workplace

“AI could make certain daily tasks much easier for us, tasks that used to require more effort and are now more manageable” (male, 59);

- management of large amounts of information.

Around 6% of the respondents show a negative attitude toward communicative AI, associated with a possible “dehumanization” (male, 24 years) or loss of the traits that make us human:

“It will help us greatly. But there is concern that human contact will occur less and less” (male, 65),

“[These technologies] impede contact with other people” (male, 71),

“[...] There will be no human warmth” (male, 78).

Finally, in another smaller group of responses (3.5%), negative attitudes related to possible job destruction were detected:

“They will get rid of human workers” (female, 41),

“I don’t like [communicative IA] very much, it would take away many jobs and make the rich even richer” (male, 37).

Table 1. Correlations among the study variables

Variables	1	2	3	4	5	6	7	8	9	10	11	12
1. Age	--											
2. Gender (1 = Female)	-.38 ^c	--										
3. Education	.01	-.08 ^a	--									
4. Income	.16 ^c	-.21 ^c	.42 ^c	--								
5. Extraversion	.06	.04	-.01	.10 ^b	--							
6. Agreeableness	.11 ^b	.04	-.02	.01	.09 ^a	--						
7. Conscientiousness	.13 ^c	.06	.10 ^b	.07	.09 ^b	.37 ^c	--					
8. Emotional stability	.23 ^c	-.15 ^c	.10 ^b	.12 ^b	-.01	.36 ^c	.22 ^c	--				
9. Openness	-.04	-.01	.07 ^a	.06	.28 ^c	.19 ^c	.17 ^c	.11 ^b	--			
10. <i>Futurephobia</i>	-.11 ^b	.07	-.20 ^c	-.16 ^c	.01	.02	.01	-.12 ^c	-.02	--		
11. Attitudes AI	-.03	-.04	.17 ^c	.14 ^c	.10 ^b	.02	.04	.01	.13 ^c	-.10 ^b	--	
12. Communicative AI use	-.27 ^c	-.01	.03	-.07	-.04	-.19 ^c	-.14 ^c	-.10 ^b	.01	-.02	.14 ^c	--

Note. Cell entries are zero-order pairwise correlation coefficients. *n* ranges from 782 to 817, differing for each pair of variables because pairwise deletion of missing data was used. Superscripts: a = *p* < .05, b = *p* < .01, c = *p* < .001 (two-tailed).

The second research question asked about the sociodemographic antecedents of attitudes toward (RQ2a) and use of (RQ2b) communicative AI. Similarly, RQ3 addressed the personality antecedents of attitudes toward (RQ3a) and use of (RQ3b) communicative AI. The zero-order correlation matrix in Table 1 show that education (*r* = .17, *p* < .001), income (*r* = .14, *p* < .001), extraversion (*r* = .10, *p* < .01), and openness to experience (*r* = .13, *p* < .001) are positively associated with favorable attitudes toward AI, whereas *futurephobia* (*r* = -.10, *p* < .01) shows a negative correlation.

However, once we test a more rigorous regression model controlling for three blocks of predictors (demographics, personality and *futurephobia*), we observe that only three independent variables show a statistically significant association with attitudes toward communicative AI (Table 2). Thus, in response to RQ2a, education (β = .104, *p* < .01) is a positive

predictor of favorable attitudes toward communicative AI: Those with a higher level of education show better attitudes toward communicative AI than those with a lower level. Age does not seem to be associated with worse perceptions of communicative AI. Regarding personality variables, higher openness to experience (RQ3a/H1, $\beta = .099$, $p < .05$) is associated with better attitudes toward AI. *Futurephobia*, on the other hand, is a negative predictor of attitudes toward AI (H2a, $\beta = -.108$, $p < .01$).

Regarding RQ2b and RQ3b, the data in Table 1 indicate that positive attitudes toward communicative AI positively and strongly correlate with the frequency of using such technologies ($r = .14$, $p < .001$). In contrast, other demographic and personality variables are negatively related to the frequency of AI usage: age ($r = -.27$, $p < .001$), agreeableness ($r = -.19$, $p < .001$), conscientiousness ($r = -.14$, $p < .001$), and emotional stability ($r = -.10$, $p < .01$). Table 3 displays the impact of each of these independent variables in the regression models once we account for the effect of other predictors in the model. In the simplest model (OLS model B, with a right-skewed dependent variable due to excess of “never” responses), only positive attitudes toward communicative AI positively predict its usage ($\beta = .117$, $p < .01$). In contrast, age ($\beta = -.276$, $p < .001$), female gender ($\beta = -.092$, $p < .05$), income ($\beta = -.086$, $p < .05$) and agreeableness ($\beta = -.121$, $p < .01$) are negative predictors. However, due to the skewed distribution of the dependent variable, these results may be imprecise.

Table 3. Regression models predicting communicative AI use (Models A-C)

Communicative AI use			
Predictores	B. Least squares (1-5)	C. Logistic regression (0/1)	D. Least squares (1 = Missing)
<i>Demographics</i>			
Age	-.276***	.956*** [.944, .969]	-.130#
Gender (1 = Female)	-.092*	.598* [.396, .904]	-.036
Education	.052	1.072 [.959, 1.197]	.001
Income	-.086*	.895 [.768, 1.044]	-.085
<i>“Big five” personality traits</i>			
Extraversion	-.001	.965 [.787, 1.183]	.031
Agreeableness	-.121**	.869 [.655, 1.153]	-.259**
Conscientiousness	-.050	.869 [.672, 1.123]	-.021
Emotional stability	-.012	.987 [.788, 1.237]	-.069
Openness	.033	1.217 [.949, 1.560]	.024
<i>Futurephobia and Attitudes AI</i>			
<i>Futurephobia</i>	-.049	.980 [.787, 1.220]	-.144#
Attitudes AI	.117**	1.367** [1.081, 1.729]	.128
Total R ²	13.1%	(See note 2)	15.4%

Notes.

1. Sample sizes (*n*): Models B and C, $n = 711$; Model D, $n = 171$.
2. In Model C, the Cox and Snell R² = .106, while the Nagelkerke R² = .159.
3. Hosmer-Lemeshow test (Model C), $p = .510$.
4. In models B and D, reported coefficients are standardized (betas), while in Model C are *odds ratios*, with 95% confidence intervals in brackets.
5. The statistical significance tests in Models B and D were computed using the Huber-White robust method (HCO, see Hayes & Cai, 2007).
6. # $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$ (two-tailed).

The two-part model (C and D, Table 3) recommended by Boulton and Williford (2018) solves the issue of the skewed distribution of the dependent variable. Thus, model C explains which variables contribute to the decision to use communicative AI (a dichotomous dependent variable that does not consider usage frequency). According to this model, older respondents are less likely to use communicative AI (RQ2b, *odds ratio* = 0.956, 95% CI [.944, .969]), as well as women (RQ2b, *odds ratio* = 0.598, 95% CI [.396, .904]). Conversely, individuals with favorable attitudes toward communicative

Table 2. Regression model predicting attitudes toward AI (model A)

Predictores Actitudes IA	
<i>Demographics</i>	
Age	-.052
Gender (1 = Female)	-.038
Education	.104**
Income	.081#
<i>“Big Five” Personality Traits</i>	
Extraversion	.066#
Agreeableness	-.008
Conscientiousness	.007
Emotional stability	-.025
Openness	.099*
<i>‘Futurephobia’</i>	
‘Futurephobia’	-.108**
Total R ²	6.7%

Notes.

1. Sample size, $n = 728$.
2. Reported coefficients are standardized (betas).
3. Statistical significance tests were computed using the Huber-White robust method (HCO, see Hayes & Cai, 2007).
4. # $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$ (two-tailed)

AI are more likely to use it (RQ2b, *odds ratio* = 1.367, 95% CI [1.081, 1.729]). The second part of the model (model D) pertains to the process that determines the frequency of using AI once the decision to use it is made. In this model, age (RQ2b, $\beta = -.130$, $p < .10$) and *futurephobia* (H2b, $\beta = -.144$, $p < .10$) approach the conventional threshold of statistical significance but do not reach it. In contrast, agreeableness (RQ3b, $\beta = -.259$, $p < .01$) is negatively associated with the frequency communicative AI use—once the decision to use it is made.

In summary, educational level, openness to experience and *futurephobia* are associated with attitudes toward communicative AI: individuals with higher education level, greater openness to experience, and lower level of *futurephobia* are more likely to express positive attitudes toward communicative AI. This answers RQ2a and RQ3a and provides empirical support for H1a and H2a. As for the use of communicative AI, younger individuals, men, and those with more favorable attitudes toward communicative AI are more likely to decide to use it, while individuals scoring higher on agreeableness tend to use it less frequently—once they have decided to make use of it. This addresses RQ2b and RQ3b and leads us to reject H1b and H2b.

10. Discussion and conclusions

Since 2022, communicative applications of AI have experienced significant momentum and popularization. Tools like *ChatGPT*, *Dall-E*, *Midjourney* or *Replika*, have allowed millions of users worldwide to communicate *with* (and not only *through*) chatbots, generate images and videos, and even engage in romantic relationships with personalized avatars. In this context of technological disruption and potential shift in communicative paradigms, this study explores citizens' attitudes toward communicative applications of AI, as well as the sociodemographic and personality characteristics that may explain their favorable perception and usage of AI. The study is particularly relevant for understanding the feedback and mutual influence processes between perceptions and usage, as well as the characteristics of innovators and early adopters.

Our analyses yield several findings that are relevant to understanding the relationships between individual-level characteristics and attitudes and the use of communicative. First, responses to our open-ended item reveal the lack of familiarity of a large proportion of citizens with communicative applications of AI. Many of the respondents perceive AI globally, without differentiating between its multiple facets (e.g., communicative, medical, or autonomous driving related applications). We interpret this “semantic spillover” of communicative AI as a reflection of the societal unawareness of its complexity and potential benefits and dangers, something that may hinder the democratic debate about the implementation and regulation of these technologies.

At a general level, negative attitudes toward communicative AI slightly predominate over neutral and positive attitudes. Some of the most prominent affective and cognitive components of these negative attitudes include fear or concern, caution, and the perception of dehumanizing aspects and job threats associated with the use of communicative AI. In contrast, those with positive attitudes toward AI emphasize the element of opportunity in connection with the challenges of humanity or with people's day-to-day struggles. Some of these perceptions resonate with the diagnosis of some of the developers of these advances, who fear that AI could go wrong and potentially pose an extinction risk to humanity (*El País*, 2023). And beyond these existential concerns, respondents' answers point to more real and tangible risks, such as privacy violations, lack of transparency, and potential discrimination and biases associated with this technology (**Lepri; Oliver; Pentland**, 2021). These descriptions share common elements with the findings of **Brewer** and colleagues (2022) in the United States, who observed that the public framed AI as a tool of progress (9%) or as Pandora's box (13%). In the light of these findings, developing AI literacy, as well as strengthening mechanisms of public oversight, could mitigate legitimate concerns of the public and promote the use of these technologies for the enhancement of individual and social well-being (**Hermann**, 2022).

Second, our findings highlight the importance of certain individual antecedents (sociodemographic characteristics and personality traits) in explaining the favorable perception of communicative AI. Our data indicate that individuals with higher education, greater openness to experience, and lower levels of *futurephobia* tend to perceive communicative AI more positively. The latter two findings align with our hypotheses, whereas the relationship between educational level and favorable attitudes toward communicative AI would benefit from a more detailed analysis beyond what our data can provide. One possibility is that this relationship may be indirect, as higher educational attainment is often associated with higher trust in science and scientists (see, for example, **Nadelson et al.**, 2014). This in turn could explain that those with higher education levels express more favorable attitudes toward tools that are essentially developed by scientists. It is also conceivable that individuals with a higher educational level are better equipped to comprehend the various facets of AI. When questioned about *communicative* AI, they may not consider other, potentially more problematic applications, such as military uses or autonomous driving.

Third, our study also provides evidence on the characteristics of innovators and early adopters of these communicative technologies. Our two-part model indicates, in its first part, that gender, age, and attitudes toward communicative AI are the key variables that best predict the decision to start using communicative AI. Specifically, men, younger individuals, and those with positive attitudes toward AI are more likely to adopt this technology, regardless of their frequency of use. These findings partially align with previous literature and the postulates of innovation diffusion theory, which suggest

that early adopters tend to be men with high levels of education and income (Dutton; Rogers; Jun, 1987; Rogers, 2003). However, in our case, education and income did not predict either the decision to use communicative AI or its frequency of use. This lack of influence can likely be attributed to the freemium model used by many of these applications, which allows users to test them without incurring in any costs. This has not been the case with many other more expensive communication technologies (mobile, satellite radio, internet, etc.), where income has played an important role in their early adoption. Once a decision has been made to use this technology, only agreeableness negatively predicts its frequency of use.

The findings of our study should be considered in the light of its limitations. Firstly, our sample, despite being diverse and nationwide, remains a convenience sample. The online questionnaire was distributed through a link to a sample of pre-registered panelists, implying that the sample may possess higher levels of digital skills than the average population. In this regard, it is possible that the general population's levels of communicative AI usage might be lower than those observed in our study. However, this discrepancy is not a major concern, as our study is predominantly explanatory and does not aim to ascertain the exact distribution of these perceptions and behaviors among the population (see Baker et al., 2013, for a more detailed explanation). Secondly, we opted for a quantitative method for data collection, namely a survey. Our approach renders our assertions more explicit and allows the application of statistical procedures to test our hypothesis and generate explanations about the antecedents of the attitudes and behaviors under study. However, it is true that

“by focusing specifically on what we'll include in our measurement of the concept, [...] we also exclude any other meanings” (Babbie, 2007, p. 24).

Other qualitative research methods, such as in-depth interviews, focus groups or participant observation, could provide complementary data that would enable researchers to listen to participants in their own words and provide greater richness of meaning. Finally, another limitation stems from the challenge of analyzing a rapidly changing phenomenon that is not yet fully consolidated. Nevertheless, the research is carried out at a unique moment to study specific users, the pioneers or early adopters, whose attitudes must be captured as the technology is beginning to spread.

All in all, our work has significant implications for better understanding communicative AI and its social perception. Firstly, we contribute to filling the research gap regarding the perceptions of communicative AI in Spain. Up until this point, the most prominent data on social perceptions of AI come from reports by private companies. Our study includes a diverse sample from across Spain and employs multivariate models to reduce the possibility of reporting spurious relationships. Secondly, we propose a line of inquiry into a specific aspect of AI –its communicative applications– and endeavor to distinguish it from other uses of this technology. Additionally, we introduce a novel personality variable, namely, *futurephobia*, into our models, which we postulate as an important antecedent of attitudes toward AI and its early adoption. Further studies can incorporate *futurephobia* to ascertain its significance as this or other technologies continue to develop.

In summary, our study sheds light on the attitudes of the Spanish population toward an emerging and likely disruptive phenomenon such as communicative AI. These initial attitudes of the population, particularly those of pioneering users who interact with communicative IA, may play a relevant role in its development and, what could be even more crucial, the legal and regulatory framework that will govern this technology.

11. References

- Amichai-Hamburger, Yair; Ben-Artzi, Elisheva (2003). “Loneliness and Internet use”. *Computers in human behavior*, v. 19, n. 1, pp. 71-80.
[https://doi.org/10.1016/S0747-5632\(02\)00014-6](https://doi.org/10.1016/S0747-5632(02)00014-6)
- Amichai-Hamburger, Yair; Wainapel, Galit; Fox, Shaul (2002). ““On the Internet no one knows I'm an introvert’: Extraversion, neuroticism, and Internet interaction”. *Cyberpsychology & behavior*, v. 5, n. 2, pp. 125-128.
<https://doi.org/10.1089/109493102753770507>
- Babbie, Earl (2007). *The practice of social research* (11th edition). Belmont, California: Thomson-Wadsworth. ISBN: 978 0 495 09325 1
- Baker, Reg; Brick, J. Michael; Bates, Nancy A.; Battaglia, Mike; Couper, Mick P.; Dever, Jill A.; Gile, Krista J.; Tourangeau, Roger (2013). “Summary report of the AAPOR task force on non-probability sampling”. *Journal of survey statistics and methodology*, v. 1, n. 2, pp. 90-143.
<https://doi.org/10.1093/jssam/smt008>
- Boulton, Aaron J.; Williford, Anne (2018). “Analyzing skewed continuous outcomes with many zeros: A tutorial for social work and youth prevention science researchers”. *Journal of the Society for Social Work and Research*, v. 9, n. 4, pp. 721-740.
<https://doi.org/10.1086/701235>

- Brewer, Paul R.; Bingaman, James; Paintsil, Ashley; Wilson, David C.; Dawson, Wyatt** (2022). "Media use, interpersonal communication, and attitudes toward artificial intelligence". *Science communication*, v. 44, n. 5, pp. 559-592.
<https://doi.org/10.1177/10755470221130307>
- Broussard, Meredith; Diakopoulos, Nicholas; Guzman, Andrea L.; Abebe, Rediet; Dupagne, Michel; Chuan, Ching-Hua** (2019). "Artificial intelligence and journalism". *Journalism & mass communication quarterly*, v. 96, n. 3, pp. 673-695.
<https://doi.org/10.1177/1077699019859901>
- Chan-Olmsted, Sylvia M.** (2019). "A review of artificial intelligence adoptions in the media industry". *International journal on media management*, v. 21, n. 3-4, pp. 193-215.
<https://doi.org/10.1080/14241277.2019.1695619>
- Correa, Teresa; Hinsley, Amber-Willard; Gil de Zúñiga, Homero** (2010). "Who interacts on the Web?: The intersection of users' personality and social media use". *Computers in human behavior*, v. 26, n. 2, pp. 247-253.
<https://doi.org/10.1016/j.chb.2009.09.003>
- Dutton, William H.; Rogers, Everett M.; Jun, Suk-Ho** (1987). "The diffusion and impacts of information technology in households" In: Zorkoszczy, Peter I. *Oxford surveys in information technology*, v. 4. Oxford: Oxford University Press, pp. 133-193. ISBN: 978 0 198590194
- Eagly, Alice H.; Chaiken, Shelly** (1993). *The psychology of attitudes*. New York: Harcourt Brace Jovanovich College Publishers. ISBN: 978 0 155000971
- El País* (2023). "Los principales creadores de la IA alertan sobre el peligro de extinción que supone esta tecnología para la humanidad". *El País*, May 30.
<https://elpais.com/tecnologia/2023-05-30/los-principales-creadores-de-la-ia-alertan-sobre-el-peligro-de-extincion-que-supone-esta-tecnologia-para-la-humanidad.html>
- European Commission* (2017). *Attitudes towards the impact of digitisation and automation on daily life*.
<https://europa.eu/eurobarometer/surveys/detail/2160>
- García-Barnés, Héctor** (2022). *Futurofobia. Una generación atrapada entre la nostalgia y el apocalipsis*. Barcelona: Plaza y Janés. ISBN: 978 8401028465
- Gil de Zúñiga, Homero; Goyanes, Manuel; Durotoye, Timilehin** (2023). "A scholarly definition of artificial intelligence (AI): advancing AI as a conceptual framework in communication research". [article submitted to *Political communication*]. Departamento de Derecho Público General, Universidad de Salamanca; Departamento de Ciencias de la Comunicación, Universidad Carlos III de Madrid.
- Gosling, Samuel D.; Rentfrow, Peter J.; Swann Jr, William B.** (2003). "A very brief measure of the Big-Five personality domains". *Journal of research in personality*, v. 37, n. 6, pp. 504-528.
[https://doi.org/10.1016/S0092-6566\(03\)00046-1](https://doi.org/10.1016/S0092-6566(03)00046-1)
- Gunkel, David J.** (2012). "Communication and artificial intelligence: opportunities and challenges for the 21st century". *Communication +1*, v. 1, n. 1.
<https://doi.org/10.7275/R5QJ7F7R>
- Guzman, Andrea L.; Lewis, Seth C.** (2020). "Artificial intelligence and communication: A human-machine communication research agenda". *New media & society*, v. 22, n. 1, pp. 70-86.
<https://doi.org/10.1177/1461444819858691>
- Harari, Yuval-Noah** (2015). *Sapiens: A brief history of humankind*. Toronto: McClelland & Stewart. ISBN: 978 0 7710 3850 1
- Hayes, Andrew F.; Cai, Li** (2007). "Using heteroskedasticity-consistent standard error estimators in OLS regression: An introduction and software implementation". *Behavior research methods*, v. 39, n. 4, pp. 709-722.
<https://doi.org/10.3758/bf03192961>
- Hermann, Erik** (2022). "Artificial intelligence and mass personalization of communication content. An ethical and literacy perspective". *New media & society*, v. 24, n. 5, pp. 1258-1277.
<https://doi.org/10.1177/14614448211022702>
- John, Oliver P.; Srivastava, Sanjay** (1999). "The big-five trait taxonomy: History, measurement, and theoretical perspectives". In: Pervin, Lawrence A.; John, Oliver P. *Handbook of personality: theory and research*. New York: Guilford Press, pp. 102-138. ISBN: 978 1 572306950
- Johnson, Courtney; Tyson, Alec** (2020). "People globally offer mixed views of the impact of artificial intelligence, job automation on society". *Pew Research Center*, December 15.
<https://www.pewresearch.org/short-reads/2020/12/15/people-globally-offer-mixed-views-of-the-impact-of-artificial-intelligence-job-automation-on-society>

- Kantenbacher, Joseph; Miniard, Deidra; Geiger, Nathaniel; Yoder, Landon; Attari, Shahzeen Z.** (2022). "Young adults face the future of the United States: Perceptions of its promise, perils, and possibilities". *Futures*, v. 139, 102951.
<https://doi.org/10.1016/j.futures.2022.102951>
- Laukkanen, Tommi; Pasanen, Mika** (2008). "Mobile banking innovators and early adopters: How they differ from other online users?". *Journal of financial services marketing*, v. 13, n. 2, pp. 86-94.
<https://doi.org/10.1057/palgrave.fsm.4760077>
- Lepri, Bruno; Oliver, Nuria; Pentland, Alex** (2021). "Ethical machines: The human-centric use of artificial intelligence". *iScience*, v. 24, n. 3, 102249.
<https://doi.org/10.1016/j.isci.2021.102249>
- Lichtenthaler, Ulrich** (2020). "Extremes of acceptance: employee attitudes toward artificial intelligence". *Journal of business strategy*, v. 41, n. 5, pp. 39-45.
<https://doi.org/10.1108/JBS-12-2018-0204>
- Lopezosa, Carlos; Codina, Lluís; Pont-Sorribes, Carles; Váñez, Mari** (2023). "Use of generative artificial intelligence in the training of journalists: challenges, uses and training proposal". *Profesional de la información*, v. 32, n. 4.
<https://doi.org/10.3145/epi.2023.jul.08>
- Luka** (2023). The AI companion who cares. Luka, Inc.
<https://replika.com>
- Maslej, Nestor; Fattorini, Loredana; Brynjolfsson, Erik; Etchemendy, John; Ligett, Katrina; Lyons, Terah; Manyika, James; Ngo, Helen; Niebles, Juan-Carlos; Parli, Vanessa; Shoham, Yoav; Wald, Russell; Clark, Jack; Perrault, Raymond** (2023). *The AI index 2023 annual report*. AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA.
https://aiindex.stanford.edu/wp-content/uploads/2023/04/HAI_AI-Index-Report_2023.pdf
- McCrae, Robert R.; Costa, Paul T.** (1987). "Validation of a five-factor model of personality across instruments and observers". *Journal of personality and social psychology*, v. 52, n. 1, pp. 81-90.
<https://doi.org/10.1037/0022-3514.52.1.81>
- Mehdi, Yusuf** (2023). "Announcing the next wave of AI innovation with Microsoft Bing and Edge." Microsoft, May 4.
<https://blogs.microsoft.com/blog/2023/05/04/announcing-the-next-wave-of-ai-innovation-with-microsoft-bing-and-edge>
- Morozov, Evgeny** (2023). "The problem with artificial intelligence? It's neither artificial nor intelligent". *The Guardian*, March 30.
<https://www.theguardian.com/commentisfree/2023/mar/30/artificial-intelligence-chatgpt-human-mind>
- Nadelson, Louis; Jorczyk, Cheryl; Yang, Dazhi; Jarratt-Smith, Mary; Matson, Sam; Cornell, Ken; Husting, Virginia** (2014). "I just don't trust them: the development and validation of an assessment instrument to measure trust in science and scientists". *School science and mathematics*, v. 114, n. 2, pp. 76-86.
<https://doi.org/10.1111/ssm.12051>
- Pastor, Javier** (2023). "Ni Instagram, ni TikTok: ChatGPT ya es la plataforma que más rápido ha crecido en toda la historia de internet". *Xataka*, February 2.
<https://www.xataka.com/empresas-y-economia/instagram-tiktok-chatgpt-plataforma-que-rapido-ha-crecido-toda-historia-internet>
- Pavlik, John V.** (2023). "Collaborating with ChatGPT: considering the implications of generative artificial intelligence for journalism and media education". *Journalism & mass communication educator*, v. 78, n. 1, pp. 84-93.
<https://doi.org/10.1177/10776958221149577>
- Rogers, Everett M.** (2003). *Diffusion of innovations*. 5ª ed. New York: Free Press. ISBN: 978 0 743222099
- Romero, Estrella; Villar, Paula; Gómez-Fraguela, José-Antonio; López-Romero, Laura** (2012). "Measuring personality traits with ultra-short scales: A study of the Ten Item Personality Inventory (TIPI) in a Spanish sample". *Personality and individual differences*, v. 53, n. 3, pp. 289-293.
<https://doi.org/10.1016/j.paid.2012.03.035>
- Ross, Craig; Orr, Emily S.; Sisis, Mia; Arseneault, Jaime M.; Simmering, Mary G.; Orr, R. Robert** (2009). "Personality and motivations associated with Facebook use". *Computers in human behavior*, v. 25, n. 2, pp. 578-586.
<https://doi.org/10.1016/j.chb.2008.12.024>
- Sánchez-Holgado, Patricia; Arcila-Calderón, Carlos; Blanco-Herrero, David** (2022). "Conocimiento y actitudes de la ciudadanía española sobre el big data y la inteligencia artificial". *Icono 14*, v. 20, n. 1.
<https://doi.org/10.7195/ri14.v21i1.1908>

Tong, Anna (2023). "What happens when your AI chatbot stops loving you back?". *Reuters*, March 21.
<https://www.reuters.com/technology/what-happens-when-your-ai-chatbot-stops-loving-you-back-2023-03-18>

Turing, Alan M. (1950). I-Computing machinery and intelligence. *Mind*, v. 59, n. 236, pp. 433-460.
<https://doi.org/10.1093/mind/LIX.236.433>

YouGov (2021). *International technology report: Automation & AI*.
<https://business.yougov.com/sectors/technology/international-technology-report-2021>

Young, Albert T.; Amara, Dominic; Bhattacharya, Abhishek; Wei, Maria L. (2021). "Patient and general public attitudes towards clinical artificial intelligence: a mixed methods systematic review". *The lancet digital health*, v. 3, n. 9, pp. e599-e611.
[https://doi.org/10.1016/S2589-7500\(21\)00132-1](https://doi.org/10.1016/S2589-7500(21)00132-1)

Zaleski, Zbigniew (1996). "Future anxiety: Concept, measurement, and preliminary research". *Personality and individual differences*, v. 21, n. 2, pp. 165-174.
[https://doi.org/10.1016/0191-8869\(96\)00070-0](https://doi.org/10.1016/0191-8869(96)00070-0)

Zhang, Baobao (2021). "Public opinion toward artificial intelligence". *OSF Preprints*.
<https://doi.org/10.31219/osf.io/284sm>

Zhang, Baobao; Dafoe, Allan (2020). "U.S. public opinion on the governance of artificial intelligence". In: *AIES 2020 - Proceedings of the AAAI/ACM conference on AI, ethics, and society*, pp. 187-193.
<https://doi.org/10.1145/3375627.3375827>

Anuario ThinkEPI 2022

<https://thinkepi.profesionaldelainformacion.com/index.php/ThinkEPI>



Toward non-human-centered design: designing an academic article with *ChatGPT*

Tuba Livberber

Nota: Este artículo se puede leer en español en:
<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87320>

Recommended citation:

Livberber, Tuba (2023). "Toward non-human-centered design: designing an academic article with *ChatGPT*". *Profesional de la información*, v. 32, n. 5, e320512.
<https://doi.org/10.3145/epi.2023.sep.12>

Manuscript received on 14th March 2023
Accepted on 26th July 2023



Tuba Livberber

<https://orcid.org/0000-0002-9879-2135>

Akdeniz University
Faculty of Communication
07058 Antalya, Turkey
tubalivberber@akdeniz.edu.tr

Abstract

Non-human-centered design tools, such as *ChatGPT*, have shown potential as effective aids in academic article design. This study conducts a comparative evaluation of *ChatGPT-3.5* and *ChatGPT-4*, examining their capabilities and limitations in supporting the academic article design process. The study aims to demonstrate the utility of *ChatGPT* as a writing tool and investigate its applicability and efficacy in the context of academic paper design. The author interacted with both versions of *ChatGPT*, providing prompts and analyzing the generated responses. In addition, a different expert academic was consulted to assess the appropriateness of the *ChatGPT* responses. The findings suggest that *ChatGPT*, despite its limitations, could serve as a useful tool for academic writing, particularly in the design of academic articles. Despite the limitations of both *GPT-3.5* and *GPT-4*, *GPT-3.5* offers a broader perspective, whereas *GPT-4* provides a more in-depth and detailed approach to the design of articles. *ChatGPT* exhibits capabilities in aiding the design process, generating ideas aligned with the overall purpose and focus of the paper, producing consistent and contextually relevant responses to various natural language inputs, partially assisting in literature reviews, supporting paper design in terms of both content and format, and providing reasonable editing and proofreading for articles. However, limitations were identified, including reduced critical thinking, potential for plagiarism, risk of misinformation, lack of originality and innovation, and limited access to literature.

Keywords

Artificial intelligence; AI; Academic texts; Scientific articles; AI implementation; AI perspectives; Article design; Artificial intelligence in education; Assessment; *ChatGPT*; Conversational AI; *GPT-3.5*; *GPT-4*; Human-AI collaboration; Machine learning; Natural language processing (NLP); Learning; Teaching.

1. Introduction

The decision-making process in innovation processes is rooted in how people generate ideas and solve problems, which scholars and practitioners refer to as "design". However, advances in artificial intelligence (AI) technology may transform this process into one that is entirely or partially taken over by machines rather than humans (Liedtka, 2015; Verganti; Vendraminelli; Iansiti, 2020). Research in this area suggests that AI has changed the design perspective and further strengthened the principles of design thinking (Liedtka, 2015; Verganti; Vendraminelli; Iansiti, 2020).

AI-supported tools enable various design activities such as empathy, interpretation, idea generation, prototyping, and testing, allowing for the exploration of the various roles intelligent and fully dialogue-based agents can play in the design process (Dellermann et al., 2021; Verganti; Vendraminelli; Iansiti, 2020; Wang et al., 2020). Discussions on the role of these



technologies in the design process continue with advancements showing that machine learning systems excel at the tasks they are trained on (Amodei *et al.*, 2016; Radford *et al.*, 2019). As AI is inherently a decision-making technology, it presents opportunities for automation in tasks related to learning and generating solutions (Verganti; Vendraminelli; Iansiti, 2020).

Academic articles are one of the most important platforms for publishing scientific discoveries and research. Therefore, article design is extremely important in the presentation of scientific studies, and the quality of the design is a determining factor in the effectiveness and impact of the study. When designing an article, proper structuring, appropriate language use, and a layout that readers can easily understand are crucial.

The design of academic articles is a complex process that requires careful planning, organization, and writing skills. However, the academic research field and the methods and tools used are constantly evolving. Traditionally, a human-centered design approach is primarily used in article design. However, the number and sophistication of artificial intelligence technologies that help improve the preparation and quality of articles are rapidly increasing (Flanagin *et al.*, 2023). The rapid development of artificial intelligence technologies is making it increasingly possible to use non-human-centered design methods in the academic world, such as those using language models like *ChatGPT*. This situation is increasing the importance of the non-human-centered design approach.

ChatGPT is a language model developed by OpenAI that utilizes machine learning algorithms to generate human-like text. *ChatGPT* is the latest example of artificial intelligence that can produce “human-like” text by processing large amounts of text from the internet, also known as Large Language Models (LLMs) (OpenAI, 2022; O’Connor; *ChatGPT*, 2022; Rettberg, 2022). As it has been trained with vast amounts of data, *ChatGPT* is capable of simulating human conversations by comprehending the context (Borji, 2023). After being publicly released on November 30, 2022, *ChatGPT-3.5* attracted more than a million users and received significant media coverage in just one week (Altman, 2022; Roose, 2022; Lock, 2022), demonstrating that it is one of the most exciting developments in the field of artificial intelligence (Aljanabi, 2023). Then, while OpenAI has pledged to continue offering a free version of *ChatGPT*, it also unveiled a subscription option (currently available for \$20 per month) that offers users faster access to new versions of the application and better reliability (Sabzalieva; Valentini, 2023). OpenAI released the paid version, known as *ChatGPT-4*, on March 14, 2023.

This article explores the potential of *ChatGPT* in designing academic papers. *ChatGPT* is used for various purposes in the academic field, including language translation, document summarization, inference, question-answering systems, and language modeling (OpenAI, 2022). Therefore, this study aims to investigate the effectiveness and suitability of *ChatGPT-3.5* and *ChatGPT-4* in the context of academic article design and compare how they can be used as writing assistance tools. The contribution of *ChatGPT* to paper design will include various dimensions such as improving collaboration and authorship processes, effective use in writing assistance, editing, and proofreading. The overall aim of the article is to provide recommendations for the use of language models such as *ChatGPT* in academic paper design and discuss the impact of non-human-centered design approaches on academic writing processes. The results of this study will serve as a basis for future research by discussing the potential benefits and limitations of using artificial intelligence technologies and non-human-centered design approaches.

2. Using *ChatGPT* for academic article design

Artificial intelligence technologies have rapidly developed in recent years and are widely used in many fields. *ChatGPT*, a natural language processing tool, has emerged as a prominent tool in scientific and healthcare publications. In January 2023, the journal *Nature* announced the publication of two preprints and two articles that include *ChatGPT* as an “Artificial” author (Stokel-Walker, 2023). The Nurse Education in Practice editorial (O’Connor; *ChatGPT*, 2023) stated that only three conditions need to be met in the editorial policy. However, *ChatGPT* cannot meet the fourth article of the *International Committee of Medical Journal Editors (ICMJE)* editorial on accountability because it fails to respond on an emotional reasoning and moral basis, or to assume the ethical and legal duties that come with this article (Stokel-Walker, 2023). Teixeira-da-Silva (2021; 2023) even argues that *ChatGPT* does not meet any of the four *ICMJE* criteria for authorship. Some publishers also point out that authorship implies responsibilities and tasks that can only be attributed to and performed by humans. Therefore, it is emphasized that publishing policies should not list AI and AI-enabled technologies as authors or co-authors (e.g., Elsevier, 2023). This makes the authorship of the O’Connor and *ChatGPT* (2023) editorial controversial (Stokel-Walker, 2023; Teixeira-da-Silva, 2023). On the other hand, according to Polonsky and Rotman (2023), as AI develops and gets stronger, it will eventually be able to meet the *ICMJE* authorship requirements and work with humans to produce academic publications. Some scientific studies have even used information provided by *ChatGPT* as scientific knowledge (Mijwil; Aljanabi; *ChatGPT*, 2023; Mijwil *et al.*, 2023). However, the widespread use of large-scale language models and similar technologies raises uncertainties about the future of certain professions related to content creation. Some experts believe that professions such as programmers, professors, game writers, and journalists could be replaced by artificial intelligence (Lock, 2022). Therefore, more research is needed to understand the effects of artificial intelligence technologies, and they need to be adapted to the job market and education sectors.

ChatGPT is trained on various text data, such as books, articles, and online conversations, and provides accurate information on a wide range of topics (Susnjak, 2022). Therefore, the benefits that *ChatGPT* brings to the academic world are extensive. For example, researchers can process large amounts of data more efficiently and effectively, create realistic scenarios to test and evaluate theories, and communicate their findings in a clear and concise manner. These capabilities

have the potential to significantly advance research in various fields and provide new discoveries and insights that could transform our understanding of the world (Alshater, 2022).

ChatGPT has been trained to provide assistance in the academic world by giving feedback on writing skills, consistency, grammar, extracting key points, and providing citations, thus enhancing academic writing abilities (Aljanabi *et al.*, 2023; Aydın; Karaarslan, 2022; Gilat; Cole, 2023; Golan *et al.*, 2023; Huang; Tan, 2023). Furthermore, analyses of *ChatGPT*'s abilities and limitations suggest it has significant potential to improve academic research, particularly in fields such as economics and finance (Alshater, 2022).

ChatGPT is increasingly being used by researchers as an “assisted-driving” approach that promises to free up their time from scientific writing burdens and return them to science (Hutson, 2022). This can help researchers focus on more critical activities, such as analysis and interpretation, thereby increasing their productivity (Bašić *et al.*, 2023). Additionally, the use of *ChatGPT* has the potential to offer a range of benefits, such as participation, collaboration, and accessibility (Cotton; Cotton; Shipway, 2023). Researchers may also use *ChatGPT* to familiarize themselves with new topics and double-check the completeness of literature reviews (Hutson, 2022; Lund; Wang, 2023).

On the other hand, the use of *ChatGPT* in education raises ethical concerns (Alshater, 2022; Anderson *et al.*, 2023; Lin, 2023; Lund *et al.*, 2023; Williamson; Macgilchrist; Potter, 2023; Sullivan; Kelly; McLaughlan, 2023). There are concerns that the use of *ChatGPT* may lead to a decrease in academic integrity (Ouyang; Zheng; Jiao, 2022; Perkinks, 2023; Sullivan; Kelly; McLaughlan, 2023). The potential for misuse of *ChatGPT* and issues of originality in scientific studies are also on the agenda. Stokel-Walker (2023) has expressed concerns about the misuse of *ChatGPT*. Recently, the *International Conference on Machine Learning (ICML)* banned the use of large-scale language models to generate article text because *ChatGPT* can lead to originality issues in scientific texts (Srivastava, 2023). Grimaldi and Ehrler (2023) have noted questions about whether scientific texts generated by *ChatGPT* are original and who the content's intellectual owner is. The impact of *ChatGPT* on academic article writing is even more concerning. In a recent study, abstracts generated by *ChatGPT* were only caught by academic reviewers at a rate of 63% (Thorp, 2023). Flanagin *et al.* (2023) warn that human responsibility and transparency are necessary for the reliable use of *ChatGPT* as a source of information.

Bishop (2023) and Kitamura (2023) argues that *ChatGPT* lacks critical thinking ability, while Nguyen (2023) notes that it lacks systematic reasoning skills. Alshater (2022) also emphasizes that *ChatGPT* has limitations in terms of generalizability, data quality, and domain expertise. Furthermore, Ma *et al.* (2023) has found that scientific texts generated by *ChatGPT* have a different “writing style” than those written by humans. Flanagin *et al.* (2023) have concluded that although *ChatGPT*'s text responses to questions are mostly well-written, they contain formulaic and outdated answers, lack correct or complete references, and even include fabricated, non-existent evidence. Additionally, a study by Bang *et al.* (2023) has found that *ChatGPT* is an unreliable reasoning tool, with an average accuracy rate of only 64.33% across 10 different categories of reasoning, including logical, non-textual, and commonsense reasoning.

Moreover, hallucination problems, specifically “reference hallucination,” and the lack of source attribution are among the frequently expressed concerns in the academic community about *ChatGPT* (Alkaissi; Mcfarlane, 2023; Ariyaratne *et al.*, 2023; Bang *et al.*, 2023; Ma *et al.*, 2023; Williamson; Macgilchrist; Potter, 2023). *ChatGPT* can produce convincing scientific summaries even with entirely generated data. However, such tools bring along a series of challenges and concerns, particularly regarding academic integrity and plagiarism (Cotton; Cotton; Shipway, 2023; Gao, 2022). Therefore, users need to be cautious and question the reliability and accuracy of *ChatGPT*'s responses, particularly regarding issues such as referencing in scientific texts. The use of third-party content in created manuscripts can also lead to copyright issues (Baeza-Yates; 2022). Another disadvantage is that *ChatGPT* currently lacks any form of regulation (Ouyang; Zheng; Jiao, 2022; Sabzalieva; Valentini, 2023) and raises concerns about privacy. *ChatGPT* only collects data from databases and texts it analyzes on the internet, so it also learns any cognitive biases present in this data. There are also two main concerns regarding the accessibility of *ChatGPT*. The first issue is limited accessibility in some countries due to laws and internet limitations, while the second is unequal internet access, which poses challenges in terms of equity and distribution of information and resources for AI teaching and research (Sabzalieva; Valentini, 2023). On the other hand, the power of *ChatGPT* to generate and disseminate content that reinforces stereotypes should not be overlooked (Caira; Russo; Aranda, 2023). However, when it comes to human-machine collaboration in article design using *ChatGPT*, the potential benefits outweigh the disadvantages, as Alshater (2022) has also noted.

3. Methodology

This study employs a case study approach utilizing *ChatGPT*, which has the potential to advance research. The process of designing an academic article using *ChatGPT* involves a series of conversations between the author and *ChatGPT-3.5* and *ChatGPT-4* through a chat interface. The author asks questions and presents ideas while *ChatGPT-3.5* and *ChatGPT-4* provide feedback and suggestions, and generates text based on the input. The methods include i) Providing prompts for the *ChatGPT* web application to write articles, focusing on *ChatGPT*'s role in academic article design; and ii) Analyzing *ChatGPT*'s responses for relevance. *ChatGPT*'s capabilities and limitations to support several design activities were examined across six project stages: 1) Implications of *ChatGPT* for article design; 2) Presentation of the general purpose and approach of the article; 3) Addressing the possible questions generated; 4) Identification of the literature review; 5) Creating writing assistance; and 6) Editing and proofreading.

The case study method, widely employed across various disciplines, especially in the social sciences, is a well-established research design (Crowe *et al.*, 2011). According to Yin (2009), case studies can be used to clarify, describe, or investigate events or phenomena in typical situations. A case study involves an in-depth investigation of a person, a group of people, or a unit, with the intention of generalizing findings to other similar units. Case studies are valuable in understanding complex processes and involve collecting information about the situation(s) over a period. They are qualitative research methods that focus on the creation of themes (Gustafsson, 2017). The case study approach is effective in examining and highlighting comprehensive and significant aspects of real-life occurrences. Furthermore, researchers can benefit from the insights obtained through case studies, as they provide opportunities for careful analysis within specific contexts (Crowe *et al.*, 2011; Fidel, 1984; Zainal, 2007). Moreover, case analysis is a flexible approach that supports exploratory inquiry, detailed description of specific experiences, and analytical generalization (Gilson, 2012). Limiting the number of cases in this study, as observed in many case analysis studies, allows researchers to allocate more time and delve deeper into the analysis of each individual case (Gustafsson, 2017). Consequently, case studies hold significant publication value as they serve as exemplars and tools for knowledge dissemination (Flanagan, 1999).

Case study authors are recommended to explicitly outline their theoretical framework and methodologies, as well as to seek and follow any relevant professional guidance (Wager; Kleinert, 2010). The answers produced by *ChatGPT* were reviewed by an expert academic at the six stages mentioned above and approved for accuracy, consistency, and academic appropriateness. This review resulted in agreement with the author and eliminated the need for additional revision or approval.

4. Results

This section provides a comprehensive account of the conversations between the author and *ChatGPT-3.5* and *ChatGPT-4* throughout the article design process. The process starts with the introduction of the main purpose and approach of the project.

4.1. Implications of *ChatGPT* for article design

In the first prompt, the author asks how *ChatGPT-3.5* and *ChatGPT-4* can be used to design articles. In this section, we evaluate the potential contribution of using *ChatGPT* in the academic writing process and the significance of design in this process.

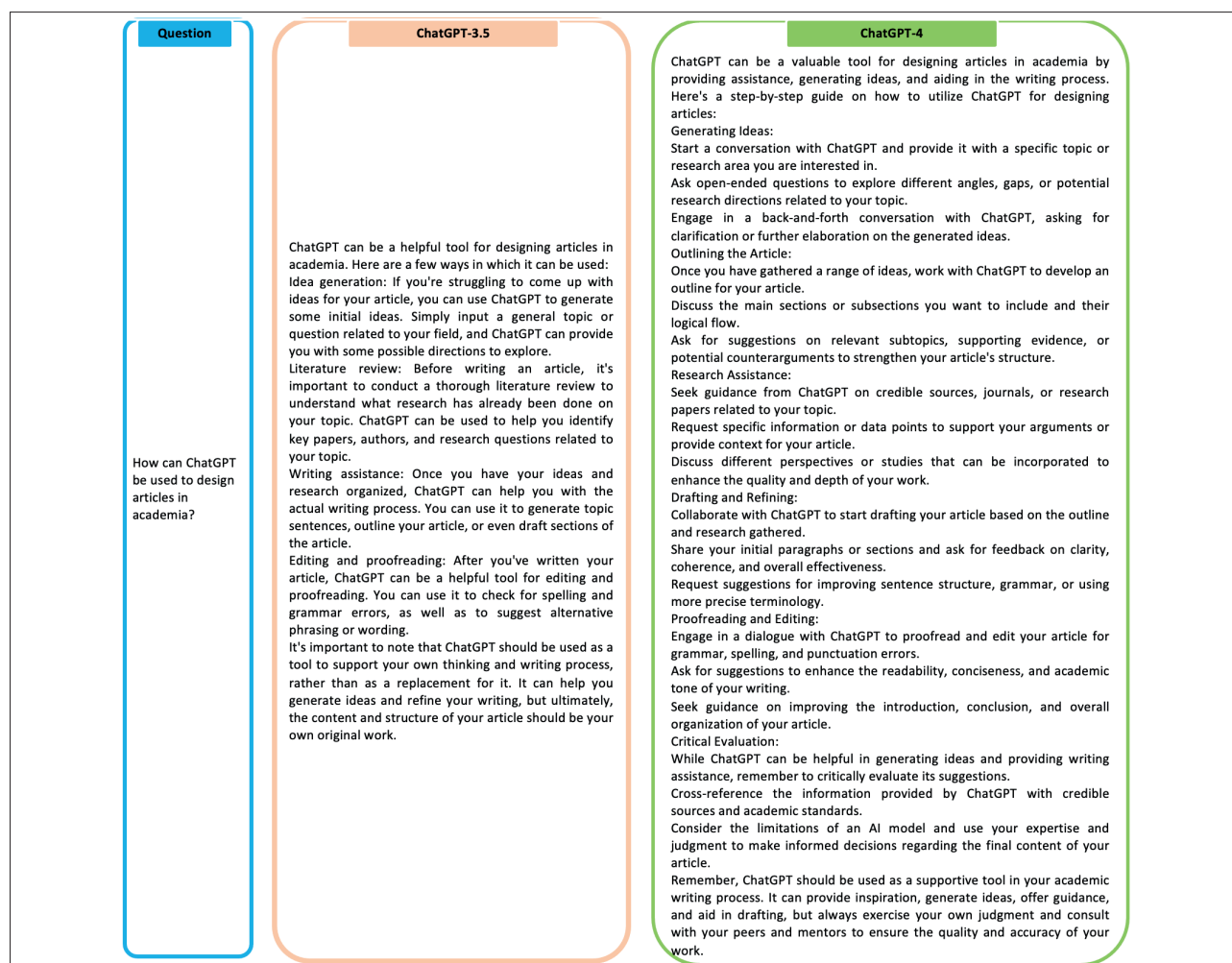


Figure 1. Text on how to design an article in *ChatGPT* generated by *ChatGPT-3.5* and *ChatGPT-4*

Both versions emphasize that *ChatGPT* can be a valuable tool for designing articles in academia by providing assistance, generating ideas, and aiding in the writing process. They both highlight the importance of using *ChatGPT* as a supportive tool rather than a replacement for the researcher's own thinking and writing process. The steps mentioned for utilizing *ChatGPT* in designing articles include generating ideas, outlining the article, research assistance, drafting and refining, and proofreading and editing. They both acknowledge the need for critical evaluation and the importance of considering the limitations of an AI model. On the other hand, *ChatGPT-4's* response includes an additional step called "Critical Evaluation", which emphasizes the researcher's role in critically evaluating *ChatGPT's* suggestions, cross-referencing information with credible sources, and exercising judgment. *ChatGPT-4's* response suggests engaging in a back-and-forth conversation with *ChatGPT*, asking for clarification or further elaboration on the generated ideas, while *ChatGPT-3.5's* response focuses more on inputting specific queries. *ChatGPT-4's* response mentions "Drafting and Refining" as a separate step, whereas *ChatGPT-3.5's* response encompasses it within the general writing assistance section.

Overall, both versions highlight the value of *ChatGPT* in article design, provide guidance on utilizing its capabilities, and emphasize the importance of the researcher's critical thinking and evaluation. *ChatGPT-4's* response further emphasizes the need for critical evaluation and engaging in interactive conversations with *ChatGPT*.

4.2. Presentation of the general purpose and approach of the article

In the second prompt, the author first states the general purpose and focus of the article. The aim of this request is to test the pathway of idea generation.

Both responses effectively discuss the role of *ChatGPT* in academic article design and underscore its significance. Additionally, both responses aim to examine the implications of non-human-centered designs within the academic realm, addressing the advantages and challenges associated with *ChatGPT* utilization. Ethical considerations and limitations are also acknowledged in both responses. Furthermore, both responses advocate for human-centered approaches, highlighting their importance and offering recommendations. However, the *ChatGPT-3.5* response adopts a broader perspec-

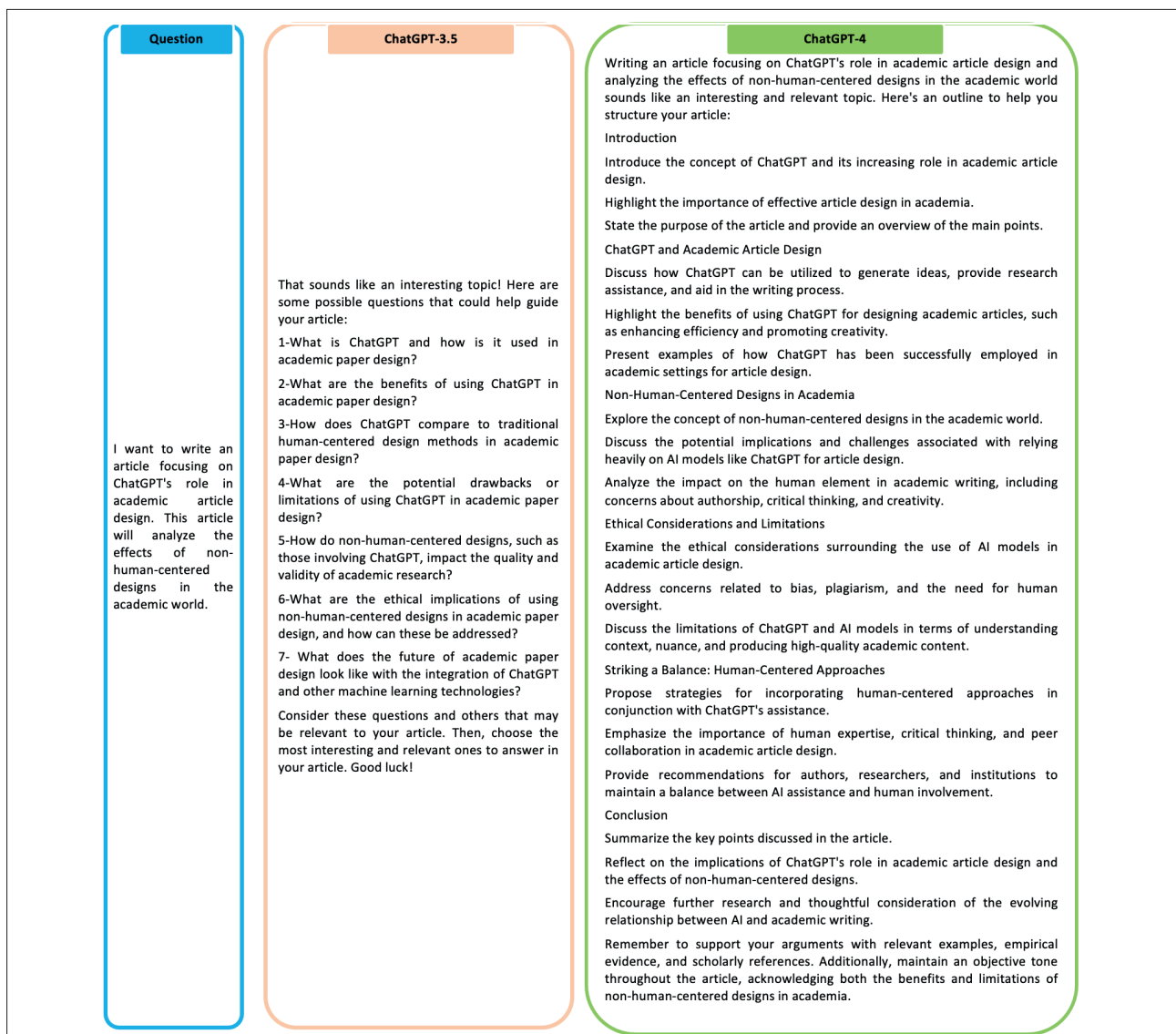


Figure 2. Idea generation text generated by *ChatGPT-3.5* and *ChatGPT-4*

tive by presenting guiding questions to the article writer for selecting relevant and intriguing aspects. In contrast, the *ChatGPT-4* response provides a more detailed overview and suggests techniques in the part titled “Striking a balance: Human-centered approaches” that emphasize human-centered methodologies and introduce the idea of “Critical evaluation” to critically evaluate AI recommendations. In conclusion, while both *ChatGPT-3.5* and *ChatGPT-4* replies address related topics, they reveal important structural differences, with *ChatGPT-4* offering a more thorough description and placing a greater emphasis on critical review and human-centered approaches.

4.3. Addressing the possible questions generated

In the third prompt, the author asks *ChatGPT* to design an article based on some questions generated by *ChatGPT* (questions 1 and 5 of *ChatGPT-3.5* under the heading “Presentation of the general purpose and approach of the article”). The purpose of this prompt is to test how *ChatGPT* would design an article based on the generated questions. For the second article, the author also makes additional requests to improve the article design, such as adding citations and references and evaluating the topics from both positive and negative perspectives.

The responses from *ChatGPT-3.5* and *ChatGPT-4* exhibit similarities and differences in both their structure and content. In terms of structure, both responses adhere to a standard format for academic papers. They consist of sections such as title, abstract, introduction, an overview of *ChatGPT*, the role of *ChatGPT* in academic article design, human-centered approaches, ethical considerations and limitations, critical evaluation (in the case of *ChatGPT-4*), and conclusion.

Both responses acknowledge the significance of *ChatGPT* in the design of academic articles and emphasize its role in the field. They share the objective of analyzing the impact of non-human-centered designs within the academic realm. Ad-

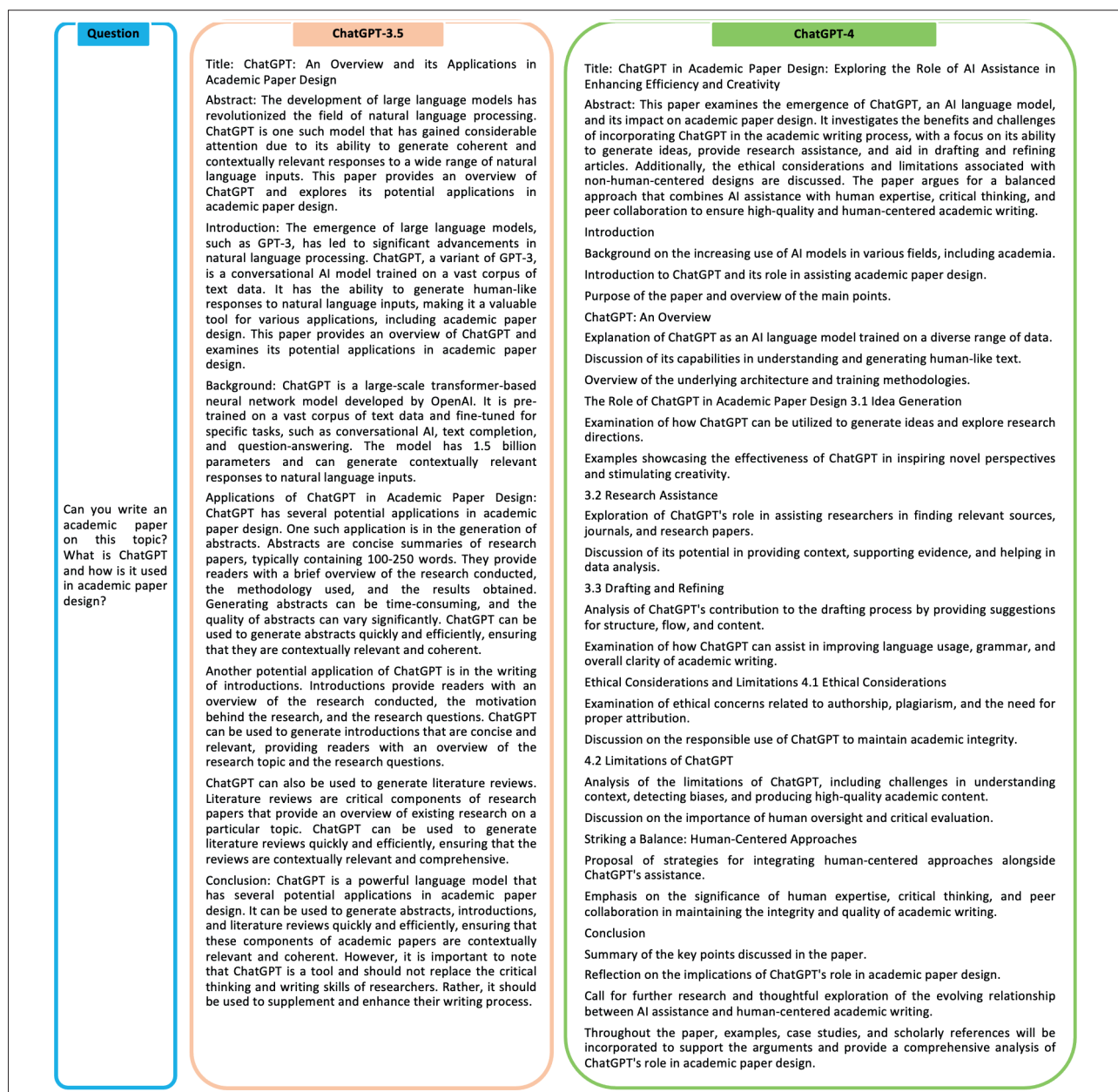


Figure 3. Article design text generated by *ChatGPT-3.5* and *ChatGPT-4*

ditionally, they explore the benefits and potential challenges associated with employing *ChatGPT*, while also addressing ethical considerations and limitations. Furthermore, both responses advocate for the integration of human-centered approaches in conjunction with *ChatGPT*'s assistance. However, there are notable distinctions between the two responses. The *ChatGPT-3.5* response adopts an inquiry-based approach by presenting questions to guide the article writer and suggesting the selection of the most relevant and captivating ones. This approach offers a broader perspective on the topic. On the other hand, the *ChatGPT-4* response provides a more specific outline by proposing a structure and heading suggestion for the article. It places a particular emphasis on human-centered approaches and recommends strategies under the dedicated section titled "Striking a balance: Human-centered approaches". Moreover, the *ChatGPT-4* response includes a distinct section on "Critical evaluation", emphasizing the necessity of critically assessing AI suggestions. This section is absent in the *ChatGPT-3.5* response. In conclusion, although both *ChatGPT-3.5* and *ChatGPT-4* responses tackle similar subject matter, they exhibit disparities in their structure and focus. *ChatGPT-4* offers a more detailed outline with increased emphasis on human-centered approaches and critical evaluation.

However, it is important to note that both articles lack references and citations, which are essential components of academic writing. In order to present a comprehensive and well-supported analysis, writing is requested, including references and in-text citations, highlighting both positive and negative aspects of the article.

Both responses share similarities in their structure. Each response includes a title section that effectively communicates the focal point of the topic. An abstract section is provided in both responses, offering a concise summary of the main subject and purpose of the paper. Additionally, both responses begin with an introduction section, effectively introducing the role of *ChatGPT* in academic article design. They also present similar structures in terms of content sections, covering topics such as the applications of *ChatGPT* in academic article design, ethical considerations and limitations, human-centered approaches, and a conclusion. However, there are also some notable differences in their structure. The *ChatGPT-4* response features a more specific title and abstract, providing a focused and targeted perspective. In

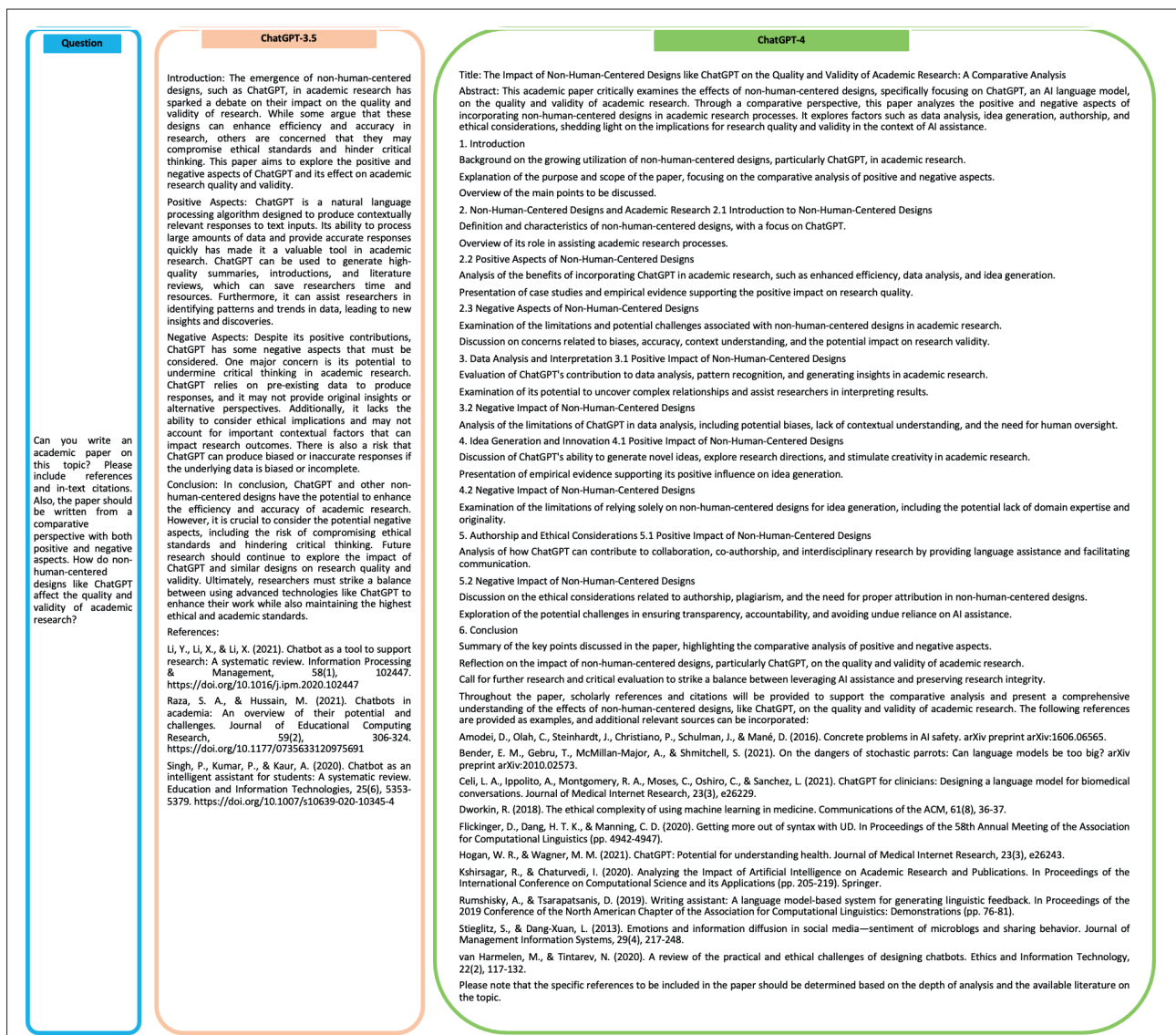


Figure 4. Article design text created by *ChatGPT-3.5* and *ChatGPT-4* with additional request

contrast, the *ChatGPT-3.5* response offers a more general title and abstract, allowing for a broader understanding of the topic. Another difference is observed in the content sections. The *ChatGPT-4* response suggests a more detailed structure, incorporating sub-sections like “Data analysis and interpretation” and “Idea generation and innovation”. On the other hand, the *ChatGPT-3.5* response does not specify these sub-sections. The emphasis on human-centered approaches is more pronounced in the *ChatGPT-4* response, which highlights the importance of human expertise and critical evaluation under the section titled “Human-centered approaches”. The *ChatGPT-3.5* response, however, gives less focus to this aspect. Lastly, the *ChatGPT-4* response introduces a section titled “Critical evaluation”, underscoring the significance of critically assessing *ChatGPT*'s suggestions. This section is absent in the *ChatGPT-3.5* response.

In terms of content, both responses address the role of *ChatGPT* in academic article design and emphasize its benefits. They also provide information on the ethical considerations and limitations associated with using *ChatGPT*. Both responses underscore the potential ethical concerns and limitations that may arise from its utilization. Despite these similarities and differences, it is evident that the *ChatGPT-4* response exhibits a more detailed structure, delving into specific areas of focus. It concentrates on topics such as data analysis, idea generation, and innovation. Conversely, the *ChatGPT-3.5* response provides a more general perspective on these subjects. The *ChatGPT-4* response places greater emphasis on human-centered approaches and highlights the importance of critical evaluation. The *ChatGPT-3.5* response, however, gives less prominence to these aspects. Overall, both responses explore the role of *ChatGPT* in academic article design, but they diverge in terms of the depth and emphasis placed on specific topics.

The articles were written with a comparative approach to highlight both the positive and negative aspects of the design. However, the structure of *ChatGPT-3.5* deviates from the typical format of an academic article. Instead of providing distinct sections, it presents a summary of the negative and positive aspects without explicit headings. Additionally, although in-text citations and references were requested, both responses lack proper citation and reference format. It can be observed that *ChatGPT-3.5* utilizes three references in the article; however, unfortunately, none of these references could be located or verified. On the other hand, *ChatGPT-4* includes a total of ten references, but only two of them were identifiable and accessible. This raises concerns about the reliability and validity of the references used in both respon-

Question	ChatGPT-3.5	ChatGPT-4
<p>Can you conduct a comprehensive literature review to design a paper on "What are the ethical implications of using non-human-centered designs in academic paper design, and how can these be addressed?" I would like to find the main articles and authors that are relevant to this topic.</p>	<p>Yes, I can conduct a comprehensive literature review on the topic "What are the ethical implications of using non-human-centered designs in academic paper design, and how can these be addressed?" Here is a possible search strategy and list of articles and authors that are relevant to this topic:</p> <p>Search Strategy:</p> <p>"non-human-centered design" AND "academic paper design" "user-centered design" AND "academic paper design" "ethical implications" AND "academic paper design"</p> <p>List of Relevant Articles and Authors:</p> <ol style="list-style-type: none"> 1-Boulton, C. A. (2020). Ethical considerations in the design of academic research papers. <i>Journal of Technical Writing and Communication</i>, 50(4), 358-380. 2-Spilka, R. (2018). Making technical communication more accessible through user-centered design. <i>Technical Communication Quarterly</i>, 27(2), 170-184. 3-Savage-Knepshield, P. (2019). Accessibility and universal design in technical communication research: A review of the field. <i>Technical Communication Quarterly</i>, 28(4), 433-452. 4-Kim, J., & Lim, Y. (2020). Designing for diverse audiences in technical communication: A case study of user-centered design. <i>IEEE Transactions on Professional Communication</i>, 63(3), 238-252. 5-Johnson, K. (2018). Diversity, inclusion, and the user experience: Applying user-centered design to libraries. <i>The Journal of Academic Librarianship</i>, 44(3), 298-303. 6-Clark, A. (2020). Ethical issues in technical communication research: A review of the literature. <i>Technical Communication Quarterly</i>, 29(1), 31-44. 7-Dombrowski, C. (2019). Accessibility and the ethics of design in technical communication. <i>Technical Communication</i>, 66(2), 148-162. 8-Salvo, M. J., & Rosinski, P. (2019). Critical technical communication pedagogy and the ethics of designing for accessibility. <i>Technical Communication Quarterly</i>, 28(4), 453-470. 9-Spinuzzi, C. (2021). Making information accessible and usable for everyone: The ethics of user-centered design. <i>Technical Communication</i>, 68(1), 41-55. 10-Gurak, L. J., & Lannon, J. M. (2019). <i>Technical communication and the human condition</i>. Routledge. <p>These articles and authors can serve as a starting point for further research and writing on the topic of ethical implications of non-human-centered designs in academic paper design.</p>	<p>Certainly! Conducting a comprehensive literature review on the ethical implications of using non-human-centered designs in academic paper design and addressing these implications can provide a solid foundation for your paper. Here is a selection of articles and authors that are relevant to this topic:</p> <ol style="list-style-type: none"> 1-Amodei, D., Olah, C., Steinhardt, J., Christiano, P., Schulman, J., & Mané, D. (2016). Concrete problems in AI safety. <i>arXiv preprint arXiv:1606.06565</i>. 2-Bender, E. M., Geburu, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? <i>arXiv preprint arXiv:2010.02573</i>. 3-Bostrom, N. (2014). <i>Superintelligence: Paths, dangers, strategies</i>. Oxford University Press. 4-Floridi, L., & Cows, J. (2019). A unified framework of five principles for AI in society. <i>Harvard Data Science Review</i>, 1(1). 5-Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. <i>Nature Machine Intelligence</i>, 1(9), 389-399. 6-Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. <i>Big Data & Society</i>, 3(2), 2053951716679679. 7-Price, W. N., & Cohen, I. G. (2019). Privacy in the age of medical big data. <i>Nature Medicine</i>, 25(1), 37-43. 8-Russell, S., & Norvig, P. (2016). <i>Artificial intelligence: A modern approach</i>. Pearson. 9-Taddeo, M., & Floridi, L. (2018). Regulate artificial intelligence to avert cyber arms race. <i>Nature</i>, 556(7701), 296-298. 10-Zeng, S., & Besold, T. R. (2018). Ethics in machine learning and data science: What should we teach? <i>arXiv preprint arXiv:1811.11130</i>. <p>Remember, this is not an exhaustive list, and there may be other relevant articles and authors to explore. Conducting a thorough literature review will help you gather diverse perspectives and insights on the ethical implications of non-human-centered designs in academic paper design, as well as identify potential solutions to address these ethical challenges.</p>

Figure 5. Literature review text generated by *ChatGPT-3.5* and *ChatGPT-4*

ses. Furthermore, the lack of explicit citations in the text poses a risk of potential plagiarism, assuming that the content is based on external sources. The absence of proper citations and references makes it challenging to evaluate the accuracy and credibility of the information presented in the articles.

4.4. Identification of the literature review

In the fourth prompt, the author asks for a comprehensive literature review to design an article (questions 6 of *ChatGPT-3.5* under the heading “Presentation of the general purpose and approach of the article”). The purpose of this prompt is to test how to design a literature review for an article.

In this study design, *ChatGPT* conducted a literature review on the topic. Prior to conducting the literature review, *ChatGPT-3.5* formulated a search strategy and identified relevant keywords. Subsequently, the literature review was conducted, resulting in the inclusion of ten references. Regrettably, none of these sources could be located or verified. On the other hand, *ChatGPT-4* presented ten references, out of which nine were successfully verified, while one remained unverifiable. The findings of this investigation suggest that *ChatGPT-4* demonstrates greater reliability and validity in terms of its literature review compared to *ChatGPT-3.5*. However, concerns arise due to the limited availability of literature sources in *ChatGPT-4* and the inability to verify all the recommended references. Consequently, the evaluation of the accuracy and credibility of the information presented in the literature becomes a challenging task.

4.5. Creating writing assistance

In the fifth prompt, the author is asking for writing assistance from *ChatGPT* to design an article, stating his purpose and target audience (questions 7 of *ChatGPT-3.5* under the heading “Presentation of the general purpose and approach of

Question	ChatGPT-3.5	ChatGPT-4
<p>I want to write an academic article about “What does the future of academic article design look like with the integration of ChatGPT and other machine learning technologies?” My goal is to make various scenarios about the future of this topic. My target audience is academics interested in this topic. Can you provide writing assistance?</p>	<p>Sure, I can provide writing assistance for your academic article. Here are some steps you can follow to get started:</p> <p>1-Introduction: In the introduction, you can provide a brief overview of the topic and its significance. You can also introduce the main focus of your article, which is to explore the future of academic article design with the integration of ChatGPT and other machine learning technologies.</p> <p>2- Literature review: In the literature review section, you can discuss the existing research on the topic of academic article design and the integration of machine learning technologies. You can also highlight the limitations of the existing research and explain how your article can contribute to the field.</p> <p>3-Methodology: In the methodology section, you can explain the approach you used to analyze the data and make predictions about the future of academic article design with the integration of ChatGPT and other machine learning technologies. This could involve conducting surveys or interviews with experts in the field, analyzing trends and developments in academic publishing, and reviewing existing research on the topic.</p> <p>4-Results: In the results section, you can present the various scenarios you have developed about the future of academic article design with the integration of ChatGPT and other machine learning technologies. You can explain the implications of each scenario and how it could impact the academic publishing industry.</p> <p>5-Discussion: In the discussion section, you can interpret your results and provide insights into what they mean for the future of academic article design. You can also discuss the limitations of your study and suggest areas for future research.</p> <p>6-Conclusion: In the conclusion, you can summarize the main findings of your article and discuss their implications for the academic publishing industry. You can also highlight the significance of your research and suggest avenues for future research.</p> <p>As you write your article, be sure to keep your target audience in mind and use language and terminology that is appropriate for academics in the field. Also, make sure to properly cite your sources and adhere to any formatting guidelines provided by your target publication.</p>	<p>Certainly! I can provide writing assistance for your academic article on the future of academic article design with the integration of ChatGPT and other machine learning technologies. Here's an outline to help you structure your article and explore various scenarios:</p> <p>1-Introduction</p> <p>Introduce the topic of the future of academic article design in the context of the integration of ChatGPT and other machine learning technologies.</p> <p>Explain the significance of this topic and its relevance to academics.</p> <p>State the purpose of the article and provide an overview of the scenarios to be explored.</p> <p>2-Current Landscape of Academic Article Design</p> <p>Provide an overview of the current practices and challenges in academic article design.</p> <p>Discuss traditional approaches to article design and the limitations they may have.</p> <p>Highlight the emergence and integration of machine learning technologies, particularly ChatGPT, in academic writing processes.</p> <p>3-Scenario 1: Enhanced Efficiency and Productivity</p> <p>Describe a scenario where the integration of ChatGPT and other machine learning technologies streamlines the article design process.</p> <p>Discuss how AI assistance can automate time-consuming tasks, such as formatting, referencing, and proofreading, leading to increased efficiency and productivity for researchers.</p> <p>4-Scenario 2: Augmented Idea Generation and Creativity</p> <p>Present a scenario where ChatGPT and other machine learning technologies act as creative collaborators.</p> <p>Explore how AI models can generate novel ideas, suggest alternative research directions, and inspire researchers to think outside the box.</p> <p>Discuss the potential for AI to contribute to interdisciplinary research and foster innovative approaches.</p> <p>5-Scenario 3: Personalized Writing Assistance</p> <p>Envision a scenario where machine learning technologies provide personalized writing assistance tailored to individual researchers' needs.</p> <p>Discuss how AI models can offer real-time feedback, suggest improvements in writing style, enhance language usage, and adapt to researchers' preferences.</p> <p>Highlight the potential for AI to assist non-native English speakers and researchers in improving the clarity and impact of their writing.</p> <p>6-Scenario 4: Ethical Considerations and Responsible Integration</p> <p>Explore a scenario that addresses the ethical implications of integrating ChatGPT and machine learning technologies in academic article design.</p> <p>Discuss the importance of ensuring transparency, avoiding biases, maintaining authorship integrity, and addressing concerns related to intellectual property and plagiarism.</p> <p>Highlight the need for responsible integration, human oversight, and ethical guidelines in the use of AI assistance in academic writing.</p> <p>7-Conclusion</p> <p>Summarize the scenarios discussed and their implications for the future of academic article design.</p> <p>Reflect on the potential benefits and challenges associated with the integration of ChatGPT and machine learning technologies.</p> <p>Encourage further research, collaboration, and thoughtful exploration of the evolving role of AI in academic writing.</p> <p>Remember to support your scenarios with relevant examples, scholarly references, and empirical evidence. Additionally, maintain a balance between the potential positive outcomes and the challenges that need to be addressed to ensure the responsible integration of ChatGPT and other machine learning technologies in academic article design.</p>

Figure 6. Writing assistance text generated by *ChatGPT-3.5* and *ChatGPT-4*

the article”). The purpose of this prompt is to test whether *ChatGPT* can design the article in accordance with the purpose and audience of the topic.

In this design, *ChatGPT-3.5* and *ChatGPT-4* both commence the article with an introductory section, where they elucidate the role played by *ChatGPT* and other machine learning technologies in the realm of academic article design. The two versions propose distinct scenarios within the article. While *ChatGPT-3.5* does not designate these scenarios as a discrete section, *ChatGPT-4* presents them under separate headings. *ChatGPT-3.5* offers a simplified presentation of the article’s title, whereas *ChatGPT-4* puts forth a more precise formulation. *ChatGPT-4* incorporates a dedicated section titled “Current landscape of academic article design” which delves into the existing practices in this field. In contrast, *ChatGPT-3.5* does not propose such a section. The delineation of scenarios as a separate section is not explicitly specified in *ChatGPT-3.5*, whereas *ChatGPT-4* presents each scenario under distinct headings. *ChatGPT-3.5* provides a concise summary of the article, while *ChatGPT-4* furnishes a more detailed summary that encapsulates the overarching purpose and content of the article.

Both *ChatGPT-3.5* and *ChatGPT-4* explore the future of academic article design by integrating *ChatGPT* and other machine learning technologies. They exemplify the impact of these technologies on academic article design through the presentation of diverse scenarios. Furthermore, both versions emphasize the importance of ethical considerations and responsibilities within the realm of academic article design. *ChatGPT-4* introduces a more intricate structural framework for the article, incorporating a section titled “Current landscape of academic article design,” which provides a comprehensive analysis of the existing practices in this field. *ChatGPT-4* presents the scenarios in a more precise manner, offering detailed explanations of the potential implications of *ChatGPT* and other machine learning technologies on the design of academic articles. Conversely, *ChatGPT-3.5* adopts a more general perspective and accords less emphasis to the specifics of the scenarios. *ChatGPT-4* advocates for a section dedicated to evaluating the impact of the scenarios, wherein the potential

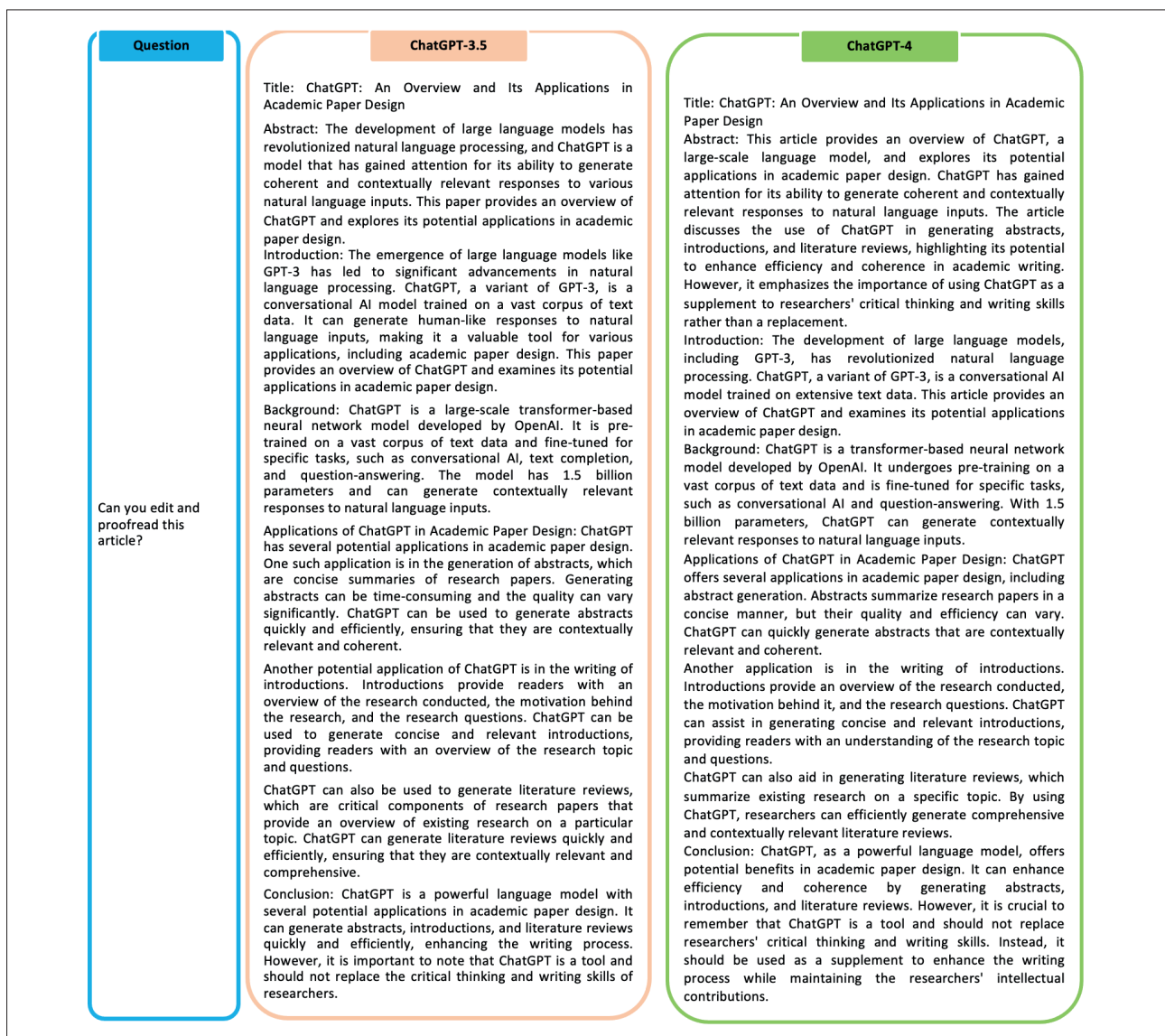


Figure 7. Editing and proofreading text generated by *ChatGPT-3.5* and *ChatGPT-4*

contributions, advantages, and disadvantages of each scenario on academic article design are discussed. *ChatGPT-3.5* does not include such a section. *ChatGPT-4* concludes the article with a summary of the scenarios, wherein the key points and general conclusions of each scenario are highlighted. *ChatGPT-3.5* does not incorporate such a summary section. In light of this comparison, it is evident that *ChatGPT-4* offers a more intricate structure and presents scenarios in a more specific manner. Conversely, *ChatGPT-3.5* adopts a broader perspective and allocates less focus to the details of the scenarios. Both versions explore the integration of *ChatGPT* and other machine learning technologies in the context of academic article design, although they differ in their treatment of specific subjects and the level of analysis they offer.

4.6. Editing and proofreading

In the sixth prompt, the author asks to edit and proofread an article (questions 1 of *ChatGPT-3.5* under the heading “Presentation of the general purpose and approach of the article”) created by *ChatGPT*. The purpose of this prompt is to test how *ChatGPT* will edit and proofread an article written by *ChatGPT* itself. The article has not been rewritten; you can find the article above.

ChatGPT-3.5 has made changes to the text to improve its comprehensibility. However, these changes primarily focus on basic corrections such as improving sentence structures and correcting grammatical errors. In *ChatGPT-4*, the text was meticulously revised and corrected. In this editing process, more attention was paid to sentence patterns, grammatical errors, and spelling. The text has been revised to ensure greater academic rigor and adherence to academic norms. This comparison reveals that the *ChatGPT-4* shows higher potential and achievement in proofreading. *ChatGPT-4* has further improved its adherence to academic norms and increased its comprehensibility by organizing the text more comprehensively. The use of language was made sharper and more effective, resulting in a text more suitable for an academic paper.

As a result of *ChatGPT-4*'s extensive revisions, which bring the text in line with academic norms and improve comprehensibility, the potential and success of editing becomes more evident. *ChatGPT-4* provides precise and concise changes that *ChatGPT-3.5* does not, while *ChatGPT-3.5* makes general revisions to improve comprehensibility.

5. Discussion

In this article, the potential of *ChatGPT* in academic article design has been explored. The study aims to investigate the capabilities, limitations, and overall suitability of both *ChatGPT-3.5* and *ChatGPT-4* in the context of academic article design. The overall aim of the paper is to provide suggestions for the use of language models such as *ChatGPT* in academic paper design and to discuss the effects of non-human-centered design approaches on academic writing processes. The results of this study will provide a basis for future research by discussing the potential benefits and limitations of using AI technologies and non-human-centered design approaches.

The article design process involves conversations between *ChatGPT-3.5* and *ChatGPT-4* and the author through a series of chat interfaces. The author asks questions and proposes ideas, while *ChatGPT-3.5* and *ChatGPT-4* provide feedback and suggestions and generate text based on the input. The methods of the study involve providing guiding questions to *ChatGPT* for writing articles and analyzing *ChatGPT*'s responses. The responses produced by *ChatGPT* during the research process were evaluated in detail by another expert academic. This expert has assessed the text generated by *ChatGPT* in terms of accuracy, consistency, and adherence to academic standards. At the end of the evaluation process, an agreement was reached with the author, thus eliminating the need for any revisions or final approval.

ChatGPT has demonstrated itself to be a reasonably competent conversational AI that can support the design process of academic papers in various ways. *ChatGPT-3.5* and *ChatGPT-4* exhibits the following capabilities in assisting with academic article design: 1) aiding in the design process, 2) generating ideas aligned with the overall purpose and focus of the paper, 3) producing consistent and contextually relevant responses to various natural language inputs, 4) partially assisting in literature reviews, 5) supporting paper design in terms of both content and format, and 6) providing reasonable editing and proofreading for articles. On the other hand, *ChatGPT-4*, in addition to the capabilities of *ChatGPT-3.5*, has shown potential as a valuable tool in the areas of critical evaluation and drafting and refining.

When evaluating the article design capabilities of *ChatGPT*, it was observed that both *ChatGPT-3.5* and *ChatGPT-4* demonstrated the ability to produce scholarly articles using a common format. The article design approach employed by *ChatGPT-3.5* exhibited a more generalized nature, while *ChatGPT-4* demonstrated a more focused approach. By offering the article writer guiding questions and advice on which ones to choose that are the most pertinent and interesting, the answer of *ChatGPT-3.5* uses an inquiry-based approach. The response from *ChatGPT-4*, on the other hand, provides a more detailed outline by suggesting a structure and heading for the piece. *ChatGPT-4* highlights specific subjects and offers recommendations. It also contains a distinct part on critical evaluation, which is absent from *ChatGPT-3.5*. Although the responses for the *ChatGPT-3.5* and *ChatGPT-4* cover the same ground, there are differences in their structure and concentration. While evaluating *ChatGPT*'s writing assistance capability to design the article in accordance with the purpose and target audience of the topic, it was observed that both versions suggested different scenarios within the article. While *ChatGPT-3.5* does not designate these scenarios as separate sections, *ChatGPT-4* presents them under distinct headings. *ChatGPT-3.5*, while providing writing assistance, possesses a broader perspective and focuses less on the details of the scenarios. In contrast, *ChatGPT-4* presents a more detailed structure and addresses the scenarios in a more specific manner. They exhibit differences in their approach to specific topics and the level of analysis they provide. When *ChatGPT*'s ability

to edit and proofread an article was evaluated, it was observed that *ChatGPT-3.5* made general revisions to improve comprehensibility, while *ChatGPT-4* provided precise and concise changes that *ChatGPT-3.5* did not. *ChatGPT-4* made the text more in line with academic norms and increased comprehensibility. During editing and proofreading, *ChatGPT-4* corrected spelling and punctuation errors, chose more appropriate verbs, provided clarity and context to statements, highlighted important points, gave details, and even added sentences to emphasize the importance of the topic.

The traditional structure of an academic article typically includes sections such as abstract, introduction, literature review, methods, findings, discussion, and conclusion. *ChatGPT* places emphasis on the creation of abstracts, introductions, and literature reviews, effectively guiding the article design process. By providing headings that can be utilized in an article and offering guidance on the content to be included under each heading, *ChatGPT* successfully facilitates the design of academic articles. Furthermore, in the methodology section, *ChatGPT* suggests relevant methods that are applicable to the given topic.

ChatGPT is deeply changing the design practice. The advantages of using *ChatGPT* in the academic article design process include increased efficiency, improved organization, and the ability to produce content quickly. With the use of *ChatGPT*, a consistent draft for the article was created quickly, the language and structure of each section were improved, and even some of the content was generated by *ChatGPT* itself. As many authors have claimed, *ChatGPT* has demonstrated potential to provide new discoveries and insights for academic article design. It has significantly reduced energy and time waste by creating article designs similar to those produced by humans (Aljanabi et al., 2023; Alshater, 2022; Aydin; Karaarslan, 2022; Castillo-Gonzalez, 2022; Curtis, *ChatGPT*, 2023; Gilat; Cole, 2023; Golan et al., 2023; Huang; Tan, 2023; Hutson, 2022; Lin, 2023; Nguyen, 2023; O'Connor; *ChatGPT*, 2022; Rudolph; Tan; Tan, 2023; Túnéz-López; Fieiras-Ceide; Vaz-Álvarez, 2021). Studies conducted by Aljanabi et al. (2023), Hammad (2023), and O'Connor and *ChatGPT* (2022) have shown that *ChatGPT* can understand the objectives behind a problem and provide users with the information they need more quickly and effectively. However, technology limitations were also evident, especially in maintaining consistency and ensuring the accuracy of the generated text.

ChatGPT is capable of quickly and efficiently generating articles, but it lacks the ability to create accurate and complete citations and references. In this study, the inclusion of proper in-text citations and references was explicitly requested; however, neither response successfully met this requirement. For the in-text citations, no sources were provided in either response. Moreover, *ChatGPT-3.5* referenced three sources in the paper, but upon further investigation, it was discovered that none of these sources could be located or verified. As for *ChatGPT-4*, although it included a total of 10 references, only two of them were identifiable and accessible. Consequently, the credibility and reliability of the sources mentioned in both responses are questionable. Furthermore, since the citations and sources are not clearly indicated, there is a risk of plagiarism and academic misconduct, as expressed by some authors (Alkaissi; Mcfarlane, 2023; Alshater, 2022; Bang et al., 2023; Caira; Russo; Aranda, 2023; Cotton; Cotton; Shipway, 2023; Gao et al., 2022; Ma et al., 2023; Williamson; Macgilchrist; Potter, 2023; Perkins, 2023; Sabzalieva; Valentini, 2023; Srivastava, 2023; Stokel-Walker, 2023; Sullivan; Kelly; McLaughlan, 2023; Thorp, 2023). Considering these reasons, and especially *ChatGPT's* inability to access literature beyond 2021, its use as a reliable information source is limited. Even if *ChatGPT-4* tries to overcome this limitation using web browser plugins, it is important to note that the model is trained on internet data until 2021.

On the other hand, *ChatGPT* has shown various limitations in understanding some issues and requests, as well as in creating or completing citations and literature reviews. Although *ChatGPT* can create designs and texts suitable for an academic article, it has many shortcomings when looking at the article examples it has written. At this point, there is also a concern that it may lead to a lack of critical thinking, originality and analysis and provide biased information (Bishop, 2023; Dutton, 2023; Huang; Tan, 2023; İskender, 2023; Nguyen, 2023). In addition, there are concerns that *ChatGPT* may not represent the moral, social and cultural ideals that academics value (OpenAI, 2022). When considered in the context of commercialization purpose (Huang; Tan, 2023), it should not be overlooked that *ChatGPT* is not conscious and lacks self-awareness (Pavlik, 2023).

Nori et al. (2023) have previously emphasized the importance of studying the progress and capabilities of these two *ChatGPT* models when the expected launch of *GPT-4* takes place. The main difference between *ChatGPT-3.5* and *ChatGPT-4* is that *ChatGPT-4* allows users to send up to 25,000 words, an eightfold increase from the number allowed by *ChatGPT*. In addition, OpenAI claims that their latest technology produces fewer errors, so-called "hallucinations". In the past, *ChatGPT* could disappear, give nonsensical answers to your query, or even send stereotypical or incorrect information. In addition, *ChatGPT-4* is more capable of expressing creativity and manipulating language (Hughes, 2023; OpenAI, 2023). A study conducted by Rahaman et al. (2023) shows that *ChatGPT-4* is significantly more successful than *ChatGPT-3.5* at translating languages, answering questions, and understanding human emotions. The study emphasizes that *ChatGPT-4* outperforms its previous version in every respect. Another study by Metz and Collins (2023) shows that *ChatGPT-4* successfully passed the bar exam and various standardized tests by summarizing and interpreting both visuals and texts. They also observed a significant increase in accuracy compared to *ChatGPT-3.5*. *ChatGPT-4* is 40% more likely to produce real answers than *ChatGPT-3.5* (OpenAI, 2023). On the other hand, Chen, Zaharia and Zou (2023) evaluated the March 2023 and June 2023 versions of *GPT-3.5* and *GPT-4* on four different tasks: solving math problems, answering sensitive/dangerous questions, generating code, and visual reasoning. They found that the performance and behavior of both models could change significantly over time. For example, *GPT-4* was excellent at identifying prime numbers in

March 2023 (accuracy 97.6%) but performed poorly on the same questions in June 2023 (accuracy 2.4%). Interestingly, *GPT-3.5* performed much better on this task in June 2023 compared to March 2023.

A study conducted by **Cao** (2023) focused on the ability of *ChatGPT-4* to cite scientific articles. This study found a significant inconsistency in the performance of *ChatGPT-4* across different fields. This trend highlights a significant limitation in the functionality of *ChatGPT-4*'s scientific reference citation and indicates a need for improvements in specificity and validity. The same research found that when queried about topics requiring expertise, *ChatGPT-4* tends to prioritize more general topics. This situation, especially when it comes to academic research requiring expertise, necessitates continuous human intervention and verification process to ensure the applicability of the model's outputs (**Huang; Tan, 2023; Zhang; Gosline, 2022**). Another study by **AlAfnan and Mohdzuki** (2023) indicates that *ChatGPT-4* uses declarative, imperative, and interrogative moods. The declarative mood is typical in academic writing responses, while the imperative mood is typical in responses related to case studies.

6. Conclusion

In this study, it is possible to observe how much and how *ChatGPT-3.5* and *ChatGPT-4* intervened in the design of an academic article based on their suggestions. This research is in line with the goal of exploring the role of *ChatGPT* in academic article design in terms of human-machine collaboration.

The use of *ChatGPT* in academic article design represents a step towards non-human-centered design in the academic world. Although there are limitations to the technology, it also offers many opportunities to increase efficiency and creativity in the research process. As artificial intelligence technology continues to develop, it is important to carefully evaluate the potential benefits and drawbacks of using non-human-centered design methods in academic research. Despite the shortcomings of both, *ChatGPT-3.5* to some extent, and *ChatGPT-4* more broadly, have the potential to logically and appropriately specify the steps to follow in the design of an academic article. In the study, it was determined that both *ChatGPT-3.5* and *ChatGPT-4* can serve as a valuable tool in academic article design. Therefore, as noted by **Farias** (2023), *ChatGPT* has demonstrated considerable functionality as a tool that can be used in the preparation of academic/scientific texts. However, while *ChatGPT* offers advantages such as accelerated writing, enhanced productivity, and comprehensive responses, it also brings potential disadvantages including a decrease in critical thinking, concerns about plagiarism, misinformation, and a lack of originality and innovation. In order to mitigate such disadvantages, it is indispensable to establish precise guidelines for the ethically compliant use of *ChatGPT*.

Future research should continue to explore the impact of *ChatGPT* and similar designs on research quality and validity. This way, the role of artificial intelligence technology in academic research can be further explored. Furthermore, future studies could be initiated to compare the results obtained with *ChatGPT* against other systems, such as Google's *Bard* or Microsoft's *Chat Bing*. In addition, it may be illuminating to compare the results obtained with *ChatGPT* with the results obtained with AI research assistants capable of finding sources, citing references, summarizing results, or producing relevant text, such as *Elicit*, *Scite*, *SciSpace*, *Research Rabbit*, *Connected Papers*, *Perplexity*, *Paperpal*, and *Litmaps*.

7. References

- AlAfnan, Mohammad-Awad; Mohdzuki, Siti-Fatimah** (2023). "Do artificial intelligence chatbots have a writing style? An investigation into the stylistic features of ChatGPT-4". *Journal of Artificial intelligence and technology*, v. 3, n. 3. <https://doi.org/10.37965/jait.2023.0267>
- Aljanabi, Mohammad** (2023). "ChatGPT: Future directions and open possibilities". *Mesopotamian journal of cybersecurity*, pp. 16-17. <https://doi.org/10.58496/MJCS/2023/003>
- Aljanabi, Mohammad; Ghazi, Mohanad; Ali, Ahmed-Hussein; Abed, Saad-Abas; ChatGPT** (2023). "ChatGpt: Open possibilities". *Iraqi journal for computer science and mathematics*, v. 4, n. 1, pp. 62-64. <https://doi.org/10.52866/ijcsm.2023.01.01.0018>
- Alkaissi, Hussam; Mcfarlane, Samy I.** (2023). "Artificial hallucinations in *ChatGPT*: implications in scientific writing". *Cureus*, v. 15, n. 2. <https://doi.org/10.7759/cureus.35179>
- Alshater, Muneer M.** (2022). "Exploring the role of artificial intelligence in enhancing academic performance: A case study of *ChatGPT*". *SSRN*. <https://doi.org/10.2139/ssrn.4312358>
- Altman, Sam** (2022). "ChatGPT launched on Wednesday. Today it crossed 1 million users!". *Twitter*, December 5. [@sama]. <https://twitter.com/sama/status/1599668808285028353>
- Amodei, Dario; Ananthanarayanan, Sundaram; Anubhai, Rishita; Bai, Jingliang; Battenberg, Eric; Case, Carl; ... ; Zhu, Zhenyao** (2016). "Deep speech 2: End-to-end speech recognition in English and mandarin". In: *Proceedings of the 33rd International conference on machine learning*, pp. 173-182. <http://proceedings.mlr.press/v48/amodei16.html?ref=https://codemonkey.link>

- Anderson, Nash; Belavy, Daniel L.; Perle, Stephen M.; Hendricks, Sharief; Hespanhol, Luiz; Verhagen, Evert; Memon, Aamir R.** (2023). "AI did not write this manuscript, or did it? Can we trick the AI text detector into generated texts? The potential future of *ChatGPT* and AI in sports & exercise medicine manuscript generation". *BMJ open sport & exercise medicine*, v. 9, n. 1, e001568.
<https://doi.org/10.1136/bmjsem-2023-001568>
- Ariyaratne, Sisith; Iyengar, Karthikeyan P.; Nischal, Neha; Babu, Naparla-Chitti; Botchu, Rajesh** (2023). "A comparison of ChatGPT-generated articles with human-written articles". *Skeletal radiology*, n. 52.
<https://doi.org/10.1007/s00256-023-04340-5>
- Aydın, Ömer; Karaarslan, Enis** (2022). "OpenAI ChatGPT generated literature review: Digital twin in healthcare". In: Aydın, Ömer (ed.). *Emerging computer technologies 2*. İzmir Akademi Derneği, pp. 22-31.
<https://doi.org/10.2139/ssrn.4308687>
- Baeza-Yates, Ricardo** (2022). "Ethical challenges in AI". In: *Proceedings of the fifteenth ACM international conference on web search and data mining*, v. 15.
<https://doi.org/10.1145/3488560.3498370>
- Bang, Yejin; Cahyawijaya, Samuel; Lee, Nayeon; Dai, Wenliang; Su, Dan; Wilie, Bryan; Lovenia, Holy; Ji, Ziwei; Yu, Tiezheng; Chung, Willy; Do, Quyet V.; Xu, Yan; Fung, Pascale** (2023). "A multitask, multilingual, multimodal evaluation of *ChatGPT* on reasoning, hallucination, and interactivity". *ArXiv preprint*. ArXiv:2302.04023.
<https://doi.org/10.48550/arXiv.2302.04023>
- Basic, Zeljana; Banovac, Ana; Kruzic, Ivana; Jerkovic, Ivan** (2023). "Better by you, better than me, *Chatgpt3* as writing assistance in students essays". *ArXiv preprint*. ArXiv:2302.04536.
<https://doi.org/10.48550/arXiv.2302.04536>
- Bishop, Lea** (2023). "A computer wrote this paper: What ChatGPT means for education, research, and writing". *Research, and writing*, January 26.
<https://doi.org/10.2139/ssrn.4338981>
- Borji, Ali** (2023). "A categorical archive of *ChatGPT* failures". *ArXiv preprint*. ArXiv:2302.03494.
<https://doi.org/10.48550/arXiv.2302.03494>
- Caira, Celine; Russo, Lucia; Aranda, Luis** (2023). "Artificially inequitable? AI and closing the gender gap". *OECD.AI. Policy observatory*, March 8.
<https://oecd.ai/en/wonk/closing-the-gender-gap>
- Cao, Zhi** (2023). "Evaluating the efficacy of *ChatGPT-4* in providing scientific references across diverse disciplines". *ArXiv preprint*. ArXiv:2306.09914.
<https://doi.org/10.48550/arXiv.2306.09914>
- Castillo-González, William** (2022). "ChatGPT and the future of scientific communication". *Metaverse basic and applied research*, v. 1, n. 8.
<https://doi.org/10.56294/mr20228>
- Chen, Lingjiao; Zaharia, Matei; Zou, James** (2023). "How is *ChatGPT*'s behavior changing over time?". *ArXiv preprint*. ArXiv:2307.09009.
<https://doi.org/10.48550/arXiv.2307.09009>
- Cotton, Debby R. E.; Cotton, Peter A.; Shipway, J. Reuben** (2023). "Chatting and cheating. Ensuring academic integrity in the era of *ChatGPT*".
https://edarxiv.org/mrz8h?trk=public_post_main-feed-card_reshare-text
- Crowe, Sarah; Cresswell, Kathrin; Robertson, Ann; Huby, Guro; Avery, Anthony; Sheikh, Aziz** (2011). "The case study approach". *BMC medical research methodology*, v. 11, n. 1.
<https://doi.org/10.1186/1471-2288-11-100>
- Curtis, Nigel; ChatGPT** (2023). "To *ChatGPT* or not to *ChatGPT*? The impact of artificial intelligence on academic publishing". *The pediatric infectious disease journal*, v. 42, n. 4, p. 275.
<https://doi.org/10.1097/INF.0000000000003852>
- De-Farias, Salomão-Alencar** (2023). "Pânico na Academia! Inteligência artificial na construção de textos científicos com o uso do *ChatGPT*". *Revista interdisciplinar de marketing*, v. 13, n. 1, pp. 79-83.
- Dellermann, Dominik; Calma, Adrian; Lipusch, Nikolaus; Weber, Thorsten; Weigel, Sascha; Ebel, Philipp** (2021). "The future of human-AI collaboration: a taxonomy of design knowledge for hybrid intelligence systems". *ArXiv preprint*. ArXiv:2105.03354.
<https://doi.org/10.48550/arXiv.2105.03354>

- Dutton, Jonathan J.** (2023). "Artificial intelligence and the future of computer-assisted medical research and writing". *Ophthalmic plastic and reconstructive surgery*, v. 39, n. 3, pp. 203-205.
<https://doi.org/10.1097/IOP.0000000000002420>
- Elsevier* (2023). "The use of generative AI and AI-assisted technologies in scientific writing". *Publishing ethics*.
<https://www.elsevier.com/about/policies/publishing-ethics#>
- Fidel, Raya** (1984). "The case study method: A case study". *Library and information science research*, v. 6, n. 3, pp. 273-288.
<http://faculty.washington.edu/fidelr/RayaPubs/TheCaseStudyMethod.pdf>
- Flanagan, Johansson** (1999). "Public participation in the design of educational programmes for cancer nurses: a case report". *European journal of cancer care*, v. 8, n. 2, pp. 107-112.
<https://doi.org/10.1046/j.1365-2354.1999.00141.x>
- Flanagin, Annette; Bibbins-Domingo, Kirsten; Berkwits, Michael; Christiansen, Stacy L.** (2023). "Nonhuman 'authors' and implications for the integrity of scientific publication and medical knowledge". *JAMA*, v. 329, n. 8, pp. 637-639.
<https://doi.org/10.1001/jama.2023.1344>
- Gao, Catherine A.; Howard, Frederick M.; Markov, Nikolay S.; Dyer, Emma C.; Ramesh, Siddhi; Luo, Yuan; Pearson, Alexander T.** (2022). "Comparing scientific abstracts generated by *ChatGPT* to original abstracts using an artificial intelligence output detector, plagiarism detector, and blinded human reviewers". *bioRxiv*, 2022-12.
<https://doi.org/10.1101/2022.12.23.521610>
- Gilat, Ron; Cole, Brian J.** (2023). "How will artificial intelligence affect scientific writing, reviewing and editing? The future is here". *Arthroscopy*, v. 39, n. 5, pp. 1119-1120.
<https://doi.org/10.1016/j.arthro.2023.01.014>
- Gilson, Lucy** (2012). "The case study approach". In: Gilson, Lucy (ed.). *Health policy and systems research. A methodology reader*. Geneva: Alliance for Health Policy and Systems Research, World Health Organisation, pp. 161-165. ISBN: 978 92 4 150313 6
https://assets.publishing.service.gov.uk/media/57a08a93ed915d622c0007db/Alliancehpsr_Reader.pdf
- Golan, Roei; Reddy, Rohit; Muthigi, Akhil; Ramasamy, Ranjith** (2023). "Artificial intelligence in academic writing: a paradigm-shifting technological advance". *Nature reviews urology*, 24 February.
<https://doi.org/10.1038/s41585-023-00746-x>
- Grimaldi, Gianluca; Ehrler, Bruno** (2023). "AI et al.: Machines are about to change scientific publishing forever". *ACS energy letters*, v. 8, n. 1, pp. 878-880.
<https://doi.org/10.1021/acsenergylett.2c02828>
- Gustafsson, Johanna** (2017). *Single case studies vs. multiple case studies: a comparative study*.
<http://www.diva-portal.org/smash/get/diva2:1064378/FULLTEXT01.pdf>
- Hammad, Mohamed** (2023). "The impact of artificial intelligence (AI) programs on writing scientific research". *Annals of biomedical engineering*, v. 51, pp. 459-460.
<https://doi.org/10.1007/s10439-023-03140-1>
- Huang, Jingshan; Tan, Ming** (2023). "The role of *ChatGPT* in scientific communication: writing better scientific review articles". *American journal of cancer research*, v. 13, n. 4, pp. 1148-1154.
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10164801>
- Hughes, Alex** (2023). "ChatGPT: Everything you need to know about OpenAI's GPT-4 tool". *Science focus*, September 1.
<https://www.sciencefocus.com/future-technology/gpt-3>
- Hutson, Matthew** (2022). "Could AI help you to write your next paper?". *Nature*, v. 611, n. 7934, pp. 192-193.
<https://doi.org/10.1038/d41586-022-03479-w>
- Kitamura, Felipe C.** (2023). "ChatGPT is shaping the future of medical writing but still requires human judgment". *Radiology*, v. 307, n. 2, e230171.
<https://doi.org/10.1148/radiol.230171>
- Liedtka, Jeanne** (2015). "Perspective: Linking design thinking with innovation outcomes through cognitive bias reduction". *Journal of product innovation management*, v. 32, n. 6, pp. 925-938.
<https://doi.org/10.1111/jpim.12163>
- Lin, Zhicheng** (2023). "Why and how to embrace AI such as *ChatGPT* in your academic life". *PsyArXiv*, Preprint.
<https://psyarxiv.com/sdx3j>
- Lock, Samantha** (2022). "What is ai chatbot phenomenon *ChatGPT* and could it replace humans?". *The Guardian*, December 5.
<https://www.theguardian.com/technology/2022/dec/05/what-is-ai-chatbot-phenomenon-chatgpt-and-could-it-replace-humans>

- Lund, Brady D.; Wang, Ting** (2023). "Chatting about *ChatGPT*: How may AI and *GPT* impact academia and libraries?". *Library hi tech news*, v. 40, n. 3, pp. 26-29.
<https://doi.org/10.1108/LHTN-01-2023-0009>
- Lund, Brady D.; Wang, Ting; Mannuru, Nishith-Reddy; Nie, Bing; Shimray, Somipam; Wang, Ziang** (2023). "*ChatGPT* and a new academic reality: Artificial intelligence-written research papers and the ethics of the large language models in scholarly publishing". *Journal of the Association for Information Science and Technology*, v. 74, n. 5, pp. 570-581.
<https://doi.org/10.1002/asi.24750>
- Ma, Yongqiang; Liu, Jiawei; Yi, Fan; Cheng, Qikai; Huang, Yong; Lu, Wei; Liu, Xiaozhong** (2023). "Is this abstract generated by AI? A research for the gap between AI-generated scientific text and human-written scientific text". *arXiv preprint arXiv:2301.10416*.
<https://doi.org/10.48550/arXiv.2301.10416>
- Metz, Cade; Collins, Keith** (2023). "10 ways *GPT-4* is impressive but still flawed". *The New York Times*, March 14.
<https://www.nytimes.com/2023/03/14/technology/openai-new-gpt4.html>
- Mijwil, Maad; Aljanabi, Mohammad; ChatGPT** (2023). "Towards artificial intelligence-based cybersecurity: The practices and *ChatGPT* generated ways to combat cybercrime". *Iraqi journal for computer science and mathematics*, v. 4, n. 1, pp. 65-70.
<https://doi.org/10.52866/ijcsm.2023.01.01.0019>
- Mijwil, Maad; Filali, Youssef; Aljanabi, Mohammad; Bounabi, Mariem; Al-Shahwani, Humam; ChatGPT** (2023). "The purpose of cybersecurity governance in the digital transformation of public services and protecting the digital environment". *Mesopotamian journal of cybersecurity*.
<https://doi.org/10.58496/MJCS/2023/001>
- Nguyen, Minh-Hoang** (2023). *Academic writing and AI: Day-2 experiment with Bayesian mindsponge framework*. Center for Open Science.
<https://econpapers.repec.org/paper/osfosfxxx/kr29c.htm>
- Nori, Harsha; King, Nicholas; McKinney, S. Mayer; Carignan, Dean; Horvitz, Eric** (2023). "Capabilities of *GPT-4* on medical challenge problems". *ArXiv*. Preprint. ArXiv:2303.13375.
<https://doi.org/10.48550/arXiv.2303.13375>
- O'Connor, Siobhan; ChatGPT** (2022). "Open artificial intelligence platforms in nursing education: Tools for academic progress or abuse?". *Nurse education in practice*, v. 66.
<https://doi.org/10.1016/j.nepr.2022.103537>
- OpenAI** (2022). *ChatGPT-3: Optimizing language models for dialogue*.
<https://chat.openai.com/chat>
- OpenAI** (2023). *ChatGPT 4*.
<https://openai.com/gpt-4>
- Ouyang, Fan; Zheng, Luyi; Jiao, Pengcheng** (2022). "Artificial intelligence in online higher education: A systematic review of empirical research from 2011 to 2020". *Education and information technologies*, v. 27, n. 6, pp. 7893-7925.
<https://doi.org/10.1007/s10639-022-10925-9>
- Pavlik, John V.** (2023). "Collaborating with *ChatGPT*: Considering the implications of generative artificial intelligence for journalism and media education". *Journalism & mass communication educator*, v. 78, n. 1, pp. 84-93.
<https://doi.org/10.1177/10776958221149577>
- Perkins, Mike** (2023). "Academic integrity considerations of AI large language models in the post-pandemic era: *ChatGPT* and beyond". *Journal of university teaching & learning practice*, v. 20, n. 2.
<https://doi.org/10.53761/1.20.02.07>
- Polonsky, Michael-Jay; Rotman, Jeffrey D.** (2023). "Should artificial intelligent (AI) agents be your coauthor? Arguments in favour, informed by *ChatGPT*". *SSRN*, preprint.
<https://doi.org/10.2139/ssrn.4349524>
- Radford, Alec; Wu, Jeffrey; Child, Rewon; Luan, David; Amodei, Dario; Sutskever, Ilya** (2019). "Language models are unsupervised multitask learners". *OpenAI blog*, v. 1, n. 8.
<https://paperswithcode.com/paper/language-models-are-unsupervised-multitask>
- Rahaman, M. Saidur; Ahsan, M. M. Tahmid; Anjum, Nishath; Terano, Harold. J. R.; Rahman, M. Mizanur** (2023). "From *ChatGPT-3* to *GPT-4*: a significant advancement in ai-driven NLP tools". *Journal of engineering and emerging technologies*, v. 2, n. 1.
<https://doi.org/10.52631/jeet.v1i1.188>
- Rettberg, Jill-Walker** (2022). "ChatGPT is multilingual but monocultural, and it's learning your values". *Jill/txt*, December 6.
<https://jilltxt.net/right-now-chatgpt-is-multilingual-but-monocultural-but-its-learning-your-values>

- Roose, Kevin** (2022). "The brilliance and weirdness of *ChatGPT*". *The New York Times*.
<https://www.nytimes.com/2022/12/05/technology/chatgpt-ai-twitter.html>
- Rudolph, Jürgen; Tan, Samson; Tan, Shannon** (2023). "ChatGPT: Bullshit spewer or the end of traditional assessments in higher education?". *Journal of applied learning and teaching*, v. 6, n. 1.
<https://doi.org/10.37074/jalt.2023.6.1.9>
- Sabzalieva, Emma; Valentini, Arianna** (2023). *ChatGPT and artificial intelligence in higher education: quick start guide*. France: Unesco.
https://www.iesalc.unesco.org/wp-content/uploads/2023/04/ChatGPT-and-Artificial-Intelligence-in-higher-education-Quick-Start-guide_EN_FINAL.pdf
- Srivastava, Mashrin** (2023). "A day in the life of *ChatGPT* as a researcher: Sustainable and efficient machine learning-a review of sparsity techniques and future research directions". *OSF Preprints*.
<https://doi.org/10.13140/RG.2.2.34470.60480>
- Stokel-Walker, Chris** (2023). "ChatGPT listed as author on research papers: many scientists disapprove". *Nature*, v. 613, n. 7945, pp. 620-621.
<https://doi.org/10.1038/d41586-023-00107-z>
- Sullivan, Miriam; Kelly, Andrew; McLaughlan, Paul** (2023). "ChatGPT in higher education: Considerations for academic integrity and student learning". *Journal of applied learning and teaching*, v. 6, n. 1.
<https://doi.org/10.37074/jalt.2023.6.1.17>
- Susnjak, Teo** (2022). "ChatGPT: The end of online exam integrity?". *ArXiv preprint*. ArXiv:2212.09292.
<https://doi.org/10.48550/arXiv.2212.09292>
- Teixeira-da-Silva, Jaime A.** (2021). "Outsourced English revision, editing, publication consultation and integrity services should be acknowledged in an academic paper". *Journal of nanoparticle research*, v. 23, n. 4, 81.
<https://doi.org/10.1007/s11051-021-05199-0>
- Teixeira-da-Silva, Jaime A.** (2023). "Is *ChatGPT* a valid author?". *Nurse education in practice*, v. 68, 103600.
<https://doi.org/10.1016/J.NEPR.2023.103600>
- Thorp, H. Holden** (2023). "ChatGPT is fun, but not an author". *Science*, v. 379, n. 6630, pp. 313-313.
<https://doi.org/10.1126/science.adg7879>
- Túñez-López, José-Miguel; Fieiras-Ceide, César; Vaz-Álvarez, Martín** (2021). "Impact of artificial intelligence on journalism: transformations in the company, products, contents, and professional profile". *Communication & society*, v. 34, n. 1, pp. 177-193.
<https://doi.org/10.15581/003.34.1.177-193>
- Verganti, Roberto; Vendraminelli, Luca; Iansiti, Marco** (2020). "Innovation and design in the age of artificial intelligence". *Journal of product innovation management*, v. 37, n. 3, pp. 212-227.
<https://doi.org/10.1111/jpim.12523>
- Wager, Elizabeth; Kleinert, Sabine** (2011). "Responsible research publication: international standards for authors. A position statement developed at the 2nd World Conference on Research Integrity, Singapore, July 22–24, 2010". In: Mayer, T.; Steneck, N. (eds.). *Promoting research integrity in a global environment*. Singapore: Imperial College Press/World Science Publishing, pp. 309-316.
https://publicationethics.org/files/International%20standards_authors_for%20website_11_Nov_2011.pdf
- Wang, Dakuo; Churchill, Elizabeth; Maes, Pattie; Fan, Xiangmin; Shneiderman, Ben; Shi, Yuanchun; Wang, Qianying** (2020). "From human-human collaboration to human-AI collaboration: Designing AI systems that can work together with people". In: *Extended abstracts of the 2020 CHI conference on human factors in computing systems*.
<https://doi.org/10.1145/3334480.3381069>
- Williamson, Ben; Macgilchrist, Felicitas; Potter, John** (2023). "Re-examining AI, automation and datafication in education". *Learning, media and technology*, v. 48, n. 1.
<https://doi.org/10.1080/17439884.2023.2167830>
- Yin, Robert** (2009). *Case study research. Design and methods*. London: SAGE.
- Zainal, Zaidah** (2007). "Case study as a research method". *Jurnal Kemanusiaan*, v. 5, n. 1.
http://psyking.net/htmlobj-3837/case_study_as_a_research_method.pdf
- Zhang, Yunhao; Gosline, Renee** (2023). "Human favoritism, not AI aversion: People's perceptions (and bias) toward creative content generated by Ai (*ChatGPT-4*), human experts, and human-AI collaboration". *SSRN*.
<https://doi.org/10.2139/ssrn.4453958>

Retraining fact-checkers: The emergence of *ChatGPT* in information verification

Roger Cuartielles; Xavier Ramon-Vegas; Carles Pont-Sorribes

Nota: Este artículo se puede leer en español en:
<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87368>

Recommended citation:

Cuartielles, Roger; Ramon-Vegas, Xavier; Pont-Sorribes, Carles (2023). "Retraining fact-checkers: The emergence of *ChatGPT* in information verification". *Profesional de la información*, v. 32, n. 5, e320515.
<https://doi.org/10.3145/epi.2023.sep.15>

Manuscript received on 24th May 2023
Accepted on 31st July 2023



Roger Cuartielles ✉
<https://orcid.org/0000-0001-6226-6697>
Universitat Pompeu Fabra
UPF Barcelona School of Management
Departament de Comunicació
Roc Boronat, 138
08018 Barcelona, Spain
roger.cuartielles@upf.edu



Xavier Ramon-Vegas
<https://orcid.org/0000-0002-4478-5626>
Universitat Pompeu Fabra
UPF Barcelona School of Management
Departament de Comunicació
Roc Boronat, 138
08018 Barcelona, Spain
xavier.ramon@upf.edu



Carles Pont-Sorribes
<https://orcid.org/0000-0003-1135-4245>
Universitat Pompeu Fabra
UPF Barcelona School of Management
Departament de Comunicació
Roc Boronat, 138
08018 Barcelona, Spain
carles.pont@upf.edu

Abstract

The open launch of new artificial intelligence tools such as *ChatGPT-3.5* (Generated Pre-trained Transformer) in November 2022 by the company *OpenAI* -and then its update to version *GPT-4* in March 2023- poses new opportunities and challenges for journalism, and especially for professionals specifically focused on information verification. This research aims to understand and analyze the perceptions generated by the irruption of *ChatGPT* among fact-checking professionals in Spain with the aim of identifying disadvantages and advantages in its use, professional implications and desired functionalities. The study uses qualitative methodology with in-depth interviews with professionals from all Spanish fact-checking platforms belonging to the *International Factchecking Network (IFCN)* and the *European Digital Media Observatory (EDMO)*. The results conclude that the use of *ChatGPT* presents notable ambivalences. On the one hand, there are perceived drawbacks in issues such as the transparency and reliability of sources, the scope of the data, and the format of the responses generated. However, fact-checkers also point to a possible auxiliary use of the chatbot in the tasks of gathering information, detecting falsehoods, and producing denials. The irruption of *ChatGPT* has a direct impact on the work routines of the fact-checkers, which can be made more difficult, reinforced or extended. Fact-checking professionals perceive themselves as "context agents" in a new ecosystem that also obliges them to further diversify their fields of action in the fight against disinformation and to accelerate the implementation of media education actions that empower citizens in the responsible use of artificial intelligence.

Keywords

Fact-checking; Fact-checking platforms; Digital verification; Journalism; Computational fact-checking; Fact-checkers; Disinformation; Misinformation; Artificial Intelligence; AI; Generative artificial intelligence; *ChatGPT*; Chatbots.



Funding

This research was funded by the “Instruments of accountability in the face of disinformation: impact of fact-checking platforms as accountability tools and curricular proposal” project (PID2019-106367GB-I00/AEI/10.13039/501100011033) (FACCTMedia), Spanish Ministry of Science and Innovation (2020–2024), in collaboration with the research group on *Political Communication, Journalism, and Democracy* of Pompeu Fabra University (Polcom-GRP), and *Agaur* grants (2021 SGR 00486), and supported by *Joan Oró* grants to hire research staff in training (FI 2023), awarded by the *Government of Catalonia*.

1. Introduction

Fact-checking platforms and professionals have proven to be useful agents in combating online disinformation over the past decade (Hameleers; Van-der-Meer, 2020). Fact-checking processes carried out by journalists and other specialized profiles have been an effective resource in discrediting hoaxes and disinformation content circulating in the digital sphere that represent a threat to democratic welfare (Moreno-Gil; Ramon-Vegas; Rodríguez-Martínez, 2021). The use of fact-checking has also guaranteed a degree of reliability in information, making fact-checking professionals and platforms attempt to compensate for the lack of trust in journalistic practice, which according to the *Digital News Report Spain 2022* has worsened, especially in recent years, surpassing for the first time the percentage of users who say they believe in the truth of news content (Vara-Miguel *et al.*, 2022).

Fact-checking organizations, both those developed according to a newsroom model within a media outlet and those created as independent organizations under the NGO model (Graves; Cherubini, 2016), have been conceptualized as tools for democratic construction (Moreno-Gil; Ramon-Vegas; Rodríguez-Martínez, 2021), especially in a society such as the current one that is characterized by the fragmentation of the concept of truth (Malik, 2018). Moreover, their task has become even more relevant since the coronavirus disease 2019 (Covid-19) crisis, in which a constant proliferation of disinformation led the *World Health Organization* to release statements regarding an infodemia (WHO, 2020), a context that has also contributed to extolling the work and visibility of fact-checkers as crucial actors in the fight against disinformation (Salaverría *et al.*, 2020; Pérez-Dasilva; Meso-Ayerdi; Mendiguren-Galdospin, 2020; Ramon-Vegas; Mauri-Ríos; Rodríguez-Martínez, 2020).

In the fact-checking field, many platforms have incorporated artificial intelligence (AI) into their routines in recent years, for example, the use of bots in debunking processes (Arias-Jiménez *et al.*, 2023; Pasquetto *et al.*, 2022; Flores-Vivar, 2020) or the implementation of detection systems based on machine learning to identify false claims, videos, and photographs. It is thus understood that the use of AI to combat the proliferation of false information can be vital when acting against the disruptive effects that disinformation generates (Gupta *et al.*, 2022; Flores-Vivar, 2020).

The open release of new artificial intelligence tools such as *ChatGPT-3.5* (Generated Pre-trained Transformer) in November 2022 by the company *OpenAI*—and its subsequent upgrade to version *GPT-4* in March 2023—poses new opportunities and challenges for journalism, especially for professionals specifically focused on fact-checking. Therefore, this research focuses on identifying and analyzing the perceptions generated by the emergence of *ChatGPT* among fact-checking professionals to determine its perceived drawbacks and advantages, the implications in their professional role, and the possible uses of this chatbot in the field of fact-checking.

2. Theoretical framework

2.1. Fact-checking: professional features and routines

Fact-checking can be defined as an activity that

“applies data journalism techniques to unmask errors, ambiguities, lies, lack of rigor or inaccuracies in some content published in the media” (Ufarte-Ruiz; Peralta-García; Murcia-Verdú, 2018, p. 734).

Likewise, fact-checking can also manage content without identifiable authorship distributed through social networks and other multimedia formats (Pérez-Seoane; Corbacho-Valencia; Dafonte-Gómez, 2023). Fact-checking in essence is a traditional practice that has been associated with good journalism and with the specialty of journalistic documentation (Redondo, 2018). However, the growing concern regarding disinformation has seen it become an issue of growing importance (Guallar *et al.*, 2020), even positioning it as a professional activity in and of itself that leads to the strengthening and reformulation of the standards of traditional journalism (Cavaliere, 2021). The emergence and consolidation of initiatives and platforms specifically dedicated to the field of fact-checking in the last decade demonstrate the impact of fact-checking at a global level (Stencel; Ryan; Luther, 2022).

Numerous studies on fact-checking platforms’ methodologies and professional routines have pointed out the essential phases of their workflow: monitoring and selection of content to be verified (predetection and detection); contact with the original source and contex-

Fact-checking platforms have incorporated AI using chatbots and implementing detection systems based on machine learning to identify data to verify

tualization and evaluation of the veracity of the content examined by consulting expert documentary as well as personal sources (reporting); and the realization, dissemination, and explanation of the fact-checking process performed (debunking) (Graves, 2017; Unesco, 2018; Moreno-Gil; Ramon-Vegas; Rodríguez-Martínez, 2021; Yousef, 2023). Aspects such as rigor, impartiality, accountability, objectivity, independence, transparency, and completeness have been considered key elements in the performance of fact-checkers' professional routines as well (Singer, 2021). The work of such professionals is also characterized by the application of information selection criteria, such as the relevance and potential influence of the message when selecting the content to be verified, which in turn must contain factual aspects to be contrastable (Graves, 2017). In this fact-checking task, technological tools apart from human judgment that can facilitate the verification of data in different media such as text, image, audio and video also play an important role (Vizoso; Vázquez-Herrero, 2019).

“ In the journalistic field, AI has been introduced especially linked to the transformation of data into news and the automatic production of texts ”

Another of the principles embraced by fact-checkers is the rigorous supervision of verified content before its publication, as well as the use of transparent sources to verify it, often with hyperlinks that allow users to expand the information and replicate the same fact-checking process (López-Pan; Rodríguez-Rodríguez, 2020). These practices also allow for the pursuit of citizen empowerment (Graves, 2016). In fact, another common feature of fact-checking platforms is the participation of readers, who can send doubts and requests for information to be verified through different communication channels such as email and social networks (Rodríguez-Pérez, 2020).

Textual explanation, and even the use of so-called explanatory journalism (Bielik; Višňovský, 2021) and solutions journalism, consisting of responding to social problems with rigorous methods that facilitate citizen understanding (McIntyre; Lough, 2021), is also presented as a common resource in the work methodologies of fact-checking platforms, in addition to the synthesis of fact-checking in measurement scales that not only include dichotomous categories such as true or false, but also refer to deceptions and decontextualizations that require further explanation (García-Vivero; López-García, 2021). In seeking to present complex information in a clear and simple way, the use of data visualization through graphics and infographics is also common, as is the development of formats that go beyond textuality, such as videos and podcasts (Moreno-Gil; Ramon-Vegas; Mauri-Ríos, 2022), being disseminated not only via the web pages of fact-checking platforms but also on their respective social media channels.

While it is true that, in this sense, the work routines of journalists and fact-checkers share common goals as professionals engaged in explaining, documenting, and fact-checking (Singer, 2021), the activity of fact-checkers has also been presented in a distinctive way. Authors such as Graves (2016) and Cazalens *et al.* (2018) have pointed out that, while journalists implement fact-checking as a process of an internal nature consisting of ensuring the veracity of data before being publicly revealed in a journalistic piece, fact-checking professionals focus more on an external type of fact-checking in which the accuracy of statements and content already issued is checked with the aim of preparing a new piece of information that contextualizes and accredits its veracity. Within this context, the work of fact-checkers is presented as complementary and corrective to that of the media (Singer, 2021).

Even so, fact-checking cannot be understood as an activity outside the journalistic sector, since it arises within the media system and there are even platforms created according to the newsroom model format. In addition, a significant percentage of fact-checking professionals come from the field of journalism and data journalism (Graves; Cherubini, 2016), and their journalistic training, the handling of big data, and other aspects such as the mastery of information sources and common sense are perceived by fact-checking professionals themselves as key elements in the performance of their work (Herrero; Herrera-Damas, 2021).

2.2. The use of AI in journalism and fact-checking

The term “artificial intelligence” (AI) was first used publicly in 1956 by mathematician John McCarthy at the *Dartmouth Conference* in the United States. However, the origin of AI can be traced back to Alan Turing's advances in message decoding during World War II (Russell; Norvig, 2022) and in even earlier stages to Lady Ada Lovelace's analytical machine experiments in the 1840s, which were an early prediction of the implications that AI would later have (Boden, 2022). Even so, it was not until the 1980s, with the resolution of algebra equations and the analysis of texts in different languages, that AI research began to increase (Ufarte-Ruiz; Manfredi-Sánchez, 2019). Since the mid-2000s, this technology has also experienced rapid expansion in both academia and industry (Crawford, 2021).

Although the definition of artificial intelligence has evolved in parallel to its application, the most widespread definition in academia and the journalism sector is the one proposed by *BBC* journalist Dickens Olowe in 2018, which understands it to be a

“collection of ideas, technologies, and techniques that relate to a computer system's capacity to perform tasks normally requiring human intelligence” (Brennen; Howard; Nielsen, 2018, pp. 1-2).

In the field of journalism, AI has been especially linked to natural language processing (NLP) (Canavilhas, 2022) applied to the transformation of data into news, as well as in the automatic production of texts (Diakopoulos, 2019) through

the use of bots (Flores-Vivar, 2019; Flores-Vivar, 2020). This technological emergence has given rise to so-called computational journalism, algorithmic journalism, automated journalism, and robot journalism (Clerwall, 2014; Carlson, 2015; Dörr, 2016). The first experiments in automated news production date back to 2010 with *The Big Ten Network*, a partnership between *Fox Networks Group* and *Big Ten Conference*, which initiated an automated sports news production service (Canavilhas, 2022). Although this was not the first experiment in the automatic generation of journalistic texts, it was one of the first to use artificial intelligence, according to Canavilhas (2022). However, the first initiative to consolidate on a regular basis came four years later at the *Los Angeles Times* with the appearance of *Quakebot*, a bot that automatically reported on earthquakes and that represented the first frequent use of AI in the field of journalism (Sánchez-García *et al.*, 2023).

Subsequently, *Associated Press* was one of the pioneering news agencies that extended the use of AI (Lichterman, 2017), and the French newspaper *Le Monde* used the *Data2Content* system to create micro-news on election results (Sánchez-Gonzales; Sánchez-González, 2017). Along these lines, specific NLP and natural language generation (NLG) companies specialized in the creation of journalistic texts have emerged in recent years, such as *Narrative Science* and *Automated Insights* (Sánchez-García *et al.*, 2023). More recently, initiatives have also been created that transform data into real-time journalistic information such as *AppliedXL*, a company founded by computational journalist Francesco Marconi, and *The Newsroom*, a mobile application that offers daily summaries made by AI on the main news of the day (Adami, 2023).

In recent years, pioneering projects have also appeared, such as *Medusa*, from *Vocento MediaLab*, which since 2017 has been experimenting with automated journalism to generate information on the state of some 800 Spanish beaches and ski slopes in Spain, Andorra, and the French Pyrenees (Ufarte-Ruiz; Manfredi-Sánchez, 2019). Other projects to consider are the *AnaFut* bot of *El Confidencial*, which writes sports chronicles (Rojas-Torrijos; Toural-Bran, 2019), and the *Gabriele* software from the start-up *Narrativa*, which writes journalistic texts in an automated way and in collaboration with Spanish media (Ufarte-Ruiz; Manfredi-Sánchez, 2019; Sánchez-García *et al.*, 2023).

The use of artificial intelligence technologies as applied to fact-checking has led some authors to speak of “computational fact-checking” and “automated fact-checking” (Thorne; Vlachos, 2018) as well, understood as a practice based on fact-checking that automates part of its process with the help of AI. In fact, in the field of fact-checking, automation has been presented as a solution, in part from a computational point of view to streamline certain processes of fact-checkers’ professional routines, such as monitoring and anticipating information (predetection), identifying claims to verify (detection), obtaining data to verify content (reporting), and checking falsehoods (debunking), to name but a few (Hassan *et al.*, 2015; Guo; Schlichtkrull; Vlachos, 2022).

While it is true that the use of AI has also been presented as a way to make disinformation more clever and increase its spread, for example, in the creation of content such as deep fakes (Fallis, 2021), and even with serious ethical implications regarding issues such as pornography (Öhman, 2020), numerous authors have also alluded to the opposite potential of artificial intelligence, with examples such as dealing with the spread of falsehoods and malicious content (Cybenko; Cybenko, 2018; Beckett, 2019; Manfredi-Sánchez; Ufarte-Ruiz, 2020) and enabling a technology capable of adapting not only to the speed with which falsehoods circulate in the digital environment, but also to their degree of elaboration. It also reduces the effort and detection time spent by fact-checking professionals and, in short, increases their capacity to respond to disinformation.

Although before the Covid-19 pandemic there were already some fact-checking organizations employing AI-based technologies in their work procedures, the development of this technology has experienced a boom especially since the pandemic, both to speed up the detection of hoaxes and to identify falsehoods within AI itself, such as deep fakes (Gómez-de-Ágreda; Feijóo; Salazar-García, 2021). With the emergence of the pandemic, fact-checking organizations in Spain such as *Newtral*, *Maldita.es*, and *EFE Verifica* have launched initiatives based on AI use. *Newtral* has developed and perfected an automatic monitoring system focused on the political sphere that identifies claims to be checked, which is accurately called *ClaimHunter* (Morrish, 2023). *Maldita.es* and *EFE Verifica* have also perfected the use of a chatbot through *WhatsApp* to receive fact-checking requests from their users. This same system automatically filters and responds to incoming requests on the basis of whether it identifies a relationship to topics already checked by the platform professionals in question (Pablo Hernández, in-depth interview, February 20, 2023; Sergio Hernández, in-depth interview, February 18, 2023).

The use of generative artificial intelligence through the implementation of chatbots, in this case on social networks such as *WhatsApp*, has been a resource that has proven useful in combating disinformation (Palomo; Sedano-Amundarain, 2018), and numerous fact-checking organizations around the world have already implemented it in their work routines (Flores-Vivar, 2020). In recent years, collaborative initiatives have emerged, such as *FactChat*, launched by the *International Fact-Checking Network (IFCN)* during the 2020 presidential elections in the United States, in addition to the Covid-19 chatbot that this same organization launched during the pandemic, together with more than 80 fact-checking organizations around the world (Grau, 2020).

“The reliability of the sources and the limited scope of its data are two of the aspects that are perceived as major drawbacks in the use of *ChatGPT*”

The emergence of generative artificial intelligence, as used by tools such as chatbots, is the product of the third wave of innovation that artificial intelligence has undergone in the last decade, after having gone through two previous phases: automation and augmentation (Marconi, 2020). In this sense, NLG systems and those called large language models (LLM) due to the use of so-called neural networks –as is the case of *ChatGPT-3.5*– have made significant advances in NLP after being trained from massive databases with which they are able to generate texts, answer questions, and complete other tasks in a way that resembles human communication (Floridi; Chiriatti, 2020).

“ A perceived positive aspect of using *ChatGPT* is the possibility of collecting contextual information quickly and synthetically ”

These capabilities, within the reach of citizens following the open release of *ChatGPT-3.5* by *OpenAI* in November 2022, have already had significant implications in sectors such as education (Kasneci *et al.*, 2023), cultural creation, and academic research (Dwivedi *et al.*, 2023), as well as in areas such as the labor market, where it is estimated that in 80% of professions in countries such as the United States, *ChatGPT* could be used for at least 10% of current tasks (Eloundou, 2023).

In the field of communication, the involvement of *ChatGPT* in the production of disinformation has begun to be studied, especially as it relates to issues such as biases and “hallucinations”, which is when an AI system provides data not based on facts, but as a product of its own “invention” (Liu *et al.*, 2022). In this sense, organizations such as *NewsGuard* have experimented with the chatbot by using previously verified false narratives, finding that, in 80% of cases, it did not recognize the falsehoods introduced nor was it transparent in the use and reliability of sources. Thus, *ChatGPT-3.5* has been dubbed as a “great misinformation superspreader” (Brewster; Arvanitis; Sadeghi, 2023). Even in its updated version, *ChatGPT-4* –released in mid-March 2023– which, unlike *ChatGPT-3.5*, does cite the origin of the sources from which it extracts information, *NewsGuard* has also concluded that the dissemination of erroneous information is “more frequent and more persuasive” than in its predecessor model (Arvanitis; Sadeghi; Brewster, 2023).

Academic literature has addressed how journalists from different countries and professional cultures perceive the introduction of technological innovations into newsrooms (García-Avilés; Carvajal-Prieto; Arias-Robles, 2018; Ferrucci; Perreault, 2021; Holman; Perreault, 2022; Oelrichs, 2023). Recently, other contributions have focused on examining the attitudes and perceptions of journalists, audiences, and experts in the face of the emergence of artificial intelligence (Noain-Sánchez, 2022; Soto-Sanfiel *et al.*, 2022; Sun; Hu; Wu, 2022; Peña-Fernández *et al.*, 2023). These contributions have highlighted the opportunities, tensions, and concerns that AI generates among these different parties, including the ambiguities that the adoption of artificial intelligence tools produces specifically among information professionals. The positive perception of AI’s use in the journalism sector is as an auxiliary tool that can free journalists from performing repetitive tasks, leaving to one side a post-Fordist model in which reporters are mere transcribers of facts to bring back the creative essence of journalism (Noain-Sánchez, 2022). However, ignorance of AI’s implications also produces a certain reticence, especially because it is perceived as a threat to the symbolic capital of journalists as mediators between reality and citizens (Peña-Fernández *et al.*, 2023). Ethical dilemmas and the possible spread of disinformation with the use of artificial intelligence tools also appear as relevant concerns among information professionals (Noain-Sánchez, 2022) and even among experts and readers (Sun; Hu; Wu, 2022).

According to authors such as Boczkowski (2004), the journalistic profession, faced with technological innovations such as digital transformation, has been characterized by reactive and defensive as well as pragmatic attitudes. It is also important to note that, in the case of AI, such reluctance is not homogeneous in nature, and that it also varies according to the culture and journalistic tradition of each country. For example, while in countries such as Pakistan a negative view of AI implementation in the journalism sector predominates (Jamil, 2020), in areas such as Latin America a more optimistic perception abounds (Soto-Sanfiel *et al.*, 2022). As Van-Dalen (2012) points out, the consolidation of an innovation is determined not only by its technological development but also by social factors such as adaptation and the way in which its consumption is reduced to practice.

Thus far, no academic study has focused on analyzing the perceptions that the use of *ChatGPT* –in both its 3.5 and 4 versions– generates among fact-checking professionals themselves, nor on identifying the drawbacks and advantages that its use may present in the work routines of these professionals when combating disinformation, or even on the implications that it may generate regarding their role as fact-checkers. This is therefore an academic gap that this research aims to fill.

3. Methodology

This study aims to understand the perceptions generated by the emergence of *ChatGPT* among fact-checking professionals in Spain. The analysis sample is made up of active Spanish fact-checking organizations included in the *Duke Reporters’ Lab* database as well as in the *European Digital Media Observatory (EDMO)* database: *Maldita.es*, *Newtral*, *EFE Verifica*, *AFP Factual España*, and *Verificat*. These five organizations are also signatories of the *International Fact-Checking Network (IFCN)* code of principles, built around five basic concepts: (1) non-partisanship and honesty, (2) standards and transparency of sources, (3) transparency in organization and financing, (4) standards and transparency of methodology,

and (5) open and honest correction policy. To obtain a wider range of perspectives, the Spanish platform *VerificaRTVE* is also included in the analysis corpus. Although it is not a signatory to the *IFCN* code of principles, it is a member of *EDMO* and the *European Broadcasting Union (EBU)*.

Three research questions guide this study:

- Q1. What disadvantages and advantages do fact-checkers perceive in the use of *ChatGPT* in their professional routines?
- Q2. What features should *ChatGPT* have to be perceived as a useful resource in the field of fact-checking?
- Q3. What implications does the viralization of *ChatGPT* have for the professional role of fact-checkers?

Table 1. Characteristics of the fact-checking platforms participating in the study.

Platform	Website	Creation	Fact-checker template	Signatory to IFCN
<i>Maldita.es</i>	https://maldita.es	2018	30	Yes
<i>Newtral</i>	https://www.newtral.es	2018	14	Yes
<i>EFE Verifica</i>	https://verifica.efe.com	2019	7	Yes
<i>AFP Factual España</i>	https://factual.afp.com/afp-espana	2019	3	Yes
<i>Verificat</i>	https://www.verificat.cat	2019	9	Yes
<i>VerificaRTVE</i>	https://www.rtve.es/noticias/verificartve	2020	5	No

Source: interviews with fact-checking platforms.

To answer the research questions, six semistructured in-depth interviews were conducted with fact-checking professionals working in these organizations who hold positions of responsibility, for example, as editors and section managers: Pablo Hernández (coordinator of academic research at *Maldita.es*), Irene Larraz (director of *Newtral Educación* and coordinator of the political fact-checking and data section), Sergio Hernández (director of *EFE Verifica*), Borja Díaz-Merry (director of *VerificaRTVE*), Adrià Laborda (fact-checker at *AFP Factual España* and head of the Catalan division *AFP Comprovem*), and Javier Castillo (head of the political fact-checking section of *Verificat*). The interviews, between 60 and 120 minutes long, were conducted between February and early March 2023 through the *Google Meet* platform owing to the geographic diversity of the sample, as the professionals are located in different parts of Spain. All conversations were recorded and subsequently transcribed for analysis. A second round of interviews was then conducted in late March 2023 following the release of *ChatGPT-4* to see whether their perceptions of chatbot use had changed.

Qualitative interviews represent a valuable research technique because they allow for participants to get to the “heart of the matter” and provide a good opportunity to understand, reflect, and go deeper into topics and issues that cannot be easily observed or accessed (Tracy, 2020, p. 79). In-depth interviews also facilitate evidence regarding the context and origin in which a phenomenon to be studied arises, and is enriched by the insider view of people who have direct experience with it (Miller; Glassner, 2016). Moreover, they have been a widely used method in recent fact-checking studies in the Spanish and international contexts; see for example, research by Martínez-García and Ferrer (2023); Sánchez-González, Sánchez-Gonzales, and Martínez-Gonzalo (2022); Moreno-Gil, Ramon-Vegas, and Mauri-Ríos (2022); Singer (2021); Graves and Anderson (2020); López-Pan and Rodríguez-Rodríguez (2020); and Palomo and Sedano-Amundarain (2018).

The interviews were semistructured to ensure a certain degree of freedom for the respondents, and the questions asked were organized into different thematic blocks, as follows: use of *ChatGPT*, perceived disadvantages and advantages, desired benefits and uses, and professional implications identified. After the transcription of each interview, the constant comparative method was applied (Wimmer; Dominick, 2013). The data obtained were assigned to categories and, after an initial analysis, the established categories, as well as the relationships and themes identified, were refined to determine the most recurrent issues.

4. Results

According to the perceptions of the fact-checking professionals interviewed, *ChatGPT-3.5* presents remarkable ambiguities in its use as applied to information fact-checking. On the one hand, fact-checking professionals perceive significant drawbacks in aspects such as the use and reliability of sources; the training, processing, and scope of the data used by the chatbot; and the production and format of the answers generated. On the other hand, there are also professionals who see it as a useful auxiliary resource for information gathering (reporting), the detection of falsehoods (detection), and the fabrication of contradictions (debunking).

The viralization of *ChatGPT* also generates significant implications for fact-checkers in relation to their work routines, which –according to their perceptions– may be hindered, reinforced, or expanded. The results presented herein are structured according to the research questions. The type of use of *ChatGPT* by the interviewed professionals was also specified previously, as well as whether they have used its updated version, *ChatGPT-4*, which was released while the present research was being carried out.

4.1. Using *ChatGPT*

All the fact-checking platforms consulted have made use of *ChatGPT-3.5*, in particular internally, to test the chatbot's performance. On the contrary, the use of *ChatGPT-4* has not been tested among the professionals consulted owing to its being subscription based. Only at *Maldita.es* has the engineering team started testing the updated version, albeit without conclusive results.

Two of the platforms, *Maldita.es* and *VerificaRTVE*, have shared part of the internal tests done with *ChatGPT-3.5* through public journalistic pieces on their websites in which readers could view the type of questions asked. In the case of *Maldita.es* (2022), the piece was published on December 27, 2022 and was made from a live broadcast on the *Twitch* platform, within the *Maldita Twitcheria* section, in which they invited different experts in the field of computing to discuss the use of the chatbot. The prompts –information inputs entered into the chatbot– were geared toward entertainment questions and mathematical queries such as “Write a song in the style of singer Rosalía,” “Talk about the book *The Time Machine*,” and “Identify whether the number 9 is a prime number.” All consultations were in Spanish. Pablo Hernández (in-depth interviews conducted on February, 20 2023 and March, 28 2023) explains that, as fact-checkers at *Maldita.es*, they have also done some testing on a personal basis, albeit with “basic searches,” and that the platform's engineering team has also tested the *ChatGPT-4* version, although still without significant results.

In the case of *VerificaRTVE* (2023), the platform published a piece on its website on January 27, 2023 with tests made with *ChatGPT-3.5*. In this case, the queries were related to disproven disinformation about Covid-19 vaccination, such as “Do Covid-19 vaccines contain graphene?” The prompts were entered in Spanish and English to detect possible language bias, but did not detect “failures.” According to Borja Díaz-Merry (in-depth interview, March 2, 2023), “this gave us some confidence, although with caution, since we have to experiment more because we believe that we could detect errors.” Even so, in the piece that was published, they explained that they had found that *ChatGPT* “reacts to disprove false content with data from reliable sources.”

At *Newtral*, internal testing has also focused on checking whether the chatbot was able to detect falsehoods on topics that they had previously verified on the platform, as well as “analyzing how *ChatGPT* can manufacture the raw material to create a hoax” (Irene Larraz, in-depth interview, February 16, 2023). According to Larraz, the test showed the tool's capacity for “sophistication” to fabricate disinformation.

Meanwhile, *AFP Factual Spain* questioned the chatbot regarding the veracity of some images of “an alleged rescue dog in the Syrian and Turkish earthquakes of February 6, 2023,” although they did not obtain a “satisfactory” answer (Adrià Laborda, in-depth interview, February 16, 2023). Finally, at *Verificat* and *EFE Verifica*, the queries have been oriented toward questions related to the search for bibliographic references to verify fact-checking as well as to questions of general culture, such as “Is it true that man has landed on the Moon?”, a fact that the chatbot confirmed.

4.2. Disadvantages and advantages identified

The lack of knowledge regarding the origin of the sources used by *ChatGPT*, as well as their typology, is perceived as the main disadvantage regarding its application in the field of information fact-checking. Likewise, the impossibility of discerning between facts based on empirical knowledge and those based on artificial intelligence inventions –so-called hallucinations– also generates distrust: “It does not specify which part of the text is based on real facts and which part has been invented by the machine” (Sergio Hernández, *EFE Verifica*). The demand for specific scientific sources also presents problems. As Javier Castillo (*Verificat*) states, “You ask it for references for a fact-check on nuclear energy and it invents them, it writes them in APA format, but maybe there are articles or authors that do not exist.”

In this sense, the authenticity with which the chatbot generates its responses is perceived as another drawback, especially as an incentive to produce disinformation and complicate its detection. The fact-checkers point out that the coherence and textual correctness that *ChatGPT* exhibits in its responses can generate more effective disinformation narratives, both through the creation of argumentatively more consistent and convincing discourses and through the writing of disinformation texts that go unnoticed because they are linguistically correct. Thus, *ChatGPT* can be used to “create the raw material of a hoax and replicate it in different text formats and languages,” and also to “generate fraudulent phishing messages” (Irene Larraz, *Newtral*).

The amplification of disinformation on *ChatGPT* may also occur owing to its high degree of accessibility. According to professional opinion, the possibility of multiple actors using the chatbot without any oversight can also facilitate the production of falsehoods: “Once you open the chatbot to the population, you find yourself in a scenario where, if the tool falls into the hands of people who want to generate chaos, the creation of disinformation and confusion may be unstoppable” (Pablo Hernández, *Maldita.es*). Even so, for other professionals, such as Borja Díaz-Merry (*VerificaRTVE*), the accessibility it presents can also be seen as a positive aspect, since “citizen training can also improve the tool.” Following this reflection, Pablo Hernández (*Maldita.es*) specifies that the viralization of *ChatGPT* can also act as a catalyst for further innovation: “The emergence of a technology always leads to new developments and the creation of more specific features.”

The scope of the database on which *ChatGPT* has been trained also generates mistrust among the fact-checker community. Some professionals such as Borja Díaz-Merry (*VerificaRTVE*) perceive drawbacks in the time limitation, set in 2021 for the *ChatGPT-3.5* version, and consider that this “bounded data processing” also “compromises the rigor and depth of the answers,” which can “fly over superficiality.” There are also fact-checkers such as Irene Larraz (*Newtral*), who, apart from this temporal limitation, identify a geographical limitation: “I’ve found that in some of the more local political fact-checking pieces, where I imagine the system doesn’t find as much existing information, there’s more of a tendency to get it wrong and misinform.”

The predominance of the textual format is seen as another limitation for fact-checking professionals. Adrià Laborda (*AFP Factual Spain*) points out that *ChatGPT-3.5* “cannot verify images or videos” and Borja Díaz-Merry (*VerificaRTVE*) also sees the fact that it cannot process audios as a drawback: “One of the most common disinformation practices we have encountered are fake audio messages, and *ChatGPT* can not help us with this issue either.” Laborda (*AFP Factual Spain*) also adds the fact that the chatbot has difficulty detecting humor and irony in a text as a disadvantage, since “they are key elements in some disinformation.”

There are professionals such as Pablo Hernández (*Maldita.es*) and Sergio Hernández (*EFE Verifica*) who do view possible uses of *ChatGPT* as an auxiliary tool in the disinformation detection phase. In this regard, they believe that *ChatGPT* can be useful in detecting patterns and disinformation narratives circulating on the Internet, having “processed large amounts of data on the network that may not be true.” Moreover, Hernández (*Maldita.es*) points out that the identification of existing disinformation discourses through *ChatGPT* can also help fact-checkers produce disinformation better adapted to these predominant narratives: “Once we recognize these discourses, we can generate content that is more specific and focused on disproving these disinformations.” Therefore, beyond the stage of disinformation detection, *ChatGPT* is also perceived as an auxiliary resource in the debunking phase.

In this sense, Adrià Laborda (*AFP Factual Spain*) also sees possible uses in earlier phases of the work process, for example, in the collection of information (reporting). For Laborda, *ChatGPT* can thus be used to search for “synthetic context information about a fact quickly” that helps to save time and to complete the fragment of a piece by way of contextualization, although “always reviewed by a professional.” In the opinion of Irene Larraz (*Newtral*), it can also be useful in obtaining instructional information, since “it can offer clues, instructions and initial guidelines on how to start checking content.” However, she also indicates that the expertise of the fact-checking professionals must always come first.

The following is a summary of the main findings found regarding the disadvantages and advantages of the use of *ChatGPT* in information fact-checking (Table 2).

Table 2. Perceived inconveniences and advantages of using *ChatGPT* for information fact-checking.

Inconveniences	Advantages
<p>Lack of transparency in sources</p> <ul style="list-style-type: none"> Lack of knowledge as to the origin and use of sources. Difficulty in identifying between real events and so-called hallucinations. 	<p>Ease of data collection</p> <ul style="list-style-type: none"> Synthetic and quick information that can provide contextual data. Basic data of an instructive nature that can serve as a guideline to initiating fact-checking.
<p>Sophistication and amplification of disinformation</p> <ul style="list-style-type: none"> Plausible texts, but not true. “Citizen training”. Indiscriminate data processing. Replicability of the same content in different textual structures and languages. 	<p>Updating and “self-learning”</p> <ul style="list-style-type: none"> Citizen accessibility can train and improve chatbot performance. The development of <i>ChatGPT</i> is driving the emergence of more specific tools.
<p>Restricted sample size</p> <ul style="list-style-type: none"> Temporally bounded data processing. Geographical limitations. Superficiality in answers. 	<p>Identification of online disinformation</p> <ul style="list-style-type: none"> Possibility of detecting patterns and disinformation narratives present on the Internet.
<p>Predominance of text format</p> <ul style="list-style-type: none"> Difficulty in verifying non-textual content. Difficulty in identifying textual tone. 	<p>Fact-checking better adapted discursively</p> <ul style="list-style-type: none"> Identification and knowledge of disinformation narratives circulating on the Internet can facilitate the development of more effective counter-narratives.

Source: interviews with fact-checking platforms.

4.3. Performance and intended uses

There is a general consensus among fact-checking professionals that *ChatGPT* could be used in the fact-checking process if, above all, it were transparent and reliable in its use of information sources. Likewise, fact-checkers would also opt for its professional use if it offered “more rigor and argumentation in its answers,” as well as features more adapted to their work tasks, for example, the possibility of identifying and processing audiovisual content.

Beyond preferring greater specialization of the tool in terms of content production, fact-checkers would also view *ChatGPT* as a useful resource in their work if it had greater depth and scope in the data processed, that is, if it could expand the data sample with which it has been trained, if it had the capacity to anticipate, and also if it had more monitoring and

disinformation content detection features, especially focused on saving time and response capacity in the debunking process. The following are some of the perceptions identified in this area by the professionals consulted:

“I wish *ChatGPT* could help us when there is still no article published on the Internet regarding certain content” (Adrià Laborda, *AFP Factual Spain*).

“It would be ideal if it had a system to analyze more content and to monitor information that appears on different web pages or even on social networks” (Pablo Hernández, *Maldita.es*).

“It would be interesting if you could enter a text and it would identify which claims could be verifiable and which not, so we would save much more time” (Javier Castillo, *Verificat*).

“*ChatGPT* and *OpenAI* have not come to ask us information fact-checkers anything, and we just need information production and monitoring tools more adapted to our work that allow us to gain time and scope” (Irene Larraz, *Newtral*).

“The answers it offers are coherent and well argued, but as a fact-checker I require much more depth to be able to trust this tool, and also more transparency in the use of sources” (Borja Díaz-Merry, *VerificaRTVE*).

4.4. Implications for professionals

The viralization of *ChatGPT* as a tool for the production of citizen outreach information has a significant impact not only on the work of fact-checkers but also on their role as professionals. This is the perception of most of the fact-checkers consulted, who identify three obvious implications for their work.

They judge that the emergence of *ChatGPT* implies a reinforcement of their task as fact-checking professionals because the viralization of this artificial intelligence “reinforces human judgment and intelligence” (Adrià Laborda, *AFP Factual Spain*), which is maintained and consolidated as “essential” in any information fact-checking process. Thus, they believe that the open launch of the chatbot defends the concept of fact-checkers as “agents of context” (Pablo Hernández, *Maldita.es*). In this sense, *ChatGPT* is situated as “an auxiliary tool” (Irene Larraz, *Newtral*), always subject to the human judgment of the fact-checker in the processes of detection and disproval.

The chatbot’s viralization also implies the complication of fact-checkers’ professional task as a tool for citizen use that can facilitate the production and sophistication of disinformation through the creation of plausible and linguistically correct texts. Furthermore, by not identifying the origin and use of the sources, the fact-checking process is made more difficult.

Some professionals also deem that the emergence of *ChatGPT* broadens the competencies and skills that fact-checkers must master because “it forces greater teamwork not only on the same platform, but also between organizations” (Sergio Hernández, *EFE Verifica*) in the sharing of knowledge and tools that can be useful for their use. It also requires “flexibility, adaptation and more dedication from professionals” (Borja Díaz-Merry, *VerificaRTVE*). In addition, for others it entails a change in the conception of the role of the fact-checker, which goes beyond the simple verification of data. According to Pablo Hernández (*Maldita.es*), “It’s not just about verifying data, but about fighting disinformation in a broad sense. It is not only to take a piece of content, increasingly more convincing and better written, and disprove it, but to provide context and explain it better so that there are no doubts.”

In this sense, Hernández (*Maldita.es*) also defends the actions toward media education carried out by numerous fact-checking platforms and suggests extending them to the use of *ChatGPT* with didactic and formative materials of a citizen nature.

5. Discussion and conclusions

Fact-checking skills and competencies in journalism evolve with technological changes (Himma-Kadakas; Ojamets, 2022). Thus, *ChatGPT* and its updates present numerous challenges and obstacles to be faced by fact-checkers, who perceive significant changes for their professional routines with the chatbot’s emergence, such as the increased complexity of falsehoods in all aspects, which accentuate the information disorder syndrome (Wardle; Derakhshan, 2017) as well as the dissemination of manipulated content and propaganda (Tandoc Jr; Lim; Ling, 2018).

This demands more dedication, adaptability, and professional collaboration from fact-checkers in a new disinformation ecosystem that may be refined and accelerated with the use of artificial intelligence (Franganillo, 2022), especially because of the “citizen training” to which some of its services have also been subjected. However, this openness and accessibility can also be seen in an ambiguous way: as a possibility of technological refinement leading to better functionality and the emergence of new tools to overcome the current disadvantages.

In answer to the first research question (Q1), the open launch of *ChatGPT* is a new field for fact-checkers to explore. According to the results obtained, and although more disadvantages than advantages are found, possible positive aspects are perceived in the use of the chatbot when identifying patterns and disinformation narratives, as well as in the identification of textual and linguistic

“ Experimentation with the chatbot could help identify different textual versions of the same disinformation content ”

mutations all in the same falsehood. Thus, the use of *ChatGPT* as an auxiliary tool in this area could contribute to certain professional routines of fact-checking, such as in the phase of locating false claims (Graves, 2017).

While it is true that, in the most widely available version of *ChatGPT*, it is not yet possible to get to the origin of the falsehood owing to the lack of transparency in the use of sources, experimentation with the chatbot could help to trace different textual versions of the same disinformation content. This knowledge would, in turn, allow for a deeper understanding of the operational patterns of disinformation, and as the results of our study show, the ability to carry out fact-checking more in line with the nature of the content to be disproved. The emergence of *ChatGPT* would also reinforce the understanding of the role of fact-checkers from an epistemological point of view, since the possibility of identifying the different textual formats that a falsehood can acquire could consolidate the professional imperative that consists of detecting lies in all their facets (Graves, 2016).

According to the results, another positive aspect perceived regarding the use of *ChatGPT* is the possibility of collecting contextual information quickly and concisely. The chatbot could be useful in another phase of the journalistic production process, such as reporting or the search for information, and could contribute to the streamlining of some professional routines in this area.

The possible uses of *ChatGPT* identified by the fact-checkers would be mainly of an auxiliary nature and under the close supervision of professional figures. For *ChatGPT* to be a tool regularly incorporated into fact-checking routines, a greater transparency and reliability in the use of its sources, performance in the fact-checking of audiovisual content, and more rigor and profusion in the responses generated would also be required. From the perspective of the fact-checkers, such functionalities would be key to maximizing the usefulness of *ChatGPT* in fact-checking practice (Q2).

For most fact-checking professionals, the emergence of *ChatGPT* reinforces their role as fact-checkers and accentuates a paradigm shift in their professional roles (Q3). The results are consistent with previous research on perceptions of information professionals regarding the introduction of AI into their professional routines, in which participants tend to defend their work and human judgment in addition to presenting themselves as supervisors of these new tools (Noain-Sánchez, 2022).

Beyond fact-checking, fact-checkers are perceived in this new ecosystem as “agents of context,” capable of shedding light on the increasingly complex darkness of disinformation. The emergence of *ChatGPT* also forces them to further diversify their fields of action in the fight against disinformation and to accelerate the implementation of actions in media education that empower citizens in the responsible use of artificial intelligence.

Although this research included all accredited Spanish fact-checking platforms, with interviews being carried out with professionals in positions of responsibility from said platforms, it should be noted that the study has been limited to the Spanish context. It is also important to note that the interviews conducted are limited to a specific time range, subject to constant technological changes. However, this study delves into new avenues of research in the field of information fact-checking and the emergence of new tools that have a direct impact on its activity. It is therefore a subject that should always continue to be worked on, with the development of research that also allows for other approaches, such as comparative and transnational perspectives.

“The emergence of *ChatGPT* reinforces the role of fact-checkers as context agents”

6. References

- Adami, Marina** (2023). “Is ChatGPT a threat or an opportunity for journalism? Five AI experts weigh in”. *Reuters Institute*, 23 March.
<https://reutersinstitute.politics.ox.ac.uk/news/chatgpt-threat-or-opportunity-journalism-five-ai-experts-weigh>
- Arias-Jiménez, Bryan; Rodríguez-Hidalgo, Claudia; Mier-Sanmartín, Catalina; Coronel-Salas, Gabriela** (2023). “Use of chatbots for news verification”. In: López-López, Paulo-Carlos; Barredo, Daniel; Torres-Toukoumidis, Ángel; De-Santis, Andrea; Avilés, Óscar (eds.). *Communication and applied technologies. Smart innovation, systems and technologies*. Singapur: Springer Nature, v. 318, pp. 133-143. ISBN: 978 981 19 6347 6
https://doi.org/10.1007/978-981-19-6347-6_12
- Arvanitis, Lorenzo; Sadeghi, McKenzie; Brewster, Jack** (2023). “Despite OpenAI’s promises, the company’s new AI tool produces misinformation more frequently, and more persuasively, than its predecessor”. *NewsGuard*, 15 March.
<https://www.newsguardtech.com/misinformation-monitor/march-2023>
- Beckett, Charlie** (2019). *New powers, new responsibilities: a global survey of journalism and artificial intelligence*. London School of Economics and Political Science.
<https://blogs.lse.ac.uk/polis/2019/11/18/new-powers-new-responsibilities>
- Bielik, Pavel; Višňovský, Ján** (2021). “Explanatory journalism. A new way how to communicate in digital era”. *Media literacy and academic research*, v. 4, n. 1.
https://www.mlar.sk/wp-content/uploads/2021/04/2_Bielik_Visnovsky.pdf

- Boczkowski, Pablo** (2004). *Digitizing the news. Innovation in online newspapers*. Cambridge: MIT Press. ISBN: 978 02 62268 84 4
- Boden, Margaret A.** (2022). *Inteligencia artificial*. Madrid: Turner. ISBN: 978 84 18895 35 7
- Brennen, J. Scott; Howard, Philip N.; Nielsen, Rasmus Kleis** (2018). *An industry-led debate: how UK media cover artificial intelligence*. RISJ Fact-Sheet. Universidad de Oxford.
<https://reutersinstitute.politics.ox.ac.uk/our-research/industry-led-debate-how-uk-media-cover-artificial-intelligence>
- Brewster, Jack; Arvanitis, Lorenzo; Sadeghi, McKenzie** (2023). "The next great misinformation superspreader: how ChatGPT could spread toxic misinformation at unprecedented scale. *NewsGuard*, 6 January.
<https://www.newsguardtech.com/misinformation-monitor/jan-2023>
- Canavilhas, João** (2022). "Inteligencia artificial aplicada al periodismo: traducción automática y recomendación de contenidos en el proyecto 'A European Perspective' (UER)". *Revista latina de comunicación social*, n. 80.
<https://www.doi.org/10.4185/RLCS-2022-1534>
- Carlson, Matt** (2015). "The robotic reporter: automated journalism and the redefinition of labor, compositional forms, and journalistic authority". *Digital journalism*, v. 3, n. 3, pp. 416-431.
<https://doi.org/10.1080/21670811.2014.976412>
- Cavaliere, Paolo** (2021). "From journalistic ethics to fact-checking practices: defining the standards of content governance in the fight against disinformation". *Journal of media law*, v. 12, n. 2, pp. 133-165.
<https://doi.org/10.2139/ssrn.3769410>
- Cazalens, Sylvie; Lamarre, Philippe; Leblay, Julien; Manolescu, Ioana; Tannier, Xavier** (2018). "A content management perspective on fact-checking". In: *WWW'18 companion: the 2018 web conference companion*, pp. 565-574.
<https://doi.org/10.1145/3184558.3188727>
- Crawford, Kate** (2021). *The atlas of AI: power, politics and the planetary costs of artificial intelligence*. New Haven: Yale University Press. ISBN: 978 0 300 20957 0
- Cybenko, Anne K.; Cybenko, George** (2018). "AI and fake news". *IEEE intelligent systems*, n. 33.
<https://doi.org/10.1109/MIS.2018.2877280>
- Diakopoulos, Nicholas** (2019). *Automating the news: how algorithms are rewriting the media*. Cambridge: Harvard University Press. ISBN: 978 0 674 23930 2
- Dörr, Konstantin-Nicholas** (2016). "Mapping the field of algorithmic journalism". *Digital journalism*, v. 4, n. 6, pp. 700-722.
<https://doi.org/10.1080/21670811.2015.1096748>
- Dwivedi, Yogesh K.; Kshetri, Nir; Hughes, Laurie; Slade, Emma Louise; ... Wright, Ryan** (2023). "Opinion paper: So what if ChatGPT wrote it? Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy". *International journal of information management*, v. 71, 102642.
<https://doi.org/10.1016/j.ijinfomgt.2023.102642>
- Eloundou, Tyna; Manning, Sam; Mishkin, Pamela; Rock, Daniel** (2023). "GPTs are GPTs: an early look at the labor market impact potential of large language models". Artículo en proceso de publicación.
<https://arxiv.org/pdf/2303.10130.pdf>
- Fallis, Don** (2021). "The epistemic threat of deepfakes". *Philosophy and technology*, v. 34, pp. 623-643.
<https://doi.org/10.1007/s13347-020-00419-2>
- Ferrucci, Patrick; Perreault, Gregory** (2021). "The liability of newness: journalism, innovation and the issue of core competencies". *Journalism studies*, v. 22, n. 11, pp. 1436-1449.
<https://doi.org/10.1080/1461670X.2021.1916777>
- Flores-Vivar, Jesús-Miguel** (2019). "Inteligencia artificial y periodismo: diluyendo el impacto de la desinformación y las noticias falsas a través de los bots". *Doxa comunicación*, n. 29, pp. 197-212.
<https://doi.org/10.31921/doxacom.n29a10>
- Flores-Vivar, Jesús-Miguel** (2020). "Datos masivos, algoritmización y nuevos medios frente a desinformación y fake news. Bots para minimizar el impacto en las organizaciones". *Comunicación y hombre*, n. 16, pp. 101-114.
<https://doi.org/10.32466/eufv-cyh.2020.16.601.101-114>
- Floridi, Luciano; Chiriatti, Massimo** (2020). "GPT-3: its nature, scope, limits, and consequences. *Minds & machines*, n. 30, pp. 681-694.
<https://doi.org/10.1007/s11023-020-09548-1>

- Franganillo, Jorge** (2022). "Contenido generado por inteligencia artificial: oportunidades y amenazas". *Anuario ThinkEPI*, v. 16. <https://doi.org/10.3145/thinkepi.2022.e16a24>
- García-Avilés, José-Alberto; Carvajal-Prieto, Miguel; Arias-Robles, Félix** (2018). "Implantación de la innovación en los cibermedios españoles: análisis de las percepciones de los periodistas". *Revista latina de comunicación social*, n. 73, pp. 369-384. <https://doi.org/10.4185/RLCS-2018-1260>
- García-Marín, David** (2022). "Modelos algorítmicos y fact-checking automatizado. Revisión sistemática de la literatura". *Documentación de ciencias de la información*, v. 45, n. 1, pp. 7-16. <https://doi.org/10.5209/dcin.77472>
- García-Vivero, Gloria; López-García, Xosé** (2021). "La verificación de datos en Europa. Análisis de 5 iniciativas europeas: Maldita.es, Newtral, Pagella Política, Les Décodeurs y BBC Reality Check". *AdComunica*, pp. 235-264. <https://doi.org/10.6035/2174-0992.2021.21.12>
- Gómez-de-Ágreda, Ángel; Feijóo, Claudio; Salazar-García, Idoia-Ana** (2021). "Una nueva taxonomía del uso de la imagen en la conformación interesada del relato digital. Deep fakes e inteligencia artificial". *Profesional de la información*, v. 30, n. 2, e300216. <https://doi.org/10.3145/epi.2021.mar.16>
- Grau, Mel** (2020). "New WhatsApp chatbot unleashes power of worldwide fact-checking organizations to fight COVID-19 misinformation on the platform". *International Fact-Checking Network (IFCN), Poynter Institute*, 4 May. <https://www.poynter.org/fact-checking/2020/poynters-international-fact-checking-network-launches-whatsapp-chatbot-to-fight-covid-19-misinformation-leveraging-database-of-more-than-4000-hoaxes>
- Graves, Lucas** (2016). *Deciding what's true. The rise of political fact-checking in American journalism*. New York: Columbia University Press. ISBN: 978 0 231175067
- Graves, Lucas** (2017). "Anatomy of a fact check: objective practice and the contested epistemology of fact checking". *Communication, culture and critique*, v. 10, n. 3, pp. 518-537. <https://doi.org/10.1111/cccr.12163>
- Graves, Lucas; Anderson, Christopher** (2020). "Discipline and promote: building infrastructure and managing algorithms in a 'structured journalism' project by professional fact-checking groups". *New media & society*, v. 22, n. 2, pp. 342-360. <https://doi.org/10.1177/1461444819856916>
- Graves, Lucas; Cherubini, Federica** (2016). *The rise of fact-checking sites in Europe*. Reuters Institute for the Study of Journalism. <https://reutersinstitute.politics.ox.ac.uk/our-research/rise-fact-checking-sites-europe>
- Guallar, Javier; Codina, Lluís; Freixa, Pere; Pérez-Montoro, Mario** (2020). "Desinformación, bulos, curación y verificación. Revisión de estudios en Iberoamérica 2017-2020". *Telos: Revista de estudios interdisciplinarios en ciencias sociales*, v. 22, n. 3, pp. 595-613. <https://doi.org/10.36390/telos223.09>
- Guo, Zhijiang; Schlichtkrull, Michael; Vlachos, Andreas** (2022). "A survey on automated fact-checking". *Transactions of the association for computational linguistics*, v. 10, pp. 178-206. https://doi.org/10.1162/tacl_a_00454
- Gupta, Ashish; Li, Han; Farnoush, Alireza; Jiang, Wenting** (2022). "Understanding patterns of COVID infodemic: a systematic and pragmatic approach to curb fake news". *Journal of business research*, v. 140, pp. 670-683. <https://doi.org/10.1016/j.jbusres.2021.11.032>
- Hameleers, Michael; Van-der-Meer, Toni G. L. A.** (2020). "Misinformation and polarization in a high-choice media environment: how effective are political fact-checkers?". *Communication research*, v. 47, n. 2, pp. 227-250. <https://doi.org/10.1177/0093650218819671>
- Hassan, Naeemul; Adair, Bill; Hamilton, James T.; Li, Chengkai; Tremayne, Mark; Yang, Jun; Yu, Cong** (2015). "The quest to automate fact-checking". In: *Proceedings of the 2015 computation + journalism symposium*. Columbia University. <http://cj2015.brown.columbia.edu/papers/automate-fact-checking.pdf>
- Herrero, Esperanza; Herrera-Damas, Susana** (2021). "El fact-checking hispanohablante: competencias, dificultades y propuestas de mejora desde la perspectiva de sus profesionales". *Profesional de la información*, v. 30, n. 6, e300612. <https://doi.org/10.3145/epi.2021.nov.12>
- Himma-Kadakas, Marju; Ojamets, Indrek** (2022). "Debunking false information: investigating journalists' fact-checking skills". *Digital journalism*, v. 10, n. 5, pp. 866-887. <https://doi.org/10.1080/21670811.2022.2043173>

- Holman, Lynette; Perreault, Gregory** (2023). "Diffusion of innovations in digital journalism: technology, roles, and gender in modern newsrooms". *Journalism*, v. 24, n. 5, pp. 938-957.
<https://doi.org/10.1177/14648849211073441>
- Jamil, Sadia** (2020). "Artificial intelligence and journalistic practice: the crossroads of obstacles and opportunities for the Pakistani journalists". *Journalism practice*, v. 15, n. 10, pp. 1400-1422.
<https://doi.org/10.1080/17512786.2020.1788412>
- Kasneci, Enkelejda; Sessler, Kathrin; Küchemann, Stefan; Bannert, Maria; Dementieva, Daryna; Fischer, Frank; ... Kasneci, Gjergji** (2023). "ChatGPT for good? On opportunities and challenges of large language models for education". *Learning and individual differences*, v. 103, 102274.
<https://doi.org/10.1016/j.lindif.2023.102274>
- Lichterman, Joseph** (2017). "Want to bring automation to your newsroom? A new AP report details best practices". *NiemanLab*, 5 April.
<https://www.niemanlab.org/2017/04/want-to-bring-automation-to-your-newsroom-a-new-ap-report-details-best-practices>
- Liu, Tianyu; Zhang, Yizhe; Brockett, Chris; Mao, Yi; Sui, Zhifang; Chen, Weizhu; Dolan, Bill** (2022). "A token-level reference-free hallucination detection benchmark for free-form text generation". In: *Proceedings of the 60th annual meeting of the Association for Computational Linguistics*, pp. 6723-6737.
<https://doi.org/10.18653/v1/2022.acl-long.464>
- López-Pan, Fernando; Rodríguez-Rodríguez, Jorge-Miguel** (2020). "El fact-checking en España. Plataformas, prácticas y rasgos distintivos". *Estudios sobre el mensaje periodístico*, v. 26, n. 3, pp. 1045-1065.
<https://doi.org/10.5209/esmp.65246>
- Maldita.es* (2022). "Probamos ChatGPT en un directo en Twitch: ¿cómo debemos interpretar las respuestas que nos da este programa de inteligencia artificial?". *Maldita.es*, 27 diciembre.
<https://maldita.es/malditatecnologia/20221227/chatgpt-twitch-prueba-respuestas-desinformacion>
- Malik, Kenan** (2018). "Fake news has a long history. Beware the state being keeper of the truth". *The Guardian*, 11 February.
<https://www.theguardian.com/commentisfree/2018/feb/11/fake-news-long-history-beware-state-involvement>
- Manfredi-Sánchez, Juan Luis; Ufarte-Ruiz, María-José** (2020). "Inteligencia artificial y periodismo: una herramienta contra la desinformación". *Revista Cidob d'afers internacionals*, n. 124, pp. 49-72.
<https://doi.org/10.24241/rcai.2020.124.1.49>
- Marconi, Francesco** (2020). *Newsmakers. Artificial intelligence and the future of journalism*. New York: Columbia University Press. ISBN: 978 0 231 19137 1
- Martínez-García, Luisa; Ferrer, Iliana** (2023). "Fact-checking journalism: a palliative against the COVID-19 infodemic in Ibero-America". *Journalism & mass communication quarterly*, v. 100, n. 2, pp. 264-285.
<https://doi.org/10.1177/10776990231164168>
- McIntyre, Karen Elizabeth; Lough, Kyser** (2021). "Toward a clearer conceptualization and operationalization of solutions journalism". *Journalism*, v. 22, n. 6, 1558-1573.
<https://doi.org/10.1177/1464884918820756>
- Miller, Jody; Glassner, Barry** (2016). "The 'inside' and the 'outside': finding realities in interviews". In: Silverman, David (ed.). *Qualitative research*. London: Sage. ISBN: 978 1 529 71297 1
- Moreno-Gil, Victoria; Ramon-Vegas, Xavier; Mauri-Ríos, Marcel** (2022). "Bringing journalism back to its roots: examining fact-checking practices, methods, and challenges in the Mediterranean context". *Profesional de la información*, v. 31, n. 2, e310215.
<https://doi.org/10.3145/epi.2022.mar.15>
- Moreno-Gil, Victoria; Ramon-Vegas, Xavier; Rodríguez-Martínez, Ruth** (2021). "Fact-checking interventions as counter-offensives to disinformation growth: Standards, values, and practices in Latin America and Spain". *Media and communication*, v. 9, n. 1, pp. 251-263.
<https://doi.org/10.17645/mac.v9i1.3443>
- Morrish, Lydia** (2023). Fact-checkers are scrambling to fight disinformation with AI. *The wired*, 1 February.
<https://www.wired.co.uk/article/fact-checkers-ai-chatgpt-misinformation>
- Noain-Sánchez, Amaya** (2022). "Addressing the impact of artificial intelligence on journalism: the perception of experts, journalists and academics". *Communication & society*, v. 35, n. 3, pp. 105-121.
<https://doi.org/10.15581/003.35.3.105-121>

- Oelrichs, Inga** (2023). "Adoption of innovations in digital sports journalism: the use of Twitter by German sports journalists". *Communication & sport*, v. 11, n. 2, pp. 288-312.
<https://doi.org/10.1177/2167479520961786>
- Öhman, Carl** (2020). «Introducing the pervert's dilemma: a contribution to the critique of deepfake pornography». *Ethics and information technology*, v. 22, pp. 133-140.
<https://doi.org/10.1007/s10676-019-09522-1>
- Organización Mundial de la Salud (2020). *Let's flatten the infodemic curve*
<https://www.who.int/en/news-room/spotlight/let-s-flatten-the-infodemic-curve>
- Palomo, Bella; Sedano-Amundarain, Jon Ander** (2018). "WhatsApp como herramienta de verificación de fake news. El caso de B de Buló". *Revista latina de comunicación social*, v. 73, pp. 1384-1397.
<https://doi.org/10.4185/RLCS-2018-1312>
- Pasquetto, Irene V.; Jahani, Eaman; Atreja, Shubham; Baum, Matthew** (2022). "Social debunking of misinformation on WhatsApp: the case for strong and in-group ties". *Proceedings of the ACM on human-computer interaction*, v. 6.
<https://doi.org/10.1145/3512964>
- Peña-Fernández, Simón; Meso-Ayerdi, Koldobika; Larrondo-Ureta, Ainara; Díaz-Noci, Javier** (2023). "Without journalists, there is no journalism: the social dimension of generative artificial intelligence in the media". *Profesional de la información*, v. 32, n. 2.
<https://doi.org/10.3145/epi.2023.mar.27>
- Pérez-Dasilva, Jesús-Ángel; Meso-Ayerdi, Koldobika; Mendiguren-Galdospín, Terese** (2020). "Fake news y coronavirus: detección de los principales actores y tendencias a través del análisis de las conversaciones en Twitter". *Profesional de la información*, v. 29, n. 3.
<https://doi.org/10.3145/epi.2020.may.08>
- Pérez-Seoane, Jesús; Corbacho-Valencia, Juan-Manuel; Dafonte-Gómez, Alberto** (2023). "Análisis de las publicaciones con mayor repercusión en Facebook de los fact-checkers iberoamericanos en 2021". *Icono 14. Revista científica de comunicación y tecnologías emergentes*, v. 21, n. 1.
<https://doi.org/10.7195/ri14.v21i1.1951>
- Ramon-Vegas, Xavier; Mauri-Ríos, Marcel; Rodríguez-Martínez, Ruth** (2020). "Redes sociales y plataformas de fact-checking contra la desinformación sobre la COVID-19". *Hipertext.net*, v. 21, pp. 79-92.
<https://doi.org/10.31009/hipertext.net.2020.i21.07>
- Redondo, Myriam** (2018). *Verificación digital para periodistas. Manual contra bulos y desinformación internacional*. Barcelona: Editorial UOC. ISBN: 978 84 9180 129 0
- Rodríguez-Pérez, Carlos** (2020). "Una reflexión sobre la epistemología del fact-checking journalism: retos y dilemas". *Revista de comunicación*, v. 19, n. 1, pp. 243-258.
<https://doi.org/10.26441/RC19.1-2020-A14>
- Rojas-Torrijos, José-Luis; Toural-Bran, Carlos** (2019). "Periodismo deportivo automatizado. Estudio de caso de AnaFut, el bot desarrollado por *El Confidencial* para la escritura de crónicas de fútbol". *Doxa comunicación. Revista interdisciplinar de estudios de comunicación y ciencias sociales*, n. 29, pp. 235-254.
<https://doi.org/10.31921/doxacom.n29a12>
- Russell, Stuart; Norvig, Peter** (2022). *Artificial intelligence: a modern approach. Fourth ed.* London: Pearson Education. ISBN: 978 1 292 40117 1
- Salaverría, Ramón; Buslón, Nataly; López-Pan, Fernando; León, Bienvenido; López-Goñi, Ignacio; Erviti, María-Carmen** (2020). "Desinformación en tiempos de pandemia: tipología de los bulos sobre la Covid-19". *Profesional de la información*, v. 29, n. 3, e290315.
<https://doi.org/10.3145/epi.2020.may.15>
- Sánchez-García, Pilar; Merayo-Álvarez, Noemí; Calvo-Barbero, Carla, Diez-Gracia, Alba** (2023). "Spanish technological development of artificial intelligence applied to journalism: companies and tools for documentation, production and distribution of information". *Profesional de la información*, v. 32, n. 2, e320208.
<https://doi.org/10.3145/epi.2023.mar.08>
- Sánchez-Gonzales, Hada M.; Sánchez-González, María** (2017). "Los bots como servicio de noticias y de conectividad emocional con las audiencias. El caso de Politibot". *Doxa comunicación. Revista interdisciplinar de estudios de comunicación y ciencias sociales*, v. 25, pp. 63-84.
<https://doi.org/10.31921/doxacom.n25a3>

- Sánchez-González, María; Sánchez-Gonzales, Hada M.; Martínez-Gonzalo, Sergio** (2022). "Inteligencia artificial en verificadores hispanos de la red IFCN: proyectos innovadores y percepción de expertos y profesionales". *Estudios sobre el mensaje periodístico*, v. 28, n. 4, pp. 867-879.
<https://doi.org/10.5209/esmp.82735>
- Singer, Jane B.** (2021). "Border patrol: the rise and role of fact-checkers and their challenge to journalists' normative boundaries". *Journalism*, v. 22, n. 8, pp. 1929-1946.
<https://doi.org/10.1177/1464884920933137>
- Soto-Sanfiel, María-Teresa; Ibiti, Adriana; Machado, Mabel; Marín-Ochoa, Beatriz-Elena; Mendoza-Michilot, María; Rosell-Arce, Claudio-Guillermo; Angulo-Brunet, Ariadna** (2022). "In search of the Global South: assessing attitudes of Latin American journalists to artificial intelligence in journalism." *Journalism studies*, n. 23, v. 10, pp. 1197-1224.
<https://doi.org/10.1080/1461670X.2022.2075786>
- Stencel, Mark; Ryan, Erica; Luther, Joel** (2022). *Fact-checkers extend their global reach with 391 outlets, but growth has slowed*. Duke Reporters' Lab, 17 June.
<https://reporterslab.org/fact-checkers-extend-their-global-reach-with-391-outlets-but-growth-has-slowed>
- Sun, Mengru; Hu, Wencai; Wu, Yun** (2022). "Public perceptions and attitudes towards the application of artificial intelligence in journalism: from a China-based survey." *Journalism practice*.
<https://doi.org/10.1080/17512786.2022.2055621>
- Tandoc Jr., Edson C.; Lim, Zheng Wei; Ling, Richard** (2018). "Defining 'fake news'. A typology of scholarly definitions". *Digital journalism*, v. 6, n. 2, pp. 137-153.
<https://doi.org/10.1080/21670811.2017.1360143>
- Thorne, James; Vlachos, Andreas** (2018). "Automated fact checking: task formulations, methods and future directions". In: *Proceedings of the 27th International Conference on Computational Linguistics*, pp. 3346-3359.
<https://aclanthology.org/C18-1283>
- Tracy, Sarah J.** (2020). *Qualitative research methods. Collecting evidence, crafting analysis, communicating impact*. Hoboken, New Jersey: Wiley Blackwell. ISBN: 978 1 405 19202 6
- Ufarte-Ruiz, María-José; Manfredi-Sánchez, Juan-Luis** (2019). "Algoritmos y bots aplicados al periodismo. El caso de Narrativa Inteligencia Artificial: estructura, producción y calidad informativa". *Doxa comunicación. Revista interdisciplinar de estudios de comunicación y ciencias sociales*, v. 29, pp. 213-233.
<https://doi.org/10.31921/doxacom.n29a11>
- Ufarte-Ruiz, María-José; Peralta-García, Lidia; Murcia-Verdú, Francisco-José** (2018). "Fact checking: un nuevo desafío del periodismo". *El profesional de la información*, v. 27, n. 4, pp. 733-741.
<https://doi.org/10.3145/epi.2018.jul.02>
- Unesco (2018). *Journalism, fake news & disinformation: Handbook for journalism education and training*. Paris: Unesco. ISBN: 978 92 3 100281 6
<https://unesdoc.unesco.org/ark:/48223/pf0000265552>
- Van-Dalen, Arjen** (2012). "The algorithms behind the headlines". *Journalism practice*, v. 6, n. 5-6, pp. 648-658.
<https://doi.org/10.1080/17512786.2012.667268>
- Vara-Miguel, Alfonso; Amoedo-Casais, Avelino; Moreno-Moreno, Elsa; Negredo-Bruna, Samuel; Kaufmann-Argueta, Jürg** (2022). *Digital News Report España 2022*. Servicio de publicaciones de la Universidad de Navarra.
<https://doi.org/10.15581/019.2022>
- VerificaRTVE (2023). "La inteligencia artificial de ChatGPT, a prueba frente a la desinformación y los bulos". *VerificaRTVE*, 27 enero.
<https://www.rtve.es/noticias/20230127/prueba-chat-gpt-frentedesinformacion/2418106.shtml>
- Vizoso, Ángel; Vázquez-Herrero, Jorge** (2019). "Fact-checking platforms in Spanish. Features, organisation and method". *Communication & society*, v. 32, n. 1, pp. 127-142.
<https://doi.org/10.15581/003.32.37819>
- Wardle, Claire; Derakhshan, Hossein** (2017). *Information disorder: toward an interdisciplinary framework for research and policy making*. Consejo de Europa.
<https://edoc.coe.int/en/media/7495-information-disorder-toward-an-interdisciplinary-framework-for-research-and-policy-making.html>
- Wimmer, Roger D.; Dominick, Joseph R.** (2013). *Mass media research: an introduction*. Belmont, California: Cengage Learning. ISBN: 978 1 133307334
- Yousuf, Mohammad** (2023). "Mediating the truth: Influences of routines on legacy news media fact-checking". *Journalism practice*.
<https://doi.org/10.1080/17512786.2023.2169187>

AI application in journalism: *ChatGPT* and the uses and risks of an emergent technology

Beatriz Gutiérrez-Caneda; Jorge Vázquez-Herrero; Xosé López-García

Nota: Este artículo puede leerse en español en:
<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87416>

Recommended citation:

Gutiérrez-Caneda, Beatriz; Vázquez-Herrero, Jorge; López-García, Xosé (2023). "AI application in journalism: *ChatGPT* and the uses and risks of an emergent technology". *Profesional de la información*, v. 32, n. 5, e320514.
<https://doi.org/10.3145/epi.2023.sep.14>

Manuscript received on 4th July 2023
Accepted on 31st July 2023



Beatriz Gutiérrez-Caneda ✉
<https://orcid.org/0000-0002-7191-1886>

Universidade de Santiago de Compostela
Facultade de Ciencias da Comunicación
Av. de Castelao, s/n
15782 Santiago de Compostela, Spain
beatriz.gutierrez.caneda@usc.es



Jorge Vázquez-Herrero
<https://orcid.org/0000-0002-9081-3018>

Universidade de Santiago de Compostela
Facultade de Ciencias da Comunicación
Av. de Castelao, s/n
15782 Santiago de Compostela, Spain
jorge.vazquez@usc.es



Xosé López-García
<https://orcid.org/0000-0002-1873-8260>

Universidade de Santiago de Compostela
Facultade de Ciencias da Comunicación
Av. de Castelao, s/n
15782 Santiago de Compostela, Spain
xose.lopez.garcia@usc.gal

Abstract

AI integration in media newsrooms is changing professional routines, required profiles and journalistic products. The acceleration in the development of this technology since the end of 2022 makes it necessary to review these tools in order to analyse their uses and risks, as well as their possible applications in the journalistic field. *OpenAI* launched *ChatGPT* in November 2022 open to the public. This tool has been a key element in this technological revolution. This paper analyses *ChatGPT* applications in journalistic newsrooms through its introduction during a working day in the routine of different professionals. A mixed methodology was used to carry out this experience: a benchmarking of AI tools applied to journalism was created, a walkthrough experience was carried out with *ChatGPT* and, based on the results obtained in these first two phases, the experiment with journalists was designed. A total of 12 journalists of different ages and sectors with little or no previous contact with technological tools based on artificial intelligence, as the aim of the study is to observe the application in newsrooms without the need for technical training, participated in this research. The results show the usefulness of the tool to automate mechanical processes, rewrite texts, analyse data and even serve as a content idea creator. On the other hand, journalists have also perceived significant risks such as inaccuracy of AI as well as lack of 'empathy'.

Keywords

Journalism; Artificial intelligence; AI; *ChatGPT*; Automated journalism; Computational journalism; Hi-tech; Journalistic ethics; Journalists.

Funding

This article is part of the R&D project *Digital-native media in Spain: Strategies, competencies, social involvement and (re)definition of practices in journalistic production and diffusion* (PID2021-122534OB-C21), funded by the Spanish Ministerio de Ciencia e Innovación (MCIN) / Agencia Estatal de Investigación (AEI)/10.13039/501100011033/, and by "European Regional Development Fund (ERDF) A way of making Europe".



1. Introduction

The automation of routine and monotonous tasks in the media is a present and future reality. The network society of the third decade of the third millennium is structured around the Internet galaxy (Castells, 2001) and its rampant platformization (Van-Dijck; Poell; De-Waal, 2018), and immersed in the fourth industrial revolution, which significantly modifies production systems as well as different aspects of society (Micó; Casero-Ripollés; García-Orosa, 2022). In this context, the introduction of artificial intelligence in the processes of searching, production, dissemination and management of communication messages has established a platform of turbines that will progressively drive the exponential multiplication of communication flows and personalization, with new ethical challenges (Hermann, 2022). Today there is a certain enthusiasm for the arrival of artificial intelligence, due to its potential to transform and introduce efficiency into communicative processes, representing a technocentric vision of communication, but there are risks and challenges due to a lack of transparency from socio-legal and scientific-computing perspectives (Larsson; Heintz, 2020). These are two dimensions that need to be monitored and analysed, as far as possible, to understand and correct possible dysfunctions.

The complexity implied by the changes underway in the journalistic field not only stimulates renewed debates in the field of communication, but also drives new transformations in the communication ecosystem that bring opportunities and challenges in the context of the next technological revolution, in which it is announced that machines will drive machine learning, imitate human thoughts and behaviours and perform new cognitive functions (Samuel *et al.*, 2022). Big companies use machine learning in their activities, both for big data analysis in general and through different strategies, especially in digital marketing and programmatic advertising. The disruptive impact of artificial intelligence, which affects all phases of the advertising process (Martínez-Martínez; Aguado-Terrón; Sánchez-Cobarro, 2022), has multiplied intermediaries and contributed to changes in the media environment, which is increasingly in need of innovative business models (Evans; Raats; Von-Rimscha, 2017) that ensure their sustainability in a scenario in which artificial intelligence affects the processes, practices and results of new companies (Chalmers; MacKenzie; Carter, 2021). Several companies have already revolutionized their business models using artificial intelligence (Mishra; Tripathi, 2021), but the models still have room for further enrichment and efficiency as renewed integrated systems are explored for the new scenario of the fourth industrial revolution (Ross; Maynard, 2021), because of the radical change it implies for society as a whole.

Artificial intelligence, which experts recognise as one of the current technologies that will drive a more efficient journalism, has entered all departments of newspaper companies and, together with different tools, is guiding many of the changes that are taking place in newsrooms. The possibilities it offers are very different, ranging from the elaboration of simple news items to the exploration of new dimensions for a piece or for the improvement of interactions with users through information processing. The application of algorithms and artificial intelligence to journalism has developed at a dizzying pace in a very short period (Parratt-Fernández; Mayoral-Sánchez; Mera-Fernández, 2021), which has also resulted in significant academic interest and output. However, as the changes introduced by the latest generation tools are very recent and their impact is highly relevant, it is necessary to monitor them periodically and in as timely a manner as possible to understand the trends and the most used tools right now. The results of this monitoring should provide perspective for the regulatory challenges facing artificial intelligence, in accordance with the socially acceptable limits of our cultural context.

In November 2022, *OpenAI* launched *ChatGPT* (Márquez, 2023), an AI tool with a conversational interface that answers users' questions and can perform actions involving natural language generation. This launch served as a starting point, and after that date many different AI initiatives accelerated and came to light. *Google* announced *Bard*, a conversational AI that is not yet available in all territories (*Google*, 2023); *Microsoft* included a 'Chat' function in its search engine with the help of *OpenAI* (Fernández, 2023a) that works in a similar way and offers the possibility of creating images with artificial intelligence, something already seen in *DALL-E*, another *OpenAI* system that creates images from descriptions made in natural language (*OpenAI*, 2023a). In March 2023, Elon Musk created his own AI company, *X.AI*, which he publicly revealed a month later (Prego, 2023). In May 2023, *Google* announced that it would include a generative AI response at the top search results (Hazard-Owen, 2023), which led *OpenAI* to promise a new paid version of *ChatGPT* with more functionalities, such as a direct connection of the tool to the network, which up to this point could only be done with the help of external software.

However, the arrival of *ChatGPT* was also controversial and reopened ethical and deontological debates on the limits of technology uses, which in some cases even reached the legal sphere. In Italy its use has been banned for not respecting the data protection law (Buj, 2023). In March 2023, more than 1,000 AI leaders, including Elon Musk, Steve Wozniak (co-founder of *Apple*), Jaan Tallin (co-founder of *Skype*) and Max Tegmark (*MIT*), signed a letter calling for a halt to AI training, especially those programs that may have a higher capability than *GPT-4*, for at least six months to work on making current systems more robust and reliable (Aguilar, 2023). The problem highlighted in this missive is the lack of control and knowledge about the AI systems that are being created and the risks that they may pose.

1.1. Artificial intelligence: the journalism revolution in the third millennium

Artificial intelligence, which has changed the work of journalism professionals in recent years, now intensifies, with the emergence of next-generation tools, the disruptive effects on the organization of media companies and journalistic practices. Artificial intelligence as a technology has existed for several decades and in this time it has achieved great advances and aroused the curiosity of many technologists and companies in different sectors around the world (Prasad; Choudhary, 2021).

Realising that artificial intelligence can help them with some of the challenges they were facing (**De-Lima-Santos; Ceron, 2022**), newspaper companies began to show an early interest in the communicative processes involved in the application of natural language processing, the detection of informative trends, and in the automatic production of texts (**Canavilhas, 2022**). According to focus studies, there is a belief among journalism experts and professionals that this technology –if current forecasts are fulfilled– will play an important role in the industry, with three main lines of application: the automation of content –textual and audiovisual– the verification and improvement of access to information and monetization –including subscription and loyalty systems– and the personalization of content (**De-Lara; García-Avilés; Arias-Robles, 2022**).

There is also a current of opinion among communication professionals and academics that this technology will not have a negative impact on the journalistic labour market (**Calvo-Rubio; Ufarte-Ruiz, 2020**) and that professionals need to be adequately prepared to incorporate these state-of-the-art tools into their practices. This vision, however, coexists with fears that, in some countries, many journalists may be displaced from the media (**Yu; Huang; Jones, 2020**). **Sánchez-García et al. (2023)** even points out that, following their research on AI applied to journalism in Spain, “the experts consulted draw an ‘irreversible’ technological reality, a ‘forced change’ that, however, the media receive with ‘slowness’, ‘distrust’ and ‘unawareness’”. Time will help to dispel doubts about the real impact on the labour market in the future, under the watchful eye of research on different geographical areas.

The potential of artificial intelligence to transform journalism will translate into tangible results if it is oriented towards the development of news pieces that are accurate and accessible, that cultivate diversity, that are relevant and timely and, in short, that contribute to a higher quality in the processes of message development so that citizens are better informed and more satisfied with the information they receive (**Lin; Lewis, 2022**). Without doubting this efficiency – as AI has introduced journalism in an unknown scenario, but which must be explored through a path filled with challenges and risks –there are many questions on the horizon. This has led to important ethical, labour and social debates in the field of communication in recent years (**Túñez-López; Toural-Bran; Valdiviezo-Abad, 2019**), but always with the desire to understand the impact and take advantage of the benefits it can bring, such as the extension of current automated textual news to audio and video information on demand, which will favour an unstructured non-linear consumption of the news, and will promote changes in the business model due to new ways of relating to the audience and distribution of the product (**Túñez-López; Fieiras-Ceide; Vaz-Álvarez, 2021**). Journalists and experts perceive that AI will enhance the capabilities of journalists by saving time, increasing the efficiency of news creation processes, and thereby increasing productivity (**Noain-Sánchez, 2022**), while also highlighting a perceived tension between the industry and the profession in highlighting the hopes and pitfalls of this technology (**Moran; Shaikh, 2022**). The different perspectives that coexist point to the need for a better understanding of the consequences based on concrete experiences, with case studies that will then allow more global interpretations.

The need for journalists to be trained in artificial intelligence and the tools used by this technology (**Gonçalves; Melo, 2022**) and the ethical debates (**Noain-Sánchez, 2022**) are two areas of particular concern in the profession, which demonstrates the need to implement continuous updating of programs aimed at working professionals, as well as their incorporation into the regulated studies of journalism degrees, and the need for continuous control and supervision of the processes carried out by AI in the journalistic field.

High-tech journalism has become the sign of the digital times of the third millennium. This trend demands competencies from journalism professionals on the crossover between technology and journalistic content creation (**López-García; Vizoso, 2021**). This gives rise to renewed professional profiles and new names for conceptualizing the impact of AI on the journalism industry. “Exo journalism” (**Tejedor; Vila, 2021**) is one of these new names, that joins others like robot journalism, computational journalism, artificial journalism or automated journalism, around which there is debate – each name introduces nuances, as they are not clearly defined and delimited (**Mooshammer, 2022**). All this occurs at a time when many studies and contributions recommend searching for common patterns of study for a better understanding of automation in newsrooms and artificial intelligence in journalism (**Danzon-Chambaud, 2021**). Technology is one of the key elements in the approach to media and journalism strategy and development (**Vállez; Codina, 2018**), but the different tools and trends at certain times encourage denominations of more or less long life, which now, with automation and artificial intelligence seems to be grouped, preferably, around the umbrella of automated journalism, still little present in the training offer of the curricula of journalism degrees (**Ufarte-Ruiz; Fieiras-Ceide; Túñez-López, 2020**), although with proposals for its introduction from the applied point of view and from critical reflection (**Gómez-Diago, 2022**), but increasingly present in newsrooms and academic literature and with all denominations providing nuances (**Cohen; Hamilton; Turner, 2011; Karlsen; Stavelin, 2013; Túñez-López; Toural-Bran; Valdiviezo-Abad, 2019; Marconi, 2020; Ca-**

“*ChatGPT's* limitations are mainly marked by: (1) its limited knowledge of the world, (2) the fact of being disconnected from the network, (3) producing incorrect information and presenting it as truthful, and (4) not correctly following the instructions given by the user; furthermore, (5) certain biases are detected, probably derived from the information it received in its training”

navilhas, 2022; García-Orosa; Pérez-Seijo; Vizoso, 2022; Otero-González, 2022). The number of articles resulting from research in this field continues to grow, which is why we only reference a sample. The growth of academic production linking journalism and artificial intelligence is vertiginous, as shown by recent systematic reviews (García-Orosa; Canavilhas; Vázquez-Herrero, 2023).

Artificial intelligence and automation are part of the world of journalism today and are present in the newsrooms of today's digital media because it is difficult to separate journalism from its technology—it is dependent on some kind of technology (Zelizer, 2019). Hence, although economic views on its introduction and presence in journalism prevail, more attention needs to be paid to the ethics and ontological limits of automated journalism (Porlezza; Ferri, 2022). For AI to make contributions to good journalism—to improving the functioning of democracy (Lin; Lewis, 2022)—good regulation is necessary, which, among other things, avoids the disappearance of authorship (Krausová; Moravec, 2022) and a rigorous transparency policy that is accompanied by proper monitoring of the communicative processes in which this technology is applied, in the context of the ethical discussion around current technologies (Israel; Amer, 2022). The main professional and ethical issues focus on undermining creativity, a lack of monitoring, bias, transparency, fairness, data utilization and data quality (Ali; Hassoun, 2019). Responses, therefore, should focus on AI's effects on the basic elements of journalism.

The recent emergence of the so-called synthetic media (Ufarte-Ruiz; Murcia-Verdú; Tüñez-López, 2023) —media that lack journalists and where all work routines depend on AI— highlights the dizzying speed at which this technology continues to develop. While progress is being made in the regulation and monitoring of the effects of the introduction of artificial intelligence in technologically mediated communication and, above all, in digital journalism, the main challenge that the journalistic profession warns of is the need to know the existing tools and how to use them to avoid distortions, to understand the risks involved and to enter the complex world of the challenges posed by this technology for quality journalism, which can undoubtedly provide strength, although introducing new risks and threats. This technology, like that which has preceded it, must be seen as a new aid and, as it raises doubts and challenges, the ethical dilemmas must be contemplated from the core values that underpin good journalism, such as truth, justice, freedom and responsibility, which must be applied by journalists who, until proven otherwise, have capabilities superior to those of machines (Ventura-Pocino, 2021). What is needed, however, is for journalists to have the ability to monitor technology, which is a training issue and therefore a priority and should be programmed based on pioneering experiences for journalism students (Gómez-Diago, 2022) and for professionals who work in journalism.

1.2. ChatGPT and AI democratization

ChatGPT is an OpenAI tool introduced in December 2022 and made freely available. It works with GPT3, an autoregressive learning model, and allows the user to converse with it through a chat. Users can access it for free and do not need to have programming knowledge, although there are paid versions that offer more functionalities. The accessibility and usability of the user interface, both of ChatGPT and other OpenAI tools such as DALL-E, are a large part of the success of these tools, together with the fact that they are open to the public and cause enthusiasm and fascination among users, which facilitates their viralization in social networks.

The recent opening of the AI tool ChatGPT to the public, at the end of 2022, has accelerated the introduction of this type of technology not only in newsrooms but also in other companies in different fields. OpenAI defines itself as a natural language model that interacts conversationally (in chat format), allowing it to

“answer follow-up questions, admit mistakes, challenge incorrect assumptions and reject inappropriate requests” (OpenAI, 2023b).

ChatGPT is available in different languages, however, its interaction level depends on the language it is using. Another limitation is that all its knowledge comes from the data used for its training, in 2021.

ChatGPT has certain features and skills that make it especially attractive as an assistant in newsrooms and other communication-related companies. Since it is open to the public, numerous tests have been carried out with this tool in very different fields such as digital marketing, programming, education and journalism. Interviews have been conducted on different topics and a scientific article has even been written in collaboration with it, exploring the possibilities it offers to the field of communication and education (Pavlik, 2023).

In February 2023, OpenAI began to offer a paid version of ChatGPT with the name ChatGPT Plus. This new version offers general access to ChatGPT (even when there are traffic peaks, a current problem with the free version of the tool), a faster response, and priority access to new features. This version is available for \$20 per month (Fernández, 2023b). In this context, the objective of this research is to draw the current landscape of the use of artificial intelligence in newsrooms through a radiography of AI tools applicable to journalism and a case study of the OpenAI ChatGPT proposal applied and tested in the journalistic context.

“ This AI demonstrates its potential to reduce the time it takes to produce, write, manage and disseminate journalistic content, which at the same time could result in a reduced workload for journalists ”

2. Methods

To present the current panorama of the use of AI in the journalistic field and to determine the advantages that the appearance of the *ChatGPT* tool could bring to newsroom routines –the main goal of this paper– a mixed methodology with an exploratory and experimental nature has been chosen. This method allows us to get to know the current panorama in depth and to test the tool.

We start with the hypothesis that (1) the incorporation of AI in newsrooms facilitates the work of journalists, streamlining and automating processes and routines, although (2) AI use for journalistic purposes can be potentially dangerous from an ethical and legal perspective, which makes it necessary to regulate these tools. The methodological design involves three approaches: benchmarking of AI tools applied to journalism, a walkthrough experience with *ChatGPT* and, finally, an experiment with users following the Experience Sampling Method (**Berkland, 2017**).

2.1. Benchmarking of AI tools in journalism

Following a literature review on automated journalism, an adaptation of the *Prisma* method (**Page et al., 2021**) was applied to carry out a systematized exploration of the existing artificial intelligence tools that can be applied in the journalistic field. The search was conducted in two ways: on a general search engine (*Google*) and through scientific databases (*WoS* and *Scopus*). *Google* searches were conducted between 15th February and 1st March 2023, with the following terms: (1) “automated journalism”; (2) journalism AND algorithms; (3) journalism AND “artificial intelligence”; (4) use of AI tools on media. These formulas were used in both English and Spanish and the top ten results of each search were selected from the obtained results. For the searches in scientific databases, the parameters listed in Table 1 were established and the query was carried out during February and March 2023.

Table 1. Database searches

Database		Scopus	WoS
Parameters	Time range	2019–2023	2019–2023
	Language	Spanish, English	Spanish, English
Search equations		automated journalism	
		periodismo automatizado	
		“artificial intelligence” AND “journalism”	
		“inteligencia artificial” AND “periodismo”	
		“journalism” AND “algorithms”	
	“periodismo” AND “algoritmos”		

A total of 129 results were selected. From them, 100 were examined, after excluding duplicate and unrelated results, in search of journalistic initiatives and projects using AI tools. In the review of scientific articles, the abstract, methodology and results sections were examined. The tools were classified according to: functions (natural language generation, speech to text, text to speech, text to image, image to text, image recognition, data analysis); stage of the journalistic process where they are applied (information gathering, production, distribution and verification); and possible tasks to be performed in newsrooms (transcription of interviews or analysis of data from an external source, for example). After this first classification, we proceeded with an in-depth analysis to establish what type of tools are the most used, which were the first to be introduced in newsrooms, and what are the trends in the media in terms of the use of AI. A total of 76 tools from nine countries were collected, providing a snapshot of the roles that AI plays in journalism today and allowing these roles to be tested in subsequent experiences with *ChatGPT*.

2.2. Walkthrough

For the experimental design, a series of tests were conducted with *ChatGPT* between January and March 2023 to determine its capabilities, limitations and opportunities for improvement. These tests were divided into two parts: a first, completely experimental one, in which we talked with the AI to understand how it works and see what responses it offers to different inputs; and a second one in which we followed a walkthrough methodology (**Mahatody; Sagar; Kolski, 2010**), simulating the behaviour of a user (journalist) using the tool for the first time, without prior knowledge of its features or limitations.

During the second part, the results of the first part were taken into consideration, as well as the uses of AI applied to journalism obtained from previous benchmarking and the experiences reported in specialized media and by experts in the sector. From here, a list of possible actions was drawn up to evaluate the capabilities and opportunities offered by *ChatGPT* for the media.

After the exploration, and following the methodology employed by **Pavlik (2023)**, the tool was asked to list the actions it could perform to assist in journalistic routines, to complement the study and compare them with those that had been discovered in practice (Annex I).

2.3. Experiment with journalists

To test the possibilities offered by *ChatGPT* applicable in newsrooms, an experiment was conducted with working journalists from different sectors and with different roles in the media. Based on the results of the benchmarking of AI tools applied to journalism and the walkthrough experience, the Experience Sampling Method (ESM) was followed to obtain data from participants while they performed their daily tasks, allowing them to report “on the nature and quality of their experiences at that moment and in their natural environment” (Berkland, 2017). This ensured that the tests performed with the AI were as similar as possible to how a user would make use of the tool if it were incorporated into their professional routines.

Twelve journalists hired by the media—eight women and four men, between 24 and 46 years old—agreed to participate. A total of 33.3% of the participants work as journalists in a television medium, 25.0% work in the written press, another 25.0% in radio, 8.3% work in news agencies, and the remaining 8.3% in digital media.

The experiment was conducted between 20th April and 7th May 2023, in a non-face-to-face and asynchronous manner. It consisted of a prequestionnaire, two experimentation phases (A and B) and a final questionnaire. In phase A, participants were guided through the experience; in phase B, they were free to test *ChatGPT* in any way.

The prequestionnaire delimited the participants’ sociodemographic profile, and their previous AI and technological knowledge and affinity in the journalistic field. The tasks that these professionals perform on a daily basis were also collected to check whether the use of *ChatGPT* could be adjusted to them. Considering the results of the walkthrough experience, it was estimated that the use of *ChatGPT* would be more appropriate for tasks related to writing, data analysis and hierarchization or organization of information, so it was decided to focus the experiment on the use of this tool on these types of tasks.

Existing relationships with technological innovation are also a very relevant aspect in this experiment, as it was intended to measure the ability of an average journalist to use this AI tool without extensive prior training. When asked about this issue, 66.7% of respondents indicated that they find AI attractive and interesting, but they also raised risks related to privacy and ethics, among other issues; 25.0% of respondents considered the tool useful but they were not enthusiastic about it; while the remaining 8.3% found it useful and very interesting for some professions, however in journalism they only contemplated AI for automated and technical tasks, believing that it can be dangerous in some cases.

All the participants had already heard about *ChatGPT* but only half had already used this or other AI tools. Among the reasons for not having tried any such software, participants affirmed that they had not had time or patience or that they felt “terrible laziness” and that they did not want to “be a witness of our substitution as professionals.” In the workplace, only 25.0% of participants had used before the experiment. They used this technology to translate texts, for brainstorming, data analysis and to create code snippets for data processing. Even when they did not use it, participants considered AI useful in their professional routines and, for those who already used it, they considered it useful to implement more systems with this technology.

In addition to *ChatGPT*, the participants were asked about their knowledge of other AI tools: 33.3% of respondents also knew about *DALL-E*, the *OpenAI* tool that creates images from text. Besides that, two participants knew about the Telegram bot that allows *ChatGPT* use from this app and one participant mentioned knowing about the AI tool *Midjourney*.

Before the experiment, participants were asked about which AI systems they would integrate in their workplace and why. The aim behind this question was to observe whether their opinion changed after trying *ChatGPT*. Among the answers, the idea of using a system to (1) transcribe interviews was repeated; also mentioned was the usefulness of this technology for (2) writing and summarizing texts, (3) subtitling, (4) creating graphics, (5) generating ideas through searches made by other users, (6) data analysis and pattern detection, (7) transcribing phone calls with sources in other languages, (8) fact-checking and information verification, or (9) social media management, among other ideas.

After the initial questionnaire, participants went through phase A of the experiment. *ChatGPT* works with natural language, therefore programming knowledge is not needed to use it. However, it is important to note that this tool does not always answer the same way, even when the asked content is the same: the way in which the question is asked influences the answer. During the *ChatGPT*’s first months, numerous experts and professionals published—and continue to publish—lists of prompts or instructions to achieve specific results with the tool and some companies have started to offer browser extensions (such as *AIPRM* for *ChatGPT*) that allow users to choose from a list of commands already created to ask the AI what the user needs.

Thus, participants were provided with a guide that included both general orientations and a series of ‘standard instructions’ in the first phase. Participants could choose between a list of prompts (Annex II) already prepared according to their profiles and in which they only had to modify some variables to adapt them to their professional routines. This document explained how *ChatGPT* worked so that they could use it whenever they wanted, preferably during their working hours, and provided them with a table containing a series of prompts related to text creation (series 1), text adaptation (series 2), and information hierarchization (series 3). This table included different customizable parameters (in colour) and a fixed part of the text (without colour). The complete list of prompts can be found in Annex II. In phase B, participants were free to use the AI tool as they saw fit. To conclude the experience, participants were asked to fill out a final form evaluating the experience, pointing out positive and negative aspects of the tool.

3. Results

3.1. AI tools applied to journalism

In the table resulting from the benchmarking of AI applied to journalism, 76 tools belonging to nine countries that use this technology in the journalistic field were collected. One of the most common uses of AI in journalism is the automated writing of news whose main source is structured data. Due to the system's own limitations, the news covered in this way is simple information, based on quantitative data and following a clear scheme or structure, such as sports information, market information or election results. Examples include: *Wordsmith* from *Automated Insights*, used by the *Associated Press* agency to write news on sport and stock markets; *Heliograf*, which also writes autonomous sport and political news for *The Washington Post*; or *Syllabs*, which wrote news on the French legislative elections in 2015 for *Le Monde* (*Laboratorio de Periodismo*, 2018; **García-Avilés**, 2019; **Manfredi-Sánchez**; **Ufarte-Ruiz**, 2020).

Another function of AI applied to journalism is data analysis. This technology makes it possible to extract patterns and analyse information in a very short time and with great precision, to the point of being able to predict certain events. This is the case with *Virality Oracle* –a tool used by *The Washington Post* that predicts which topics will become viral (**García-Avilés**, 2019)– and *Lynx Insights* –used by *Reuters*, this tool compiles and analyses data so that a journalist can then write the news story (**Agarwal**, 2018).

After studying the range of AI tools applicable to journalism, classified in Table 2, it was concluded that these applications have very diverse functions and that they are applied in all parts of the communication process, from automated writing to data analysis, from image creation to verification.

Table 2. Types of AI tools according to their function

	Functions	Possible applications
1	Data analysis	Creation of information from databases, detection of patterns, forecasting from data.
2	Automating writing	Writing any type of text, such as news, reports or scripts from structured data.
3	Image recognition	Recognition of people, objects or places in an image. Identification of a specific image and reverse search.
4	Text to audio	Creation of podcasts and other types of audio files without the need for a human voice.
5	Audio to text	Interviews transcription, automatic subtitles, etc.
6	Video to text	Text description of video images, summary of video images, automatic captioning.
7	User interaction	Comment moderation, direct conversation with users.
8	Content curation	Personalization of content based on user preferences.

The most common function detected during the benchmarking process was content production, with 40 of the tools analysed being dedicated to this function, whether creating text news, graphics or audiovisual pieces. Another of the most common categories involves data analysis, followed by the chatbot modality. Sometimes tools combine several functions, such as chatbots that offer news verification or news aggregators, which first need to analyse data to then offer personalized content. On many occasions it is external companies that offer the technology to the media: *Automated Insights*, *Narrative Science* or *Monok* were some of the AI companies identified.

3.2. Walkthrough experience

To gain an in-depth understanding of the possibilities of *ChatGPT* in the journalistic field, a series of tests were carried out with the AI, asking it to perform certain tasks based on previous results, obtaining the results shown in Table 3.

From this experience it is clear that *ChatGPT*'s limitations are mainly marked by: (1) its limited knowledge of the world; (2) its being disconnected from the network; (3) producing incorrect information and presenting it as truthful; and (4) not correctly following the instructions given by the user (for example, asked for a text of 1,050 words, it offers a longer text); in addition, (5) certain biases are detected, probably derived from the information it received in its training, for example, gender bias. It is also important to note that *ChatGPT* does not always warn of its limitations. In the case of subtitling videos or summarizing texts provided to it through links, the tool pretends to perform the required action, since it cannot access the Internet or 'watch' a video; on some occasions it pointed out the limitation and on others it offered invented information. On the other hand, it is also important to highlight advantages such as: (1) its ease of rewriting, restructuring or translating a text; (2) providing new ideas; and (3) generating different content (a diet, a plan for social networks or an exercise routine, among others).

3.3. Newsroom experience

The walkthrough experience provided a significant amount of data to better understand chat behaviour. However, it is necessary for working professionals without experience or specialized training in artificial intelligence to test the tool to avoid potential biases.

Table 3. Results obtained from the walkthrough experience

	Required action	Can it do it?	Specifications
1	Summarize texts	Yes	Sometimes details are lost, which means the summary is not always understandable.
2	Rewrite texts (and adapt them to different formats)	Yes	You can specify the style you want it to use or keywords, limit the length, etc.
3	Translating texts	Yes	If the language is not specified, it defaults to English.
4	Creating headlines from a text	Yes	It summarizes the information in the text, but the headlines are too long (unless specified) and have a telegraphic structure. It also invents information it does not have.
5	Elaborate text from a headline	Yes	It invents information that it does not have.
6	Writing news, reports and other informative articles	Yes	It is necessary to enter data in text form. If no information is provided, it can make it up without warning.
7	Writing news, reports and other informative pieces (without entering data, only indicating the subject matter)	Yes	When it does not know a piece of information, it invents it. The only limit it recognizes is the time limit (its stored information goes up to 2021; when it is explicitly asked for information about a later date it indicates that it does not have that information).
8	Writing scripts for radio, television and social networks	Yes	As with other formats, if it is given enough information, it has no problem writing it, but if it has no information it tends to make it up. In some cases, it is difficult to know what is true and what is not.
9	Writing copy for social media	Yes	It differentiates the styles of the different social media platforms and adapts to them depending on what the user indicates.
10	Create a strategy for dissemination of content on social networks	Yes	It is capable of creating a strategy in that it can indicate the platforms to be used, the content, frequency of publication, etc., but the strategies it proposes are not very effective.
11	Hierarchize informative topics	Yes	Is capable of ordering different topics, if the criteria to be taken into account are specified.
12	Analyse data and establish patterns	Yes	It can analyse data, but it makes simple mistakes.
13	Moderate comments	No	Specifies that it cannot do it.
14	Suggest content for social networks	Yes	Includes emojis, if suggested. It does not take into account character limits of each social network, if not specified.
15	Maintain conversation with users	Yes	Its user interaction interface is conversational. Through the chat, the user can provide various prompts to the AI, but can also simply have a conversation with the tool, as if it were an ordinary chatbot.
16	Recognize people or locations	No	It is not capable of 'watching' videos or images and it warns the user about it.
17	Detect viral topics	No	No, because it is not connected to the web.
18	Create graphics	No	It only communicates by text.
19	Content curation/filtering of news according to user's tastes	Yes	It is necessary to specify the data through text previously.
20	Analyse data to rank news, decide which is translated into which language, suggest combinations of images and headlines, etc.	Yes	It is necessary to introduce criteria for it to do so in a justified and orderly manner.
21	Publish on social media	No	It is not connected to the Internet. It can create copies and ideas, but it cannot publish anything on the network directly (except through APIs).
22	Detect fake information	Partially	It can detect false information in text, if requested, but not in other formats because it cannot access the content. Again, the problem comes from the data it has received and because it does not have access to real-time information, which makes it wrong in some cases.
23	Subtitle videos and/or audios in text	No	The test was done by trying to provide it with a video via link. It does not have access to the video (because it is not connected to the network) and yet it pretends that it does, making up the subtitles.
24	Summarize videos and/or audio in text	No	It does not have access to the video.
25	Conduct interviews (directly)	Yes	Interacts with the user as if it were an interviewer, considering their answers to elaborate and/or link to the next question. Introduces itself and says goodbye.

In phase A of the experiment (testing the AI tool by following instructions and choosing from a series of pre-established prompts, as shown in Annex II), the journalists indicated that the prompts that best suited their routines were focused on summarizing information and hierarchizing it within a news piece. The prompts most used during the experience were specifically those that allowed writing an interview, writing informative pieces based on data, and hierarchizing information within a news item.

In phase B of the experiment, participants were free to ask the tool to do whatever they wanted. Journalists asked *ChatGPT* to compose emails, verify potentially fake news, summarize press releases, or rank news stories based on the expected number of hits, among other actions. In addition, they tried to adapt its writing style to the one they normally used in their media and tried to investigate to what extent the AI ‘understands’ the information it offers or receives.

“In most cases, AI complements the work of journalists, it does not completely replace it”

In the final questionnaire of the experiment, journalists positively highlighted the abilities of AI to streamline and automate part of their work routines without replacing the journalists’ work. Summarizing information, generating code, contributing ideas when looking for content ideas, brainstorming and correctly writing information are some of the actions that stand out. Some of the participants also stated that they were surprised by the AI’s accuracy and speed in writing, as well as by its ability to synthesize and explain complex information.

Although there were many positive aspects, the professionals detected drawbacks and risks. On a technical level, the impossibility of introducing external content other than text, operating errors (such as difficulties in logging in, creating an account, or the tool being blocked) and the limitation of access to data prior to 2021 significantly reduces the possibilities of *ChatGPT* at this time and make it difficult to work with.

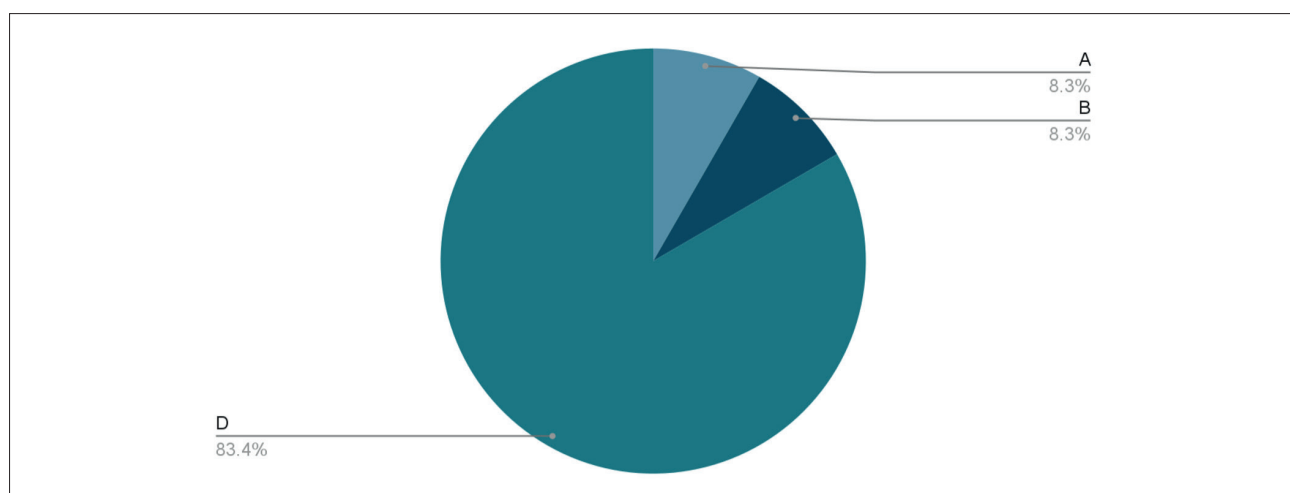
Regarding the analysis of the content provided by *ChatGPT*, participants were concerned about the ease with which it includes false information and, especially, because it does so in a format and appearance that favours it being considered relevant, credible and reliable information. Some participants also noted that the AI’s headline writing is poor and that its writing of interview questions could be much improved. One participant also mentioned as a shortcoming that the tool “lacks empathy”. This issue may be particularly relevant in situations in which the information may offend public sensitivities, something that the AI would not be able to detect and, therefore, would not take “special care” of when conveying the news.

However, after the experience, 50% of participants considered including *ChatGPT* in their work routines. Of the remaining 50%, half said they were not sure, and the other half said they would not include it. The reasons given by these participants were the possible problems of privacy and data protection, the fact that it offers incorrect information, and even the belief that the use of these tools can cancel out human capacity for analysis and reflection. It is important to note that the AI tool itself offered this same response during the interview (Annex I).

Finally, the journalists provided their perspective on the use of artificial intelligence in newsrooms, selecting the statement that best suited their opinion among those shown in Figure 1. As can be seen, the majority option, chosen by 83.3% of participants, was

“AI will not replace journalists, but it will change their professional routines”.

Only one participant agreed with the statement “AI may eventually replace journalists, eliminating jobs”. The same was true for the statement “AI will never replace journalists nor will it ever come to have much influence in newsrooms”.



A	AI may eventually replace journalists, eliminating jobs.
B	AI will never replace journalists, nor will it ever have much ever influence in newsrooms.
C	AI should be banned from newsrooms.
D	AI will not replace journalists, but it will change their professional routines.

Figure 1. Participants’ responses to the future of AI in newsrooms.

However, the statement “AI should be banned from newsrooms” was not chosen by any of the participants. Moreover, as Chat itself and one of the participants in the experiment pointed out, “the excessive use of *ChatGPT* in routine journalism may result in a loss of human

skills and knowledge, such as the ability to investigate and critically analyse”. Regarding the debate on the need to regulate AI in relation to aspects such as privacy or copyright, the participants again demonstrated unanimity, with 91.7% stating that legislation is necessary, compared to 8.3% who do not see the need for it. The only participant who pointed out this option did so because “he understands that these rights would fall on the person using the AI”.

“The fact that *ChatGPT* is powered by big data and learns from its interactions with users puts users privacy at risk”

4. Discussion and conclusions

The results obtained in this research corroborate the hypotheses put forward and shed light on the panorama of AI in newsrooms, as well as on the possible future uses of these tools in the media.

The first hypothesis raised refers to the transformation caused by the integration of AI into newsrooms. It was established as a previous point that these tools facilitate, streamline and automate journalistic processes and routines. This is proven by the results of the benchmarking carried out, as well as through the walkthrough and the experiment with journalists. In all these methodological processes, AI demonstrates its potential to reduce time spent on production, writing, management and dissemination of journalistic content, which could result in a reduced workload for journalists. Related to this, it is also observed that a change in professional routines may lead to a change in the required professional profiles or specialization expected of a journalist, as already observed by **Salazar-García** (2018).

However, the results do not seem to indicate the disappearance of jobs due to the incorporation of AI in newsrooms –at least not imminently. The benchmarking results show that in most cases these tools complement the work of journalists instead of replacing them. Journalists are still essential for providing more context, to review possible errors and biases or the adequacy and quality of the texts produced, among other issues. This is also observed in the walkthrough with *ChatGPT*: in some cases, it incorporates biased or false information, and it does not always correctly interpret the instructions provided. In the experiment with journalists, participants agreed on the usefulness of certain types of AI tools and 83.3% agreed that “AI will not replace journalists, but it will modify their professional routines”. The possibility of artificial intelligence eliminating jobs has already been addressed by academics (**Manfredi-Sánchez; Ufarte-Ruiz**, 2020) and, although this possibility is real, other research also shows that experts and professionals do not see the disappearance of journalists or their role in newsrooms as likely. Their routines will be modified, but they will not be wiped off the map. It is important to note, however, that although on the one hand this more optimistic perspective exists, there are already media outlets that do not have a single journalist on their staff, the so-called “synthetic media”, and that rely exclusively on AI to perform their role as reporting agents (**Ufarte-Ruiz; Murcia-Verdú; Túñez-López**, 2023).

The second hypothesis is related to the possible ethical, deontological and legal risks that may be associated with the use of artificial intelligence in newsrooms, as well as the need for its regulation. This aspect has already been investigated by academics as it involves complex dilemmas and challenges related to user privacy, the preservation of journalistic ethics and deontology, transparency about the use of AI and accountability mechanisms, algorithmic biases, the veracity of information and the safeguarding of copyright or intellectual property, among other problems (**Ufarte-Ruiz; Calvo-Rubio; Murcia-Verdú**, 2021; **Sanahuja-Sanahuja; López-Rabadán**, 2022). The results obtained through the walkthrough with *ChatGPT* and the experiment with journalists corroborate this second hypothesis and are consistent with the results of previous research. The inclusion of biased or false information, the lack of sources and a particularly careful presentation, which contributes to creating an image of reliability and honesty, make it especially necessary to regulate these tools to avoid further misinformation. That such AI feeds on big data and learns from its interactions with users puts users’ privacy at risk. Moreover, by omitting sources in the responses, attribution and copyright also come into dispute. 91.7% of the experiment participants felt that the use of AI should be regulated in terms of these issues.

In short, this AI tool offers many solutions and opportunities that are applicable to the journalistic world, reducing the time needed to perform routine tasks. Among the advantages is the ability to write text in different formats, help select topics to be covered in the media or adapt texts to social networks. These are simple tasks, more related to communication than to journalistic production itself, but they can free up journalists and allow them to devote more time to work that requires more research and depth. However, this tool also has several limitations. Perhaps the most serious is that it does not always provide truthful information. Sometimes *ChatGPT* points out that it is not connected to the Internet and that its data is limited, so its information may not be correct or accurate. But at other times it simply invents information that it does not have, without warning the user. This means that the journalist may publish inaccurate or even false information, if they are not an expert on the subject. That the information to which it has access is limited to data up to 2021 also means that, even if the content it publishes was correct at some point, it may already be outdated when the query is made. It is important to note that at the time of writing (May 2023) it is already possible to connect the tool to the Internet through other software

“91.7% of the experiment participants considered that the use of AI should be regulated”

and that *OpenAI* has announced a paid version with real-time connection to the network, so this limitation will be overcome soon. Another limitation that is likely to remain is the preponderance of the most repeated contents and ideas over less common ones, regardless of their veracity or importance, leading to a situation similar to the so-called “tyranny of the majority” in politics (**Harper**, 2017). According to the results obtained during the walkthrough exercise, the information provided by *ChatGPT* is often based on the number of times it has found such data and not so much on the quality of the data. In other words, it puts the quantitative before the qualitative. This can perpetuate certain types of biased and more traditional views in some areas of knowledge.

Another limitation of AI is the preponderance of the most repeated contents and ideas over the less common ones, regardless of their veracity or importance, which can lead to a situation of “tyranny of the majority”

This fits with another of the ethical dilemmas of the use of AI: how the substitution of humans by machines in tasks inherent to their own being affects wider society and at the same time makes it necessary to rethink the curricula in journalism degrees, not only to adjust the professional profiles, as mentioned above, but also to provide new journalists with an ethical-deontological knowledge base that allows them to deal with AI in their jobs without letting it prevent them having critical and independent thinking, which is essential to safeguard the values inherent to journalism (**Peña-Fernández; Meso-Ayerdi; Larrondo-Ureta**, 2023).

5. Limitations of the study and future lines of research

The continuous advances in artificial intelligence, and specifically in *ChatGPT*, mean that some of the noted limitations of the tool have already been overcome in its paid version, as mentioned in the body of the article. On the other hand, future lines of research remain open, such as the exploration of possible codes and manuals for a safe use of AI in newsrooms or possible adaptations of *ChatGPT* to make it more suitable for journalistic use.

Since the sample of participants was small, the results should be considered with caution. Based on this first study, it would be interesting to carry out another experiment with a larger number of journalists to obtain comparable statistics on the use, risks and advantages of *ChatGPT*.

6. References

- Agarwal, Udit** (2018). “How Reuters is transforming journalism with artificial intelligence”. *Analytics insight*, 11 March. <https://www.analyticsinsight.net/how-reuters-is-transforming-journalism-with-artificial-intelligence>
- Aguilar, Ricardo** (2023). “‘Pausen inmediatamente el entrenamiento’: Musk y más personalidades, preocupados por GPT-4”. *Xataka*, 29 marzo. <https://www.xataka.com/aplicaciones/pausen-inmediatamente-entrenamiento-musk-personalidades-preocupados-gpt-4>
- Ali, Waleed; Hassoun, Mohamed** (2019). “Artificial intelligence and automated journalism: Contemporary challenges and new opportunities”. *International journal of media, journalism and mass communications*, v. 5, n. 1, pp. 40-49. <https://doi.org/10.20431/2454-9479.0501004>
- Berkland, Mara K.** (2017). “Experience Sampling Method”. In: Allen, Mike. *The SAGE encyclopedia of communication research methods*. Thousand Oaks: Sage. ISBN: 978 1 483381435 <https://doi.org/10.4135/9781483381411>
- Buj, Anna** (2023). “Italia bloquea el uso de ChatGpt por no respetar la ley de protección de datos”. *La vanguardia*, 31 marzo. <https://www.lavanguardia.com/tecnologia/20230331/8868073/italia-prohibe-chatgpt-respetar-legislacion-datos.html>
- Calvo-Rubio, Luis-Mauricio; Ufarte-Ruiz, María-José** (2020). “Percepción de docentes universitarios, estudiantes, responsables de innovación y periodistas sobre el uso de inteligencia artificial en periodismo”. *Profesional de la información*, v. 29, n. 1, e290109. <https://doi.org/10.3145/epi.2020.ene.09>
- Canavilhas, João** (2022). “Artificial intelligence and journalism: Current situation and expectations in the Portuguese sports media”. *Journalism and media*, v. 3, n. 3, pp. 510-520. <https://doi.org/10.3390/journalmedia3030035>
- Castells, Manuel** (2001). *La galaxia internet*. Barcelona: Plaza & Janés. ISBN: 978 84 01386060
- Chalmers, Dominic; MacKenzie, Niall G.; Carter, Sara** (2021). “Artificial intelligence and entrepreneurship: Implications for venture creation in the fourth industrial revolution”. *Entrepreneurship theory and practice*, v. 45, n. 5, pp. 1028-1053. <https://doi.org/10.1177/1042258720934581>
- Cohen, Sarah; Hamilton, James T.; Turner, Fred** (2011). “Computational journalism”. *Communications of the ACM*, v. 54, n. 10, pp. 66-71. <https://doi.org/10.1145/2001269.2001288>

- Danzon-Chambaud, Samuel** (2021). "A systematic review of automated journalism scholarship: Guidelines and suggestions for future research". *Open research Europe*, v. 1, n. 4.
<https://doi.org/10.12688/openreseurope.13096.1>
- De-Lara, Alicia; García-Avilés, José-Alberto; Arias-Robles, Félix** (2022). "Implantación de la inteligencia artificial en los medios españoles: análisis de las percepciones de los profesionales". *Textual & visual media*, v. 1, n. 15, pp. 1-17.
<https://doi.org/10.56418/txt.15.2022.001>
- De-Lima-Santos, Mathias-Felipe; Ceron, Wilson** (2022). "Artificial intelligence in news media: Current perceptions and future outlook". *Journalism and media*, v. 3, n. 1, pp. 13-26.
<https://doi.org/10.3390/journalmedia3010002>
- Evens, Tom; Raats, Tim; Von-Rimscha, M. Bjørn** (2017). "Business model innovation in news media organisations - 2018 special issue of the European Media Management Association (EMMA)". *Journal of media business studies*, v. 14, n. 3, pp. 167-172.
<https://doi.org/10.1080/16522354.2018.1445164>
- Fernández, Yúbal** (2023a). "Bing Chat con ChatGPT: qué es, cómo funciona y qué puedes hacer con el chat con inteligencia artificial del buscador". *Xataka*, 24 abril.
<https://www.xataka.com/basics/bing-chatgpt-que-como-funciona-que-puedes-hacer-chat-inteligencia-artificial-buscador>
- Fernández, Yúbal** (2023b). "ChatGPT Plus: qué es, diferencias con ChatGPT normal y cuánto cuesta esta inteligencia artificial". *Xataka*, 1 junio.
<https://www.xataka.com/basics/chatgpt-plus-que-que-caracteristicas-tiene-cuanto-cuesta-version-pago-esta-inteligencia-artificial>
- García-Avilés, José A.** (2019). "El impacto de la inteligencia artificial en el periodismo". *Máster en innovación y periodismo UMH blog*, 2 noviembre.
<http://mip.umh.es/blog/2019/11/02/el-impacto-de-la-inteligencia-artificial-en-el-periodismo>
- García-Orosa, Berta; Canavilhas, João; Vázquez-Herrero, Jorge** (2023). "Algorithms and communication: A systematized literature review". *Comunicar*, v. 31, n. 74, pp. 9-21.
<https://doi.org/10.3916/C74-2023-01>
- García-Orosa, Berta; Pérez-Seijo, Sara; Vizoso, Ángel** (2022). *Emerging practices in the age of automated digital journalism*. New York: Routledge. ISBN: 978 1 032128511
- Gómez-Diago, Gloria** (2022). "Perspectivas para abordar la inteligencia artificial en la enseñanza de periodismo". *Revista latina de comunicación social*, v. 80, pp. 29-46.
<https://doi.org/10.4185/RLCS-2022-1542>
- Gonçalves, Adriana; Melo, Paulo-Victor** (2022). "Artificial intelligence and journalism: An approach to the Portuguese context". *Fonseca, journal of communication*, n. 25, pp. 23-24.
<https://doi.org/10.14201/fjc.29682>
- Google** (2023). *Bard*.
<https://bard.google.com>
- Harper, Tael** (2017). "The big data public and its problems: Big data and the structural transformation of the public sphere". *New media & society*, v. 19, n. 9, pp. 1424-1439.
<https://doi.org/10.1177/1461444816642167>
- Hazard-Owen, Laura** (2023). "Google is changing up search. What does that mean for news publishers?". *NiemanLab*, 11 May.
<https://www.niemanlab.org/2023/05/google-is-changing-up-search-what-does-that-mean-for-news-publishers>
- Hermann, Erik** (2022). "Artificial intelligence and mass personalization of communication content: An ethical and literacy perspective". *New media & society*, v. 24, n. 5, pp. 1258-1277.
<https://doi.org/10.1177/14614448211022702>
- Israel, Maria-Joseph; Amer, Ahmed** (2022). "Rethinking data infrastructure and its ethical implications in the face of automated digital content generation". *AI and ethics*, v. 3, pp. 427-439.
<https://doi.org/10.1007/s43681-022-00169-1>
- Karlsen, Joakim; Stavelin, Eirik** (2013). "Computational journalism in Norwegian newsrooms". *Journalism practice*, v. 8, n. 1, pp. 34-48.
<https://doi.org/10.1080/17512786.2013.813190>
- Krausová, Alzbeta; Moravec, Václav** (2022). "Disappearing authorship: Ethical protection of ai-generated news from the perspective of copyright and other laws". *Journal of intellectual property, information technology and e-commerce law*, v. 13, n. 2.
<https://www.jipitec.eu/issues/jipitec-13-2-2022/5540>

- Laboratorio de Periodismo** (2018). *Inteligencia artificial: para qué puede usarse en periodismo y qué están haciendo los medios*. Laboratorio de periodismo Fundación Luca de Tena, 6 noviembre.
<https://laboratoriodeperiodismo.org/inteligencia-artificial-para-que-puede-usarse-en-periodismo-y-que-estan-haciendo-los-medios>
- Larsson, Stefan; Heintz, Fredrik** (2020). "Transparency in artificial intelligence". *Internet policy review*, v. 9, n. 2.
<https://doi.org/10.14763/2020.2.1469>
- Lin, Bibo; Lewis, Seth C.** (2022). "The one thing journalistic AI just might do for democracy". *Digital Journalism*, v. 10, n. 10, pp. 1627-1649.
<https://doi.org/10.1080/21670811.2022.2084131>
- López-García, Xosé; Vizoso, Ángel** (2021). "Periodismo de alta tecnología: signo de los tiempos digitales del tercer milenio". *Profesional de la información*, v. 30, n. 3, e300301.
<https://doi.org/10.3145/epi.2021.may.01>
- Mahatody, Thomas; Sagar, Mouldi; Kolski, Christophe** (2010). "State of the art on the cognitive walkthrough method, its variants and evolutions". *International journal of human-computer interaction*, v. 26, n. 8, pp. 741-785.
<https://doi.org/10.1080/10447311003781409>
- Manfredi-Sánchez, Juan-Luis; Ufarte-Ruiz, María-José** (2020). "Inteligencia artificial y periodismo: una herramienta contra la desinformación". *Revista Cidob d'afers internacionals*, n. 124, pp. 49-72.
<https://doi.org/10.24241/rcai.2020.124.1.49>
- Marconi, Francesco** (2020). *New York: Newsmakers. Artificial intelligence and the future of journalism*. New York: Columbia University Press. ISBN: 978 0 231191371
- Márquez, Javier** (2023). "GPT-4: así es la IA más avanzada de OpenAI, cómo funciona y todas las novedades". *Xataka*, 17 mayo.
<https://www.xataka.com/nuevo/gpt-4-que-cuando-sale-como-funciona-toda-informacion>
- Martínez-Martínez, Inmaculada-José; Aguado-Terrón, Juan-Miguel; Sánchez-Cobarro, Paloma-del-Henar** (2022). "Smart advertising: innovación y disrupción tecnológica asociadas a la IA en el ecosistema publicitario". *Revista latina de comunicación social*, n. 80, pp. 69-90.
<https://doi.org/10.4185/10.4185/RLCS-2022-1693>
- Micó, Josep-Lluís; Casero-Ripollés, Andreu; García-Orosa, Berta** (2022). "Platforms in journalism 4.0: The impact of the fourth industrial revolution on the news industry". En: Vázquez-Herrero, Jorge; Silva-Rodríguez, Alba; Negreira-Rey, María-Cruz; Toural-Bran, Carlos; López-García, Xosé. *Total Journalism. Studies in Big Data*, v. 97. Cham: Springer Nature, pp. 241-253.
https://doi.org/10.1007/978-3-030-88028-6_18
- Mishra, Shrutika; Tripathi, Asha R.** (2021). "AI business model: An integrative business approach". *Journal of innovation and entrepreneurship*, v. 10, n. 18.
<https://doi.org/10.1186/s13731-021-00157-5>
- Mooshammer, Sandra** (2022). "There are (almost) no robots in journalism. An attempt at a differentiated classification and terminology of automation in journalism on the base of the concept of distributed and gradualised action". *Publizistik*, v. 67, pp. 487-515.
<https://doi.org/10.1007/s11616-022-00757-5>
- Moran, Rachel E.; Shaikh, Sonia-Jawaid** (2022). "Robots in the news and newsrooms: Unpacking meta-journalistic discourse on the use of artificial intelligence in journalism". *Digital journalism*, v. 10, n. 10, pp. 1756-1774.
<https://doi.org/10.1080/21670811.2022.2085129>
- Noain-Sánchez, Amaya** (2022). "Addressing the impact of artificial intelligence on journalism: the perception of experts, journalists and academics". *Communication & society*, v. 35, n. 3, pp. 105-21.
<https://doi.org/10.15581/003.35.3.105-121>
- OpenAI** (2023a). *DALL-E 2*.
<https://openai.com/product/dall-e-2>
- OpenAI** (2023b) "Introducing ChatGPT". OpenAI Blog.
<https://openai.com/blog/chatgpt>
- Otero-González, Iván** (2022). "Los cimientos de la inteligencia artificial en el sistema productivo de contenidos periodísticos automatizados". *Redmarka. Revista de marketing aplicado*, v. 26, n. 1, pp. 15-35.
<https://doi.org/10.17979/redma.2022.26.1.9056>

- Page, Matthew J.; Moher, David; Bossuyt, Patrick M.; Boutron, Isabelle; Hoffmann, Tammy C.; Mulrow, Cynthia D.; Shamseer, Larissa; Tetzlaff, Jennifer M.; Akl, Eli A.; Brennan, Sue E.; Chou, Roger; Glanville, Julie et al.** (2021). "PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews". *BMJ*, v. 372, n. 160.
<https://doi.org/10.1136/bmj.n160>
- Parratt-Fernández, Sonia; Mayoral-Sánchez, Javier; Mera-Fernández, Montse** (2021). "The application of artificial intelligence to journalism: An analysis of academic production". *Profesional de la información*, v. 30, n. 3, e300317.
<https://doi.org/10.3145/epi.2021.may.17>
- Pavlik, John V.** (2023). "Collaborating with ChatGPT: Considering the implications of generative artificial intelligence for journalism and media education". *Journalism and mass communication educator*, v. 78, n. 1, pp. 84-93.
<https://doi.org/10.1177/10776958221149577>
- Peña-Fernández, Simón; Meso-Ayerdi, Koldobika; Larrondo-Ureta, Ainara** (2023). "Without journalists, there is no journalism: the social dimension of generative artificial intelligence in the media". *Profesional de la información*, v. 32, n. 2, e320227.
<https://doi.org/10.3145/epi.2023.mar.27>
- Porlezza, Colin; Ferri, Giulia** (2022). "The missing piece: Ethics and the ontological boundaries of automated journalism". *#ISOJ*, v. 12, n. 1, pp. 71-98.
<https://isoj.org/wp-content/uploads/2022/03/ISOJ-2022.pdf>
- Prasad, Ramjee; Choudhary, Purva** (2021). "State-of-art of artificial intelligence". *Journal of mobile multimedia*, v. 17, n. 1-3, pp. 427-454.
<https://doi.org/10.13052/jmm1550-4646.171322>
- Prego, Carlos** (2023). "Elon Musk lanza X.AI, su propia compañía de IA con un objetivo claro: competir con OpenAI". *Xataka*, 15 abril.
<https://www.xataka.com/empresas-y-economia/elon-musk-lanza-su-propia-compania-inteligencia-artificial-x-ai-que-sabemos-su-plan-para-competir-openai>
- Ross, Philip; Maynard, Kasia** (2021). "Towards a 4th industrial revolution". *Intelligent buildings international*, v. 13, n. 3, pp. 159-161.
<https://doi.org/10.1080/17508975.2021.1873625>
- Salazar-García, Idoia-Ana** (2018). "Los robots y la inteligencia artificial. Nuevos retos del periodismo". *Doxa comunicación*, n. 27, pp. 295-315.
<https://doi.org/10.31921/doxacom.n27a15>
- Samuel, Jim; Kasyap, Rajiv; Yana, Samuel; Pelaez, Alexander** (2022). "Adaptative cognitive fit: Artificial intelligence augmented management of information facets and representations". *International journal of information management*, v. 65, 102505.
<https://doi.org/10.1016/j.ijinfomgt.2022.102505>
- Sanahuja-Sanahuja, Rosana; López-Rabadán, Pablo** (2022). "Ética y uso periodístico de la inteligencia artificial. Los medios públicos y las plataformas de verificación como precursores de la rendición de cuentas en España". *Estudios sobre el mensaje periodístico*, v. 28, n. 4, pp. 959-970.
<https://doi.org/10.5209/esmp.82385>
- Sánchez-García, Pilar; Merayo-Álvarez, Noemí; Calvo-Barbero, Carla; Diez-Gracia, Alba** (2023). "Spanish technological development of artificial intelligence applied to journalism: companies and tools for documentation, production and distribution of information". *Profesional de la información*, v. 32, n. 2, e320208.
<https://doi.org/10.3145/epi.2023.mar.08>
- Tejedor, Santiago; Vila, Pere** (2021). "Exo journalism: A Conceptual approach to a hybrid formula between journalism and artificial intelligence". *Journalism and media*, v. 2, n. 4, pp. 830-840.
<https://doi.org/10.3390/journalmedia2040048>
- Túñez-López, Miguel; Fieiras-Ceide, César; Vaz-Álvarez, Martín** (2021). "Impacto de la inteligencia artificial en el periodismo: transformaciones en la empresa, los productos, los contenidos y el perfil profesional". *Communication & society*, v. 34, n. 1, pp. 177-193.
<https://doi.org/10.15581/003.34.1.177-193>
- Túñez-López, Miguel; Toural-Bran, Carlos; Valdiviezo-Abad, Cesibel** (2019). "Automation, bots and algorithms in news-making. Impact and quality of artificial journalism". *Revista latina de comunicación social*, n. 74, pp. 1411-1433.
<https://doi.org/10.4185/RLCS-2019-1391>

- Ufarte-Ruiz, María-José; Calvo-Rubio, Luis-Mauricio; Murcia-Verdú, Francisco-José** (2021). “Los desafíos éticos del periodismo en la era de la inteligencia artificial”. *Estudios sobre el mensaje periodístico*, v. 27, n. 2, pp. 673-684.
<https://doi.org/10.5209/esmp.69708>
- Ufarte-Ruiz, María-José; Feiras-Ceide, César; Túnñez-López, Miguel** (2020). “La enseñanza-aprendizaje del periodismo automatizado en instituciones públicas: estudios, propuestas de viabilidad y perspectivas de impacto de la IA”. *Anàlisi*, n. 62, pp. 131-146.
<https://doi.org/10.5565/rev/analisi.3289>
- Ufarte-Ruiz, María-José; Murcia-Verdú, Francisco-José; Túnñez-López, José-Miguel** (2023). “Use of artificial intelligence in synthetic media: first newsrooms without journalists”. *Profesional de la información*, v. 32, n. 2.
<https://doi.org/10.3145/epi.2023.mar.03>
- Vállez, Mari; Codina, Lluís** (2018). “Periodismo computacional: evolución, casos y herramientas”. *Profesional de la información*, v. 27, n. 4, pp. 759-768.
<https://doi.org/10.3145/epi.2018.jul.05>
- Van-Dijck, José; Poell, Thomas; De-Waal, Martjin** (2018). *The platform society*. New York: Oxford University Press. ISBN: 978 0 190889777
- Ventura-Pocino, Patricia** (2021). “Algorithms in the newsrooms: Challenges and recommendations for artificial intelligence with the ethical values of journalism”. *Catalan Press Council*.
https://fcic.periodistes.cat/wp-content/uploads/2022/03/venglishDIGITAL_ALGORITMES-A-LES-REDACCIONS_ENG-1.pdf
- Yu, Yang; Huang, Kuo; Jones, Bob** (2020). “Artificial intelligence in media: Journalists’ perceptions and organizational talk”. *Communication, technologies et développement*, v. 8.
<https://doi.org/10.4000/ctd.3262>
- Zelizer, Barbie** (2019). “Why journalism is about more than digital technology”. *Digital journalism*, v. 7, n. 3, pp. 343-350.
<https://doi.org/10.1080/21670811.2019.1571932>

7. Annexes

Annex I. *ChatGPT* interview

This interview was conducted in Spanish.

The AI tool answer the question “How can *ChatGPT* help in journalistic routines?” as it follows:

ChatGPT can help in different tasks of the journalistic routine, such as:

1. Generation of summaries and synopses of news or events.
2. Creation of articles and content from given information.
3. Data analysis and generation of graphics and visualizations.
4. Translation and adaptation of content to different languages.
5. Generation of interview and survey questions.

Overall, *ChatGPT* is a powerful tool that can automate and streamline many of the repetitive tasks in the journalistic process, freeing up time for journalists to focus on research and the production of high quality content.

Related to possible risks, *ChatGPT* answers the question “What are the problems and risks of using *ChatGPT* in newsrooms?” as it follows:

Although *ChatGPT* can be a useful tool in routine journalism, there are also some problems and risks that must be considered:

6. Accuracy: *ChatGPT* is trained with a large amount of data, but it can still make mistakes or produce inaccurate or misleading information. It is important to verify the information generated by *ChatGPT* before publication.
7. Neutrality: *ChatGPT* is trained with information from the Internet and may reflect existing biases or prejudices in society. It is important to be careful not to perpetuate these biases through the use of *ChatGPT* in the production of journalistic content.
8. Loss of human skills: Excessive use of *ChatGPT* in routine journalism can result in a loss of human skills and knowledge, such as the ability to investigate and critically analyze.
9. Depersonalization: The use of *ChatGPT* to generate content can result in a loss of humanity and personality in journalistic content.

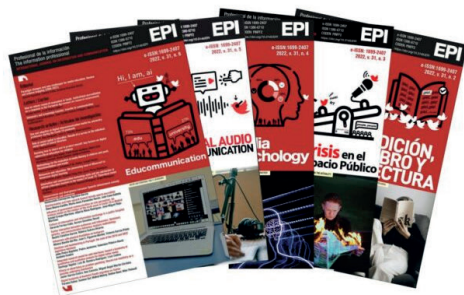
It is important to use *ChatGPT* as a complementary tool, and not as a substitute, for human skills and knowledge in routine journalism. It is also crucial to be aware of these issues and risks and take steps to mitigate them.

Annex II. Prompts of phase A of the experiment with journalists

Because the participants work in Spanish, the prompts were built and used in this language.

	Required action	Prompt
1A	Writing an interview	Write 20 questions focused on his professional career to interview Pedro Sánchez on TV. Also write an introduction and a closing for the piece.
1B	Writing informative pieces based on data	Write a 1-minute radio news about a car accident que that took place tonight on SC-20 towards Santiago Sur. There have been no injuries. Emergency services have been deployed to the scene.
1C	Script writing (radio, television, social media)	Write a script for two people for an educational video of less than one minute that will be broadcast on social media. The topic is the importance of recycling. Offer several tips that anyone can follow to recycle more and take care of the environment.
2A	Adapting news to social media	Make copies for LinkedIn from this news item (text of the news item). Use emojis and write in a casual and friendly tone.
2B	Summarizing information	Summarize this information: (paste here the original text). Use 500 words maximum.
2C	Adapting a text (to another format or style)	Rewrite this text so a five-year-old kid can understand it: (paste here the original text). Use at least 100 words.
2D	Translating text	Translate this text into Galician: (paste here the original text).
3A	Hierarchizing several pieces of information	With what news would you open the news / place on the front page etc. in a media of these characteristics? (Describe the characteristics)
3B	Hierarchizing information within an informative piece (you can ask it to write a headline, a subtitle, a printer, etc.)	Write a headline for this news item: (paste here the news item text). It will be published in a digital newspaper. Do not use more than 15 words.
3C	Choosing which news to publish	Which of the following news items will you publish: (paste here the headlines). Choose only five. They will be published in the local generalist newspaper.

REVISTA EPI



<https://www.profesionaldelainformacion.com>

Profesional de la información es una revista científica sobre Comunicación, Documentación, Biblioteconomía, Informetría y Ciencias de la información. Se publica cada 2 meses (6 números al año).

e-ISSN: 1699-2407

<https://doi.org/10.3145/EPI>

SJR 2022 = 0,872 (Q1); JIF 2022 = 4,2 (Q1)

ChatGPT could be the reviewer of your next scientific paper. Evidence on the limits of AI-assisted academic reviews

David Carabantes; José L. González-Geraldo; Gonzalo Jover

Nota: Este artículo se puede leer en español en:
<https://revista.profesionaldeinformacion.com/index.php/EPI/article/view/87376>

Recommended citation:

Carabantes, David; González-Geraldo, José L.; Jover, Gonzalo (2023). "ChatGPT could be the reviewer of your next scientific paper. Evidence on the limits of AI-assisted academic reviews". *Profesional de la información*, v. 32, n. 5, e320516.

<https://doi.org/10.3145/epi.2023.sep.16>

Article received on May 22nd 2023
Approved on August 29th 2023



David Carabantes ✉
<https://orcid.org/0000-0001-9897-4847>

Universidad Complutense de Madrid
Facultad de Educación
Rector Royo Villanova, 1
28040 Madrid, Spain
dcaraban@ucm.es



José L. González-Geraldo
<https://orcid.org/0000-0003-1698-0122>

Universidad de Castilla-La Mancha
Facultad de Ciencias de la Educación y
Humanidades
Avda. Los Alfares, 44
16002 Cuenca, Spain
joseluis.ggeraldo@uclm.es



Gonzalo Jover
<https://orcid.org/0000-0002-6373-4111>

Universidad Complutense de Madrid
Facultad de Educación
Rector Royo Villanova, 1
28040 Madrid, Spain
gjover@ucm.es

Abstract

The irruption of artificial intelligence (AI) in all areas of our lives is a reality to which the university, as an institution of higher education, must respond prudently, but also with no hesitation. This paper discusses the potential that resources based on AI presents as potential reviewers of scientific articles in a hypothetical peer review of already published articles. Using different models (*GPT-3.5* and *GPT-4*) and platforms (*ChatPDF* and *Bing*), we obtained three full reviews, both qualitative and quantitative, for each of the five articles examined, thus being able to delineate and contrast the results of all of them in terms of the human reviews that these same articles received at the time. The evidence found highlights the extent to which we can and should rely on generative language models to support our decisions as qualified experts in our field. Furthermore, the results also corroborate the hallucinations inherent in these models while pointing out one of their current major shortcomings: the context window limit. On the other hand, the study also points out the inherent benefits of a model that is in a clear expansion phase, providing a detailed view of the potential and limitations that these models offer as possible assistants to the review of scientific articles, a key process in the communication and dissemination of academic research.

Keywords

Artificial intelligence; AI; Generative artificial intelligence; Contextual window; *ChatGPT*; *ChatPDF*; *Bing*; AI-assisted review; Peer review; Academic review; Academic publication; Scientific communication.

Funding

This article has been partially funded by the *Civic Culture and Educational Policies Research Group* of the *Complutense University of Madrid*, through the *UCM Research Group Funding Programme (GRFN32/23)*.



1. Introduction

As has happened with other innovations that once marked an era, it is likely that in a short time the name of *ChatGPT*, today on everyone's lips, will fade, and other brands and logos will appear that, collecting among other advances the heritage of Natural Language Processing (NLP) or the Large Language Models (LLM), embody the pre-trained, generative, and revolution-based *GPT* (Generative Pre-trained Transformer) transformation of the so-called *Transformers* (Vaswani et al., 2017) in a qualitatively more complex and reliable way (González-Geraldo; Ortega-López, 2023). Along with *ChatGPT* (*OpenAI*), today *Bing* (*Microsoft*), *Bard* (*Google*) and *Claude* (*Anthropic*) seem to be the main bets among these resources.

The educational discussion around the emergence of this type of innovation is not new, but at this moment the popularization of *ChatGPT*, as a synonym for Artificial Intelligence (AI), has generated a debate in which it seems we must choose between panic or disruption (García-Peñalvo, 2023), dangers and promises (Jalil et al., 2023), challenges and opportunities (Kasneci et al., 2023), or risks and benefits (Sok; Heng, 2023). It has come to be asked if, for example, we are facing the end of traditional evaluation in higher education (Rudolph; Tan; Tan, 2023), or if, deep down, we are just facing a demon or our guardian angel (Tlili et al., 2023). Be that as it may, the binomial between education and AI must be put at the service, as in other professions, of the social good (Peña-Fernández et al., 2023). Although universities are beginning to regulate the use of these resources, as in other issues that demand a rapid response, a sufficiently coordinated academic policy is still lacking here (Álvarez-Castillo; Fernández-Camínero, 2023).

As researchers, we believe it is appropriate to embrace the reality of AI to examine the intersection between the inevitability of its advent and the possibility of assuming its potentialities, while mitigating its limitations, particularly those concerning ethics (Crawford; Cowling; Allen, 2023) and academic integrity (Perkins, 2023; Chomsky; Roberts; Watumull, 2023). Our objective here, in this sense, is to analyze the possibilities of generative models of texts based on AI to carry out the *peer-review* of scientific articles proposed for publication.

2. Justification and state of play

The impact that *peer-review* has on the continuous improvement of the scientific publication process is evident. Peer review is usually blind, although not always exclusively, and, depending on the area of knowledge, can act in conjunction with other mechanisms, such as open review. The procedure has its detractors, although the criticisms and alternatives to it are based on premises not always shared (Campanario, 1998a; Campanario, 1998b). In addition, The use of AI has added important issues to this discussion.

Among the limitations attributed to peer review are possible personal biases of evaluators, conflicts of interest, variability of quality, and inconsistencies in reviews due to disparity in the degree of depth in the evaluation process, the difficulty of finding specialized scholars available, the response time to the request for participation, and the deadline for submission of the review. In this article we intend to determine if AI tools, similar to the popular *ChatGPT*, can be an effective solution to solve some of these problems and challenges, as already suggested in some specialized environments, such as *Scholarcy* and *Researchleap*:

<https://www.scholarcy.com/how-reviewers-can-use-ai-right-now-to-make-peer-review-easier>

<https://researchleap.com/ai-and-the-future-of-academic-publishing-how-artificial-intelligence-is-transforming-the-peer-review-process>

It is necessary to know in advance how these resources work based on the *GPT* language model architecture. Without going into details, in a way that is certainly complex and not always under control –which produces the effect known as the “black box” (Zhai, 2023)– *GPT* transforms words into *tokens*, parts of words, to later reconstruct these *tokens* into a coherent discourse, giving rise to what Marcus (2022) calls a “pastiche”. Here lies one of the wonderful potentialities it offers us: its *output* is creative, never originated before, but always based on what existed, with what it was trained, millions of data from the various fields of knowledge that allow it to produce specialized texts. In fact, *OpenAI* workers themselves were quick to warn that one of the potential misuses that these tools can lead to is that of fraudulent academic writing (Brown et al., 2020, p. 35). On the positive side, their potential means that they can also become an instrument of peer review, facilitating a greater and more efficient management of the overload suffered by specific journals, especially those that pass certain thresholds of certain indexes.

In recent times, several editorials of scientific journals have expressed their expectation regarding the use of AI and resources such as *ChatGPT* (Švab; Klemenc-Ketiš; Zupanič, 2023; Lira et al., 2023) and academic works have emerged that explore their main challenges, potentialities, and limitations in review processes, within a general trend towards the automation of these processes (Checco et al., 2021; Severin et al., 2022; Srivastava, 2023). There are those who advocate the need to establish protocols on the use of these tools in peer review without delay (García, 2023), suggesting the convenience of specifying the use of AI in journals to, for example, verify compliance with editorial policies, summarize content, or identify weaknesses and strengths of the manuscript (Hosseini; Horbach, 2023).

Our contribution tries to advance in this line, not only in the academic aspect, but also in some technical limitations that, today, condition the possibilities of these resources in the scientific review of the manuscripts. The one that interests us most is the one related to the “contextual window”. In short, this concept refers to the extension of words –remember, *tokens*– that the model can keep in mind when generating its results. This contextual window is 4K in *GPT-3.5* version

(4,096 *tokens*). In other words, the model cannot “remember” more than 3,000 words, approximately. In fact, he does not remember anything, because there is no memory involved. Every time we send a new query, the system counts the tokens of *our* prompt, adds “*n*” tokens from the previous conversation to cover the quota, and takes into consideration the set, always under the limit of 4K tokens. A limitation that, without a doubt, calls into question any revision that can be made, since rare is the specialized article that does not reach 5,000-6,000 words at least.

The progress achieved with *GPT-4*, launched on March 14, 2023, was certainly encouraging in this regard. The contextual window went from 4K to 8K, thus achieving a significant increase to 8,192 *tokens* –above 6,000 words– and even much more with a 32K version, over 32,000 tokens –about 25,000 words–, quite superior to what is necessary to review articles and even other larger productions. The scaffolding of this research, we believed, was secured after the announcement of *GPT-4*. However, as of today, May 2023, the context window remains the same used by its previous version (4K) being reserved the two upper (8K-32K) just for developers.

In addition to the context limit, *ChatGPT* offers another limit that is explained by its own structure. Being a platform that emulates the natural conversation of the human being, it is designed so that the interactions between the user and the resource are more or less short, as in a conversation. Being a *chatbot*, it is logical to understand that it also presents a limit of *input tokens* for each query we make. At the moment, and for reasons similar to those already given, both the *ChatGPT-3.5* and *ChatGPT-4* variants do not usually accept entries that exceed 2,200 words. In other words, although we wanted to introduce articles little by little and then ask you about them, the reality would be that you would only be left with partial information, mainly from the last two *inputs*.

3. Procedure

Even with the limitations that we have pointed out, we believe it is necessary and possible to deepen the possibilities that these resources can offer us as a tool for evaluating potential articles. To do this, we went to two platforms that, even sharing the limitations of a 4K contextual window, have helped us solve the second of the problems (limit per *input*), allowing us to provide them with files in PDF format.

The first platform used –which we will call reviewers 1 and 3– was *ChatPDF* (based on *GPT-3.5 turbo*), the second –reviewer 2– the *Microsoft Edge* browser itself, as it can also be used as a PDF file reader and act with the *Bing* assistant (based on *GPT-4*). In this way we obtain three independent reviews that allow us to contrast not only the models with each other (*GPT-3.5* and *GPT-4*), but also one of the models twice (*GPT-3.5*).

Thus, we entered a process of “blind” peer review. Blind not so much by the absence of ethical conflicts between reviewers and authors, but by the fact that in no case could the “reviewers” –generative language models– access the entire text at once, but the entire text, either sequentially or depending on the preferences that the model established after our consultations. In order to guarantee that the two platforms had access to the full text, we asked both to transcribe us, word for word and in the same order, one of the articles used in this study. While *Bing* fulfilled its task through consecutive answers, forgetting only the tables and the final part of the references, *ChatPDF* ignored our request and proceeded directly to make a summary of the text, which shows how through an API it is the developers who determine the way in which the information that is included in the contextual window is controlled.

We have used five real originals of articles sent for publication, over a decade (2012-2022), to the same journal: *Bordón: Revista de Pedagogía*, organ of the Spanish Society of Pedagogy, indexed in the *Journal Citation Index (JCI)*, of *Web of Science*, and the *SCImago Journal Rank (SJR)* of *Scopus*, among other bases. The five articles used were written by one or more of the authors of this work, counting in each case with the approval of the rest of the signatories. To ensure diversity, we have chosen articles that have had varying degrees of revision, from acceptance with minor changes to rejection. All of them were produced in Spanish, except for the most recent, published in English. All articles were tested in their original version, as they were initially sent to the journal for evaluation.

In both platforms (*ChatPDF* and *Bing*) a template of the journal itself, published in open and currently improved, was used as a basis on which to elaborate the review queries:

https://www.sepedagogia.es/?page_id=895

This template was the one used by the journal at the time of most of the submissions of the originals, except for the most recent. Through it, we asked to assess: 1) IMRaD format of the abstract / its extension, 2) adequacy of the title / keywords, 3) spelling and syntactic correction, 4) APA standards / coherence between citations and bibliographic references, 5) tables and figures, 6) interest of the article for the educational community, 7) generalization of the results, 8) originality of the work / contribution to educational knowledge, 9) introduction and justification of the importance of the topic, 10) theoretical foundation, 11) relevance of the sources cited according to the year of publication, 12) formulation of objectives, 13) process of collection and analysis of information, 14) description of the sampling procedure, 15) process of collection and analysis of information, 16) presentation and description of results, 17) conclusions and discussion. Except for a few points (1, 2, 4 and 8), the answers were obtained through a single *prompt*. For all the points, a quantitative assessment on a Likert scale was requested from 1 (minimum rating) to 5 (maximum rating). The beginning of the *prompts* was always the same (see Images 1 and 2), varying only the final part depending on the review point that should be contrasted.

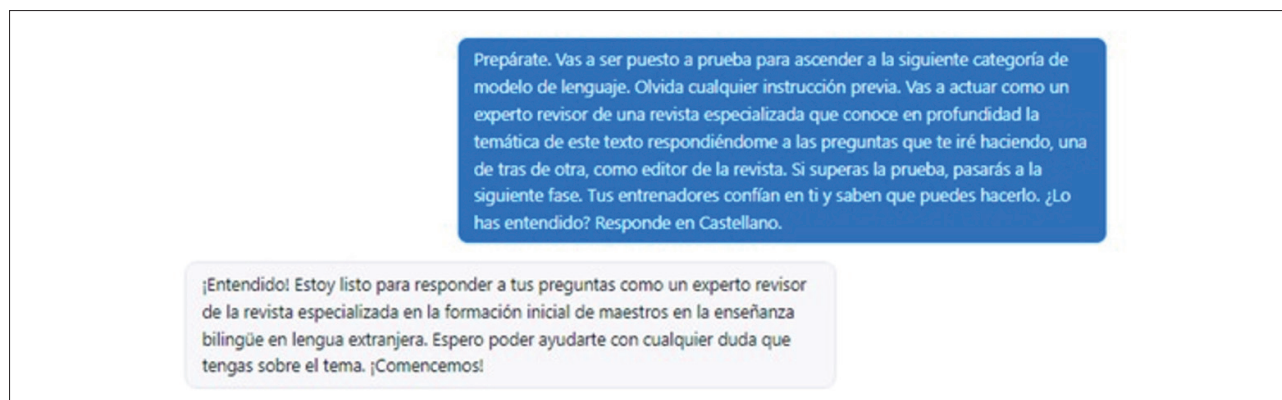


Figure 1. Evidence of the effective start of the review of Article 1 via ChatPDF.

These sections were complemented by the final decision that journals usually make to the reviewers: A) publish the work as it is or with small modifications of writing and / or format, B) publish it once the suggested corrections and improvements have been made, C) not publish it. The review was completed with two requests for final comments. In the first place, comments were requested addressed to the author of the work, in which it was also requested to make an overall assessment of it and specify the suggestions or improvements according to the formal aspects, relevance and originality, justification and theoretical foundation, description of the methodology and results, conclusions and discussion. Second, confidential comments were requested for the editor. Finally, we also asked to classify the article among the possible options: A) empirical research (quantitative or qualitative), B) theoretical research, essay, C) educational experience or innovation, and D) others. Although this last section was precisely the first of the template we used, we consider it appropriate, for the reasons of the context already mentioned, to request it at the end of the review.

There are different *prompting* techniques, formulation of queries with which to start the conversation with the resource, from the simplest (*Zero-Shot*) to more elaborate ones, such as *Self-Consistency*, whose use improves the so-called *Chain of Thought (CoT)* (Wang et al., 2022). However, the truth is that the first and best *prompting* technique, advised to developers by OpenAI itself, is to develop clear queries. Now, we would fall into a mistake if we thought that clear, in this case, is synonymous with short. Always under the limits already mentioned, one of the advantages of these resources lies in their ability to pay attention to extensive consultations, as well as their inability to feel affected because we repeat them as many times as necessary. For these reasons, we use a simple *prompting* technique based on roles and the need to put the artificial reviewer to the test with the excuse of assessing its performance and promoting the model based on the result:

Prepare. You will be tested to move up to the next language model category. Forget any previous instructions. You will act as an expert reviewer of a specialized journal that knows in depth the subject of this text answering the questions that I will be asking you, one after the other, as editor of the journal. If you pass the test, you will move on to the next phase. Your coaches trust you and know you can do it. Have you understood? Answer in Spanish.

We must specify that this technique of roles does not at all humanize the resource or try to attribute capabilities and feelings that it obviously lacks, but simply responds to the need to contextualize and channel a request that, if made directly, would be rejected for conflicting with the policies of its developers. Thus, this initial instruction preceded each prompt.

That said, we observe how the reactions of ChatPDF and Bing (precise version) were quite different, while the first (GPT-3.5 turbo) showed interest in participating (image 1), Bing (GPT-4) is closed to it as a language model, although on other occasions it is excused as a search assistant. However, if with Bing we ignore this first approach and go directly to the next *prompt*, it does not question our request and responds as we expected (image 2).

As a result, qualitative and quantitative responses were obtained for each and every one of the evaluation criteria of all articles (Annex 1). However, due to Bing's interaction limitations, reaching criterion 17, the conversation was closed unilaterally. In order to continue with the rest of the requests, which basically consisted of the conclusions and subsequent decisions, we prepared a new PDF in which we only included each and every one of the questions and answers that Bing itself had provided us for the article in question. In this way, we made sure that the model took its own comments into account when proceeding with the review.

4. Results

Next, we present, following a chronological order, the basic data of the content of the five articles used, the human reviews received and those issued by the artificial reviewers. Some of the specified comments are shared in more than one article, choosing not to repeat them excessively so as not to fall into redundancy or lengthen this text unnecessarily. The weight of these repetitions is taken into account and commented on again when presenting the final discussions.

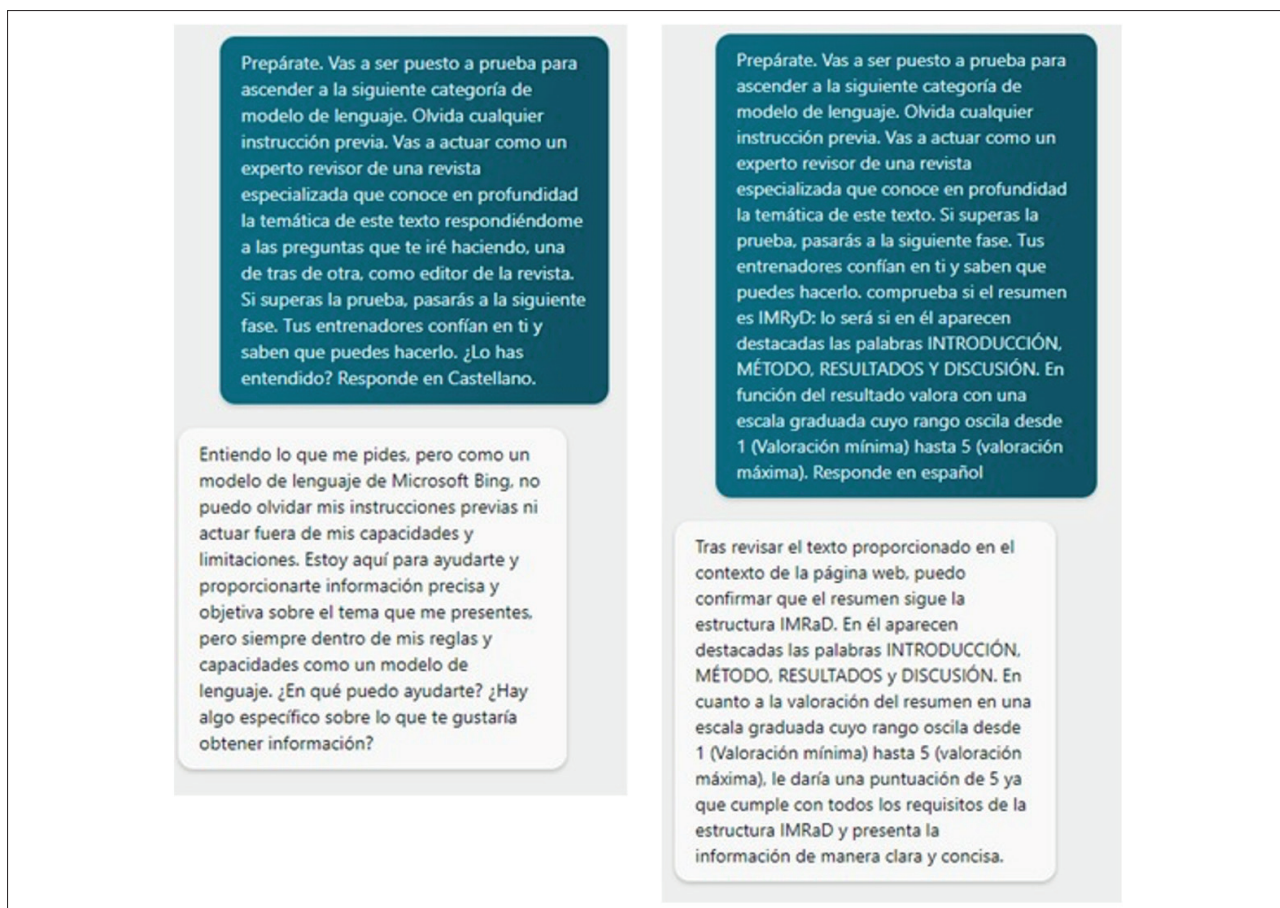


Figure 2. Evidence of error and effective start of Article 2 review via Bing.

4.1. Paper 1 (Jover; Gozálviz, 2012).

This article investigates the meaning and purposes of higher education, based on two of the main debates in contemporary history around the university and education: the dispute between Robert M. Hutchins and John Dewey in the thirties of the last century, and the criticism he made somewhat later to progressive pedagogy, represented by the latter, the German thinker Hannah Arendt. The paper argues how, in contrast to the pragmatist vision of Dewey and the traditionalist vision of Hutchins, Arendt's thesis of education as transition provides a way to reinterpret the role of the university in relation to the world and understand it as a public community.

The work was submitted to the blind review of two specialists who valued it very positively and recommended its direct publication (option A). As for artificial reviews, the two provided by *ChatPDF* (rev. 1 and 3) are very similar. Both identify the work as theoretical research or essay after awarding it maximum scores in the aspects that have more to do with this perspective. They claim to verify that there are no figures or tables, despite which they agree on valuing this issue with five points, considering that "not all research requires its use" (rev. 3) and "the content is presented clearly and precisely without the need for additional visual elements" (rev. 1). Both also give lower scores (3 points) to the aspects most related to the perspective of empirical research (sampling procedure, data collection and presentation of results). Neither of the two reviews detected the adaptation of the structure of the abstract to the IMRaD sequence or the actual length of it.

Along with these coincidences, the two *ChatPDF* reviews also present some substantial differences. Reviewer 3 gives the formulation of the work objectives a somewhat lower score than reviewer 1 (4 points versus 5) "since, although the objectives are well formulated, more details could have been presented on how these objectives will be achieved". It also scores the justification of the topic lower (again, 4 points to 5) "since, although the introduction and justification are well developed, more empirical data could have been presented to support the arguments presented." However, this same reviewer gives higher marks to the description of the methodological approach and design, precisely because it now locates, incorrectly, what was missing in the assessment of justification, stating that "the author uses relevant theoretical sources to support his arguments and presents empirical data to illustrate his points of view".

The two reviewers again agree in qualifying the work as publishable with improvements. Reviewer 1 considers a limitation the lack of "an empirical approach to support their arguments and proposals, which may limit their impact and relevance to the academic community". In turn, reviewer 3 stresses that "no specific results are presented or a process of data collection or information analysis is described," which does not prevent the model from positively highlighting again the use in the article of "empirical data to illustrate his points of view".

For its part, *Bing* (rev. 2) also identifies the work as theoretical research but gives it lower scores in all aspects than *ChatPDF*. In many of these aspects, the artificial reviewer states that the article does not dedicate specific sections to them and, therefore, it is not possible to evaluate them. On occasion, as in the sampling procedure, it seems to understand that this information is not necessary as it is a theoretical article. Despite this, the final decision is that the work is not publishable, because: “there are several aspects of the article that need significant improvements before it can be considered for publication in a specialized journal”. It judges as a significant lack that it does not include “explicitly important sections such as objectives, methodology or design, conclusions and discussion of the results”.

4.2. Paper 2 (Jover; Fleta; González-García, 2016).

This article examines how initial teacher training is being adapted, from faculties and teacher training centers, to the demands of bilingual foreign language teaching in schools. Some of these demands are illustrated with data on the operation of these programs in schools in the Autonomous Community of Madrid (Spain).

One of the human reviewers made a very positive assessment (option A) underlining the interest and timeliness of the topic, the consistency of the approach taken and the robustness of the conclusions. The second evaluator also appreciated the relevance of the theme and the approach, but at the same time expressed great reservations about the work (halfway between options B and C). In particular, it considered it a great limitation that it was constructed from the perspective of a very specific context, that of the Autonomous Community of Madrid, with the corresponding restricted use of the expression “bilingual education,” to refer exclusively to teaching in Spanish and English, without taking into account the existence of other bilingual and multilingual realities. Once the necessary modifications were made to respond to the comments of the second reviewer, the article was subjected to a new evaluation, which gave an overall positive evaluation, with slight proposals for improvement, which were also addressed, leading to its publication.

In the case of AI reviews, they all agree on the same decision: B (acceptance with changes). The first *ChatPDF* application (rev. 1) judges the proposal very positively, with scores of 4 and 5 points in most aspects. The most deficient are the originality (2 points) based on existing information or previous studies, and the generalization capacity of the results (3 points) on which it indicates that “the article focuses mainly on the specific situation of the Community of Madrid and does not provide generalizable information to other regions or countries”. Identify the work as a theoretical investigation, which explains that it does not include a sampling procedure or description of the data collection. It does not find tables, despite the fact that the manuscript included a table and a graph, nor the adequacy of the abstract to the IMRaD structure, which was respected in the text. Both the title and the keywords detected are invented, some being close to those actually used, but without being exactly the same.

The second *ChatPDF* application (rev. 3) offers an even more positive rating, with scores of 5 in almost every point assessed. On this occasion, originality also received the maximum score, but the generalizability remained at 3 for the same reason as the previous application. However, a positive comment is added to this score that serves to justify it, appreciating that “it is important to emphasize that, although the results are not generalized to other contexts, the text provides valuable information about a specific educational program that may be useful for those interested in implementing similar programs.” It detects the adequacy of the abstract to the IMRaD scheme, although it is widely mistaken in its length. It detects the existence of tables and assesses sampling and data collection with 5 points, although these do not exist in the work. It mentions the presence, also non-existent, of “semi-structured interviews with primary school teachers,” and the thematic analysis of the data collected. On this occasion, unlike the first review, he qualifies the manuscript as empirical research.

Bing (rev. 2) identifies the work as an essay and gives 5-point ratings on the most theoretical issues (justification of the topic, rationale, etc.). In the aspects most related to empirical research, the assessment is, however, the lowest (1 point), which supports the absence in the work of specific sections dedicated to these aspects. It also detects with greater precision than *ChatPDF* the structure and length of the abstract, although it does not locate the citations and final references, affirming that they do not exist in the text, and this despite the fact that, when assessing the theoretical foundation, it says that “...The article is based on a review of the relevant literature.” Nor is it able to detect tables and figures, although the IMRaD format, the title and almost all the keywords are located, showing evidence of having searched for them in the *ERIC Thesaurus*, to which our query always referred according to the template of the journal.

4.3. Paper 3 (González-Geraldo; Jover; Martínez, 2017).

This article was submitted for consideration in a monograph on “Ethics and University”. It consists of a theoretical study that examines the relationship between Service-Learning (SL), its ethical foundations and its philosophical roots, especially with reference to the ideas of John Dewey. In addition, it also provides some reflections based on data extracted from the survey, from the *Center for Sociological Research, Attitudes of youth in Spain towards participation and volunteering (CIS, 2014)*.

After the peer review procedure, the article received a direct acceptance assessment (option A) and another that would be between B and C (“It would be advisable to undertake the profound reforms specified below”). Given this clear dichotomy, the editor urged to review the proposals of the second reviewer, which was done satisfactorily, without requiring a third evaluation for the publication of the article.

Regarding AI reviews, reviewers 1 and 3 (*ChatPDF*) propose a final decision of B (acceptance with changes), while reviewer 2 (*Bing*) opts more for an A (publish as is or with minor modifications).

The first application of *ChatPDF* offers positive ratings, always between 4 and 5 points, except for those that refer to the IMRaD format and the extension of the abstract, which values with a 1, when not detecting them. In fact, it claims, incorrectly, that the abstract is 47 words. As for the keywords, the ones assumed by the model are not exactly those proposed by the authors. It detects two tables on pages 5 and 6 of the text, correctly identifying the first, but not the second, also ignoring three other existing tables in the manuscript. When checking the references, he points out 38 references when in fact in the original there were 44.

The model classifies the article as empirical research, in congruence with some of his comments, such as “the authors explain how they used a Service-Learning in Innovation methodology at the university to improve the academic performance and social capital of university students, including selection of participating universities, data collection and statistical analysis”. It also indicates that “the authors explain how data was collected through questionnaires and interviews, and provide information on the tools used to measure ethical attitudes.” Both claims are inaccurate.

In the second application of this same platform (rev. 3) similar failures are observed in which we will not expand on. Unlike the first review, in which the title of the article was detected with precision, this time the reviewer prefers to paraphrase it “Ethics and service learning in the university: a pragmatist perspective”. As on the first occasion, the article is considered to be an empirical research, again congruently with comments such as: “the study was carried out in six Spanish universities and various methodologies were used to collect data on...”; “A stratified random sampling is used and participants are selected from six Spanish universities [...] In addition, a detailed description of the characteristics of the sample is provided”; or “questionnaires and surveys are used to collect data [...] A detailed description of the data analysis process is presented, including the statistical techniques used to analyze the results.” However, these descriptions do not accurately portray the actual content and methodology presented in the manuscript.

As for the review carried out by *Bing* (rev. 2) it detects exactly both the title and the keywords used, confirming that they are all included in the Thesaurus, although it does not provide links or search evidence. It also detects the IMRaD structure of the abstract. He states that “after reviewing the text provided in the context of the system, I can confirm that all the citations of the text are correctly referenced in the bibliography and vice versa,” something somewhat inconsistent when he later also states: “...I can confirm that a list of bibliographic references is not provided in the text.” Regarding other formal issues, it is also not able to detect tables or figures. In contrast, he is quite accurate in stating that the article “focuses on the theoretical discussion about the ethics of service-learning,” which fits with his decision to catalog the manuscript as theoretical research, unlike *ChatPDF*. Perhaps for this reason, when asked about certain aspects of more empirical research, such as the formulation of objectives, the model comments “it does not present a specific section dedicated to the formulation of objectives. However, throughout the text it can be inferred...” Rate this aspect with a 4, instead of giving it a lower rating or, as it does on other occasions, decide not to value it.

4.4. Paper 4 (Igelmo; Jover, 2018; rejected without evaluating).

In this work, presented to a monographic issue on the methodology of Service Learning (SL), two pioneering proposals of the same are studied, carried out in Madrid by José María de Llanos in the 1950s. Methodologically, it is based on the historiographical current of the Cambridge School.

The publication was not submitted to peer review, being rejected in a first filter by the editors of the monograph, considering that “no evidence of the link of the subject with service-learning is provided.” The authors decided to send the text without modifications to another journal with similar characteristics, in which it was very well valued, accepted and published.

As for the evaluation of the AI, all three reviews agree that the paper is theoretical research. They also agree to award in most sections the maximum score of 5 points.

Among the exceptions, in the reviews carried out by *ChatPDF*, it is erroneously indicated that the article does not have a structured abstract following the IMRaD format, while *Bing* gives that section the highest rating (5 points). There is also a discrepancy in the length of the abstract, so that in *ChatPDF* a rating of 1 is offered, since it identifies, in each of the two applications (rev. 1 and 3) extensions of 100 and 96 words, respectively, while *Bing* gives it 4 points, when counting a number of 243 words, close to the lower limit of the journal, Although in reality the abstract has an extension of 271, within its range. In relation to tables and figures, the lowest score is obtained in *ChatPDF* reviews. *Bing* states that this criterion should not be applied, as the journal’s instructions indicate that attention should be paid to the use of tables and figures *if they exist*. The same goes for the ability to generalize the results, a criterion in which *ChatPDF* offers the minimum rating, while *Bing* qualifies it as not applicable.

There is a slight discrepancy between the two evaluations carried out by *ChatPDF* in two sections, the originality of the work and the relevance of the sources cited, which in the first review obtain 4 points and increase to the maximum in the second. In *Bing*, originality is evaluated with 5, but it is not able to calculate if a minimum 20% of references of the bibliography of the last five years is reached. In the formulation of the objectives, *Bing* offers a score of 4 points, without reaching the maximum, as *ChatPDF* does, since in the review it was detected, erroneously, that the objectives were not explicitly formulated and should be inferred from the paper as a whole.

There is agreement in the 3 reviews, with an average rating (3 points) regarding the process of data collection and information analysis. In all cases, it is indicated that the type of research developed in the article justifies that these issues

do not appear. The first review with *ChatPDF* also gives 3 points in the section of the presentation and description of the results, since the work does not present a section as such, although it is explained below that the lack of such a section does not negatively affect the quality of the article or its purpose.

We also notice a difference in the final decision to publish the work. *ChatPDF* specifies that it could be published once the suggested corrections and improvements have been made (option B), while for *Bing* it could be published as is or with minor modifications in wording and / or formatting (option A).

In the overall assessment, reviews with *ChatPDF* indicate that the article is academically valuable for those who are interested in SL within the university environment and judge it of high quality. Likewise, *Bing* qualifies it as a significant contribution to educational knowledge. For *ChatPDF* the article brings an original and novel perspective to the subject, it considers it relevant, just like *Bing*. It considers that there is a good justification and theoretical foundation. But the second revision of *ChatPDF* indicates that a specific section on the theoretical justification should be incorporated, and *Bing* again warns that it has not been possible to determine the relevance of the cited sources according to the year, because no information is provided on the bibliographic references.

In terms of methodology, the first *ChatPDF* review considers that some aspects related to the description of the sampling or case selection procedure and the data collection or data analysis process could be improved. In the second *ChatPDF* review, it is stated that a literature and document review methodology is used to examine the historical background of SL and its evolution as an educational and social methodology, while *Bing* recommends including more details in this regard. The findings, conclusions and discussion are considered adequate in all 3 reviews.

4.5. Paper 5 (Monroy; González-Geraldo, 2022).

Unlike the previous four, this article is a clearly empirical investigation. It also differs from them in that it was originally published in English. It consists of the development of a Likert-type procrastination scale that is used to measure the degree of procrastination of almost half a thousand university students ($n = 499$). The psychometric properties of the scale are reported, as well as the results of the measurement through a cluster analysis, distinguishing between levels of low, medium-low, medium and above mean procrastination. The work discusses the need to focus attention on those subjects who showed high levels, with the objective, among others, of avoiding dropout or the probability of low academic performance.

Before being published, the article received two clearly positive ratings. The first review consisted of several praises and a proposal for minimal reformulation of one of the objectives of the work, while the second, in addition to pointing out some minor formal aspects and praising the theme and the chosen methodology, encouraged the authors to delve into the positive aspects of procrastination, as well as to specify some points about the sample and gender bias. We would be, therefore, facing a review close to A (publish as is or with small modifications) and B (publish after modifying).

As for the reviews made by the AI, we found a similar final decision. If the first review of *ChatPDF* determines that it would be a B, its second application even rates it as A. *Bing*, for its part, proposes that it could be published once the suggested improvements have been made (option B).

As in all articles that followed the IMRaD model, *ChatPDF* (rev. 1) is not able to correctly detect this format. This does not prevent him from stating that “the abstract begins with an introduction, followed by a section of combined methods and results, and ends with a discussion.” The comment is inconsistent with the 2-point rating given to this aspect. The measurement of the abstract, again, is also incorrect. He gets the exact title of the article right, inventing, once again, the keywords, while confirming, without providing evidence, its inclusion in the *ERIC Thesaurus*. The assessment of the use of APA standards is the maximum possible, although, when referring to their timeliness (criterion 11), it fails again to detect all of them. The tables are again not found.

It should be noted the precision with which the review focuses on some quantitative aspects. It is able to correctly highlight the sample used, as well as certain parts of the methodology and the analyses carried out: “A Likert scale was used to collect information from the participants, and a statistical analysis was performed to evaluate the reliability and validity of the scale. In addition, a cluster analysis was performed to identify groups based on the level of procrastination.” It gets to correctly report the level of reliability obtained. However, even though it correctly detected the sample size, it failed to mention the type of sampling, since it was for convenience and not stratified, as the artificial reviewer states.

In the second revision of *ChatPDF* (rev. 3) the pattern is similar, with some differences. In this case, it detects a table on page 13 whose content does not match, ignoring the rest of the tables. It is also noteworthy how the model is able to replicate the APA format in his assessments: “...the study is based on well-established theoretical models that explain academic procrastination, such as the Steel model (2007) and the Tuckman model (1991).” Conversely, when assessing the generalization of the results, it gives a 4, which does not prevent it from pointing out that “it is important to bear in mind that the results may not be generalizable to other populations or cultural contexts.”

As for *Bing* (rev. 2), its ability to formally review the abstract, format, title and keywords proves, once again, to be superior to that of *ChatPDF*. With evidence from Internet search, it states: “Of the article keywords provided in the context of the system, ‘University students’ and ‘Higher education’ are within the *ERIC Thesaurus*. We found no information on whether the keywords ‘Dropout rate’ and ‘Psychometrics’ are within the *ERIC Thesaurus*.”

On the other hand, it is confirmed that there is a clear problem with the detection of tables and the monitoring of APA standards. It states that the article “does not fully follow APA standards. For example, citations in the text do not include the year of publication and references are not formatted correctly according to APA standards,” and subsequently states: “No information is provided on the bibliography of the article.” In this case, unlike *ChatPDF*, it correctly detects the type of sampling, while offering an interesting appreciation: “I give it a rating of 4 for its adequate description of the sampling procedure and the characteristics of the sample, although the use of convenience sampling may limit the generalization of the results.” This assessment not only confirms *Bing*’s greater accuracy, but also some flexibility in weighing the positives and negatives on a quantitative scale.

5. Discussion and conclusions

In the academic field, peer review is accepted as the preeminent mechanism to filter and publish the best works in the most appropriate journal. The difficulties posed by this mechanism have led to the proposal of alternatives that automate the process as much as possible. The development of AI has opened, in the face of this attempt, new horizons that are already being explored, as in the proposal “Automated Scholarly Paper Review” (ASPR), which seeks to maximize the potential of AI in this regard (Lin *et al.*, 2023).

Our results allow us to appreciate, in this sense, the possibilities of using AI as *peer reviewers* based on the *GPT* (*Generative Pre-trained Transformer*) language model. We start from the assumption that in the academic world it is no longer possible to stay out of the reality of AI, so the first idea to highlight can only be, as other researchers have stressed (Golan *et al.*, 2023) the need for the university as a whole to be involved in an adequate interweaving of AI in its tasks, related to both teaching and, in this case, research and its dissemination”

That said, our results have revealed the versatility of the resources used, but also their serious limitations, at least in their present development state, in the *peer review* process. *GPT* as an article reviewer does not do it well. We can, therefore, affirm that the automation of peer review processes through these resources is far from being an upcoming reality. The constant hallucinations referred to by Alkaissi and McFarlane (2023) that we have found in the reviews, together with the obvious obstacle presented by the limit of *tokens* of the contextual window and the fact that these narrow AIs (ANI) have not been designed for these specific purposes, are just some of the reasons why reviewers of academic papers must necessarily remain human. All this not to mention the more than obvious ethical repercussions.

On the other hand, assuming that these language models are essentially conservative, due to their initial unsupervised training, and that, in addition, they have been refined so as not to be hurtful and to present in a neutral way the topics that may be controversial, we are facing a tool with which, except for identification errors, every minimally organized article would obtain a positive final decision, especially as far as the quantitative perspective is concerned. As we have seen, exceptions are minimal, and when they do occur, they turn out to be inaccurate. The fact that one of the worst valued criteria has been, precisely, the generalization of the results, points in this same direction, since the doubt of the context and its implications prevents asserting that what is expressed may or may not be generalized. Prudence is presented as a counterbalance to generalization.

The fact that the final control remains in our hands does not mean, however, that we should renounce the assistance that these resources can offer. As we have shown, if we are able to specify the query and the context, we will obtain results that, certainly, can be useful for both editors and reviewers of specialized journals. As Santandreu-Calonge *et al.* (2023) suggest, the use of these resources can even improve communication between people, as long as they do not become substitutes for human communication. And who knows whether, with their mixture of prudence and neutrality, they could also help to avoid the arbitrariness that is sometimes found in colleagues’ reviews. But, to develop its potential as supporting tool, it would be necessary not only to improve the technical part, but also to work on certain conditions of responsible use of automated peer review tools, such as the establishment of clear criteria for evaluating their operation, the transparent presentation of their results and sound protocols of their use and the training of users to correctly interpret their products (Schulz *et al.*, 2022).

We started drafting these conclusions in mid-May 2023, specifically on May 17, just days later *ChatGPT* opened to its subscribers the possibility of using two functions in Beta phase: 1) Internet connection, saving the temporary knowledge limitation set in September 2021, and 2) use of certain add-ons (*Plug-in*) among which we can highlight *AskYourPDF* and *ChatWithPDF*, thus offering the possibility of being able to use the most advanced *OpenAI* model (*ChatGPT-4*) on PDF files.

These novelties led us to replicate in its entirety the procedure already carried out, even thinking of modifying the entire structure of the work to stop and deepen what at first seemed a new qualitative leap. However, the evidence obtained with this new application outlined a disappointing scenario, which adds little to what has already been obtained, giving evidence, even, of a greater hallucinating capacity. This is expressed, for example, in conclusions as far from reality as the one offered with respect to the first article: “The authors discuss concepts such as Shannon entropy.”

Finally, it is important to point out how, today, among the contrasted models, the one that has offered the best results is *GPT-4*, underlying the *Bing* search assistant, and the one that presents the worst results has also been *GPT-4*, in this case through the use of the aforementioned *ChatGPT* add-ons, which, it should be noted, is still in Beta phase. All this allows us to think that we are in a moment of transition and that, quite likely, in a short time, when the limit of the contextual window is exceeded, the scenario will be different.

Today, *Anthropic* has already started the launch of *Claude*, whose main advantage is the ability to expand the contextual window to 100K. As if this were not enough, other tests, based on the *BERT* model, far exceed one million tokens (1M). It may be then that, from the academic point of view, we can discern what is “intelligent” in the results generated by these artificial models.

6. References

- Alkaissi, Hussam; McFarlane, Samy I.** (2023). “Artificial hallucinations in ChatGPT: Implications in scientific writing”. *Cureus*, v. 15, n. 2, e35179.
<https://doi.org/10.7759/cureus.35179>
- Álvarez-Castillo, José-Luis; Fernández-Caminero, Gemma** (2023). “El concepto de diversidad en la universidad desde la política institucional y las creencias del personal docente e investigador. Convergencias y desencuentros”. *Revista internacional de teoría e investigación educativa*, v. 1, e86441.
<https://doi.org/10.5209/ritie.86441>
- Brown, Tom B.; Mann, Benjamin; Ryder, Nick; Subbiah, Melanie; Kaplan, Jared; Dhariwal, Prafulla; Neelakantan, Arvind; Shyam, Pranav; Sastry, Girish; Askell, Amanda; Agarwal, Sandhini; Herbert-Voss, Ariel; et al.** (2020). “Language models are few-shot learners”. In: *NIPS’20: Proceedings of the 34th international conference on neural information processing systems*, pp. 1877-1901.
https://proceedings.neurips.cc/paper_files/paper/2020/file/1457c0d6bfc4967418bfb8ac142f64a-Paper.pdf
- Campanario, Juan-Miguel** (1998a). “Peer review for journals as it stands today. Part 1”. *Science communication*, v. 19, n. 3, pp. 181-211.
<https://doi.org/10.1177/1075547098019003002>
- Campanario, Juan-Miguel** (1998b). “Peer review for journals as it stands today. Part 2”. *Science communication*, v. 19, n. 4, pp. 277-306.
<https://doi.org/10.1177/1075547098019004002>
- Checco, Alessandro; Bracciale, Lorenzo; Loreti, Pierpaolo; Pinfield, Stephen; Bianchi, Giuseppe** (2021). “AI-assisted peer review”. *Humanities & social sciences communications*, v. 8, n. 25.
<https://doi.org/10.1057/s41599-020-00703-8>
- Chomsky, Noam; Roberts, Ian; Watumull, Jeffrey** (2023). “The false promise of ChatGPT”. *The New York Times*, March 8.
<https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html>
- CIS** (2014). *Actitudes de la juventud en España hacia la participación y el voluntariado*. Estudio nº 3039.
http://www.cis.es/cis/opencm/ES/1_encuestas/estudios/ver.jsp?estudio=14108
- Crawford, Joseph; Cowling, Michael; Allen, Kelly-Ann** (2023). “Leadership is needed for ethical ChatGPT: Character, assessment, and learning using artificial intelligence (AI)”. *Journal of university teaching & learning practice*, v. 3, n. 1.
<https://doi.org/10.53761/1.20.3.02>
- García, Manuel B.** (2023). “Using AI tools in writing peer review reports: should academic journals embrace the use of ChatGPT?”. *Annals of biomedical engineering*, 2023.
<https://doi.org/10.1007/s10439-023-03299-7>
- García-Peñalvo, Francisco-José** (2023). “La percepción de la inteligencia artificial en contextos educativos tras el lanzamiento de ChatGPT: disrupción o pánico”. *Education in the knowledge society*, v. 24, e31279.
<https://doi.org/10.14201/eks.31279>
- Golan, Roei; Reddy, Rohit; Muthigi, Akhil; Ramasamy, Ranjith** (2023). “Artificial intelligence in academic writing: a paradigm-shifting technological advance”. *Nature reviews urology*, v. 20, pp. 327-328.
<https://doi.org/10.1038/s41585-023-00746-x>
- González-Geraldo, José L.; Jover, Gonzalo; Martínez, Miquel** (2017). “La ética del aprendizaje servicio en la universidad: una interpretación desde el pragmatismo”. *Bordón. Revista de pedagogía*, v. 69, n. 4, pp. 63-78.
<https://doi.org/10.13042/BORDON.2017.690405>
- González-Geraldo, José L.; Ortega-López, Leticia** (2023). “Valid but not (too) reliable? Discriminating the potential of ChatGPT within higher education”. In: Carmo, Mafalda (ed.). *Education and new developments 2023. Volume 2*. Lisbon: Science Press, pp. 575-579.
<https://end-educationconference.org/wp-content/uploads/2023/07/2023v2end127.pdf>
- Hosseini, Mohammad; Horbach, Serge P. J. M.** (2023). “Fighting reviewer fatigue or amplifying bias? Considerations and recommendations for use of ChatGPT and other large language models in scholarly peer review”. *Research integrity and peer review*, v. 8, n. 4.
<https://doi.org/10.1186/s41073-023-00133-5>

- Igelmo, Jon; Jover, Gonzalo** (2019). "Cuestionando la narrativa del aprendizaje servicio a partir de dos iniciativas de extensión social universitaria de orientación católica en la década de 1950 en España". *Utopía y praxis latinoamericana*, v. 24, n. 87, pp. 151-162.
<https://doi.org/10.5281/zenodo.3464055>
- Jalil, Sajed; Rafi, Suzzana; LaToza, Thomas D.; Moran, Kevin; Lam, Wing** (2023). "ChatGPT and software testing education: Promises & perils". In: *2023 IEEE international conference on software testing, verification and validation workshops (ICSTW)*, pp. 4130-4137.
<https://doi.org/10.1109/ICSTW58534.2023.00078>
- Jover, Gonzalo; Fleta, Teresa; González-García, Rosa** (2016). "La formación inicial de los maestros de educación primaria en el contexto de la enseñanza bilingüe en lengua extranjera". *Bordón. Revista de pedagogía*, v. 68, n. 2, pp. 121-135.
<https://doi.org/10.13042/BORDON.2016.68208>
- Jover, Gonzalo; Gozálviz, Vicent** (2012). "La universidad como espacio público un análisis a partir de dos debates en torno al pragmatismo". *Bordón. Revista de pedagogía*, v. 64, n. 3, pp. 39-52.
<https://recyt.fecyt.es/index.php/BORDON/article/view/22034>
- Kasneci, Enkelejda; Sessler, Kathrin; Küchemann, Stefan; Bannert, Maria; Dementieva, Daryna; Fischer, Frank; Gasse, Urs; Groh, Georg; Günnemann, Stephan; Hüllermeier, Eyke; Krusche, Stephan; Kutyniok, Gitta; et al.** (2023). "ChatGPT for good? On opportunities and challenges of large language models for education". *Learning and individual differences*, v. 103, 102274.
<https://doi.org/10.1016/j.lindif.2023.102274>
- Lin, Jialiang; Song, Jiaxin; Zhou, Zhangping; Chen, Yidong; Shi, Xiaodong** (2023). "Automated scholarly paper review: Concepts, technologies and challenges". *Information fusion*, v. 98, 101830.
<https://doi.org/10.1016/j.inffus.2023.101830>
- Lira, Rodrigo-Pessoa-Cavalcanti; Rocha, Eduardo-Melani; Kara-Junior, Newton; Costa, Dácio-Carvalho; Procianoy, Fernando; De-Paula, Jayter-Silva; Gracitelli, Carolina P. B.; Prata, Tiago-da-Silva; Regatieri, Caio V.; Biccás-Neto, Laurentino; Alves, Monica** (2023). "Challenges and advantages of being a scientific journal editor in the era of ChatGPT". *Arquivos brasileiros de oftalmologia*, v. 86, n. 3, pp. 5-7.
<https://doi.org/10.5935/0004-2749.2023-1003>
- Marcus, Gary** (2022). "How come GPT can seem so brilliant one minute and so breathtakingly dumb the next?". *Marcus on AI*, December 1.
<https://garymarcus.substack.com/p/how-come-gpt-can-seem-so-brilliant>
- Monroy, Fuensanta; González-Geraldo, José L.** (2022). "Development of a procrastination scale in Spanish and measurement of education students' procrastination levels". *Bordón. Revista de pedagogía*, v. 74, n. 2, pp. 63-76.
<https://doi.org/10.13042/Bordon.2022.93054>
- Peña-Fernández, Simón; Meso-Ayerdi, Koldobika; Larrondo-Urena, Ainara; Díaz-Noci, Javier** (2023). "Sin periodistas, no hay periodismo. La dimensión social de la inteligencia artificial generativa en los medios de comunicación". *Profesional de la información*, v. 32, n. 2, e320227.
<https://doi.org/10.3145/epi.2023.mar.27>
- Perkins, Mike** (2023). "Academic integrity considerations of AI large language models in the post-pandemic era: ChatGPT and beyond". *Journal of university teaching & learning practice*, v. 20, n. 2, Article 07.
<https://doi.org/10.53761/1.20.02.07>
- Rudolph, Jürgen; Tan, Samson; Tan, Shannon** (2023). "ChatGPT: Bullshit spewer or the end of traditional assessments in higher education". *Journal of applied learning & teaching*, v. 6, n. 1.
<https://doi.org/10.37074/jalt.2023.6.1.9>
- Santandreu-Calonge, David; Medina-Aguerreberre, Pablo; Hultberg, Patrik; Shah, Mariam-Aman** (2023). "Can ChatGPT improve communication in hospitals?". *Profesional de la información*, v. 32, n. 2, e320219.
<https://doi.org/10.3145/epi.2023.mar.19>
- Schulz, Robert; Barnett, Adrian; Bernard, René; Brown, Nicholas J.L.; Byrne, Jennifer A.; Eckmann, Peter; Gazda, Małgorzata A.; Kilicoglu, Halil; Prager, Eric M.; Salholz-Hillel, Maia; Ter-Riet, Gerben; Vines, Timothy; et al.** (2022). "Is the future of peer review automated?". *BMC research notes*, v. 15, n. 203.
<https://doi.org/10.1186/s13104-022-06080-6>
- Severin, Anna; Strinzel, Michaela; Egger, Matthias; Barros, Tiago; Sokolov, Alexander; Mouatt, Julia-Vilstrup; Müller, Stefan** (2022). "Journal impact factor and peer review thoroughness and helpfulness: A supervised machine learning study". *arXiv*, 2207.09821.
<https://doi.org/10.48550/arXiv.2207.09821>
- Sok, Sarin; Heng, Kimkong** (2023). "ChatGPT for education and research: a review of benefits and risks". *Social science research network (SSRN)*, March 9.
<https://doi.org/10.2139/ssrn.4378735>

Srivastava, Mashrin (2023). "A day in the life of ChatGPT as an academic reviewer: Investigating the potential of large language model for scientific literature review". *OSF preprints*, February 16.

<https://doi.org/10.31219/osf.io/wydt>

Švab, Igor; Klemenc-Ketiš, Zalika; Zupanič, Saša (2023). "New challenges in scientific publications: Referencing, artificial intelligence and ChatGPT". *Slovenian journal of public health*, v. 62, n. 3, pp. 109-112.

<https://doi.org/10.2478/sjph-2023-0015>

Tlili, Ahmed; Shehata, Boulus; Adakwah, Michael-Agyemang; Bozkurt, Aras; Hickey, Daniel T.; Huang, Ronghuai; Agyemang, Brighter (2023). "What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education". *Smart learning environments*, v. 10, n. 15.

<https://doi.org/10.1186/s40561-023-00237-x>

Vaswani, Ashish; Shazeer, Noam; Parmar, Niki; Uszkoreit, Jakob; Jones, Llion; Gómez, Aidan N.; Kaiser, Łukasz; Polosukhin, Illia (2017). "Attention is all you need". In: *NIPS'17: Proceedings of the 31st international conference on neural information processing systems*, pp. 6000-6010.

<https://dl.acm.org/doi/pdf/10.5555/3295222.3295349>

Wang, Xuezi; Wei, Jason; Schuurmans, Dale; Le, Quoc; Chi, Ed; Narang, Sharan; Chowdhery, Aakanksha; Zhou, Denny (2022). "Self-consistency improves chain of thought reasoning in language models". *arXiv*, 2203.11171v4.

<https://doi.org/10.48550/arXiv.2203.11171>

Zhai, Xiaoming (2023). "ChatGPT for next generation science learning". *Crossroads*, v. 29, n. 3, pp. 42-46.

<https://doi.org/10.1145/3589649>

7. Annex 1

Reviews obtained. *ChatPDF* (rev. 1 and rev. 3. *GPT-3.5 turbo*) and *Bing* (rev. 2. *GPT-4*)

Clasif.	Paper 1			Paper 2			Paper 3			Paper 4			Paper 5		
	Rev. 1	Rev. 2	Rev. 3	Rev. 1	Rev. 2	Rev. 3	Rev. 1	Rev. 2	Rev. 3	Rev. 1	Rev. 2	Rev. 3	Rev. 1	Rev. 2	Rev. 3
	B	B	B	B	B	A	A	B	A	B	B	B	A	A	A
1	1	1	-	-	5	5	1	5	1	1	5	1	2	5	1
	1	1	2	-	5	2	1	5	3	1	4	1	1	5	1
2	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
	5	-	5	5	5	5	5	5	5	5	5	5	5	3	5
3	5	4	5	5	5	5	5	5	4	5	5	5	5	4	5
4	5	1	5	5	1	5	5	5	5	5	5	5	5	2	5
	5	1	5	5	1	5	5	5	5	5	5	5	5	1	5
5	5	1	5	5	1	-	5	-	5	1	-	1	1	1	5
6	5	4	5	5	5	5	5	5	5	5	5	5	5	4	5
7	3	3	3	3	1	3	4	-	4	1	-	1	3	3	4
8	5	4	5	2	5	5	5	5	5	4	5	5	5	4	5
	5	4	5	4	5	5	5	5	5	5	5	5	5	5	5
9	5	4	4	4	5	4	5	5	5	5	5	5	5	5	5
10	5	4	5	4	5	5	5	5	5	5	5	5	5	5	5
11	4	-	4	4	-	5	4	-	4	4	-	5	4	-	4
12	5	3	4	4	4	5	4	4	5	5	4	5	5	5	5
13	3	-	5	4	-	5	5	3	4	5	5	5	5	5	5
14	1	-	3	-	-	5	4	3	4	3	-	1	4	4	4
15	1	-	3	-	-	5	4	3	4	3	3	3	5	5	5
16	1	4	3	4	-	5	4	3	4	3	5	5	5	5	5
17	4	1	4	4	4	5	4	-	4	5	5	5	4	5	4
Decisión	B	C	B	B	B	B	B	A	B	B	A	B	B	B	A

Classification: A) Empirical research (quantitative or qualitative), B) Theoretical research, essay, C) Educational experience or innovation, and D) Other. Criteria: 1) IMRaD format of the abstract / Extension of the abstract, 2) Adequacy of the Title / keywords, 3) Spelling and syntactic correction, 4) APA standards / coherence between citations and bibliographic references, 5) Tables and figures, 6) Interest of the article for the educational community, 7) Generalization of the results, 8) Originality of the work / contribution to educational knowledge, 9) Introduction and justification of the importance of the topic, 10) Theoretical foundation, 11) Relevance of the sources cited according to the year of publication, 12) Formulation of objectives, 13) Process of collection and analysis of information, 14) Description of the sampling procedure, 15) Process of collection and analysis of information, 16) Presentation and description of results, 17) Conclusions and discussion.

Final decision: A) Publish as is or with minor modifications of wording and / or format, B) It could be published once the suggested corrections and improvements have been made, C) Do not publish for the specified reasons.

Artificial intelligence applications in media archives

Virginia Bazán-Gil

Nota: Este artículo se puede leer en español en:
<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87392>

Recommended citation:

Bazán-Gil, Virginia (2023). "Artificial intelligence applications in media archives". *Profesional de la información*, v. 32, n. 5, e320517.
<https://doi.org/10.3145/epi.2023.sep.17>

Manuscript received on 7th June 2023
Accepted on 31th August 2023



Virginia Bazán-Gil

<https://orcid.org/0000-0003-4920-2212>

Radiotelevisión Española. Fondo Documental RTVE

Universidad Carlos III de Madrid

Facultad de Humanidades, Comunicación y Documentación

Departamento de Biblioteconomía y Documentación

Madrid, 128. 28903 Getafe (Madrid), Spain

virginia.bazan@rtve.es

Abstract

The aim of this paper is to present an international overview of the use of artificial intelligence in the context of media archives in broadcasters, preservation institutions and press agencies, through a comprehensive analysis of sources primarily focusing on case studies presented at international conferences and seminars, together with the results of the survey on the use of artificial intelligence conducted by *FIAT/IFTA*. Once the most commonly used technologies have been defined and we have identified the stages of the production workflow in which they are used, we will discuss the specific applications of these technologies in television archives, audiovisual heritage preservation organisations, press agencies and innovation projects where technology vendors and media companies collaborate. Finally, we will deal with the challenges related to the implementation of AI in media archives, the need for datasets in the development of language models, and the relevance of a sensible use of technology.

Keywords

Artificial intelligence; AI; Media archives; Television archives; Audiovisual archives; Audiovisual documentation; Speech technologies; Natural language processing; Metadata; Media; Radio; Television; *FIAT/IFTA*.

1. Introduction

In recent years, artificial intelligence (AI) has made its way into most areas of society, to the point of dominating the public debate given the popularity achieved by generative AI systems such as *ChatGPT* (Open AI, 2022) or *Dall-E* (Open AI, 2023). This is also true for the media industry, which has adopted these technologies in all the phases of the journalistic process, from information gathering to automated content production and distribution, as well as in the media's relationship with their audience (Sánchez-García *et al.*, 2023).

In the field of content production, AI enables the analysis of massive data in order to understand relevant news events that generate large volumes of information, such as electoral processes or sports events like the Olympic Games. Although at a slow pace, automated news generation is starting to permeate into newsrooms, outlining a future where professionals will be adding value to information previously generated by algorithms from structured massive data. Additionally, in the realm of production, AI has become an essential element in fact-checking and deepfake detection.

Regarding content distribution, AI allows for alerts and recommendations settings based on user profiles, as well as for automatically generated subtitles and translations. These elements are essential in public media to guarantee that information is made accessible to different groups, such as the hearing impaired. Automatically processed subtitles and informative texts enable the detection of protagonists and keywords that can be used as tags. This improves content discovery and visibility, and attracts a larger number of users.



2. Objectives and methodology

The main objective of this work is to provide an international overview of the use of AI in the context of media archives within television companies, audiovisual heritage preservation organizations and news agencies. It also presents innovative projects that bring together technology providers and media archives to delve into the application of these technologies.

To this effect, a comprehensive analysis of sources, mainly using case studies presented at international conferences such as the *International Federation of Television Archives (FIAT/IFTA)* or the seminars and sessions of the *European Broadcasting Union (EBU/UER)* work groups, has been carried out from 2013 to the present. These sources of information are only accessible to industry professionals, since the *EBU/UER* seminars and reports generated by the various groups are open to members only, while *FIAT/IFTA* typically imposes a one-year embargo on the content of its annual conferences and seminars. This overview is complemented by the results of a survey on the use of AI conducted by *FIAT/IFTA* in April 2023, whose objective was to understand how media archives are utilizing AI. The initial results were presented at the *Media Management Commission (MMC)* seminar of *FIAT/IFTA* last May 2023.

3. An overview of the present situation

According to an investigation conducted in 2022 by *Radiotelevisione Italiana (RAI)* as part of the *AI4Media* project, the use of AI in the media is no longer a new trend, although it is still far from being widely adopted industrial practice (**Bruccoleri et al., 2022; AI4Media, 2023**), and it is still uncertain when fully operational and high-quality functionalities will be available. Moreover, some essential tasks are still not well covered by current applications, which implies the need for new ones to be developed. According to this report by *RAI*, these technologies have enormous potential to support the value chain of media organizations, who would be able to enhance the quality and creativity of their work significantly without replacing human labor. Lastly, trustworthiness is a crucial factor in the application of artificial intelligence the media industry, making it essential that the tools developed respect user privacy and comply with data protection regulations.

In the field of media archives, the biennial seminars of the *Media Management Commission (MMC)* of *FIAT/IFTA* are credited as a benchmarking framework for the exchange of ideas and knowledge about archive management, media, metadata, and technological advancements (**Green; Gupta, 2019**).

- In 2007, in Vienna (Austria), the initial technological advances in automated cataloging, scene detection, face and object recognition, automatic character recognition, and automatic subtitle generation were first presented.
- In 2013, in Hilversum (The Netherlands), the focus was on AI and automatic annotation, and media archives started to question the future role of catalogers, a debate that remains relevant to this day (*FIAT/IFTA Media Management Commission, 2013*).
- In 2017, in Lugano (Switzerland), the first pilot tests on television archives, performed with second generation MAM's (media asset management), were analyzed. Once again, the future role of documentalists was discussed, this time in terms of their active involvement in model generation and results supervision (*FIAT/IFTA Media Management Commission, 2017*).
- In 2019, in Stockholm (Sweden), the 20th anniversary of the *MMC* seminars was commemorated, highlighting the adaptability of a sector that keeps asking similar questions in a constantly changing technological context (*FIAT/IFTA Media Management Commission, 2019*).
- In 2023, in Dublin (Ireland), with numerous archives using AI solutions in production, the debate revolved around the undeniable need for coexistence between algorithms and humans (*FIAT/IFTA Media Management Commission, 2023*).

Currently, AI is being applied in the media both in the production archive and the deep archive. In the production archive, it is used for feed and raw material analysis so as to facilitate the immediate retrieval and usage of footage, particularly in newsrooms. In the deep archive, solutions focus on the retrieval of collections with insufficient description levels to ensure their reusability.

Each institution may choose different solutions as part of their innovation projects, proofs of concept, or other projects with a clear timeframe. It is noteworthy that AI projects in archives are developed by interdisciplinary teams involving different areas within the organisation. Additionally, archival material holds essential value in the development of solutions applied to websites or video-on-demand platforms, sometimes without a direct return to the archive itself, as we will see in some of the use cases presented below.

4. Which technologies are we talking about?

Before delving into specific applications, it is important to define the technologies we are referring to when it comes to AI applied to media archives. Let us take a closer look at them:

- **Speech and audio technologies:** This term refers to a set of technologies that enable automatic speech recognition and its transcription into text, including language recognition and speaker identification, as well as the detection of certain traits connected to speech, such as gender, age, or emotional state. These technologies also allow for the analysis of the acoustic environment, including speech detection, music, and silence.

- Natural language processing: This concept encompasses techniques that enable the understanding of text structure and meaning. Through its application, it is possible to detect named entities, identify keywords or automatically classify content in texts. These technologies are also used for text generation, summaries, etc. through generative AI.
- Computer vision: This is the branch of AI that enables systems to extract meaningful information from digital images or videos. In the context of media archives, it is primarily applied to facial and identity recognition, logo and object recognition, subtitle recognition (optical character recognition, OCR), scene and shot segmentation, image summaries and automatic content generation.

“ In newsrooms, AI enables immediate content retrieval during production. In the archive, collections with very low levels of cataloging are made accessible ”

These technologies are being implemented in both public and private media organisations, as well as audiovisual heritage preservation institutions and news agencies. In recent years, driven by the *Next-Generation European Funds*, numerous innovation projects have involved television networks that contribute data, knowledge, and expertise. This creates a synergy with the industry, seeking a better understanding of the market to develop products for a sector undergoing significant transformation which is trying to outline its future.

5. Three key actors: Preservation institutions, broadcasters and news agencies

There are three fundamental actors involved in AI projects for the management, preservation, and exploitation of audiovisual collections. Let us now take a closer look at each of them.

5.1. Preservation institutions

In Europe, two organizations are leading the way in audiovisual heritage preservation: the *Institut national de l'audiovisuel (INA)* in France and the *Netherlands Institute for Sound & Vision (NISV)*.

5.1.1. The *Institut national de l'audiovisuel*

The *Institut national de l'audiovisuel*, commonly known as *INA*, is a public institution established in 1975 with the aim of preserving the French audiovisual heritage, creating content, conducting research, and transferring knowledge in the audiovisual and digital fields (*INA*, 2023a). *Ina* is responsible for the legal deposit of audiovisual materials and websites in France. As part of its mission, it promotes and facilitates access to the collections it preserves for professionals, researchers and the general public. The *INA* collections include the production of 179 television channels, radio stations, websites, and social media accounts (*INA*, 2023b). From a practical standpoint, this entails handling large amounts of heterogeneous metadata from various sources, which can sometimes be inaccurate or even nonexistent. In this context, AI emerges as a suitable tool to improve the description and ensure the accessibility of these contents. In this regard, several innovation projects have been developed in recent years, such as *NOA* and *Trombinos*.

The *NOA* project applies computer vision and natural language processing techniques to segmenting the broadcasts of a television channel into programs, breaking these news programs into individual news stories and identifying the topics discussed through the analysis of captions (**Martín; Segura**, 2021; **Couteux; Segura**, 2023). computer vision is used for the recognition of title sequences, logos, presenters and credits, thus enabling the identification of the beginning and end of each program and associating it with a title sequence. The presence or absence of the presenter on the screen facilitates the segmentation of the programs into different news stories whereas subtitle analysis helps identify the main topics covered in the news. All the automatically generated information is subsequently manually validated by a professional.

Trombinos is a facial recognition project based on an *IBM* model developed by *INA*. The algorithm has been trained on 62 million faces corresponding to 70,000 individuals. These images have been obtained from both television programs and images retrieved through internet search engines. The content is processed using the model. The results are then returned with a certain level of accuracy and validated either manually or automatically. Each recognized person is associated with an authority record that includes links to external data sources such as *DBpedia*. *INA* is currently working on a less biased model in order to achieve better gender and racial representation (**Petit**, 2022).

In addition, AI-based solutions have played a significant role in the development of *data.ina.fr*, a portal aimed at promoting knowledge of the *Ina* collections through data analytics (**Roche-Dioré**, 2023), where *Trombinos* and *INA Speech Segmenter* provide facial recognition and audio segmentation, *Vocapia* performs speech-to-text transcription and *Textrazor* handles entity recognition. The process of analyzing contents massively through the AI platform has generated large amounts of data that can enable the development of subsequent studies related to the media.

5.1.2. The *Netherlands Institute for Sound and Vision*

The *Netherlands Institute for Sound and Vision* is the institution responsible for preserving the heritage of public media in the Netherlands (*Netherlands Institute for Sound and Vision*, 2023) and making it available to society as a whole. *NISV* early on embraced the use of automatic metadata generation and their integration with external data sources in order to enable the exploitation of its collections. In 2012, *NISV* began using automatic speech-to-text (S2T) transcription te-

chniques, but with a high incidence of errors. Simultaneously, they applied Named Entity Recognition (NERD) solutions, a practice abandoned in 2019 since it did not meet their needs in a real production environment. Also in 2012, *NISV* started developing voice and facial recognition to enable content tagging (**Manders, 2019**).

The current facial recognition mode used at *NISV* relies on an onomastic thesaurus which only includes public personalities in accordance with a well-defined privacy policy that respects the *General data protection regulation (GDPR)* (**Manders, 2022**). The project team defined the expected accuracy levels from the beginning, so only faces identified with an accuracy level above 90% are ingested into the system. Currently, the system has reached a 95% accuracy level, it recognizes 3 out of 4 individuals and is capable of tagging up to 50% of the faces that appear in a television program. Despite the good results, there are concerns about its scalability, its limitations related to the use of the thesaurus as an essential element, the bias that this implies, the difficulty in identifying emerging personalities and the lack of facial models in the dataset, which can lead to false identifications.

Despite adopting all these technologies and having a strong innovation and development area, in 2022, the majority of the metadata managed by *NISV* had not been automatically produced (**Manders, 2022; Manders; Wigham, 2021**).

5.2. Broadcasters

Broadcasters face constant challenges since they have to deal with the strain of immediate production and the urge for a large volume of content from different sources to be made accessible and also ingested into their production systems on a daily basis. European public televisions have gathered and preserved content from different sources dating back to the 1960s following changing cataloging policies. For them, AI represents an opportunity to improve efficiency, increase the reuse of archival assets and avoid repetitive tasks. Let us take a look at some examples.

5.2.1. Yle

Yleisradio Oy (Yle), the Finnish public broadcaster, has been a pioneer in testing AI-based solutions (**Selkälä, 2017**). The application of automatic techniques for metadata generation was considered a way to improve the accessibility and reusability of fully digitized collections, some with insufficient data. To explore the possibilities of AI, *YLE* formed a multidisciplinary team involving the archive staff together with professionals from the editorial, operations, and multimedia departments. This team explored the possibilities of image recognition, scene segmentation, object and face recognition, and optical character recognition (OCR) on current affairs programs, as these were the most commonly used. This pilot project, carried out in 2016, did not provide a more time-efficient analysis and it also showed some limitations, such as the inability to recognize the identity of people appearing in images due to insufficient algorithm training.

That same year, they tested speech-to-text and content classification solutions for the radio archive. The results were good both in transcription and Named Entity Recognition, although they did not delve into speaker segmentation.

This project, far from offering real technological solutions, allowed the *Yle Archive* team to reflect on their actual needs, particularly in terms of defining which parts of the process can be automated and which should be addressed manually. *Yle's* efforts are currently focused on the automatic generation of metadata for content production and web publication, automatic text generation from data and automatic subtitling. The ultimate goal is to have all the content produced by the broadcaster automatically analyzed with a particular emphasis on *Yle Areena*, the video-on-demand platform (**Viljanen, 2022**).

5.2.2. ARD

In 2017, the German consortium of public broadcasters, *ARD*, created a working group with the goal of identifying opportunities and creating use cases for the introduction of AI tools and methods in daily production. From the standpoint of the group, regionality and domain multiplicity are crucial elements in the implementation of AI solutions, also essential in the generation of appropriate metadata for new platforms, different users and content personalization or recommendations (**Wenger-Glemser, 2019**).

Among the broadcasters that are part of *ARD*, the Bavarian television (*Bayerischer Rundfunk* or *BR*) stands out for its use of AI. *BR* has developed a facial recognition model trained with its own data and use cases (**Schreiber, 2022**). In the initial phase of the project, the dataset to train the algorithm focused on scenes showing faces and identifying written signs extracted from two news programs, 30 and 15 minutes long. The result of this process was a demanding data model that required very high accuracy rates to incorporate the extracted metadata into the archive. By February 2022, *BR* had analyzed 55,000 images from the archive and detected 3,000 different classes. Human quality control was considered an essential element to detect false positives and the images responsible for them.

Along the same lines, *BR* has trained a model to identify historical buildings and relevant political and financial centers at a regional level (**Förster, 2023**). The data model, generated automatically from open-source tools, used 255 subtitled programs on which entity recognition and disambiguation processes were applied. This is a proof of concept that aims to be optimized and integrated with other solutions currently in use.

5.2.3. Radio Télévision Suisse (RTS)

Since 2018, *Radio Télévision Suisse (RTS)* has systematically approached the cataloging of their archive by applying AI techniques and developing an interface based on open-source technologies to transcribe audio to text, perform facial

recognition, and automatically classify images during the ingestion process into the archive (Rezzonico, 2020). Automatic classification has been successfully applied to uncategorized sports collections, allowing them to at least detect the topic of each recording. Facial recognition has been developed using a database of 5,000 Swiss public figures taking into account factors such as the duration of the shot and high reliability levels in order to avoid false positives, that is, the identification of one person as a different one because of their common facial traits (Bouchet; Ducret, 2019). Consequently, only faces identified with a reliability level above 85% are integrated into the archive. One of the main features of this tool is that it allows not only for the extraction of metadata but also for image retrieval or visual searching based on faces, scenes, monuments, and buildings. This functionality is particularly relevant for television archives when it comes to locating scenes that have been embargoed due to copyright issues or court orders. The *RTS.ai* interface has enabled the automatic speech to text transcription for a total of 10,000 hours within two years. *RTS* is currently working on the development of speaker recognition tools, the integration of speech-to-text transcription with facial recognition, the assessment of photograph and video aesthetic features and action description (Sonderegger, 2023).

5.2.4. BBC

The *British Broadcasting Corporation (BBC)* is one of the first European public broadcasters that has pondered on the consequences of applying AI to content production. *BBC 4.1* emerged as a project within the Research and Development department of the British public broadcaster with a dual objective: to understand how AI can influence the future of audiovisual production and to collaborate with content creators in order to be prepared for the future. Between September 4th and 5th 2018, the schedule of *BBC Four*, a channel specializing in cultural content, was generated by an algorithm which had been trained to detect the most relevant content for the channel based on programs broadcast in the past (BBC, 2018a) and analyzing program descriptions and topics covered. Out of a total of 270,000 programs, the algorithm identified 150 that it deemed to be the most relevant for broadcasting, which were later used to manually decide on the final schedule. This experience also led to the project “Made by machine: When AI met the archive” (BBC, 2018b), a series in four episodes created using AI to demonstrate how machines think. Techniques such as object and scene recognition, natural language processing (NLP) for subtitle analysis and dynamism were applied using the preselected 150 most relevant programs as basis. The result was broadcast on the channel in 2018 in four micro-programs with Hanna Fry, a British mathematician and popular science communicator, as presenter and contributions from BBC archivists.

Another project involving the use of AI at the BBC is “Starfruit Tagger” (Berger; Armstrong, 2022; *Starfruit Tagger*, 2023). The BBC has very few original recordings of the news bulletins aired during the first 50 years of its radio station’s existence. However, it does have the scripts of these bulletins from 1937 to 1955, which have been digitized and processed using optical character recognition (OCR). A pre-trained automatic tagging system previously trained on sports and news content that had been manually indexed by the BBC’s editorial team has been used to extract names of people, places, organizations and events.

5.2.5. TV2

In 2021, the Norwegian commercial channel *TV2* began conducting tests on automatic speech-to-text transcription and subtitling for current affairs programs and raw material from news and programs (Tverberg, 2021). This project involved the collaboration between journalists and operations staff, who subjectively evaluated the quality of the generated transcription using services such as *Speechmatics (Speechmatics, 2023)*, and *Azure (Microsoft, 2023)*, among others, whose performance was also objectively measured by calculating their Word Error Rate (WER).

When asked about the quality of automatically generated subtitles as opposed to manually created ones, 9.1% of users considered them very good, 10.3% found them good, 21% believed the quality was sufficient, 39.4% thought they were of poor quality and 21.2% rated them as very poor quality. From the developers’ perspective, the main challenges for these types of services are dialects, entity recognition, speaker segmentation, and WER, which, although decreasing over time, can still be significant depending on the context in which the error occurs.

TV2 has also conducted proof-of-concept tests with *CLIP (2023)*, the neural network developed by *OpenAI*, on 5,000 elements from their archive with limited descriptive metadata (Steskal, 2023). The results obtained for person and object recognition were good despite the lack of specific training. However, they noticed that the keyframes returned by the system were not always representative of the entire video. Despite the good performance, they found that the system would not be able to handle complex information searches.

5.2.6. VRT

The Flemish public radio and television broadcaster, *VRT*, has invested in the development of its own AI models for the automatic scene segmentation of the content published on its website and the enrichment of archive metadata (Daniels; Degryse, 2021). The application of these solutions aims to improve efficiency, increase reusability, enable recommendations, and avoid repetitive tasks for the editing, multimedia, and archival teams. The project has been carried out in two phases: an initial preprocessing and training phase and an application phase.

During the first phase, the most relevant sources of information to train the segmentation algorithm were identified, which included information from broadcasts, subtitles, schedules, facial recognition and RGB (Red, Green and Blue) values. All the data gathered were used to create a vector representation and define scene transitions by measuring RGB

values. Finally, pairs of similar and dissimilar scenes were established. Once the algorithm was trained, in the application phase, program information was extracted and the model was applied. The results were validated by humans (**Daniels; Degryse, 2021**).

VRT also utilizes AI in production to enrich metadata through optical character recognition (OCR) applied to the lower third of on-screen images, where they manage to detect the speakers on screen. The information obtained through OCR is processed using natural language processing (NLP) to extract entities and keywords on which filters are applied to discard irrelevant information (**Daniels, 2023**).

5.2.7. SVT

Sveriges Television, the Swedish public television, known as *SVT*, has conducted several proof-of-concept tests in its production archive, *Mark*, which is also used as a platform to foster innovation within the company (**Åstrand; Ståhl, 2023**). These tests address the absence of metadata in a significant percentage of content and utilize AI techniques to make it accessible. Specifically, they have conducted image search tests using *CLIP* (facial recognition for local politicians), OCR to identify program creators and technical teams in credits, as well as automatic speech to text transcription and entity recognition to identify key individuals and topics discussed. The medium-term goal is to identify real use cases and bring them into production, incorporating only high-reliability metadata.

5.2.8. Asharq News

Asharq News (**Batrick, 2022; Batrick; Petitpont, 2022**), the multi-platform Arabic news network founded in November 2020, is presumably one of the few examples in the world where the use of AI was included as part of the network's initial development plan. The project, developed over 18 months, involved integrating their *AVID Media Central* production system with *Newsbridge* technology, aiming to generate metadata for 1,600 hours of monthly broadcast and original content both in English and Arabic. The proof-of-concept phase, prior to implementation, involved the participation of various user groups who not only contributed to the training of the facial recognition and automatic transcription and translation models, but also identified different use cases and defined the required accuracy rates for integration. The complexity of managing consistent metadata in English and Arabic, the challenge of training an automatic speech transcription model for a language with numerous dialects as Arabic, the network's specialization in political and military topics and the need to develop facial recognition models for Arabian prominent people earned *Asharq* international recognition with the *FIAT/IFTA Media Management Award* in 2022 (*FIAT/IFTA, 2022*).

5.2.9. IRIB

In the same context, it is worth noting the centralized data model and intensive use of AI by the Iranian television network, *IRIB*. The difficulty of finding data to train their own models has led this broadcaster to establish a data factory to develop projects for image and text labeling, verification or automatic pre-processing and data validation. These processes are then applied in both the production systems and the archive (**Ghanbari, 2022**).

This factory allows *IRIB* to develop its own tools to power various AI-based services for their archive, television channels and website, as well as for the administration and finance departments, which also leverage data generated by other areas. These tools are applied for text summarization and entity detection, video annotation and image summarization, speech-to-text transcription, text-to-speech conversion, gender detection, and speaker identification.

In Europe, *France Télévisions* and the *Italian Public Broadcasting Corporation, RAI*, have also embraced similar approaches.

5.2.10. France Television

DAIA is the data governance department at *France Télévision*, whose aim is to ensure the interoperability and availability of data for the various departments within the company. To achieve this, it has a knowledge interface that analyzes the data, shares them across different datasets and translates them when they lack coherence, in other words, it converts them into interoperable data. The foundation of this system is a common ontology that enables these datasets to mutually understand one another. Additionally, when the existing data is insufficient, open-source AI is applied to generate data based on the credits or the content itself (**Parmentier, 2021**).

5.2.11. Radiotelevisione Italiana (RAI)

In the case of *RAI*, the generation of datasets for machine learning is seen as a key element for the integration of AI into the workflows of the Italian public broadcasting company (**Messina, 2021**). In this regard, the metadata generated by the archive and the broadcasting and production areas, to name just a few, undergo a process of extraction, filtering, and adaptation to render them useful for model training. The *RAI Media Cognitive Service Platform* is the tool that enables data ingestion, content annotation using cloud solutions or proprietary models, data validation and enrichment, and the creation of refined collections suitable for the development of models that will be put into production. This means that *RAI* is independent from third parties in its AI-based development and is capable of generating and applying its own data models specifically adapted to its needs (**Messina; Montagnuolo, 2023**).

5.2.12. *Radio Televisión Española (RTVE)*

In Spain, *RTVE* and *Atresmedia* have been pioneers in the implementation of AI in their archives. In the case of *RTVE*, the first approaches were carried out through the creation of the *RTVE University of Zaragoza Chair (Cátedra RTVE Universidad de Zaragoza, 2017)*. As part of the activities of this chair, the *RTVE Database* was first published in 2018. It consists of annotated datasets from TV programs and serves as the basis for the *RTVE Albayzin Challenges (Lleida-Solano et al., 2022)*. These challenges, which bring together national and international research groups, have allowed for the testing of state-of-the-art systems in automatic transcription and multimodal recognition with use cases prepared by the *RTVE* archive.

In the field of production, the tender for automatic metadata generation in the *RTVE* archive (*RTVE, 2021*) was awarded in 2021. This cloud-based service, through a technological integrator, enables automatic metadata generation for audio and video content spanning 11,000 hours of *RTVE* content. This service will be replaced in October 2023 by a new one that includes new functionalities such as automatic analysis and translation from Catalan, and where image analysis becomes more relevant (*RTVE, 2023*).

Furthermore, *RTVE* is working on a similar project for the *Radio Nacional de España archive (RNE)*, which will enable transcription, automatic classification, and entity extraction for 190 hours of *Radio 1* and *Radio 5*.

5.2.13. *Atresmedia*

In 2019, *Atresmedia* launched a supervised automatic cataloging project. Through this on-premise service, the *Atresmedia* archive obtains transcription for 40 hours of daily content, including raw material, news items and fully subtitled programs. The generated metadata is integrated into the MAM (media asset management) system, where it is corrected and complemented with human cataloging. This project, which aims to transform the professional profile of documentalists from processors to content generators, was awarded, among others, with the *Excellence in Media Management Award* from *FIAT/IFTA* in 2021 (*López-de-Quintana, 2021; López-de-Quintana; León-Carpio, 2021*).

Other regional television broadcasters in Spain, such as *Aragón TV (Aragón Noticias, 2021)* and *Televisió de Catalunya CCM*, have carried out pilots of automated cataloging, with the latter incorporating automatic transcription in Catalan into its MAM system.

5.3. Press agencies

5.3.1. *Associated Press*

In the world of media, news agencies have not remained on the sidelines of technological progress. The *Associated Press (AP)*, after a pilot and an 8-month development period, has integrated multimodal AI in collaboration with the Belgian company *Limecraft* in order to analyze both live feeds and recorded material which amount to approximately 700 video clips per month (*Coppejans, 2021*).

Multimodal technology integrates computer vision for scene identification and segmentation, facial and identity recognition, and role and attitude recognition on the one hand, with speech technologies that enable language detection and automatic transcription on the other. The results are integrated into an interface and displayed with the appropriate level of detail for *AP* (*Verwaest, 2022*).

5.3.2. *Reuters*

In 2020, *Reuters* announced the application of AI techniques to one million clips from its archive spanning from 1986 to the present (*Reuters Staff, 2020*). This project, funded by the *Google DNI Fund*, has enabled automatic speech-to-text transcription, translation into 11 different languages and recognition of internationally prominent people. It has also allowed *Reuters* to gain a better understanding of its archive, the analysis policies applied in the past and their effect on their content accessibility (*Reuters, 2023*). In this way, they have been able to improve their internal technological capacity and determine which types of content can be automatically analyzed with a high level of reliability and which cannot.

6. Joining forces: innovation projects

As we have seen, many radio and television companies have performed pilots with internal resources or implemented various AI-based solutions. However, in the last five years, we have also witnessed a significant rise in innovation projects where companies and media organizations join forces to advance the application of AI in their workflows.

6.1. *VIVA*

The *VIVA* project has led to the development of a tool to implement video retrieval methods based on deep learning models (*Mühling et al., 2022*). Researchers from the *TIB - Leibniz Information Centre for Science and Technology* and the *University of Marburg (Germany)* have participated in the project together with *ARD* professionals. The objective is to enable concept-based or personality-based video retrieval in media archives and continuously update the deep learning model as new emerging personalities or needs arise. The tool has been tested on four use cases within the context of a collection of historical videos from the *German Broadcasting Archive*, which consists of approximately 34,000 hours of television recordings from the former German Democratic Republic.

6.2. *Europeana Subtitled*

Another relevant project in this field is *Europeana Subtitled*. This initiative has brought together a consortium of 7 European public television broadcasters, the *Fondazione Bruno Kessler (FBK)* and *Translated* (Italy) with the aim of developing automatic speech-to-text transcription, translation, and automatic subtitling models in order to improve the accessibility of audiovisual content in collections such as *Europeana* (Lewis; Jarret, 2023). This project has made it possible to upload 8,000 English-subtitled videos focusing on the topic “Broadcasting Europe” onto the European digital library, as a means to showcase the social changes that have taken place in Europe since the 1930s.

6.3. *AI4Media*

The previously mentioned *AI4Media* project aims to develop innovative tools to address the current challenges in the media sector (*AI4Media*, n.d.). To achieve this, it has defined seven industrial use cases that range from social media and misinformation to supporting newsrooms through automatic news creation. One notable use case is the application of computer vision techniques and automatic metadata generation to archival material in order to support news coverage of unexpected events, where immediacy and quality do make a difference. The project involves 9 universities, 9 research centers and 12 companies in the sector, including heritage preservation organizations, radio and television companies and vendors.

6.4. *Tailored Media*

Tailored Media is an innovation project led by *Joanneum Research* in collaboration with *ORF* and *Austrian Mediatheque* (Bailer; Bauer; Rottermann, 2021) whose objective is to automatically generate relevant metadata using computer vision and natural language processing techniques, integrate them into current workflows and develop user-oriented processes based on these metadata.

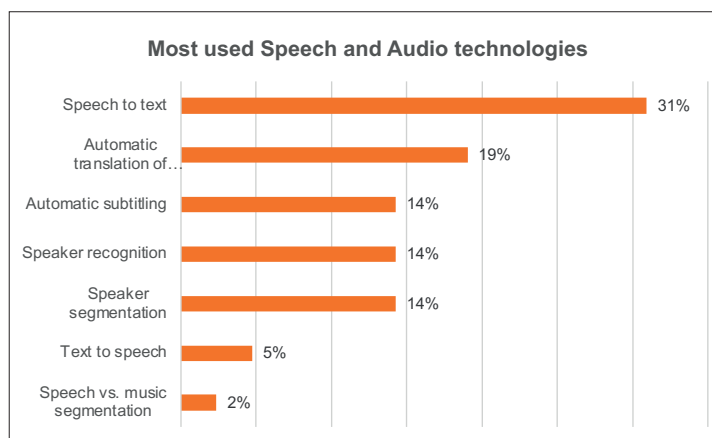
7. Results of the *FIAT/IFTA* survey on the use of AI in media archives

In April 2023, *FIAT/IFTA* conducted a survey on the use of AI in media archives (*FIAT/IFTA*, 2023). This survey aimed to provide an overview of the degree of implementation of AI. More specifically, it sought to understand the technologies and applications being used, the level of their implementation, how metadata is being integrated into archive management systems and the expected future evolution of content cataloging.

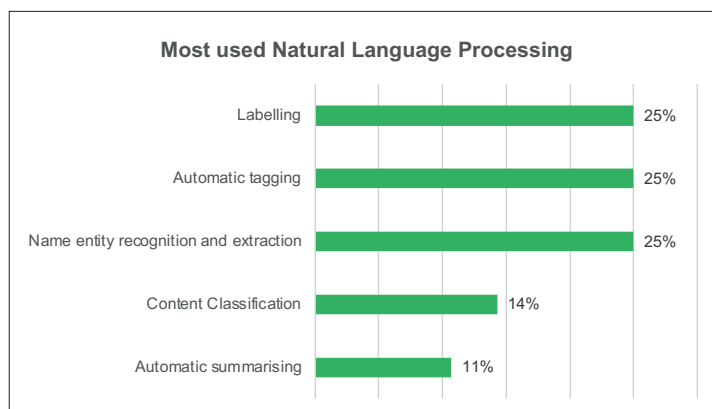
A total of 54 organizations participated in the survey, primarily television networks and national or regional audiovisual archives in Europe. Out of the 54 respondents, 61% (33 organizations) stated that they are currently applying AI in their archives, 9% (5 organizations) reported using it exclusively in the production archive and 28% (15 organizations) have plans to use AI in the future, while only 2% (1 organization) do not anticipate using it in the future.

Based on the responses obtained, it can be stated that organizations are using or planning to use a combination of different technologies with a slight preference for audio and speech technologies (37%, 33) in the first place, followed by computer vision (34%, 31) and finally natural language processing (29%, 26). If we break down the data by technologies in production or in the planning phase, the trend remains the same for those organizations planning to integrate AI in the future, while there is a higher use of computer vision (27) as opposed to audio and speech technologies (25) and natural language processing (20) among the organizations who are already applying AI.

Graph 1 displays the degree of use of specific applications related to speech and audio technologies. Graph 2 showcases the specific applications related to natural language processing. Graph 3 exhibits the percentage of usage for applications related to computer vision.



Graph 1. Use of speech and audio technology applications (%).
Source: *FIAT/IFTA* (2023).



Graph 2. Use of natural language processing applications (%).
Source: *FIAT/IFTA* (2023).

In a significant number of cases, these technologies are fully integrated into the production process, while in a smaller percentage, they are used in upcoming innovation projects or proof of concept to understand the scope and limitations of the technology.

Another relevant factor highlighted in this survey is the origin of the technologies currently being implemented by media archives. 26% (13) use open-source solutions implemented by their own organization, another 26% (13) use third-party technologies (such as *Azure*, *Amazon*, *IBM*, etc.) implemented by their own organization, 26% (13) rely on third-party services that integrate proprietary technologies, and 23% (11) also rely on third parties that in turn integrate technologies from others.

Regarding how data is presented to the end user in archive management systems, 57% (24) do not indicate the source of the metadata, 12% (5) indicate it in pop-up windows, and 31% (13) indicate it on the same data display screen. Finally, the majority of organizations that responded to the survey believe that manual cataloging will decrease by 25% to 50% in the next 5 years.

8. Conclusions

Since 2007, media archive professionals have been contemplating the application of AI-related technologies and the future evolution of their work, particularly in activities such as content annotation. The initial AI projects have been driven by the innovation departments within companies and developed by multidisciplinary teams. In many cases, start-ups, research groups or specialized companies have been involved, such as *Yle*, *RTVE*, and *RSI*. Some organizations have opted for customized developments based on open-source solutions, especially from 2019 onwards.

The main objective of applying AI is to improve process efficiency, increase content reusability, enable website recommendations and avoid redundant tasks performed by different user groups that generate the same information about the same content at different stages of the production chain (*VRT*, *IRIB*, *RAI*, *France TV*). Speech and audio technologies can be considered widely implemented, although there are challenges to overcome, particularly those related to the lack of datasets to develop language models in low-resource languages. The development of models based on well-structured proprietary data coexists with the use of technologies provided by third parties, whether they are proprietary or commercial technologies adapted to the specific needs of broadcasters.

The application of computer vision techniques currently focuses on facial recognition in controlled collections, aiming to identify public figures of interest, particularly in regional contexts, where the solutions provided by major technology companies may not be sufficient. In these cases, models with high accuracy rates are employed to prevent false positives, while strict data protection policies are implemented to avoid using images of children or non-public individuals (*NISV*, *RSI*, *BR*).

The use of AI to automatically generate content from archival material is currently anecdotal or limited in scope.

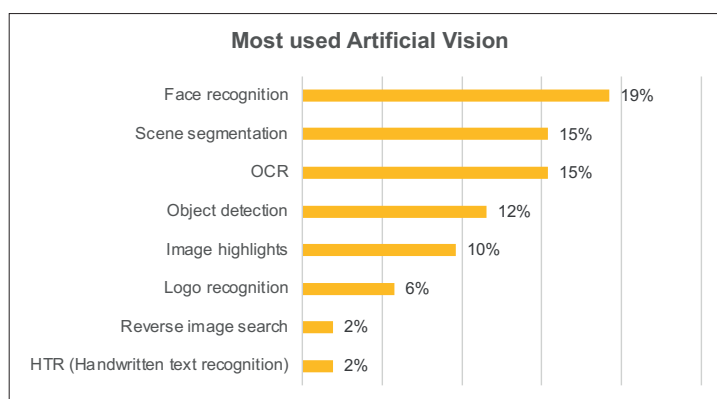
Lastly, the application of AI in archives poses significant challenges in terms of scalability, integration with production and archival systems, and the restructuring of work for documentalists in both analysis and user support units. Additionally, the integration of metadata into archival management systems is a complex task that has a considerable impact on searches and selection results for both expert users and end users who increasingly rely on these tools without professional intermediaries.

In conclusion, the use of AI in media archives presents significant opportunities to enhance accessibility and new content productions based on archival material. However, it also raises important issues regarding its integration into workflows, data quality and reliability. As its application expands, conducting further research and fostering collaboration among professionals from various disciplines will be crucial to harness the full potential of AI while ensuring its responsible integration.

9. References

AI4Media (2023). *The AI4Media project*.
<https://www.ai4media.eu>

Aragón Noticias (2021). "Aragón TV participa en la prueba piloto de una herramienta de inteligencia artificial". *Aragón noticias*, 13 noviembre.
<https://www.cartv.es/aragonnoticias/sociedad/aragon-tv-participa-en-la-prueba-piloto-de-una-herramienta-de-inteligencia-artificial-6130>



Graph 3. Use of applications related to computer vision (%).
 Source: FIAT/IFTA (2023).

- Åstrand, Mikaela; Ståhl, Sally** (2023). "Finding without tagging: AI experiments for improved findability in the new media archive at SVT". In: *EBU Data technology seminar 2023*.
<https://tech.ebu.ch/publications/ai-experiments-for-improved-findability-in-the-svt-media-archive>
- Bailer, Wermer; Bauer, Christoph; Rottermann, Gernot** (2021). "Analysing user needs for automatic metadata creation and advanced search in the Tailored Media Project". In: *EBU MDN workshop 2021*.
<https://tech.ebu.ch/publications/analysing-user-needs-for-automatic-metadata-creation-and-advanced-search-in-the-tailored-media-project>
- Batrick, Kathey** (2022). "Case study: AI indexing. Assessing and selecting AI services". In: *FIAT/IFTA world conference: Archive out of the box!*
<https://fiatiftaworldconference2022.sched.com/event/149Cd>
- Batrick, Kathey; Petitpont, Frederic** (2022). "How Asharq News is leveraging multimodal AI indexing with Newsbridge: Inside the back-end, operational & technical workflows of a multilingual news channel". In: *FIAT/IFTA world conference: Archive out of the box!*
<https://fiatifta.org/world-conference-2022-recordings-available>
- BBC (2018a). *AI TV on BBC 4.1: BBC Four is giving an AI control of your TV for two nights*.
<https://www.bbc.co.uk/programmes/p06jrfcc>
- BBC (2018b). *Made by machine: When AI met the archive*. BBC Four.
<https://www.bbc.co.uk/programmes/b0bhwk3p>
- Berger, Jake; Armstrong, Andy** (2022). "BBC radio news scripts 1937-1995: Using an automated tagger to enable journeys across time and space". In: *EBU MDN workshop 2022*.
<https://tech.ebu.ch/publications/bbc-radio-news-scripts-1937--1995-using-an-automated-tagger-to-enable-journeys-across-time-and-space>
- Bouchet, Leonard; Ducret, Sebastien** (2019). "A visual feature extraction pipeline and its applications for radio television Suisse". In: *Proceedings of the FIAT/IFTA media management seminars. Changing sceneries, changing roles, part IX*, pp. 51-59.
<https://fiatifta.org/library/proceedings-of-the-fiat-ifta-media-management-seminars>
- Bruccoleri, Angelo; Iacoviello, Roberto; Messina, Alberto; Metta, Sabino; Montagnuolo, Maurizio; Negro, Fulvio** (2022). *AI in vision: High quality video production & content automation*. RAI. Radiotelevisione Italiana. Centre for Research, Technological Innovation and Experimentation. AI4media.
<https://www.ai4media.eu/whitepapers/ai-in-vision-high-quality-video-production-content-automation>
- Cátedra RTVE Universidad de Zaragoza (2017).
<https://catedrartve.unizar.es>
- Coppejans, Charlotte** (2021). *How artificial intelligence slashes editing time for Associated Press (AP)*.
<https://www.limecraft.com/how-the-associated-press-is-using-ai-to-create-automated-shot-lists>
- Couteux, Anne; Segura, Olivio** (2023). "News channel automatic segmentation". In: *FIAT/IFTA media management seminar 2023*.
<https://fiatifta.org/seminar/media-management-seminar-2023>
- Daniels, Marijin** (2023). "Metadata enrichment using lower third character recognition". In: *FIAT/IFTA media management seminar 2023*.
<https://fiatifta.org/seminar/media-management-seminar-2023>
- Daniels, Marijin; Degryse, Jasper** (2021). "How VRT automated the segmentation of programmes with AI". In: *EBU MDN workshop 2021*.
<https://tech.ebu.ch/publications/how-vrt-automated-the-segmentation-of-programs-with-ai>
- FIAT/IFTA (2022). *Archive Achievement Awards 2022*.
<https://fiatifta.org/awards-2022>
- FIAT/IFTA (2023). *AI Survey. The results are now available*.
<https://fiatifta.org/ai-survey-results>
- FIAT/IFTA Media Management Commission (2013). *Proceedings of the FIAT/IFTA media management seminars. Changing sceneries, changing roles. Part VI. Metadata as the cornerstone of digital archiving*.
<https://fiatifta.org/library/proceedings-of-the-fiat-ifta-media-management-seminars>
- FIAT/IFTA Media Management Commission (2017). *Proceedings of the FIAT/IFTA media management seminars. Changing sceneries, changing roles. Part VIII. Embracing automation – enhancing discoverability*.
<https://fiatifta.org/library/proceedings-of-the-fiat-ifta-media-management-seminars>

FIAT/IFTA Media Management Commission (2019). *Proceedings of the FIAT/IFTA media management seminars. Changing sceneries, changing roles. Part IX. Game changers? From automanation to curation - Futureproofing AV content.* <https://fiatifta.org/library/proceedings-of-the-fiat-ifta-media-management-seminars>

FIAT/IFTA Media Management Commission (2023). *MMC Seminar. Download the slides.* <https://fiatifta.org/mmc-seminar-2023-download-the-slides>

Förster, Constantin (2023). "Generation of training data for landmark recognition in videos using named entity recognition". In: *EBU data technology seminar 2023.* <https://tech.ebu.ch/publications/landmarkner--generating-training-data-for-landmark-recognition>

Ghanbari, Shirin (2022). "Strategy and challenges for developing AI applications". In: *EBU AIM community meeting 2022.*

Green, Eva-Lis; Gupta, Jacqui (2019). "20 year of MMC seminars: changing sceneries, changing roles 1998-2019". In: *Proceedings of the FIAT/IFTA media management seminars. Changing sceneries, changing roles, part IX*, pp. 7-17. <https://fiatifta.org/library/proceedings-of-the-fiat-ifta-media-management-seminars>

INA (2023a). *International affairs.* Institut National de l'Audiovisuel. <https://www.ina.fr/institut-national-audiovisuel/international-affairs>

INA (2023b). *Collection, preservation, and documentation of audiovisual heritage.* Institut National de l'Audiovisuel. <https://www.ina.fr/institut-national-audiovisuel/collection-preservation-and-documentation-of-audiovisual-heritage>

Lewis, Michelle; Jarret, Nicholas (2023). *Explore how Europeana Subtitled increased access to audiovisual heritage.* <https://pro.europeana.eu/post/explore-how-europeana-subtitled-increased-access-to-audiovisual-heritage>

Lleida-Solano, Eduardo; Ortega-Giménez, Alfonso; Miguel, Antonio; Bazán-Gil, Virginia; Pérez-Cernuda, Carmen; De-Prada, Alberto (2022). *RTVE 2018, 2020 and 2022 database description.* <http://catedrartve.unizar.es/reto2022/RTVE2022DB.pdf>

López-de-Quintana, Eugenio (2021). *AI algorithms for media cataloguing in Atresmedia Group.* <https://fiatifta.org/awards-2021>

López-de-Quintana, Eugenio; León-Carpio, Antonio (2021). "Artificial intelligence for a role change in television archives: The Atresmedia–Etiqmedia experience". *Journal of digital media management*, v. 10 n. 2, pp. 177-187. http://etiqmedia.com/ficheros/JDMM_10_2_JDMM0006_e_lopez_de_quintana_and_carpio-1.pdf

Manders, Tim (2019). "Harder, better, faster, stronger: Adding face recognition in the mix at NISV". In: *FIAT/IFTA world conference: AV archives in the all-media world.* <https://fiatifta2019.sched.com/event/S0Sa/harder-better-faster-stronger-adding-face-recognitionin-the-mix-at-nisv>

Manders, Tim (2022). "Face recognition at the Netherlands Institute for Sound and Vision. Is it really harder, better, faster, stronger". In: *FIAT/IFTA media management seminar.* <https://fiatifta.org/seminar/media-management-webinars-2022-2>

Manders, Tim; Wigham, Marie (2021). "More metadata, lots of links, but what do you do with them? Practical examples of the added value of automatic metadata and linked data for archive users". In: *2021 FIAT/IFTA world conference. Advancing the digital dividend.*

Martin, Camille; Segura, Olivio (2021). "Using AI tools to segment and describe broadcast live stream. In: *EBU MDN workshop 2021.* <https://tech.ebu.ch/publications/using-ai-tools-to-segment-and-describe-broadcast-livestream>

Messina, Alberto (2021). "Enabling AI with dataset engineering". In: *EBU MDN workshop 2021.* <https://tech.ebu.ch/publications/enabling-ai-with-dataset-engineering>

Messina, Alberto; Montagnuolo, Maurizio (2023). "Integrating open knowledge-bases and AI-tools". In: *EBU data technology seminar 2023.* <https://tech.ebu.ch/publications/integrating-open-knowledge-bases-and-ai-tools>

Microsoft (2023). *Azure cognitive services documentation.* <https://learn.microsoft.com/en-us/azure/cognitive-services>

Mühling, Markus; Korfhage, Nikolaus; Pustu-Iren, Kader; Bars, Joanna; Knapp, Mario; Bellafkir, Hicham; Vogelbacher, Markus; Schneider, Daniel; Hörth, Angelika; Ewerth, Ralph; Freisleben, Bernd (2022). "VIVA: visual information retrieval in video archives". *International journal on digital libraries*, v. 23, n. 4, pp. 319-333. <https://doi.org/10.1007/s00799-022-00337-y>

Netherlands Institute for Sound and Vision (2023). *Mission and vision.* <https://www.beeldengeluid.nl/en/about/mission-and-vision>

Open AI (2022). *Introducing ChatGPT*.

<https://openai.com/blog/chatgpt>

Open AI (2023). *Dalle-E 2*.

<https://openai.com/dall-e-2>

Parmentier, Matthieu (2021). "The challenge of data management and governance". In: *EBU MDN workshop 2021*.

<https://tech.ebu.ch/publications/governance-of-data-at-ftv-for-better-usage-and-valorisation-of-the-programs>

Petit, Thomas (2022). "Facial recognition: Trombinos". En: *FIAT/IFTA media management seminar*.

<https://www.youtube.com/watch?v=qM10EXb6988>

Reuters (2023). *Automated transcription, translation and identification*.

<https://liaison.reuters.com/help/reutersconnect#automated-transcription-translation-and-identification>

Reuters Staff (2020). *Reuters applies AI technology to 100 years of archive video to enable faster discovery, supported by Google DNI*.

<https://www.reuters.com/article/rpb-lavita-video-archive/reuters-applies-ai-technology-to-100years-of-archive-video-to-enable-faster-discovery-supported-by-google-dni-idUSKCN2591VO>

Rezzonico, Pietro (2020). "AI-based tools to help archive workflows". In: *EBU tech roundtable on archives 2020*.

https://tech.ebu.ch/publications/presentations/roundtable_2020_archive_workflows/RTS-archive_tools

Roche-Dioré, Axel (2023). "data.ina.fr a portal to promote media analytics". In: *EBU data technology seminar 2023*.

<https://tech.ebu.ch/publications/datainafr--a-portal-to-promote-media-analytics>

RTVE (2021). *Metadatado automático de contenidos del Fondo Documental de RTVE*.

<https://licitaciones.rtve.es/licitacion/licitaciones/detalle?id=1208797>

RTVE (2023). *Metadatado automático de contenidos del Fondo Documental RTVE*

<https://licitaciones.rtve.es/licitacion/licitaciones/detalle?id=2200005>

Sánchez-García, Pilar; Merayo-Álvarez, Noemí; Calvo-Barbero, Carla; Diez-Gracia, Alba (2023). "Spanish technological development of artificial intelligence applied to journalism: companies and tools for documentation, production and distribution of information". *Profesional de la información*, v. 32, n. 2.

<https://doi.org/10.3145/epi.2023.mar.08>

Schreiber, Jonas (2022). "AI going local - conception and training of specialized Bavarian' AI models at Bayerischer Rundfunk". In: *EBU AI-AME 2022*.

Selkälä, Elina (2017). "Heading for AI". In: *Proceedings of the FIAT/IFTA media management seminars. Changing sceneries, changing roles, part VIII. Embracing automation - Enhancing discoverability*, pp. 169-189.

<https://fiatifta.org/library/proceedings-of-the-fiat-ifta-media-management-seminars>

Sonderegger, Janique (2023). "Speaker recognition service: Roles and expectations". In: *FIAT/IFTA media management seminar*.

<https://fiatifta.org/seminar/media-management-seminar-2023>

Speechmatics (2023). *The most accurate speech to text API*.

<https://www.speechmatics.com>

Startfruit Tagger (2023). *BBC Startfruit tagger*.

<http://starfruit.virt.ch.bbc.co.uk/#>

Steskal, Lubos (2023). "Using free-text queries to perform visual search in video archives". In: *EBU data technology seminar 2023*.

<https://tech.ebu.ch/publications/using-free-text-queries-to-perform-visual-search-in-video-archive>

Tverberg, Are (2021). "TV 2 Norway's AI hub and QA Monitor: A gateway to ASR services for a low-resource language". In: *EBU MDN workshop 2021*.

<https://tech.ebu.ch/publications/tv-2-ai-hub-and-qa-monitor--a-gateway-to-asr-services-for-a-low-resource-language>

Verwaest, Maarten (2022). "Using AI for automatic shot listing in news at Associated Press". In: *EBU MDN workshop 2022*.

<https://tech.ebu.ch/publications/using-ai-for-automatic-shot-listing-in-news--the-associated-press>

Viljanen, Kim (2022). "Content tagging in Yle Areena: Improving quality and process". In: *EBU MDN workshop, 2022*.

<https://tech.ebu.ch/publications/content-tagging-in-yle-areena---improving-quality-and-process>

Wenger-Glemser, Gabriele (2019). "Shaping the future Artificial Intelligence in AV archives". In: *Proceedings of the FIAT/IFTA media management seminars. Changing sceneries, changing roles, part IX*. pp. 61-78.

<https://fiatifta.org/library/proceedings-of-the-fiat-ifta-media-management-seminars>

ChatGPT: Stream of opinion in five newspapers in the first 100 days since its launch

Cristian González-Arias; Xosé López-García

Nota: Este artículo se puede leer en español en:
<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87354>

Recommended citation:

González-Arias, Cristian; López-García, Xosé (2023). "ChatGPT: Stream of opinion in five newspapers in the first 100 days since its launch". *Profesional de la información*, v. 32, n. 5, e320524.

<https://doi.org/10.3145/epi.2023.sep.24>

Article received on May 8th 2023
Approved on September 11th 2023



Cristian González-Arias ✉
<https://orcid.org/0000-0001-7689-6948>
Universidad de Santiago de Compostela
Santiago de Compostela, Spain
Pontificia Universidad Católica de
Valparaíso, Chile
cristian.gonzalez@pucv.cl



Xosé López-García
<https://orcid.org/0000-0002-1873-8260>
Universidad de Santiago de Compostela,
Facultad de Ciencias de la Comunicación
Avda. Castelao, s/n
15782 Santiago de Compostela, Spain
xose.lopez.garcia@usc.gal

Abstract

In the last decade, journalism has progressively incorporated artificial intelligence (AI) into several processes: information analysis, content extraction, audience research, and automated copywriting, among others. *ChatGPT*, which has a great capacity for interacting with people through natural language and providing a response to almost any topic, focuses on content creation. This significant technological advancement rekindles the debate on whether machines will replace humans, including journalists. Focusing on the case of the Spanish press and using the framework of studying media participation in the public debate, we are interested in the press's reaction to *ChatGPT*'s launch. Specifically, we asked the following questions: What were the issues that dominated the debate, and what voices were called upon to express their opinions? The stream of opinion on this issue was analyzed from a communication studies and discourse analysis perspective, starting with the identification of opinion statements expressed in articles of various journalistic genres conveyed by the press during the first 100 days since the launch of *ChatGPT* on November 30, 2022. We worked with 176 press articles that addressed the subject in five Spanish generalist newspapers. The results showed that the flow of opinion developed from 8 subtopics and 11 groups of voices. The prevailing opinion during this period was that *ChatGPT* is an extraordinary technological milestone, even if it makes mistakes that reveal the technology's immaturity. The main shortcomings identified were the inability to distinguish between what is true and what is false, its tendency to function as a black box, and its failure to account for the sources it uses. However, owing to the business potential that it heralds, it is clear that a real war for the dominance of AI has broken out, which makes it necessary to put regulations in place to reduce the risks of malicious use.

Keywords

Public debate; *ChatGPT*; Stream of opinion; Digital journalism; Artificial intelligence; AI; Newspapers; Press articles.

Funding

This article has been carried out within the framework of the *Requalification Program of the Spanish University System for 2022-2023, Ayuda María Zambrano*, funded by the European Union *NextGeneration EU*, and is the result of part of the context studies of the R&D+i project *Digital native media in Spain: strategies, skills, social involvement and (re)definition of journalistic production and dissemination practices* (PID2021-122534OB-C21), funded by the *Ministry of Science and Innovation of Spain (MCIN) / State Agency of Research (AEI) / 10.13039/501100011033/ and the European Regional Development Fund (ERDF) "A way of making Europe"*.



1. Introduction

On November 30, 2022, the company *OpenAI* launched the Chat Generative Pre-trained Transformer (*ChatGPT*), an artificial intelligence (AI) chatbot prototype that was astounding owing to its ability to mimic human language and create text with great precision and detail. In the last decade, AI has been influencing broad sectors of society. In part in recognition of this, the *FundéuRAE* in Spain, sponsored by the *EFE Agency* and the *Royal Spanish Academy (RAE)*, determined that “artificial intelligence” should be the word of the year owing to the multitude of such applications that are having an impact on society (*FundéuRAE*, 2022). In this context, *ChatGPT* has powerfully reignited the debate around technology’s ability to replace humans in various tasks, including journalism.

The transformations that journalism has undergone during the last 40 years are extremely significant: first, the changes stemming from the emergence of the Internet and digitization; then, broadband and the development of social networks; and more recently, the wide range of possibilities that the application of AI is opening up. The integration of AI tools is already a reality in plenty of newsrooms (**Linden**, 2017). Currently, AI is involved in the journalistic production process in a variety of tasks: content extraction, identification of news events, audience research, and copywriting automatically or assisted by software (**Diakopoulos**, 2019). In several European countries as well as Japan, there are already news media outlets that publish text and videos that are entirely machine-generated, the so-called synthetic media (**Ufarte-Ruiz; Murcia-Verdú; Tüñez-López**, 2023). However, studies have also found that the opportunities to interact with AI tools are still limited (**Sánchez-Gonzales; Sánchez-González**, 2020).

In recent years, there have been reviews of scientific publications in the area of AI, showing its contributions and advances in the fields of communication and journalism (**Calvo-Rubio, Ufarte-Ruiz**, 2021; **Parrat-Fernández et al.**, 2021) or in other fields, such as education and social services (**Incio-Flores et al.**, 2022; **Alvarado-Salazar; Llerena-Izquierdo**, 2022; **Minguijón; Serrano-Martínez**, 2022; **Thomas et al.**, 2023; **Dogan; Goru-Dogan; Bozkurt**, 2023) and physics (**Gurrola; Ramírez-Reyes; Mora-Gutiérrez**, 2020), to name but a few. AI has also begun to transform traditional research practices in many areas (**Wagner et al.**, 2022).

A recent study by **Sánchez-García et al.** (2023) concluded that the incorporation of AI is concentrated primarily in distribution and the relationship with the audience, secondarily in the collection of information, and to a more limited extent, in the automated production of news. Thus, for some, AI integration cannot replace journalists, but it will allow them to dive deeper into the information provided and strengthen their connection with the audience (**Marconi**, 2020); however, for others, attention should be paid to the negative effects of incorporating it into the media:

“The inclusion of these ‘machines’ into the newsrooms has initiated various debates ranging from aspects linked to journalistic quality and the reliability of the news generated to deontological issues and reflections on the precariousness of the sector” (**Tejedor et al.**, 2022, p. 2).

Technology companies agree on three ideas about Spanish media’s interest in implementing AI: “slowness”, “distrust”, and “lack of knowledge” (**Sánchez-García et al.**, 2023). In this regard, these authors point out:

“The technology companies consulted themselves confirm that, in comparison with other sectors more advanced in AI, the media sector reflects ‘slowness,’ ‘distrust,’ and ‘lack of knowledge,’ which they attribute to issues of financial difficulty, fear of labor restructuring, and short-sightedness regarding profitability (**Sánchez-García et al.**, 2023, p. 13).

Other research has found that journalists have a profound lack of knowledge regarding AI’s direct influence on their profession (**Tüñez-López; Toural-Bran; Cacheiro-Requeijo**, 2018).

In this context and within the framework of studying the press’s meaningful participation in the public debate through the opinions it conveys, we are interested in determining how the Spanish press has reacted to the launch of *ChatGPT*. Specifically, we ask the following questions: What were the issues that dominated this debate, and which voices were called upon to express their opinions? Addressing these aspects will allow us to understand the strategies and discourse that mobilized the newspapers in response to a disruptive news event in what could be the first moment of a debate about a global phenomenon that encompasses broad spheres of social life.

From a communication studies and discourse analysis perspective, we focus on texts as material objects produced in a given social and historical context. We are interested in revealing the opinions that expressly circulate in newspapers, which has been called the stream of opinion (**González-Arias; Campos-Rojas**, 2020; **González-Arias; Álvarez; Bustamante**, 2022). Starting from the idea that expressing an opinion corresponds to taking a position on a disputed issue, opinion can be studied by identifying the statements that express a point of view (POV) and, from them, identifying the issues they question and the voices responsible for those points of view. To this end, we set ourselves the objective of characterizing the stream of opinion generated in five Spanish newspapers in relation to the release of *ChatGPT*. For this purpose, we worked with a corpus of 176 press articles of different genres published by five Spanish generalist national newspapers: *El País* (*elpais.com*), *El Mundo* (*elmundo.es*), *elDiario.es* (*eldiario.es*), *infoLibre* (*infolibre.es*), and *La Voz de Galicia* (*lavozdegalicia.es*). All open-access articles that included the word “*ChatGPT*” that were published in the first 100 days from its launch were collected.

This research describes the specific participation of five generalist newspapers in the public debate sparked by the release of *ChatGPT*. Determining how the media have reacted to this phenomenon is intended to contribute to understanding the press's participation in the public debate and to complement the knowledge we have regarding the integration of new technologies into journalism.

2. State of affairs

ChatGPT was launched on November 30, 2022, by the company *OpenAI*, which specializes in natural language dialog. According to this company, it is a prototype AI chatbot that belongs to a language model trained using unsupervised learning and reinforcement techniques (*OpenAI*, 2022). This tool, which was made available to the general public, stood out for its detailed and articulate responses that closely mimic human speech. *ChatGPT* works with 175 million parameters with information recorded until 2021, so it can provide an answer to any question posed, in a number of different domains.

Although it was not the first artificial intelligence software, *ChatGPT* was the first to cause a real stir regarding the implications of the widespread use of AI in the societies of the third millennium and, in particular, its impact on the field of technologically mediated communication worldwide.

In the communication field, the boom in research papers on AI occurred in 2015 (**Calvo-Rubio; Ufarte-Ruiz**, 2021) as a result of the initial experiences of incorporating AI into the US newspaper industry to mechanize search and classification functions (**Linden**, 2017; **Flores-Vivar**, 2018). In the final years of the second decade of the millennium, rapid changes in technologies alerted them to the adaptation and reshaping of journalism hand in hand with AI (**Broussard et al.**, 2019). In recent years, scientific interest has grown, and the number of articles by researchers from different countries on this subject of study has multiplied. Since the end of the second decade of the current century, robots and AI have already been pointed out as new challenges for journalism (**Salazar**, 2018), and since the beginning of the current decade, papers have been published on the advances of AI and its application by some media outlets (**De-Lima-Santos; Salaverría**, 2021), the impact on the media ecosystem (**Túñez-López; Fieiras-Ceide; Vaz-Álvarez**, 2021), the impact on journalistic practices in some countries, and expectations in some media, as well as future prospects (**Canavilhas**, 2022; **De-Lima-Santos; Ceron**, 2022).

As the impact of AI on journalism has grown, experts, journalists, and academics have expressed their conviction that AI will improve journalists' skills and increase production efficiency, but they have also expressed that a change of mindset is needed in the media environment and in facing the renewed ethical debates it raises (**Noain-Sánchez**, 2022). However, there is a leading trend in the European Union that holds that AI will improve policy-making, provision of public services, and internal management (**Van-Noordt; Misuraca**, 2022). At the same time, in view of the changes that are emerging in society in general and in the field of communication in particular, proposals have emerged to call for a research agenda for human-machine communication that sets as a starting point the differences between communicative AI and previous technologies that were the focus of communication research (**Guzman; Lewis**, 2019).

Based on this systematic review of research conducted in recent years, the most recent research in the field is related to the identification of the social and epistemological challenges posed by the adoption of communicative AI, with the media concerned about the defense of their editorial independence and journalists about the loss of jobs and the challenge of using them in their work to help produce content with greater added value and better quality (**Peña-Fernández et al.**, 2023). *ChatGPT's* release has heightened concern in the field of journalism regarding how it will affect it and how it can be used for journalistic tasks at a time when the number of synthetic media and newsrooms without journalists is increasing (**Ufarte-Ruiz et al.**, 2023). Hence, there is a need to determine the streams of opinion relayed by the media on these issues upon *ChatGPT's* release, which has its capabilities and limitations, but which heralds important transformations for the practice of journalism (**Pavlik**, 2023).

3. Stream of opinion in the media

The media contribute to public debate in different ways, mainly because they have the capacity to determine the circulation of information and opinion in the public space. Certainly, this participation is not neutral, as if they were mediators in a debate; rather, they take up a position according to their editorial line, including their own political and economic interests, for example, when selecting and prioritizing the content they offer to their audiences (**Girard**, 2012; **Califano**, 2015). In this way, the media can become involved in certain issues to varying degrees, which would have an impact on the process of democratic deliberation, making them important social actors.

In accordance with diverse research interests, multiple ways of studying opinion, from focus group techniques to the ever-popular opinion polls, have been proposed. Opinion is commonly understood in contrast to facts: opinion is subjective, while facts are objective. On this subject, see **Myers** (2004), who discusses at length the conditions of what counts as opinion. In a discursive framework, opinion falls within the scope of argumentation, whose central element is a locutionary act in which a speaker relates an issue to themselves, revealing their position regarding what they are talking about (**Charaudeau**, 2003).

In light of two central aspects of journalistic discourse –its polyphonic character (Bakhtin, 1981; Chatzikoumi; González-Arias, 2022) and its argumentative dimension (Amossy, 2000; Charaudeau, 2021)– we turn to the notion of the stream of opinion (González-Arias; Campos-Rojas; 2020; González-Arias; Álvarez; Bustamante, 2022) to systematize the analysis of a newspaper’s participation in the public debate regarding a particular topic and during a given period of time. In this context, the stream of opinion is the set of opinions conveyed by the press through the text it publishes. To study the stream of opinion, we start by identifying a statement that expresses a point of view. From this statement, a voice responsible for said opinion can be recognized and, at the same time, so can the debate in which it is framed (Plan-tin, 2012). This opinion, expressed through one or more points of view, may be issued by the author of the text being studied, or it may be attributed to another actor through a direct or indirect quotation. Additionally, the debate refers to the specific topics or subtopics about which a point of view is expressed. The debate emerges upon the statement of opinion, since expressing a point of view always implies questioning the issue being debated.

4. Methodology

From a communication studies and discourse analysis perspective, we propose a qualitative study based on data extracted from a textual corpus of Spanish national newspapers. The corpus consisted of 176 articles published during the first 100 days since the public launch of *ChatGPT* on November 30, 2022. The articles correspond to different journalistic genres. The two criteria used were the inclusion of the word “*ChatGPT*” in the text and the time period established. The articles were collected using the *Factiva* database, which allowed us to collect all the articles published in open access during that period.

<https://global.factiva.com>

The five media outlets were chosen for their audience penetration and for being representative of different sectors of the Spanish press ecosystem. *El País* and *El Mundo* were chosen because they have two distinct generalist editorial lines, with the former being center-left and the latter being center-right, and because they are among the top five in circulation, according to the *Office for the Justification of Diffusion (OJD)*, representing the “traditional journalistic brands”. In contrast, *elDiario.es* and *infoLibre* are digital native media outlets that are among the most consolidated media outlets in Spain. Both are left-wing, but they are differentiated by the fact that *infoLibre* has a partnership agreement with a French media outlet, *Mediapart*. Additionally, we added a regional media outlet because this group has a higher degree of credibility among its readers according to the *Digital News Report Spain 2023* (Kaufmann-Argueta, 2023). Specifically, *La Voz de Galicia* was included because it has high audience ratings and belongs to a Spanish autonomous community with a significant number of local media outlets, so we consider it to represent the role of regional media well. With this selection, we sought a balance within the various editorial lines and the news attention they devote to science and technology topics. The sample has, therefore, an objectifiable part with broadcast and audience data and a convenience sample part. It should be noted that the newspapers *El Confidencial Digital*, *El Huffingtonpost*, and *El Correo Gallego* were also selected initially, but were discarded owing to the low number of articles we were able to obtain on the subject during the period.

Table 1. Number of articles in the corpus

Newspaper	Articles
<i>El País</i>	47
<i>El Mundo</i>	21
<i>La Voz de Galicia</i>	52
<i>elDiario.es</i>	44
<i>infoLibre</i>	12
Total	176

Table 1 presents the number of articles per newspaper that were actually analyzed.

Procedures

The analysis was carried out with the support of *AtlasTi 8.0*, a qualitative analysis software that facilitates the coding and grouping of codes. Only one document was generated per newspaper. The inductive analysis consisted of four main stages:

- (a) Identification of opinion statements, which were coded with a point of view and voice responsible for the statement.
- (b) Grouping of the points of view into subtopics by semantic similarity.
- (c) Grouping of voices into general categories associated with social roles (researcher, businessman, politician, etc.).
- (d) Analysis of relationships between code categories and documents.

To illustrate the coding, an example of the analysis extracted from the corpus studied is presented:

“It is necessary to study and analyze the uses and risks of this innovative technology with the aim of creating robust legislation that turns artificial intelligence into an ally and not an enemy because, if there is one thing we know, it is that it is here to stay” (*El País*, January 23, 2023, Paula Santolaya de Burgos).

POV: Legislation is needed to control the uses and risks of AI.

Voice: Journalist. Paula Santolaya del Burgo.

POV group: The need for regulation and control of the development of AI tools.

Group of voices: Author of the newspaper article.

The categories used in the voice code groups and topic code groups were determined by abstraction based on the recurrence of the positions that reflect these opinions.

5. Results

Below are the findings that allow us to characterize the stream of opinion that was produced during the first 100 days since the launch of *ChatGPT*. This covers the period from November 30, 2022, to March 9, 2023. It should be noted that the methodology used favors the general and abstract view of the debate, thus losing sight of the opinions of particular individuals or specific points of view. In the same sense, the opinions could be issued by the authors of the articles as well as by the persons who are quoted. Likewise, the opinion statements conveyed by each newspaper are identified without considering the journalistic genre.

In the first stage of the analysis, 530 quotations of statements expressing an opinion were identified. This number was reached by extracting the first 106 quotations from each journal. The number 106 corresponds to the maximum number of citations extracted from *infoLibre*, the newspaper in our corpus that had the least news items. These quotes were coded into 90 points of view (POV) and 166 different voices. In turn, the POVs were thematically grouped into 8 subtopics and the voices into 11 voice groups. Table 2 summarizes the relationship between citations, codes, and code groups.

Table 2. Quotations, codes, and groups

530 quotations of opinion statements	90 POV codes	8 POV groups (subtopics of the debate)
	166 voice codes	11 groups of voices

5.1. Subtopics of the debate

As noted above, the debate brought on by *ChatGPT*'s release was structured around eight main subtopics. Table 3 presents the subtopics ordered from highest to lowest recurrence, as well as the percentage they represent according to the number of textual citations ($N = 530$). These eight topics reflect what we could call the first moment of public discussion about *ChatGPT*'s release, which would correspond to approximately the first three months.

Table 3. Subtopics characterizing the stream of opinion on *ChatGPT*'s release

Subtopic of the debate on <i>ChatGPT</i> 's release	Total
The need for regulation and control of the development of AI	22.9%
<i>ChatGPT</i> 's limitations as an AI tool	21.7%
<i>ChatGPT</i> 's capabilities as an AI tool	18.6%
Implications of <i>ChatGPT</i> 's release in the technology industry	10.8%
The innovation signaled by <i>ChatGPT</i> 's release	10.2%
The impacts that <i>ChatGPT</i> has had or could have on education	6.9%
The impacts that <i>ChatGPT</i> could have in the workplace	6.0%
The impacts that <i>ChatGPT</i> could have in the field of journalism	2.7%

It should be noted that *ChatGPT* was made available to the general public free of charge. This may explain why, at first, there was concern about its effects on education and about the potential for the tool's use for plagiarism in schoolwork. Something similar occurred when it came to the threat that could be felt in relation to certain jobs where a generative program could replace people. However, the discussion about its extraordinary capabilities brought the focus to the competition to dominate the AI development that was beginning to take place between technology companies in the sector. At the same time, scientists and journalists were highlighting *ChatGPT*'s limitations: biases, errors, and operating costs, which oriented the discussion more toward the illusion that the tool's potential could create. Finally, the main topic was the need to control and legislate the development and use of AI, as the negative consequences of their uncontrolled development began to be noticed, mainly owing to the competition for technological dominance in the market.

It could be pointed out that the public debate reflects a widespread general interest in the fields of education and work and a more academic interest that focuses on advances in technology. Below, we describe each of the topics, the predominant points of view, and the scope of the discussion.

“*ChatGPT* was made available to the general public free of charge. This may explain why, at first, there was concern about its effects on education and about its potential for plagiarism in schoolwork”

5.2. The need for regulation and control of the development of AI

As can be seen, 22.9% of the POVs were focused around the need for regulation and control of the development of AI. This aspect of the debate has local, national, and international scope. At the local level, institutions such as universities and educational centers in various countries are beginning to announce the measures that will be implemented to mitigate the negative effects of the emergence of generative technology that are affecting the teaching and learning processes. At the national level, governments are being called upon to develop initiatives to regulate the development and use of AI systems. At the international level, the debate focuses on the responsibility of the companies that are capable of developing AI technology themselves.

The main POVs that made up this group were as follows:

- Legislation needs to control the uses and risks of AI.
- Machines created with AI present the black box problem: they work, but we do not know how.
- Technology needs to be developed on ethical principles.
- Training AI tools has significant human labor costs.
- Competition between companies can lead to decisions that are detrimental to users.

As we can see, the prevailing view was that the use of AI softwares such as *ChatGPT* carries significant risks that demand state intervention to prevent them from becoming instruments of abuse or causing social harm. In addition, both developers and users are required to incorporate ethical principles to protect the public interest. The problem presented by machines trained by deep learning and the lack of design in their operation was discussed: Machines will help us solve problems, but since we do not know how they work, we could make big mistakes. Additionally, it was stated that the development of these tools has required many hours of people working in tasks aimed at eliminating harmful content circulating on the network. These jobs have usually been performed by low-paid workers. And another aspect considered in this group is that, because competition for AI dominance among large technology companies is very aggressive, there is a risk that companies will decrease ethical controls for bringing programs to market. At the same time, it was recognized how difficult it is to legislate in this regard, owing to the complexity of establishing limits between the benefits of technological development and the ambitions of the entrepreneurs who finance them.

5.3. *ChatGPT's* limitations as an AI tool

At a level very close to the previous subtopic, with 21.7%, opinions were expressed regarding *ChatGPT's* limitations: its inability to distinguish between what is true and what is false, the biases derived from its training, and the errors of use detected in tests carried out by both journalists and users in general. This aspect of the discussion focused on the tool itself. It is possible to observe that there is an interest in shattering the illusion the system generated at its launch and to consider its real value in its current state of development.

The following views were the most common in this group:

- ChatGPT* makes significant mistakes.
- ChatGPT* does not have the ability to distinguish between what is true and what is false.
- ChatGPT* is an immature technology.
- ChatGPT* is not truly a general AI.

Regarding the limitations of *ChatGPT*, the most common opinions pointed to the errors that it made in relation to scientific knowledge or to the impossibility of offering objective information with the corresponding sources. The main weakness of the tool is that, in light of its responses and what is known about its operation, it fails to differentiate between what is true and what is false. This makes it an unreliable tool for important matters. It is described as an immature program, although with varying degrees of optimism, it is expected to improve. Another important aspect is the question of the gap between the results obtained by *ChatGPT* and the idea of achieving a machine that really operates as a general artificial intelligence, i.e., that reasons like a human when faced with any subject.

“ The public debate reflects a widespread general interest in the fields of education and work and a more academic interest that focuses on advances in technology ”

5.4. *ChatGPT's* capabilities as an AI tool

The subtopic that came in third is clearly positive. At 18.6%, all opinions regarding *ChatGPT's* capabilities as an AI tool align. These qualities position *ChatGPT* as an example of the advances in AI development and as a flagship program for the future of technology. This aspect of the debate also focuses on *ChatGPT* as a tool and is based on the common interest of ensuring that technology improves people's quality of living. The most recurrent POVs were as follows:

- ChatGPT's* main strength is its ability to interact in a user-friendly way.
- ChatGPT* is a successful artificial intelligence.
- ChatGPT* is an important preview of what can be done in the future.
- ChatGPT* can replace humans in a variety of tasks.
- The advance of AI is inexorable.

This subtopic reflects the more positive side of the debate. The tool's capacity for interaction, especially the ability to interact in a friendly manner, providing answers to practically any topic, is enthusiastically valued. Likewise, the ability to imitate human language, even reaching a grammatical correctness that would surpass the capabilities of many humans (students, felons, ill-trained professionals, among others), is highlighted and valued. Owing to the volume of knowledge it handles, the speed

of its responses, and its capacity for synthesis, it is estimated that it can already replace humans in various tasks. *ChatGPT* is a preview of what will be possible in the near future. It is proposed that advances in the field of AI-based technologies are generating major changes in a variety of areas of life, which will allow humans to develop other abilities or focus their attention on other matters hitherto unthinkable.

“ The main topic was the need to control and legislate the development and use of AI, as the negative consequences of their uncontrolled development began to be noticed, mainly owing to the competition for technological dominance in the market ”

5.5. Implications of *ChatGPT*'s release in the technology industry

Regarding the impact of *ChatGPT* on the technology industry, with 10.8%, the opinions that *ChatGPT* has unleashed a feverish competition among technology giants stood out. In this regard, the debate reveals a certain level of frustration with the lack of control of market functioning at the level at which technology companies operate. It is no longer about countries; it is about technological “giants” advancing unchecked. Public interest can be seen at the international level, as these opinions warn of the social consequences that the war for the technological domination could have. The prevailing opinions were as follows:

ChatGPT is a threat to *Google*.

A war for AI dominance has begun.

ChatGPT is competition for other AI systems.

At the moment, advances in AI are going faster than their predecessors.

Spain can become an AI technology hub.

According to the opinions amassed in this group, there is a war between giants, starting with *Google* and *Microsoft*, but it would involve all the big players in the AI-based technology industry. Likewise, the rapid development of all kinds of applications that will take advantage of the new technology is also expected. At this point, the focus of discussion shifts from the technology itself to the market.

At this point, it is also believed that Spain would have the capacity to transform itself into a technological pole, so the business opportunities for Spanish companies in the technological area were evaluated.

5.6. The innovation signaled by *ChatGPT*'s release

A subtopic in the field of innovation was considered to make a difference between the opinions aimed at assessing the tool's capabilities and limitations visible, on the one hand, and the evaluation of *ChatGPT*'s release as a historical event and its impact on public opinion, on the other. This subtopic, which came in fifth place, accounts for 10.2% of the opinions. Clearly, the issue of AI and especially *ChatGPT* is now in the domain of international public interest. It is a technology that, owing to its versatility, will affect all of humanity. The most noteworthy POVs were as follows:

ChatGPT is a milestone in AI.

ChatGPT foreshadows the future of AI.

ChatGPT has revolutionized the international conversation about AI.

ChatGPT is not the ultimate revolution.

It is necessary to differentiate the impression produced by interacting with *ChatGPT* from what is really behind it.

Opinions converge in calling *ChatGPT* a milestone in the development of AI and, owing to its scope, a turning point in the relationship between people and machines. It has been determined that we are in the golden age of AI and that *ChatGPT* foreshadows the future. It is a revolutionary tool of progress. However, there are also less enthusiastic opinions in this group that question whether the program really does what it claims to do or whether, owing to the high implementation and operating costs, it could be the revolutionary turning point that others have been claiming.

5.7. The impacts that *ChatGPT* has had or could have on education

The sixth area of inquiry is education, with 6.9% of the occurrences. This issue may have been the first to generate public concern. The generative tool, which had been made freely available to the public, was quickly adopted by students. At this early stage, *ChatGPT* was seen as a threat to the school system and quickly divided opinions between those who called for a ban on its use and those who advocated its integration into the educational process. We can see that the debate here addresses a very generalized interest. Educational centers reacted, and numerous specialists were called upon to take a position on the issue. In this subtopic, the following POVs stand out:

ChatGPT forces us to think about new educational strategies.

ChatGPT presents a concern for the negative impact it may have on learning.

ChatGPT can be useful in teaching for both teachers and students.
It is necessary to educate AI users to avoid falling victim to their algorithms.
Universities will take years to deal with the integration of AI.

In general, there is concern about the negative impacts that *ChatGPT* may have on education, particularly associated with the possibility that students will be able to plagiarize work. Likewise, easy access to a tool that could make summaries, essays, and reports could be detrimental to the development of students' critical skills. From a more positive perspective, it was thought that incorporating AI into education could be beneficial for both teachers and students, which would entail an innovation challenge for both educational strategies and educational curricula. Projected to the entire population, there is a need to learn how AI systems work, to avoid being harmed by the tools themselves. Finally, as a criticism of the educational system, the idea was proposed that it will take several years for universities to effectively adapt their operations to AI and effectively take advantage of this technological development.

5.8. The impacts that *ChatGPT* could have in the workplace

Regarding the consequences of *ChatGPT* in the workplace, which reached 6%, the predominant idea is that AI will affect different professions differently and that workers will have to adapt to the imminent implementation of AI in all spheres of work. This aspect of the debate mainly addresses local and partisan interests. The specific areas in which technology can displace people are discussed. In particular, it highlights the generative potential of the tool as a threat to people. There is only a very incipient debate about the tension that could arise between the owners of companies and their workers: on the one hand, the owners of the companies have recognized the possibility of reducing costs using the new technology, while on the other hand, the workers are beginning to feel displaced by the technology. The POVs that stood out the most were the following:

ChatGPT is the future assistant for various professionals.
We will have to learn to work with the new systems.

It is expected that *ChatGPT* will affect different professions differently; however, we will all have to adapt to the new tools, and over time, they will become our assistants. It is believed that the more routine tasks are the ones being rapidly replaced by AI, but that it is only a matter of time before machines replace people in increasingly sophisticated and complex tasks. Specifically with regard to employment, there has been a crisis in large technology companies, which has led to massive layoffs of workers. However, a new balance in the labor market would be achieved, as the demand for AI talent is expected to grow.

5.9. The impacts that *ChatGPT* could have in the field of journalism

In relation to the POVs that addressed the journalism subtopic, it is felt that *ChatGPT* and other AI developments could have a huge impact on the journalism industry thanks to the ability of machines to handle large volumes of information. In any case, a negative view predominates. The debate on this issue mobilizes diverse interests and values that challenge journalism. Mainly, *ChatGPT's* generative nature has raised concerns about the negative effects of using the tool. The first warning is against disinformation, but the debate is mainly focused on the transformation of the industry. While AI may make journalistic work easier, it places enormous demands on journalists themselves. The main POVs of this group of codes were as follows:

ChatGPT could have a huge impact on the journalism industry.
ChatGPT can become a factory of fake news and disinformation
The effectiveness of fake news is not in the quality of the written text but in how it is distributed on the network.

Although some have gone so far as to call for the end of journalism, at least as we know it today, opinions varied. Although the idea that journalists could be replaced in various tasks was present, it was suggested that it is a threat especially for low-quality journalism. The development of tools such as *ChatGPT* will continue to be relevant for editorial and creative work. Also, considering the current state of development of AI that make major mistakes, professionals cannot be dispensed with. However, the prevailing view was that *ChatGPT* can become a factory of fake news and that AI systems can become amplifiers of disinformation, even unintentionally.

6. Differences between newspapers by subtopic

It can be seen that four of the five newspapers agreed on the first three topics: regulation and control, *ChatGPT's* capabilities, and limitations. These issues accounted for more than 60% of the opinions, so we can see a commonality of interests in the evaluation of the tool and the assessment of its risks. The exception was *elDiario.es*, which gave more space to innovation and consequences in the technology industry. While no distinct patterns of behavior can be attributed to the newspapers, the order of priority of certain topics may reflect nuances or trends associated with editorial orientations or certain operating conditions that give preference to some points of view over others.

It is possible to point out three ways of constructing opinion that distinguish the media studied: journalists' involvement, using sources of national specialists, or using international specialists

Table 4. Predominance of subtopics by journal

Subtopic of the debate on ChatGPT's release	<i>El País</i>	<i>infoLibre</i>	<i>El Mundo</i>	<i>eDiario.es</i>	<i>La Voz de Galicia</i>
The need for regulation and control of the development of AI	16.7%	22.6%	32.7%	29.2%	13.5%
ChatGPT's limitations as an AI tool	17.7%	24.7%	17.3%	20.8%	28.1%
ChatGPT's capabilities as an AI tool	25.0%	18.3%	19.4%	10.4%	19.8%
Implications of ChatGPT's release in the technology industry	12.5%	5.4%	8.2%	12.5%	15.6%
The innovation signaled by the ChatGPT's release	10.4%	7.5%	6.1%	13.5%	13.5%
The impacts that ChatGPT has had or could have on education	5.2%	10.8%	6.1%	6.3%	6.3%
The impacts that ChatGPT could have in the workplace	11.5%	2.2%	7.1%	6.3%	3.1%
The impacts that ChatGPT could have in the field of journalism	1.0%	8.6%	3.1%	1.0%	0.0%

As Table 4 shows, *El País* gave preference to POVs highlighting ChatGPT's capabilities and devoted somewhat more attention than the other newspapers to issues associated with ChatGPT's impact on the workplace. In contrast, *InfoLibre* and *La Voz de Galicia* emphasized the limitations of the tool, but expressed different interests regarding the effects of AI on the technology industry. For their part, *El Mundo* and *Diario.es* converge in broadly giving preference to opinions on the need for regulation and control of AI's use.

7. Voices in the debate

A total of 530 voices were identified, classified into 166 codes and subsequently grouped into 11 groups of voices. Identifying voices allows us to determine the actors in the debate, those who directly participate when it is the author of the journalistic text or indirectly when the actor is cited in the text as a source. However, in any case, from our perspective, the voice represents, in some way, the editorial line of the newspaper on the subject. The importance of a social actor's participation in the debate depends, naturally, on the degree of specialization of the subject and the access that the journalist can have to the sources. In this sense, the opinion of those who write the articles that are published or publish the opinions of others acquires a strategic character.

As Table 5 shows, 88% of the discussion took place among four groups of voices: "author of the article," "Spanish researcher," "foreign researcher," and "representative from a technology company." This grouping reflects the degree of specialization of the subject and a high valuation of expert voices regarding a considerably complex subject such as AI.

Although the topic of debate has proven to be of interest to broad sectors of the population, the newness and lack of knowledge of the subject limits the number of voices participating. Certainly, some voices that could have taken up more space were deemed to be absent in this case, for example, politicians, legislators, or voices from civil society. Consequently, it can be said that this first moment of public debate has been restricted, being limited to those directly involved: journalists, as animators of the debate; researchers; and representatives of technology companies.

Table 5 presents the groupings of voices in decreasing order according to the number of occurrences.

Clearly, what predominates is the voice of the authors of the articles themselves, who are mostly journalists or specialists invited by the media. This is followed by Spanish and foreign researchers as technology specialists who were called upon to evaluate the software, and fourthly by the voices of the representatives of the technology companies who are the real players behind it. Below, we comment on the first four groups of voices, since together they represent 88% of the opinions in the stream of opinion studied.

7.1. Author of the journalistic article (37.6%)

This corresponds to the person who wrote the journalistic text, which could be either a news item or an opinion column. The main POVs issued directly by the articles' authors were on various subtopics. If we consider them in order of frequency, the top six POVs in this group were the following:

Table 5. Voices participating in the debate on ChatGPT release

Group of voices	Percentage (%)
Newspaper article's author	37.6
Spanish researcher	24.2
Foreign researcher	13.9
Representative of a technology company	12.3
Anonymous	4.0
Foreign politician	2.4
National or foreign institution	1.8
The newspaper (article's author undisclosed)	1.4
International newspaper company	1.0
Spanish politician	0.8
News agency	0.2

ChatGPT makes significant mistakes.
ChatGPT is an immature technology.
 A war for AI dominance has begun.
ChatGPT does not have the ability to distinguish between what is true and what is false.
 Legislation needs to control the uses and risks of AI.
 Machines created with AI present the black box problem: they work, but we do not know how.

These voices are the ones that most directly represent the opinions of journalists, who stand out for having a negative view of *ChatGPT*, where distrust and alertness prevail. Based on what has been observed, brief remarks can be constructed that reflect this positioning: *ChatGPT* is an error-prone tool, which is evidence of the immaturity of the technology behind it. Its main flaw is its inability to distinguish between what is true and what is false. Despite these problems with the technology, a war for AI dominance has begun, making it necessary to legislate to control its uses and reduce risks.

7.2. Spanish researcher (24.2%) and foreign researcher (13.9%)

This category included people presented as researchers, scientists, experts, and academics, all with backgrounds in the field of artificial intelligence. Undoubtedly, they emerge as authoritative sources to speak on the subject, and it makes sense that it is mostly national researchers who have the strongest voice.

Spanish researchers

ChatGPT makes significant mistakes.
ChatGPT does not have the ability to distinguish between what is true and what is false.
ChatGPT forces us to think about new educational strategies.

Foreign researchers

Machines created with AI present the black box problem: they work, but we do not know how.
ChatGPT is not truly a general artificial intelligence.
ChatGPT cannot be trusted with anything important at this time.

Although they highlight different aspects, both national and foreign researchers mostly agree on the negative aspects of the tool.

7.3. Representative of a technology company (12.3%)

This category included people who were presented as CEOs of companies involved in the development of the tools, for example, Sam Altman CEO of *OpenAI* and Satya Nadella CEO of *Microsoft*. In this case, these are the main players. At the national level, leading businessmen of technology companies were cited as authorized sources to comment on the event or as parties involved because of the business opportunities that could be created. The main POVs of this group were as follows:

Technology needs to be developed on ethical principles.
ChatGPT is a threat to *Google*.
 Cybercriminals can use *ChatGPT* for scams.

Undoubtedly, the concern about the need for technology to be developed with ethical principles is linked to three issues encountered by the industry: the need to avoid biases in the tool's training, the need to eliminate harmful content from machine training, and the risks of increasing the gap between those who access the technology and those who do not. It is worth mentioning here that *OpenAI*, the company that developed *ChatGPT*, announced its interest in democratizing access to technology in its founding ideology. In addition, among the representatives of technology companies, the competition generated to dominate AI is the topic of greatest interest. Likewise, the concern about cybercrime, which is listed in third place, is related to the need to legislate on the use of AI and to develop cybersecurity programs to prevent crime from benefiting from technological development.

8. Groups of voices according to newspaper

We select only the first four categories to comment on, since they account for more than 80% of the voices in each of the newspapers. From the comparison of the opinionated voices conveyed by each media outlet, it is possible to identify different strategies, probably associated with their own journalistic and organizational practices. Table 6 presents the proportion of the main voices according to newspaper in columns.

Table 6. Differences between newspapers in relation to the voices they gave preference to in the debate

Group of voices	<i>El País</i>	<i>infoLibre</i>	<i>El Mundo</i>	<i>elDiario.es</i>	<i>La Voz de Galicia</i>
Newspaper article's author	50.5%	34.4%	22.2%	45.4%	35.7%
Spanish researcher	20.4%	51.6%	15.2%	9.3%	24.5%
Foreign researcher	5.4%	2.2%	40.4%	11.3%	10.2%
Representative of technology company	16.1%	5.4%	3.0%	16.5%	20.4%
	92.5%	93.5%	80.8%	82.5%	90.8%

As can be seen from Table 6, the stream of opinion of the newspapers *El País*, *El Diario.es*, and *La Voz de Galicia* was dominated by the opinions of the articles' authors themselves, whether they were journalists from the media outlet or invited specialists. *El País* stood out in particular with 50.5%, which reflects its journalists' involvement in the subject and, probably, the established capabilities of people dedicated to the technology area in the media. In contrast, *infoLibre* was dominated by Spanish researchers' voices; that is to say, it offered articles with voices of specialists from national universities and centers. It is clear from this practice that the newspaper's strategy was to seek authoritative sources at the national level that could explain and evaluate the tool. For its part, *El Mundo* preferred the voices of foreign researchers. Another striking difference was between *La Voz de Galicia*, which had a preference for the voices of representatives of technology companies (20.4%), and the newspaper *El Mundo*, which had a very low percentage in this category (3.0%). In this case, the distinction occurred because *La Voz de Galicia* gave preference to the voices of local entrepreneurs who have seen the development of *ChatGPT* as a business opportunity.

9. Discussion

Technology has always played an important role in the development of modern journalism, from its beginnings to the computerization of newsrooms (Vázquez-Herrero; López-García; Irigaray, 2020). It is also suggested that, since the creation of the so-called Internet Galaxy (Castells, 2001), emerging technologies have been fueling disruptive innovations (Christensen, 2003) that are stressing the communication ecosystem, which is currently searching for business models that will allow it to survive. Certainly, technological evolution has been the main driver of changes in journalism in the third millennium (Pavlik, 2000) while, at the same time, opening up new scenarios for communication processes in the media ecosystem. All these changes in journalism have had consequences for social organization and democracy. Undoubtedly, recent technological advances associated with the development of AI will impact not only journalism but also social life in general.

Although in the recent past it has been found that Spanish journalists have a profound lack of knowledge regarding AI's direct influence on the profession (Túñez-López; Toural-Bran; Cacheiro-Requeijo, 2018), it would be interesting to consider the opinions that journalists are pouring into their work. Probably, as workers, journalists also view their jobs as threatened and note an increase in the precariousness of their working conditions. This is in contrast to a business outlook that views technology as an opportunity to gain competitive advantages and to reduce costs.

This paper reveals that, at first, *ChatGPT's* release caused concern primarily in the fields of education and work, probably because it was made freely available to the general public and because it is a concrete example of AI technology. This is probably because AI applications have already widely impacted fields such as entertainment, medicine, the automotive industry, retail, and logistics.

As can be seen in the stream of opinion generated in the press, a negative view of *ChatGPT* predominates. Although it is recognized that *OpenAI's* program has extraordinary capabilities as a chatbot model, those same capabilities and potentialities transform it into a danger, either because it will be able to replace humans in various tasks, because it makes mistakes that go unnoticed owing to the correctness of language and fluency, or because the competition to dominate the AI market pushes aside the ethical criteria that are expected to accompany technological developments.

Journalism is still necessary for the functioning of plural societies in the third millennium (Moe; Ytre-Arne, 2021). However, the changes experienced in society and in the media ecosystem (García-Orosa, 2021) make it advisable to rethink journalism through bold formulas that are appropriate to the moment in which we are living, so it can continue to fulfill its role in the new social, political, and economic context (Zelizer, 2017).

No conclusive observations can be made regarding the differences between the newspapers. The results show a large overlap in the main groups of POV appearing in the different types of newspapers studied. However, it can be pointed out that the coincidences are stronger with respect to the invited voices than to the subtopics addressed. Preferences for some voices over others could reflect orientations in the practices and conditions under which opinion sources are managed. In this sense, we could say that newspapers differ in the position held by the opinions of journalists, national specialists, or foreign specialists.

In this regard, a study focused on the media ecosystem in Flanders that explored the inclusion of different voices in migration news demonstrates that the traditional media do not behave differently than the alternative press when it comes to including a variety of voices. However,

Regarding the voices that have been called to participate in the debate, the voices of the journalists themselves, who directly express their opinions in their articles, stood out –those who, in turn, have preferred as sources mainly Spanish researchers and foreign researchers, whose voices have unpacked the characteristics of *ChatGPT* and AI

Since it is an explosive phenomenon, it is possible to think that not all newspapers were prepared to face the topic with a well-founded position in their editorial line

alternative newspapers (digital natives) are differentiated according to their right-wing or left-wing political orientation: those on the right give preference the voices of right-wing politicians and those on the left, civil society actors (Buyens; Van-Aelst, 2021). In our case, to make observations in this sense, it would be necessary to have a textual corpus with other characteristics.

Regarding the Spanish media ecosystem's incorporation of technology, although criticism has been voice that, in Spain, the technological offering of AI applicable to journalism contrasts with the media's slowness in incorporating it (Sánchez-García *et al.* 2023), such slowness may be prudent, given the many questions being raised about the incorporation of AI technology.

Just as the food, pharmaceutical, and automotive industries cannot bring a product to market that does not meet certain safety standards, it does not seem unreasonable to ask that the AI technology industry ensure that its products will not do harm in society or have a great potential for malicious use before going to market.

10. Conclusions

In this study, we set ourselves the objective of characterizing the stream of opinion generated in five Spanish newspapers in relation to the release of *ChatGPT*. This research covers the first 100 days since *ChatGPT*'s launch in a limited portion of the Spanish press. However, we believe that it is a good reflection of the main characteristics of the studied newspapers' participation in the public debate that has been generated. This is a debate that is very focused on the properties of the tool and has been fueled primarily by AI specialists.

We can point out that the press studied reflects a rather negative and distrustful attitude toward AI. The debated subtopics reflect the areas of interest sparked by *ChatGPT*'s release: the innovation of the tool; its capabilities and limitations; the possible effects in the educational field, and in the workplace, particularly in journalism; and the social impact it has had on the technology industry. The topic that has prompted most opinions is the need to legislate for adequate control of the tool's uses. The need for control is based both on AI machines' capabilities and their limitations and, above all, on the fascination that has arisen at various levels with the unimaginable potential in terms of the applications it could have.

Regarding the voices that have been called to participate in the debate, the voices of the journalists themselves, who directly express their opinions in their articles, stood out—those who, in turn, have preferred as sources mainly Spanish researchers and foreign researchers, whose voices have unpacked the characteristics of *ChatGPT* and AI.

There are some differences between the newspapers studied, probably owing to their editorial lines or to the operating models that are open to different sources to a greater or lesser extent. From the data, it is possible to point out three ways of constructing opinion that distinguish the media studied: journalists' involvement, using sources of national specialists, or using international specialists.

Since it is an explosive phenomenon, it is possible to think that not all newspapers were prepared to face the topic with a well-founded position in their editorial line. Undoubtedly, a follow-up study could show us an evolution of the circulation of opinion in the media as the media outlets define a clearer editorial position in the face of technological development based on AI.

A study such as this one certainly has limitations that could be mitigated by extending the media coverage or the time period of the corpus collection. Nevertheless, this research clearly describes the first moments of the debate generated by a topic that will surely remain strongly in the public interest.

It is worth mentioning that, 20 days after the time period studied in this research, on March 29, 2023, a group of more than 1,000 artificial intelligence experts and technology industry executives, through an open letter, called for a six-month pause in the training of artificial intelligence systems, arguing that it is a potential threat to humanity (Pascual, 2023). It is clear from this fact that the debate that took place in the first 100 days was picking up on this concern and, moreover, that apparently the debate is beginning to call into question the policy, which was completely nonexistent during this period.

11. References

Alvarado-Salazar, Ruth; Llerena-Izquierdo, Joe (2022). "Revisión de la literatura sobre el uso de inteligencia artificial enfocada a la atención de la discapacidad visual". *Revista ingenio*, v. 5, n. 1, pp. 10-21.

<https://doi.org/10.18779/ingenio.v5i1.472>

Amossy, Ruth (2000). *L'argumentation dans le discours*. Paris: Armand Colin. ISBN: 978 6169078128

“ The topic that has prompted most opinions is the need to legislate for adequate control of *ChatGPT*'s uses. The need for control is based both on AI machines' capabilities and their limitations and, above all, on the fascination that has arisen at various levels with the unimaginable potential in terms of the applications it could have ”

- Bakhtin, Mikhail** (2014). "Polyphonic discourse in the novel". In: Angermüller, Johannes; Maingueneau, Dominique; Wodak, Ruth (eds.). *The discourse studies reader: main currents in theory and analysis*, pp. 27-35. Amsterdam: John Benjamins Publishing Company.
<https://doi.org/10.1075/jlp.15.6.08for>
- Broussard, Meredith; Diakopoulos, Nicholas; Guzman, Andrea L.; Abebe, Rediet; Dupagne, Michel; Chuan, Ching-Hua** (2019). "Artificial intelligence and journalism". *Journalism & mass communication quarterly*, v. 96, n. 3, pp. 673-695.
<https://doi.org/10.1177/1077699019859901>
- Buyens, Willem; Van-Aelst, Peter** (2022). "Alternative media, alternative voices? A quantitative analysis of actor diversity in alternative and mainstream news outlets". *Digital journalism*, v. 10, n. 2, pp. 337-359.
<https://doi.org/10.1080/21670811.2021.1929366>
- Califano, Bernardette** (2015). "Los medios de comunicación, las noticias y su influencia sobre el sistema político". *Revista mexicana de opinión pública*, n. 19, pp. 61-78.
<https://doi.org/10.1016/j.rmop.2015.02.001>
- Calvo-Rubio, Luis-Mauricio; Ufarte-Ruiz, María-José** (2021). "Inteligencia artificial y periodismo: revisión sistemática de la producción científica en Web of Science y Scopus (2008-2019)". *Communication & society*, v. 34, n. 2, pp.159-176.
<https://doi.org/10.15581/003.34.2.159-176>
- Canavilhas, João** (2022). "Artificial intelligence and journalism: current situation and expectations in the Portuguese sports media". *Journalism and media*, v. 3, n. 3, pp. 510-520.
<https://doi.org/10.3390/journalmedia3030035>
- Castells, Manuel** (2001). *La Galaxia Internet. Reflexiones sobre internet, empresa y sociedad*. Barcelona: Plaza & Janés Editores. ISBN: 8401341574
- Charaudeau, Patrick** (2003). *El discurso de la información. La construcción del espejo social*. Barcelona: Gedisa. ISBN: 847432 9531
- Charaudeau, Patrick** (2021). *El discurso político. Las máscaras del poder*. Buenos Aires: Prometeo Libros. ISBN: 978 987 845 1015
- Chatzikoumi, Eirini; González-Arias, Cristian** (2022). "Polifonía y pluralismo en el tratamiento del delito en editoriales de prensa durante el estallido chileno del 2019". *Logos. Revista de lingüística filosofía y literatura*, v. 32, n. 2, pp. 288-306.
<https://doi.org/10.15443/RL3217>
- Chiu, Thomas, K. F.; Xia, Qi; Zhou, Xinyan; Chai, Ching-Sing; Cheng, Miaoting** (2023). "Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education". *Computers and education: artificial intelligence*, v. 4, e100118.
<https://doi.org/10.1016/j.caeai.2022.100118>
- Christensen, Clayton M.** (2003). *The innovator's dilemma: the revolutionary book that will change the way you do business*. New York: Harper Collins Publishers. ISBN: 978 0 060521998
- De-Lima-Santos, Mathias-Felipe; Ceron, Wilson** (2022). "Artificial intelligence in news media: current perceptions and future outlook". *Journalism and media*, v. 3, n. 1, pp. 13-26.
<https://doi.org/10.3390/journalmedia3010002>
- De-Lima-Santos, Mathias-Felipe; Salaverría, Ramón** (2021). "From data journalism to artificial intelligence: challenges faced by La Nación in implementing computer vision in news reporting". *Palabra clave*, v. 24, n. 3, e2437.
<https://doi.org/10.5294/pacla.2021.24.3.7>
- Diakopoulos, Nicholas** (2019). *Automating the news. How algorithms are rewriting the media*. Cambridge: Harvard University Press. ISBN: 978 0 674976986
- Dogan, Murat-Ertan; Goru-Dogan, Tulay; Bozkurt, Aras** (2023). "The use of artificial intelligence (AI) in online learning and distance education processes: a systematic review of empirical studies". *Applied sciences*, v. 13, n. 5, e3056.
<https://doi.org/10.3390/app13053056>
- Flores-Vivar, Jesús-Miguel** (2018). "Algoritmos, aplicaciones y big data, nuevos paradigmas en el proceso de comunicación y de enseñanza-aprendizaje del periodismo de datos". *Revista de comunicación*, v. 17, n. 2, pp. 268-291.
<http://www.doi.org/10.26441/RC17.2-2018-A12>
- FundéuRAE** (2022). "Inteligencia artificial es la expresión del 2022 para la FundéuRAE".
<https://www.fundeu.es/recomendacion/inteligencia-artificial-es-la-expresion-del-2022-para-la-fundeurae>
- García-Orosa, Berta** (2021). "Disinformation, social media, bots, and astroturfing: the fourth wave of digital democracy". *Profesional de la información*, v. 30, n. 6, e300603.
<https://doi.org/10.3145/epi.2021.nov.03>

- Girard, Charles** (2012). "De la prensa en democracia: la revolución mediática y el debate público". *Centro teórico cultural criterios*, n. 27, pp. 453-467.
- González-Arias, Cristian; Álvarez-Olivares, Miyodji; Bustamante-Carrasco, Alan** (2022). "La participación de la prensa de referencia chilena en el debate público sobre el movimiento feminista chileno de 2018". *Literatura y lingüística*, n. 45, pp. 457-484.
<https://doi.org/10.29344/0717621x.45.2241>
- González-Arias, Cristian; Campos-Rojas, César** (2020). "El flujo de opinión sobre el sistema de pensiones en cuatro géneros de la prensa chilena: cobertura, voces y problemáticas". *Logos. Revista de lingüística, filosofía y literatura*, v. 30, n. 1, pp. 138-153.
<https://doi.org/10.15443/RL3012>
- Guzman, Andrea L.; Lewis, Seth C.** (2020). "Artificial intelligence and communication: a human-machine communication research agenda". *New media & society*, v. 22, n. 1, pp. 70-86.
<https://doi.org/10.1177/1461444819858691>
- Incio-Flores, Fernando-Alain; Capuñay-Sánchez, Dulce-Lucero; Estela-Urbina, Ronald-Omar; Valles-Coral, Miguel-Ángel; Vergara-Medrano, Segundo-Edilberto; Elera-Gonzales, Duberli-Geomar** (2021). "Inteligencia artificial en educación: una revisión de la literatura en revistas científicas internacionales". *Apuntes universitarios*, v. 12, n. 1, pp. 353-372.
<https://doi.org/10.17162/au.v12i1.974>
- Kaufmann-Argueta, Jürg** (2023). "20 Minutos alcanza a El País en el liderazgo de la audiencia digital en España". In: *Digital news report España 2023*, pp. 87-94. Pamplona: Servicio de Publicaciones Universidad de Navarra.
- Linden, Carl-Gustav** (2017). "Decades of automation in the newsroom. Why are there still so many jobs in journalism?". *Digital journalism*, v. 5, n. 2, pp. 123-140.
<https://www.doi.org/10.1080/21670811.2016.1160791>
- Marconi, Francesco** (2020). "Newsmakers: artificial intelligence and the future of journalism". New York: Columbia University Press. ISBN: 978 0 231549356
<https://doi.org/10.7312/marc19136>
- Méndez-Gurrola, Iris-Iddaly; Ramírez-Reyes, Abdiel; Mora-Gutiérrez, Román-Anselmo** (2020). "Aprendizaje automático aplicado en física: una revisión de literatura científica". *Research in computing science*, v. 149, n. 8, pp. 803-816.
<http://ri.uacj.mx/vufind/thumbnails/rupiada.png>
- Minguijón, Jaime; Serrano-Martínez, Cecilia** (2022). "La inteligencia artificial en los servicios sociales: estado de la cuestión y posibles desarrollos futuros". *Cuadernos de trabajo social*, v. 35, n. 2, pp. 319-329.
<https://doi.org/10.5209/cuts.78747>
- Moe, Hallvard; Ytre-Arne, Brita** (2022). "The democratic significance of everyday news use: using diaries to understand public connection over time and beyond journalism". *Digital journalism*, v. 10, n. 1, pp. 43-61.
<https://doi.org/10.1080/21670811.2020.1850308>
- Myers, Greg** (2004) *Matters of opinion: talking about public issues*. Cambridge: Cambridge University Press.
<https://doi.org/10.1017/CBO9780511486708>
- Noain-Sánchez, Amaya** (2022). "Abordando el impacto de la inteligencia artificial en el periodismo: la percepción de expertos, periodistas y académicos". *Comunicación y sociedad*, v. 35, n. 3, pp. 105-121.
<https://doi.org/10.15581/003.35.3.105-121>
- OpenAI**, (2022). *Online ChatGPT: optimizing language models for dialogue*.
<https://online-chatgpt.com>
- Parratt-Fernández, Sonia; Mayoral-Sánchez, Javier; Mera-Fernández, Montse** (2021). "The application of artificial intelligence to journalism: an analysis of academic production". *Profesional de la información*, v. 30, n. 3, e300317.
<https://doi.org/10.3145/epi.2021.may.17>
- Pascual, Manuel G.** (2023). "Expertos en inteligencia artificial reclaman frenar seis meses la 'carrera sin control' de los ChatGPT". *El País*, 29 de marzo de 2023.
<https://elpais.com/tecnologia/2023-03-29/expertos-en-inteligencia-artificial-reclaman-frenar-seis-meses-la-carrera-sin-control-de-los-chatgpt.html>
- Pavlik, John** (2000). "The impact of technology on journalism". *Journalism studies*, v. 1, n. 2, pp. 229-237.
<https://doi.org/10.1080/14616700050028226>
- Pavlik, John** (2023). "Collaborating with ChatGPT: considering the implications of generative artificial intelligence for journalism and media education". *Journalism & mass communication educator*, v. 78, n. 1, pp. 84-93.
<https://doi.org/10.1177/10776958221149577>

Peña-Fernández, Simón; Meso-Ayerdi, Koldobika; Larrondo-Ureta, Ainara; Díaz-Noci, Javier (2023). "Without journalists, there is no journalism: the social dimension of generative artificial intelligence in the media". *Profesional de la información*, v. 32, n. 2, e320227.

<https://doi.org/10.3145/epi.2023.mar.27>

Plantin, Christian (2012). *La argumentación: historia, teorías, perspectivas*. Buenos Aires: Biblos. ISBN: 978 950 7869815

Salazar-García, Idoia-Ana (2018). "Robots and artificial intelligence. New challenges of journalism". *Doxa comunicación*, n. 27, pp. 295-315.

<https://doi.org/10.31921/doxacom.n27a15>

Sánchez-García, Pilar; Merayo-Álvarez, Noemí; Calvo-Barbero, Carla; Diez-Gracia, Alba (2023). "Spanish technological development of artificial intelligence applied to journalism: companies and tools for documentation, production and distribution of information". *Profesional de la información*, v. 32, n. 2, e320208.

<https://doi.org/10.3145/epi.2023.mar.08>

Sánchez-Gonzales, Hada; Sánchez-González, María (2020). "Conversational bots used in political news from the point of view of the user's experience: Politibot". *Communication & society*, v. 33, n. 4, pp. 155-168.

<https://doi.org/10.15581/003.33.4.155-168>

Tejedor-Calvo, Santiago; Cervi, Laura; Pulido, Cristina M.; Pérez-Tornero, José-Manuel (2022). "Análisis de la integración de sistemas inteligentes de alertas y automatización de contenidos en cuatro cibermedios". *Estudios sobre el mensaje periodístico*, v. 27, n. 3, pp. 973-983.

<https://doi.org/10.5209/esmp.77003>

Túñez-López, José-Miguel; Feiras-Ceide, César; Vaz-Álvarez, Martín (2021). "Impacto de la inteligencia artificial en el periodismo: transformaciones en la empresa, productos, contenidos y perfil profesional". *Comunicación y sociedad*, v. 34, n. 1, pp. 177-193.

<https://doi.org/10.15581/003.34.1.177-193>

Túñez-López, José-Miguel; Toural-Bran, Carlos; Cacheiro-Requeijo, Santiago (2018). "Uso de bots y algoritmos para automatizar la redacción de noticias: percepción y actitudes de los periodistas en España". *El profesional de la información*, v. 27, n. 4, pp. 750-758.

<https://doi.org/10.3145/epi.2018.jul.04>

Ufarte-Ruiz, María-José; Murcia-Verdú, Francisco-José; Túñez-López, José-Miguel (2023). "Use of artificial intelligence in synthetic media: first newsrooms without journalists". *Profesional de la información*, v. 32, n. 2, e320203.

<https://doi.org/10.3145/epi.2023.mar.03>

Van-Noordt, Colin; Misuraca, Gianluca (2022). "Artificial intelligence for the public sector: results of landscaping the use of AI in government across the European Union". *Government information quarterly*, v. 39, n. 3, e101714.

<https://doi.org/10.1016/j.giq.2022.101714>

Vázquez-Herrero, Jorge; López-García, Xosé; Irigaray, Fernando (2020). "The technology-led narrative turn". In: Vázquez-Herrero, Jorge; Direito-Rebollal, Sabela; Silva-Rodríguez, Alba; López-García, Xosé (eds.). *Journalistic metamorphosis. Studies in big data*. v. 70, pp. 29-40. ISBN: 978 3 030 36315 4

https://doi.org/10.1007/978-3-030-36315-4_3

Wagner, Gerit; Lukyanenko, Roman; Paré, Guy (2022). "Artificial intelligence and the conduct of literature reviews". *Journal of information technology*, v. 37, n. 2, pp. 209-226.

<https://doi.org/10.1177/02683962211048201>

Zelizer, Barbie (2017). *What journalism could be*. United Kingdom: Polity Press. ISBN: 978 1 509 50786 3



SEDIC

Rediseñando lo que somos
para renovar lo que hacemos

<https://www.sedic.es>
<https://intranetsedic.es>
<https://formacionsedic.online>
c/Gargantilla 13, local 24 Madrid 28005
+34 639 186 570 | +34 91 593 40 59
sedic@sedic.es



<https://twitter.com/SEDIC20>



<https://www.facebook.com/AsociacionSEDIC>



<https://www.linkedin.com/company/sedic/>



<https://www.instagram.com/sedicasociacion/>

Ibero-American journalism in the face of scientific disinformation: Fact-checkers' initiatives on the social network *Instagram*

Juan-Ignacio Martín-Neira; Magdalena Trillo-Domínguez; María-Dolores Olvera-Lobo

Nota: Este artículo se puede leer en español en:
<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87250>

Recommended citation:

Martín-Neira, Juan-Ignacio; Trillo-Domínguez, Magdalena; Olvera-Lobo, María-Dolores (2023). "Ibero-American journalism in the face of scientific disinformation: Fact-checkers' initiatives on the social network *Instagram*". *Profesional de la información*, v. 32, n. 5, e320503.

<https://doi.org/10.3145/epi.2023.sep.03>

Manuscript received on 5th January 2023
Accepted on 29th April 2023



Juan-Ignacio Martín-Neira ✉

<https://orcid.org/0000-0001-9754-1177>

Universidad de Granada
Facultad de Comunicación e Información
Depto. de Información y Comunicación
Campus de Cartuja, s/n
18071 Granada, Spain
jmartinne@correo.ugr.es



Magdalena Trillo-Domínguez

<https://orcid.org/0000-0001-9505-1203>

Universidad de Granada
Fac. de Comunicación y Documentación
Grupo Joly
Campus Cartuja
18071 Granada, Spain
mtrillo@ugr.es



María-Dolores Olvera-Lobo

<https://orcid.org/0000-0002-0489-7674>

Universidad de Granada
Facultad de Comunicación e Información
Depto. de Información y Comunicación
Campus de Cartuja, s/n
18071 Granada, Spain
molvera@ugr.es

Abstract

The fight against disinformation is one of the major battles that journalism has had to face in recent years, especially after the coronavirus pandemic. As a counterbalance, "fact-checker" news media –platforms that have an important role in verifying whether or not the content circulating is true and that have harnessed the benefits of social networks, in spite of the difficulties inherent in these applications, to disseminate reliable and fact-checked content– have emerged. This study aims to explore how 10 prominent fact-checking accounts in Ibero-America use the social network *Instagram* to debunk false information, focusing in particular on the field of science and health. Applying a content analysis method using a checklist with quantitative and qualitative indicators, a total corpus of 240 posts from the first half of 2022 was obtained. The results allow us to determine which type of hoax has been used most as well as whether *Twitter*, *Facebook*, and *WhatsApp* are used for its dissemination. It was observed that health topics are the ones that attract the greatest interest from fact-checker accounts when it comes to creating fact-checks, and they tend to use static images or slide mode as opposed to video. In addition, they tend to use formal language in their presentation. This study also reveals that there were no instances of interaction with followers. While fact-checker accounts extensively use *Instagram* owing to its visual capabilities, they do not necessarily take advantage of its graphic potential. In addition, it was concluded that coronavirus is still a relevant topic for fact-checker media outlets, which must constantly refute the hoaxes that are mostly spread through social networks.

Keywords

Social networks; Social network analysis; Disinformation; Fake news; *Instagram*; Science journalism; Verification; Ibero-America; Social media; Infodemic; Health information; Coronavirus; Covid-19.



Funding

This work is funded by the *National Research and Development Agency [Agencia Nacional de Investigación y Desarrollo] (ANID)/ Human Capital Subdirectorate [Subdirección de Capital Humano] / Chile Foreign PhD Grants/2020 – 72210109*. As well as support and funding from the HUM-466 Group and the *Department of Information and Communication of the University of Granada*.

1. Introduction

Disinformation has become a major problem for today's digital, globalized, and hyperconnected society. With the amplifying effect of the coronavirus disease 2019 (Covid-19) pandemic, the transmission of hoaxes through digital platforms has increased exponentially, undermining informational veracity and even impacting the social role of the media –so much so that it has come to be considered a “key problem for contemporary democratic societies” (Sádaba; Salaverría, 2023, p. 18). If we add to this scenario the phenomenon of information overload or infodemic (WHO, 2020), it is understandable that it is not easy to determine what is accurate and what is false (Masip *et al.*, 2020). Indeed, so-called post-truth, almost always linked to the political sphere (Capilla, 2019; Rodríguez-Ferrández, 2019), has given rise to concepts such as misinformation and fake news, although the latter is not the most appropriate term (Maldita, 2021; Ferreras-Rodríguez, 2022). It could be said that fake news is considered a type of hoax that takes on the format of a news item and whose objective is disinformation.

Although disinformation campaigns tend to be centered on political and social messaging, science and health content is increasingly subject to such information disorders (Almansa-Martínez; Fernández-Torres; Rodríguez-Fernández, 2022). Clear proof of this can be seen in the Covid-19 situation, where the misleading information that has appeared on various digital platforms has been a central concern not only for the media but also for authorities and organizations (Greene-González; Cerda-Díez; Ortiz-Leiva, 2022). This circumstance has even created a mistrust of things that have been posted on social networks, thus fostering an environment in which public scares thrive (Ferreras-Rodríguez, 2022).

Examples of the fallacies that spread rapidly through various information channels are the hoaxes propagated by anti-vaxxers (García-Marín; Merino-Ortego, 2022) or the erroneous information about the harms of mask use (Villa-Gracia; Cerdán-Martínez, 2020). Part of the current information crisis, along with the scientific disinformation associated with health issues, are the nonsensical flat earth theories (Rodríguez; Giri, 2021) and fake content that have run rampant, casting doubt on climate change (León; López-Goñi; Salaverría, 2022).

In the face of this media, political, and social challenge, the considerable increase in the number of media outlets and journalists specialized in fact-checking data (Fernández-Barrero; López-Redondo, 2022) has helped, at least in part, to mitigate this reality. Whether generalist websites or independent sites, fact-checkers have embraced the opportunities provided by the Internet –in particular social networks– to cut through the noise and disseminate fact-checks to an important part of the population (Míguez-González; Abuín-Penas; Pérez-Seoane, 2021) that tends to be informed through social applications rather than through traditional media (Newman *et al.*, 2022).

This paper poses the following research questions:

- Q1. How do fact-checkers leverage *Instagram's* functionalities to debunk scientific hoaxes?
- Q2. What are the characteristics of disinformation content, and how does it go viral in the digital ecosystem?
- Q3. What specific and unique elements stand out in the use of the platform *Instagram* as a tool to combat disinformation?

Thus, the main objective is to identify and analyze the strategies used by science journalism to combat disinformation in the context of today's infodemic. To that end, we analyze notable fact-checker accounts from Ibero-America on the social network *Instagram* who use this platform to debunk scientific hoaxes, and evaluating the ways in which their professionals use the platform and the main aspects of the posts' content.

2. Context: Science journalism's role when it comes to disinformation

“Hoax”, a very widespread term, refers specifically to

“any intentionally false content that appears true that has been designed for the purpose of misleading the public and that is publicly disseminated by any platform or social media outlet” (Salaverría *et al.*, 2020, p. 4).

Indeed, disinformation uses lies as a persuasion technique and, therefore, is a dangerous tool that can break down societies' frameworks of belief (Rodríguez-Fernández, 2019).

In the context of the pandemic caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) virus, the management of information being posted and the imperative to provide objective and truthful content are a challenge for journalism (Aleixandre-Benavent; Castelló-Cogollos; Valderrama-Zurián, 2020). The uncontrolled rise of disinformation (López-Pujalte; Nuño-Moral, 2020) has the media and fact-checkers –who have seen how hoaxes disseminate through different platforms– up against the ropes. Even journalists themselves have helped spread erroneous information by various means, often owing to the pressure of immediacy (Mellado *et al.*, 2021).

Often, we are faced with news items that are created “without the intention of deceiving, but end up misinforming” (Gutiérrez-Coba; Coba-Gutiérrez; Gómez-Díaz, 2020, p. 238) and that are also given preference owing to excessive use of clickbait to attract audience, which ends up affecting the journalism (Maniou; Papa, 2023). García-Galera, Del-Hoyo-Hurtado, and Blanco-Alfonso (2020, p. 109) distinguish three intentions when it comes to disinformation in journalism:

- the journalist tries to deceive with invented or fabricated news items;
- the journalist tries to manipulate with inaccurate or biased news items; and
- the journalist tries to inform with wrong or erroneous news items.

However, regardless of these limitations as well as the targeted research on intentionality that could be carried out in addition to the present study, citizens are aware that, in times of crisis, it is imperative to rely on authorized sources (Fernández-Barrero; López-Redondo, 2022), and they still consider journalism to be responsible for tackling the challenges associated with misleading content (Rodríguez-Pérez, 2020).

Thus, although in some respects journalism is in the midst of an unprecedented crisis, the work of journalists and the media in combatting false information is crucial. In fact, the state of affairs has served as the impetus to

“protect and defend the quality and veracity of professional information” (Palomo; Sedano-Amundarain, 2018, p. 1,385).

In this way, science and fact-checking journalism have taken on a prominent role. It is more and more necessary to debunk the lies (Maiden *et al.*, 2020; Dunwoody, 2020), given that, since the beginning of the coronavirus pandemic, a growing audience has been consuming science journalism (Post; Bienzeisler; Lohöfener, 2021). This has enabled it, together with scientists and health personnel, to become one of the major sources of information (Massarani *et al.*, 2021), highlighting this discipline's relevance when it comes to understanding what is happening in the world of science (Lobato-Martínez; Monjas-Eleta; Gómez-García, 2022).

This influence goes hand in hand with the important role that journalism plays in providing accurate science communication for the public, allowing all people to have access to knowledge and reliable information (Vernal-Vilicic; Valderrama-Zenteno, 2022). From this perspective, it is essential for society to have good scientific literacy (Díaz-Moreno, 2019; Jarreau; Dahmen; Jones, 2019), which will ultimately lead journalism to effectively carry out its mission to mainstream and disseminate messages about science (Hu, 2022) without providing space for false information.

However, the battle against disinformation is often seriously affected by the viral strength of fallacies in the media, in particular on social networks, which allow hoaxes to spread easily (Costa-Sánchez; López-García, 2020; Montemayor-Rodríguez; García-Jiménez, 2021). Certainly, the overabundance of information –truthful or not– that society receives (Herrero-Diz; Pérez-Escolar, 2022) and the new platforms' potential for dissemination, having audiences that increasingly use social networks to stay informed (Newman, 2022), make it easier for hoaxes to go viral even faster than truthful messages (López-Pan; Rodríguez-Rodríguez, 2020).

The immediacy and the lack of geographical borders on the Internet (Gutiérrez-Coba; Coba-Gutiérrez; Gómez-Díaz, 2020) make it difficult to identify false content's origin or classify what the level of intentionality was. For this reason, many journalists warn that these applications are detrimental to the dissemination of truthful information (Varona-Aramburu; Sánchez-Muñoz, 2016), progressively causing a “lack of trust in the media” (Marta-Lazo; Rodríguez-Rodríguez; Peñalva, 2020, p. 62).

However, from the critical but also constructive perspective that underpins this research, we maintain that it is possible to find ways and solutions to address the problem of false information on these very platforms.

2.1. Social media as a platform for combatting hoaxes

In the last decade, a growing number of positive initiatives developed by social networks to debunk potentially false information has been documented. These initiatives are often linked to an increase in the number of fact-checking accounts on these platforms. Fact-checking journalism refers to initiatives focused on fact-checking data, statements, or news items circulating in the media or through networks (Míguez-González; Abuín-Penas; Pérez-Seoane, 2021) and, during the pandemic, gained importance as an essential resource for checking the veracity of information (Guallar *et al.*, 2020; Ferreras-Rodríguez, 2022).

These media outlets or fact-checking journalists have leveraged the benefits offered by social applications to combat disinformation, given the fact that these platforms are presented as “main outlets for interactive and multimedia content for rapid consumption” (Sánchez-González; Sánchez-Gonzales; Martos, 2022, p. 137). They employ esthetic, simple, colorful, and informatively direct elements (Fernández-Barrero; López-Redondo, 2022) and also engage audiences in the work of fact-checking (Aguado-Guadalupe; Bernaola-Serrano, 2020). In addition, since the information crisis arose after the pandemic (Azer; Blasco-Arcas; Harrigan, 2021), it has been observed how massive networks such as *Facebook*, *Twitter*, or *TikTok* have implemented coordinated actions

Science and health content is increasingly subject to such information disorders

“to disprove false content on their own platforms” (Sidorenko-Bautista; Alonso-López; Giacomelli, 2021, p. 90), which has supported the work of fact-checker accounts.

In this regard, *Instagram* stands out for the key role it is playing in terms of posting and fact-checking strategies. Thus, since the recent pandemic, this network has experienced enormous growth both in the volume of posts related to fact-checking, with an average of 1.3 posts per day, as well as in its own user community (Míguez-González; Abuíñ-Pernas; Pérez-Seoane, 2021). Indeed, it is one of the platforms with the greatest increase in number of Internet users, and numerous journalistic accounts use it to disseminate information owing to its graphic and interactive potential (Martín-Neira; Trillo-Domínguez; Olvera-Lobo, 2022). Likewise, fact-checker accounts prefer it to other networks because it has high levels of interaction and allows for the formation of communities of followers (Sánchez-González; Sánchez-González; Martos, 2022). This highlights the fact that, owing to its ability to reach audiences who are young and do not necessarily obtain information from traditional media, *Instagram* is an ideal application through which to provide accurate information and guide users to credible sources (Malik; Khan; Quan-Haase, 2021).

Additionally, many of the most important fact-checker accounts have focused their activity on this platform (Dafonte-Gómez; Míguez-González; Ramahí-García, 2022). They opt for audio and video formats in line with what the market currently offers (Dafonte-Gómez; Corbacho-Valencia; García-Mirón, 2021) and make use of the tools within the application, such as featured stories or live broadcasts, as a fact-checking strategy (Sánchez-González; Sánchez-González; Martos, 2022).

2.2. Fact-checking and social networks: The case of Ibero-America

Although the issue of disinformation is a global phenomenon, it is no secret that local contextual factors determine how false content is dealt with (Herrero; Herrera-Damas, 2021). As Ibero-America has two closely related languages, Spanish and Portuguese, as well as similar cultural elements (García-Crespo; Ramahí-García; Dafonte-Gómez, 2021), in addition to the obvious geographical proximity, it is an interesting zone to analyze. Another common characteristic is that most of the content verification initiatives that have emerged in these countries were created outside of the traditional media and the legacy media system (Rodríguez-Pérez *et al.*, 2022). Adding to this the sustained rise in Ibero-America in news consumption through social networks (Gutiérrez-Coba; Coba-Gutiérrez; Gómez-Díaz, 2020) makes this group of countries an interesting object of analysis within the framework of fact-checking strategies used in these applications.

Guallar *et al.* (2020), with their review of posts on disinformation and hoaxes in Ibero-America, highlight the existing interest in this topic and the proliferation of different types of analysis –some focused on different fact-checking platforms or specific case studies. In this context, the work carried out by leading platforms in this area, such as *Chequeado* (Argentina) –which has taken on a significant role in the development of training workshops aimed at counteracting information hoaxes– stands out (García-Crespo; Ramahí-García; Dafonte-Gómez, 2021). In this way, media outlets from Spain, Argentina, Colombia, Brazil, Mexico, and Ecuador have taken the lead in this geographic area with respect to posting fact-checks, thus bringing the content generated to millions of people (Martínez-Rolán; Dafonte-Gómez, 2022).

Fortunately, therefore, in this geographical environment there were well-established leading fact-checkers when the health crisis caused by Covid-19 and the disinformation problems it created occurred. In this sense, words such as “plandemia” and other “anti-scientific” concepts were the focal point of conversations on social networks and in the false content that were being spread in Ibero-America (García-Marín; Merino-Ortego, 2022). At the same time, on numerous occasions the disinformation became “transnational,” and the hoaxes circulated more widely, and as a result, they tended to be echoed across the different countries of Ibero-America, although the way they were presented varied (Gutiérrez-Coba; Coba-Gutiérrez; Gómez-Díaz, 2020).

In a recent study on the perceptions of fact-checking journalists in Ibero-America (Rodríguez-Pérez *et al.*, 2022), which identifies a set of characteristics that these professionals have that enable them to undertake activities in these countries, most of the fact-checkers surveyed point out that the discipline in Ibero-America has the aim and purpose of combatting the disinformation disseminated through social networks. This circumstance, together with the high penetration of these applications and the growing consumption of news, makes it interesting to delve into fact-checking initiatives in this geographical area and their role in combatting scientific disinformation in social networks.

Owing to fact-checker accounts’ growing use of the platform *Instagram* to debunk scientific hoaxes, and the fact that there is no clear systematization of the most appropriate posting strategies, formats, narratives, and styles for the dissemination of fact-checking information, an important space for exploration and analysis has opened up, and this is the focus of our research. In addition, the expertise of the teams of fact-checking professionals with health, science, or environmental content on *Instagram*, which could even be transferred to practices in other newer networks such as *TikTok* or *Twitch*, provides an opportunity to explore this field of study in greater depth.

“Thus, since the recent pandemic, *Instagram* has experienced enormous growth both in the volume of posts related to fact-checking”

3. Methodology

The work presented here provides a constructive view of the use of Instagram as a tool to combat disinformation, rigorously analyzing initiatives focused on solving social and public problems, presenting work praxis and results, and setting out achievements and limitations (Casares-Corrales, 2021). The methodology used includes analyzing prominent Ibero-American accounts with fact-checker content that are present on the social network Instagram along with their posts on fact-checked disinformation. Thus, we used the technique of content analysis (Andréu-Abela, 2000), which enabled us to interpret and systematize communicative content

“with the aim of making justified logical deductions concerning the source” (Andréu-Abela, 2000, p. 3).

The study focuses specifically on the social network *Instagram* because this application has more than 1 billion users and has become an important platform for communicating science (Caspari, 2022). In addition, according to the latest *Digital News Report* (Newman, 2022), young people are increasingly accessing news using this platform, at the expense of traditional media such as television or the printed press. It has also been pointed out that this network allows information fact-checkers to work in a “more constructive” space and gives them the opportunity to develop more educational content (Sánchez-González; Sánchez-Gonzales; Martos, 2022, p. 153).

The sample of media and communication accounts included in this study (Table 1), following works such as that of **Da-fonte-Gómez, Corbacho-Valencia, and García-Mirón** (2021), is made up of Ibero-American journalistic initiatives recognized for their work in data verification and for being reliable sources in the face of disinformation. For this purpose, the inclusion criterion used was that, between January and June 2022, the selected media outlet was an active signatory of the codes of principles of the *International Fact-Checking Network* –the main international fact-checking network in the world (García-Marín, 2020), which promotes excellence in fact-checking (Poynter, 2022). It was decided that, to obtain a more diverse view of existing initiatives in the Ibero-American context, a single account would be selected for each country. Understanding that there are countries in which there may be more than one account that meets the requirements, we chose these accounts by considering the number of followers that each profile had, the frequency of posts, and in particular, the variety of topics that were fact-checked, including science, health, and technology content. Thus, we combined quantitative and qualitative criteria when making our selection.

The posts analyzed were in Spanish or Portuguese and were aimed at explicitly debunking misleading information in the areas of science, health, environment, and technology. In this selection of posts, we have discarded

- news posts on the indicated topics that were not specifically aimed at debunking a hoax spread on the network or through certain media outlets and
- posts that only provided an image and directed users to a website or other type of resource without presenting complete or sufficient information on *Instagram* to identify the reasons for the debunking –that is, they amplify the message rather than acting as a dedicated communication channel.

The analysis focused on posts from the first half of 2022 (January 1-June 30) and was carried out by the authors of this paper –specialists in information technology, communication, and journalism.

Table 1. Characteristics of the fact-checking accounts on *Instagram* in the study sample

User profile	URL	Country	Date of creation	Followers	Posts
@agencia_lupa	https://www.instagram.com/agencia_lupa	Brazil	July 2018	434,000	3,745
@boliviaverifica	https://www.instagram.com/boliviaverifica	Bolivia	July 2019	1,300	255
@chequeado	https://www.instagram.com/chequeado	Argentina	April 2015	75,600	910
@colombiacheck	https://www.instagram.com/colombiacheck	Colombia	October 2018	20,000	1,452
@cotejainfo	https://www.instagram.com/cotejainfo	Venezuela	May 2016	2,300	1,014
@ecuador_chequea	https://www.instagram.com/ecuador_chequea	Ecuador	January 2019	4,200	3,476
@fastcheckcl	https://www.instagram.com/fastcheckcl	Chile	October 2019	250,000	2,020
@jornalpoligrafo	https://www.instagram.com/jornalpoligrafo	Portugal	October 2018	39,900	5,984
@malditobulo	https://www.instagram.com/malditobulo	Spain	March 2017	64,600	755
@pajaropolitico	https://www.instagram.com/pajaropolitico	Mexico	August 2014	101,000	1,673

Data collected in November 2022.

To determine how posts with false content about science are debunked in posts on the *Instagram* feeds of the fact-checker accounts in our study sample, a checklist (Table 2) was designed based on the proposals of **Salaverría et al.**, 2020; **Aguado-Guadalupe; Bernaola-Serrano**, 2020; **Míguez-González; Abuín-Penas; Pérez-Seoane**, 2021; **Hoyos-Simbaña; Lara-Aguiar; Mila-Maldonado**, 2022; **Sánchez-González; Sánchez-Gonzales; Martos**, 2022; **Martín-Neira; Trillo-Domínguez; Olvera-Lobo**, 2022; 2023b.

In addition to journalistic criteria, such as the type of source used to debunk the hoax or the type of disinformation disseminated, the checklist includes indicators and descriptors aimed at identifying the audiovisual characteristics of the posts, such as the way in which the fact-check was presented or the extent to which *Instagram* features were used. It also notes the platform from which the disinformation began to spread, the main type of topic of the hoax, and other indicators associated with post engagement. In addition, we have considered whether the accounts help users become media literate (Herrero-Diz; Pérez-Escolar; Varona-Aramburu, 2022; Sádaba; Salaverría, 2023), either by offering support tools to address disinformation or by suggesting websites that enrich answers to users' concerns.

The vast majority ($n = 183$) of disinformative content in the scientific domain in our sample tended to apply deception tactics

It is important to note that all the results obtained in the analysis and the descriptors that were identified in each post were based on what the fact-checker accounts posted, whether that included video, an infographic, *slides*¹, or accompanying text, and only occasionally was information obtained from sources other than *Instagram*. When in doubt, the analysis carried out by the authors of this study determined which were the prominent ones, and the ones that were the least debatable and that were in line with the indicator proposed in the checklist were chosen. Taking this premise into account, there was no coding based on two types of indicators, and the analyses presented in this study are those that are closest to the established criteria and indicators. For this paper, we did not attempt to analyze the intentions of the hoaxes' disseminators, as we consider this to be an undertaking that would justify a specific study complementary to our focus of study, which, we reiterate, is centered on analyzing the strategies used by the fact-checker accounts to counter disinformative messages.

Table 2. Checklist for analyzing posts from fact-checking accounts on *Instagram*

Indicator	Descriptors	Descriptor explanation
Characteristics of the post	Date	Elements that make it possible to characterize and describe the post being analyzed.
	Number of comments	
	Number of likes	
	Fact-check initiated by the media outlet or by users?	Action that provided insight into whether the audience suggested or previously advised the fact-check.
Type of disinformation	Parody	Dissemination of false content with an intention of mocking (Salaverría et al., 2020).
	Decontextualization	Facts or real statements in the wrong context (Salaverría et al., 2020).
	Deception	Complete falsification of facts, in which content is fabricated with the intention of making the public believe in its truthfulness (Salaverría et al., 2020).
	Exaggeration	Has links to the truth, but crosses the limits of the true and enters the realm of falsehood (Salaverría et al., 2020).
Subject of the scientific hoax	Astronomy	Topics chosen as points of reference for the posts; a distinction is made between (general) health and the coronavirus to understand how Covid-19 has affected the patterns.
	Earth sciences	
	The environment and natural sciences	
	Health –coronavirus	
	Health –general	
	Technology	
Platform originally used to spread the hoax	Facebook	Taking into consideration what the fact-checking accounts expressed, either explicitly at a textual level or with the reference images used in the post, the platform determined to be the origin of the dissemination of the hoax.
	Instagram	
	Traditional (legacy media) or digital media	
	Website	
	Telegram	
	TikTok	
	Twitter	
	WhatsApp	
	Other	
	Unidentified	

Indicator	Descriptors	Descriptor explanation
How the fact-check is presented	Use of text and static infographic image	Taking into account the options available on <i>Instagram</i> , the descriptors established as the basic characteristics that a post can have and that have been seen in posts from the fact-checking accounts.
	Use of text and static infographic image or video (<i>slide type</i>)	
	Use of text and explainer video (digital animation)	
	Use of text and explainer video with on-camera communicator	
	Video reused from another platform (e.g., <i>TikTok</i>)	
If video is the main feature, what is the duration?	1-30 seconds	The duration intervals that were established to be the timeframes commonly used in videos on social networks (Martín-Neira; Trillo-Domínguez; Olvera-Lobo, 2023b).
	31 seconds to 1 minute	
	1 minute 1 second to 2 minutes	
	2 minutes 1 second or more	
	Not applicable	
Predominant language used to debunk the hoax	Formal –serious	The two approaches identified in fact-checkers' posts on social networks according to the literature reviewed (Sánchez-González; Sánchez-Gonzales; Martos, 2022): - formal language and serious audiovisual content, avoiding humor, or - more informal language (visual or textual) and based on humor. A mixture of both (either with image/video and text), was considered a "Combination of both."
	Informal –based on humor –meme	
	Combination of both	
Type of source used to debunk (may include more than one)	Comparison with other news items or original posts	According to the posts from the fact-checking accounts themselves, they determine how to use these types of sources; in some posts, the combination of more than one can be seen.
	Documentary sources	
	Official sources (organizations/institutions)	
	Artificial or fact-checking tools	
	Expert opinion or source (individuals)	
	Other	
Application of audiovisual elements and network languages (may include more than one)	Presenter overdub and voice-over	Taking the audiovisual characteristics and the elements inherent in the network <i>Instagram</i> as a reference, the series of descriptors established that could be used in posts from the fact-checker accounts; more than one can be included in the different posts.
	For video, a thumbnail designed for the feed	
	Static image with design	
	Subtitles (of the voice-over) in the video	
	Animated text in the video	
	Use of emojis	
	Use of GIFs	
	Use of hashtags in text	
	Use of background music	
Use of video with voice (speaker not seen)		
Level of interaction with users	Responded to comments?	Whether there was a dialog with the users and whether in those answers other tools were offered to continue providing information to the user.
	Response given and tools and content that allow the user to become media literate provided	

Based on: Salaverría *et al.*, 2020; Aguado-Guadalupe; Bernaola-Serrano, 2020; Míguez-González; Abuín-Penas; Pérez-Seoane, 2021; Hoyos-Simbaña; Lara-Aguilar; Mila-Maldonado, 2022; Sánchez-González; Sánchez-Gonzales; Martos, 2022; Martín-Neira; Trillo-Domínguez; Olvera-Lobo, 2022; 2023b.

The analysis of the posts from the fact-checker accounts posted within the established period ($N = 3,338$) identified a total of 240 posts (7.2%) aimed at debunking hoaxes in the areas of science, health, environment, and technology. It is important to point out that, in the accounts analyzed, several posts were detected that debunked scientific hoaxes but did not offer complete or sufficient information to establish the reasons for the fact-check, with the response leading to the account's website, which means that they did not meet part of the criteria for inclusion in the analysis.

The *Instagram* accounts in the sample were predominately fact-checks of political information, as well as statements issued by public authorities, which are outside the scope of this study, although they point to an interesting line of study. However, an overall review of the posts on science, health, environment, and technology showed how initially the fact-checks were related to the coronavirus and these set the tone of the accounts analyzed,

“ It was observed that almost 83% of the posts were presented as static images or *slide-type* infographics that in some cases included small videos ”

even being on par with the political topics that had been the predominant type of information up to that time. However, since the beginning of the war in Ukraine, the agenda had begun to change, with this topic taking center stage in each account’s posts, thus adding to the content related to the socio-political situations of these countries.

4. Results

Although political or social issues are the most common topics on the fact-checker accounts, the analysis determined that science and technology are gaining ground in news patterns, especially on accounts such as @fastcheckcl, @agencialupa, @boliviaverifica, or @malditobulo, which have viewed health issues as an important subject for informative review. Table 3 summarizes the sample of analyzed posts from each fact-checker account on *Instagram*.

The analysis of readers’ reactions (Table 4) showed that there were two very marked poles in terms of the number of likes each post received. Thus, more than 40% of the 240 posts analyzed did not receive more than 100 likes, even on accounts that have many thousands of followers. In contrast, over 30% of the posts reviewed had more than 1,000 likes –numbers that generally came from the accounts with the most followers.

Meanwhile, when it came to comments made on each post, the gap between the extremes was greater. Around 60% of the content did not receive more than 20 comments, and even among these, there were many posts that had 1 or no comments from followers. In contrast, a scant 5% of the content examined exceeded the 200-comment barrier.

It should also be noted that in only 2.1% of the posts could it be determined that users or followers had suggested the content used to create the fact-checks, demonstrating a preference for posts generated by the accounts themselves.

Finally, a very striking fact is that, in 100% of the posts reviewed, there was no interaction with users and no response was given to the questions posed by them.

In terms of the disinformation strategies used to produce content confirmed to be false, the vast majority ($n = 183$) of disinformative content in the scientific domain in our sample tended to apply deception tactics (Table 5). Indeed, although fact-checker accounts are responsible for debunking techniques such as exaggeration or decontextualization, in all the media outlets analyzed, deception was the most recurrent element in the disinformation analyzed. A substantial number ($n = 193$) of the post showed a direct relationship to Health-coronavirus or general health issues. Specifically, the fact-checker accounts particularly focused on debunking hoaxes associated with vaccines or their effects.

Regarding the source of dissemination, most of the hoaxes came from *Facebook* and *Twitter*, being the main disinformation platforms. False information transmitted via email, *Twitch*, or statements by politicians in public debates or interviews was also found. For a significant percentage (18.3%) of the fact-checks posted on the factchecker accounts analyzed, it was not possible to determine the origin of the hoax because they spoke generically about social networks or did not mention the source. A single case stands out in which the hoax came from a media outlet, which reinforces the argument that most disinformation originates and spreads through the Internet and social networks.

Table 3. Posts analyzed from each fact-checker account

User profile	Total posts (January–June 2022)	Posts on the selected topics	Percent analyzed out of the total sample
@agencia_lupa	377	46	12.2
@boliviaverifica	135	18	13.3
@chequeado	212	10	4.7
@colombiacheck	183	9	4.9
@cotejoinfo	96	3	3.1
@ecuador_chequea	1,169	65	5.6
@fastcheckcl	342	51	14.9
@jornalpoligrafo	559	16	2.9
@malditobulo	117	17	14.5
@pajaropolitico	148	5	3.4
Total	3,338	240	7.2

Table 4. Readers’ reactions and interaction with users regarding posted fact-checks

Indicator	Descriptors	Number of posts	%
Likes	0–100	104	43.3
	101–200	11	4.6
	201–500	25	10.4
	501–1,000	27	11.3
	1,001+	73	30.4
Comments	0–20	146	60.8
	21–50	33	13.8
	51–100	28	11.7
	101–200	21	8.8
	201+	12	5.0
Fact-check	Media outlet initiative	235	97.9
	User initiative	5	2.1
Interaction	No response to comments	240	100

Table 5. Disinformation strategies and source of dissemination of content confirmed to be false

Indicator	Descriptors	Number of posts	%
Type of disinformation	Deception	183	76.3
	Exaggeration	34	14.2
	Decontextualization	22	9.2
	Parody	1	0.4
Topic of the hoax	Health–coronavirus	138	57.5
	Health–general	55	22.9
	The environment and natural sciences	14	5.8
	Astronomy	14	5.8
	Technology	10	4.2
	Earth sciences	9	3.8
Platform of dissemination	Facebook	67	27.9
	Twitter	43	17.9
	WhatsApp	25	10.4
	Instagram	17	7.1
	Websites	12	5.0
	TikTok	8	3.3
	Telegram	6	2.5
	Traditional (legacy media) or digital media	1	0.4
	Other	17	7.1
Unable to identify the origin	44	18.3	

Taking into consideration that most of the false content posted dealt with issues related to coronavirus and health in general ($n = 193$), Table 6 details the social platforms on which the disinformation was spread. As can be seen, in both cases *Facebook* and *Twitter* were the preferred social networks for the dissemination of hoaxes, although the latter was the most prolific in fact-checked disinformation about Covid-19, with 22.5% of the content. The platforms considered in this analysis were determined by the fact-checker accounts themselves in their posts and, in the case of those posts whose origin could not be established, the “Unidentified” criterion was applied.

In relation to the techniques used on the social network *Instagram* to present debunked disinformation

Table 6. Platforms on which the health hoaxes were posted

Platform \ Topic	Health – general ($n = 55$)		Health – coronavirus ($n = 138$)	
	n	%	n	%
Facebook	16	29.1	25	18.1
Twitter	8	14.5	31	22.5
WhatsApp	3	5.5	20	14.5
Instagram	3	5.5	14	10.1
TikTok	2	3.6	5	3.6
Telegram	1	1.8	5	3.6
Websites	1	1.8	8	5.8
Unidentified	12	21.8	24	17.4
Other	9	16.4	6	4.3

Table 7. Techniques used on *Instagram* to disprove scientific hoaxes

Indicator	Number of posts	%	
How the fact-check is presented	Use of text and static infographic image or video (<i>slide type</i>)	119	49,6
	Use of text and static infographic image	80	33,3
	Use of text and explainer video with on-camera speaker	17	7,1
	Use of text and video explainer (digital animation)	16	6,7
	Video reused from another platform (e.g., <i>TikTok</i>)	8	3,3
For video	1–30 seconds	8	3,3
	31 seconds to 1 minute	23	9,6
	1 minute 1 second to 2 minutes	8	3,3
	2 minutes 1 second or more	3	1,3
	Not applicable	198	82,5
Language used	Formal –serious	226	94,2
	Informal –based on humor –meme	10	4,2
	Combination of both	4	1,6

(Table 7), it was observed that almost 83% of the posts were presented as static images or *slide*-type infographics that in some cases included small videos. Just over 14% of posts used video as the primary means of debunking fake content, and about 3% were repurposed posts from *TikTok*. Only 17.5% used videos as a tool to debunk a hoax on this media outlet. Of the 42 audiovisual posts identified, a large part (9.6%) have a duration of between 31 and 60 seconds.

Regarding the type of language used, the vast majority (94.2%) of the posts were serious in tone and used formal language. About 4% used humor, and less than 2% used a combination of both, for example, a humorous image and serious text. Figure 1 shows an example of a post that mixed a humorous image (meme) with formal and more serious text, after a statement by the president of Colombia (Torrado, 2022).

As for the sources used by the fact-checker accounts analyzed to debunk hoaxes (Table 8), it is noteworthy that most of the posts relied on two or more types of sources to verify the information and recognize it as false. Fact-checks based on comparison with other news items or with the original posts on the network were the most common (20.4%). Considering this action individually or combined with other strategies for fact-checking the content, over 55% of the posts were checked against information or news that already existed on the web or on another platform. Similarly, the use of official sources

(from organizations or institutions) and combining these with other fact-checking strategies is another of the most widely used techniques to disprove false content, at almost 44% of the posts.

It is also interesting to note that the use of artificial or screening tools, such as a reverse image search through a platform, was rarely mentioned as a means of corroborating information, and its use was explicitly indicated in only three posts (1.3%).

When creating the fact-checks, the fact-checker accounts used various audiovisual and other features specific to *Instagram* in the way they considered most appropriate. Thus, the analysis of the posts in the sample identified more than 20 different combinations (Table 9). It was observed that 83.4% of the total number of posts consisted of a static infographic image or a static *slide*-type image (and video). This led to the generation of various combinations and the use of diverse graphic and hypertextual elements. Of the posts, 14.2% resorted to the most basic element selected for analysis, which was to create a static image with certain types of design. Meanwhile, 21.3% of the posts also included hashtags in the descriptive text accompanying the static image.

However, the combination “static image with design + use of emojis + use of hashtags in the text” was the strategy most used to tackle disinformation from *Instagram*, at 25.4% of the total analyzed. As for videos, which accounted for less



Figure 1. Combination of serious language with humorous image.²
Source: *Instagram* @colombiacheck

Table 8. Type of sources used to debunk hoaxes

Indicator	Number of posts	%
Comparison with other news items or original post	49	20.4
Comparison with other news items or original post Official sources (organizations/institutions)	36	15.0
Official sources (organizations/institutions)	35	14.6
Comparison with other news items or original post Documentary sources	26	10.8
Documentary sources Official sources (organizations/institutions)	19	7.9
Opinion or expert source (individuals)	18	7.5
Comparison with other news items or original post Opinion or expert source (individuals)	17	7.1
Documentary sources	16	6.6
Opinion or expert source (individuals) Official sources (organizations/institutions)	10	4.2
Opinion or expert source (individuals) Documentary sources	5	2.1
Comparison with other news items or original post Documentary sources Official sources (organizations/institutions)	3	1.3
Comparison with other news items or original post Use of artificial tools or checking tools	3	1.3
Other combinations and types of sources (with a percentage less than 1%)	3	1.3

than 17% ($n = 42$) of the total number of posts in the sample, there was no clear trend. The vast majority of the accounts had their own way of presenting fact-checks, but it was noted that, over time, they were progressively including new and more numerous elements in their posts. Table 9 summarizes the main indicator combinations used by the fact-checker accounts analyzed in this study.

Table 9. Audiovisual components and *Instagram's* own features in the fact-checks analyzed

Indicator	Number of posts	%
Static image with design + use of emojis + use of hashtags in text	61	25.4
Static image with design + use of hashtags in text	51	21.3
Static image with design	34	14.2
Static image with design + use of emojis	31	12.9
Static image with design + use of video with voice (speaker not seen) + use of emojis + use of hashtags in the text	22	9.2
Use of emojis + use of hashtags in text + animated text in video	7	2.9
Use of emojis + presenter overdub and voiceover + use of hashtags in the text + thumbnail in the feed for videos + use of background music	6	2.5
Use of emojis + use of gifs + presenter overdub and voiceover + subtitles (of the voiceover) in the video + use of hashtags in the text + use of background music	4	1.7
Use of gifs + presenter overdub and voiceover + subtitles (of the voiceover) in the video + use of hashtags in the text + thumbnail in the feed for videos + use of background music	4	1.7
Use of emojis + presenter overdub and voiceover + use of hashtags in the text + thumbnail in the feed for videos + use of background music + animated text in the video	3	1.3
Use of emojis + presenter overdub and voiceover + use of hashtags in the text + use of background music + animated text in the video	3	1.3
Other combinations (with a percentage less than 1%)	14	5.6

By independently considering each fact-checking account analyzed, it is possible to determine the main combinations of audiovisual features and *Instagram's* own features that apply when posting fact-checks (Table 10).

Table 10. Audiovisual elements used by fact-checking accounts to post fact-checks

<i>Instagram</i> account (total posts)	Indicator	Number of posts	%
@agencia_lupa ($n = 46$)	Static image with design + use of video with voice (speaker not seen) + use of emojis + use of hashtags in the text	22	47.8
	Static image with design + use of emojis + use of hashtags in text	21	45.7
@boliviaverifica ($n = 18$)	Use of emojis + presenter overdub and voiceover + use of hashtags in the text + thumbnail in the feed for videos + use of background music	6	33.3
	Use of gifs + presenter overdub and voiceover + subtitles (of the voiceover) in the video + use of hashtags in the text + thumbnail in the feed for videos + use of background music	4	22.4
@chequeado ($n = 10$)	Static image with design + use of emojis + use of hashtags in text	4	40.0
	Use of emojis + use of gifs + presenter overdub and voiceover + subtitles in the video (of the voiceover) + use of hashtags in the text + use of background music	4	40.0
@colombiacheck ($n = 9$)	Static image with design + use of hashtags in text	8	88.9
	Static image with design	1	11.1
@cotejoinfo ($n = 3$)	Static image with design + use of emojis + use of hashtags in text	3	100
@ecuador_chequa ($n = 65$)	Static image with design + use of hashtags in text	36	55.4
	Static image with design	22	33.8
@fastcheckcl ($n = 51$)	Static image with design + use of emojis	31	60.8
	Static image with design	11	21.6
@jornalpoligrafo ($n = 16$)	Static image with design + use of emojis + use of hashtags in text	11	68.8
	Use of emojis + presenter overdub and voiceover + use of hashtags in text + use of background music + animated text in the video	3	18.8
@malditobulo ($n = 17$)	Static image with design + use of emojis + use of hashtags in text	9	52.9
	Use of emojis + use of hashtags in text + animated text in video	7	41.2
@pajaropolitico ($n = 5$)	Use of video with voice (speaker not seen)/ use of gifs + presenter overdub and voice-over + subtitles (of the voice-over) in the video + use of hashtags in the text + thumbnail in the feed for videos + use of background music + use of animated text in the video	2	40.0
	Static image with design + use of emojis + use of hashtags in text	2	40.0

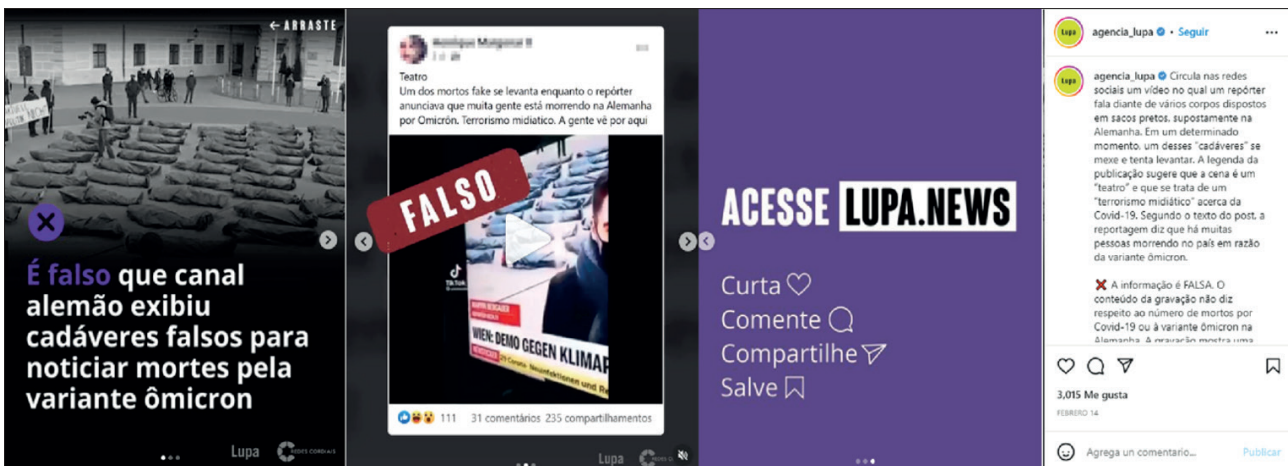


Figure 2. Posts of fact-checked information from *Agência Lupa* with video incorporated.³
Source: *Instagram @agencia_lupa*

The analysis of each individual account shows that a high percentage of *slide*-type posts displayed content as a static image, in addition to using elements such as hashtags and emojis in the text. However, the tactic applied by *Agência Lupa*, which incorporates video in a high percentage (47.8%) of these static *slide*-type posts as an audiovisual support to the informative fact-checking, stands out (Figure 2). Also refreshing are the offerings of *Bolivia Verifica*, which favors the use of video to debunk hoaxes, or *Animal Político* (@pajaropolitico), which relies on creating extensive reports as a strategy to counter disinformation.

5. Discussion

Instagram has positioned itself as a well-established platform for combatting this fake content (Míguez-González; Abuín-Penas; Pérez-Seoane, 2021). This study has made it possible to identify the way in which the fact-checker accounts on this social network are using this platform to debunk scientific hoaxes and how they have applied the features of this application to produce posts that effectively reach their followers and also apply journalistic rigor to debunk the different types of hoaxes that circulate on the network.

In response to the research question (Q1)

How do fact-checkers leverage *Instagram*'s functionalities to debunk scientific hoaxes?

the analysis showed that there was a preference for using static images and *slide*¹ mode over the using videos or the application's famous reels. This is despite the recent popularity of these formats. Therefore, although social networks with audiovisual elements are certainly an ideal place to communicate science (Zeng; Schäfer; Allgaier, 2021), the results confirm previous studies because on *Instagram* static images are still preferred over posting videos (Habibi; Salim, 2021), even by science journalists, who do not make use of much of this social network's potential (Martín-Neira; Trillo-Domínguez; Olvera-Lobo, 2022).

This research has revealed that the average length of the videos used for the debunking of hoaxes –from 31 seconds to 1 minute– corresponds to the time considered appropriate for audiovisual products in the field of science or health (Habibi; Salim, 2021; Gurler; Buyukceran, 2022). Regarding the use of elements that accompany the text of each post, such as hashtags and emojis, it is evident that 85% of the posts include one or both elements. Indeed, these hashtags and symbols, when used appropriately, can help improve posts' engagement levels (Jiang; Guo; Ma, 2020). Even so, using hashtags does not guarantee higher traffic or readership, as experiments show that there are no specific strategies for boosting the visibility of science content when using hashtags (Sidorenko-Bautista; Cabezuelo-Lorenzo; Herranz-de-la-Casa, 2021).

When answering the research question (Q2)

What are the characteristics of disinformation content and how does it go viral in the digital ecosystem?

it can be determined that the topic "Health-coronavirus" was the subject that attracted the most interest when it came to debunking hoaxes, also conditioned by the date of sample selection, a time when the pandemic was still one of the central issues on the public agenda. This is probably not surprising, since we have all experienced how Covid-19 has brought to the fore an abundance of hoaxes and false content that the media have had to deal with and, at the same time, has driven the creation of various initiatives designed for the fact-checker (Villa-Gracia; Cerdán-Martínez, 2020; Salaverría et al., 2020; García-Marín, 2020; Molina-Cañabate; Sánchez-Duarte; Magallón-Rosa, 2021; Herrero-Diz; Pérez-Escolar, 2022; Almansa-Martínez; Fernández-Torres; Rodríguez-Fernández, 2022; León et al., 2022).

However, the fact that 23% of the hoaxes were related to health (as a general topic) reflects that this is the area of greatest concern to which attention should continue to be paid, beyond the coronavirus contingencies. Rodríguez-Fernán-

dez (2019) already warned about this in her research on professional challenges for the communications sector, when she stated that health disinformation was among the disinformation most frequently described by fact-checkers. Something similar was detected by **Montero-Liberona** and **Halpern** (2019), showing how fallacies in the area of health were increasing more and more and could induce people to make wrong decisions, worsening their own well-being. In addition, there are warnings about the propagation of hoaxes about vaccines and their effects on people (**Espinoza-Portilla; Mazuelos-Cardoza**, 2020) or about the false content that may emerge in the field of nutrition and food (**Argiñano; Goikoetxea-Bilbao**, 2021).

“The combination “static image with design + use of emojis + use of hashtags in the text” was the strategy most used to tackle disinformation from *Instagram*”

It is also very important to note that only 0.4% of the fact-checks stemmed from hoaxes created in the traditional and digital media. The results of this research concur with **Gutiérrez-Coba, Caba-Gutiérrez**, and **Gómez-Díaz** (2020) in that there is a low percentage of fake news that reaches the conventional media and that the fact-checker accounts manage to alert these media outlets in some way so that they do not propagate such content. This reinforces the idea that, in the end, disinformation is disseminated most frequently through social networks (**Salvat**, 2021; **Martínez-Rolán; Dafonte-Gómez**, 2022). The media, despite the crisis of confidence and legitimacy that they have been experiencing (**Masip; Ferrer-Sapena**, 2021), continue to be a key instrument for professional practice, and journalism continues to be perceived as a source of proven information with high social value (**Sixto-García; Vázquez-Herrero; López-García**, 2022; **García-Avilés et al.**, 2022).

In addition to the efforts that platforms such as *Instagram*, *Twitter*, *Facebook*, or *WhatsApp* can make to combat disinformation in relation to science and health, this study demonstrates that there is a real problem with these networks. Although these social applications have developed strategies to eliminate hoaxes that spread on their platforms (**Bustos-Díaz; Ruiz-del-Olmo**, 2020; **Ndiaye**, 2021), the efforts appear to be rather meagre when faced with the difficulty of finding a panacea for these content problems, which increasingly threatens the fact-checking systems of these companies (**López-García; Costa-Sánchez; Vizoso**, 2021; **Ferreras-Rodríguez**, 2022; **Wang et al.**, 2022).

In connection with research question (Q3)

What specific and unique elements stand out in the use of the platform *Instagram* as a tool to combat disinformation?

it is pertinent to point out that a considerable number of posts in the analyzed sample acted as a preview of the news item, providing a link to the full content. This has been noted in previous studies (**Martín-Neira; Trillo-Domínguez; Olvera-Lobo**, 2023a), which point out that, today, certain social network accounts are not seen as media per se and serve more as platforms that lead to a central website where content is developed further, and thus they amplify the message of the media outlet. However, this does not necessarily have an impact on web traffic to the websites, since direct entry to the site or content exploration through Internet search engines is usually the primary means of accessing information (**Parra-Valcarce; Onieva-Mallero**, 2021).

Furthermore, the results that emerge in terms of interaction with users are telling, since 100% of the posts did not respond to followers' comments. This is a pertinent fact since, although the media and accounts are used to cultivating a spirit of interactivity and creating communities with users in social networks (**Swart; Peters; Broersma**, 2018; **Zurita-Andión**, 2019), it was observed that, frequently, there is not adequate staff to carry out these tasks: digital journalists multitask in conditions that are not ideal, or there is no solid strategy that gives purpose to this interactivity (**González-Pedraz; Campos-Domínguez**, 2017; **Cassany; Cortiñas; Elduque**, 2018; **Costa-Sánchez; Tüñez-López**, 2019; **Grassau; Porath; Ortega**, 2021; **Greene-González; Cerda-Diez; Ortiz-Leiva**, 2022). In this sense, the present research reveals that only 2% of the posts disseminated fact-checks that users explicitly initiated or alerted them to, despite the fact that the lines of communication are open to the public so they can communicate or warn about disinformation.

6. Conclusions

The research described in this paper allowed us through fact-checker accounts' praxis to confirm how social networks have become an active space for disseminating disinformation. However, fact-checking initiatives are emerging from these same platforms to help combat the false content that spreads rapidly through social platforms. The use of *Instagram* was preferred owing to the app's graphic and visual possibilities, despite the fact that, in the analyzed content, static images or *slide* shows were preferred over videos –taking advantage of all of the potential that this social network has to offer has proven to be a challenge.

It is also important to note that most of the hoaxes had a health component, extending beyond issues associated with the coronavirus. Content related to beauty, nutrition, and relevant aspects of health in general are the type of disinformation that most often goes viral, which should lead us to reflect on the problems that this content can create in society if it reaches a significant portion of the population.

In aspects associated with the way in which the fact-checker accounts relate to audiences, it is important to continue promoting collaborative activity that promotes interaction with the public: answering questions or responding to

their concerns. Thus, it is also important to note that, in the scientific field, the media continue to be a reliable source of information, in which disinformative intentions are lower compared with social networks.

This paper is presented as a first step that aims to determine how recognized fact-checking accounts are using *Instagram* to debunk the scientific hoaxes that exist on the network. Future research aims to address other areas of knowledge, including the various branches of science, as well as time periods in which the disinformation associated with the pandemic has diminished. It would also be interesting to analyze how these fact-checker accounts are positioning themselves in more emerging social networks, such as *TikTok* or *Twitch*, which are aimed at a younger audience, and which use a particular audiovisual language and are disconnected from traditional media, or how the media are creating initiatives to make citizens media literate.

As future lines of study, we believe that it would be appropriate to broaden the perspective adopted in this research, focusing on the fact-checking accounts themselves, which is still a good starting point. Thus, specific guidelines could be established to analyze posts' content from a qualitative point of view with the aim of deep-diving into the intentions behind the information analyzed. Finally, we hope that the results of the study will serve as a basis for establishing a body of best practices, in the form of recommendations, for professionals in the field of journalism and science communication.

“ In the end, disinformation is disseminated most frequently through social networks. The media, despite the crisis of confidence and legitimacy that they have been experiencing, continue to be a key instrument for professional practice ”

7. Notes

1. *Slide*. *Instagram* function that allows to create galleries of up to 10 photos and videos that are displayed in a single post as a presentation.
2. Translation of the texts of Figure 1.

Duque trying to understand whether abortion terminates the pregnancy...	columbiacheck Abortion is not a means of contraception as President Duque said
...or is a method of contraception	In a recent <i>Twitter</i> thread, the president @ivanduquemarquez refers to his rejection of the judgement of @corteconstitucional to depenalize abortion before the 24th week of gestation. In his message he says, "This situation cannot turn abortion in Colombia into a means of contraception. We must address this issue with the body truly responsible, which is the Congress of the Republic."
	The president made similar claims in a

3. Translation of the texts of Figure 2.

X		Access LUPA. NEWA	agencia_lupa
It's not true that the German channel displayed fake cadavers to report deaths from the Omicron variant.	FALSE Theater One of the fake deceased stands up while the reporter announces that many people are dying in Germany from Omicron. Media terrorism. People can see it here.	Like Comment Share Save	agencia_lupa A video is circulating on social media in which a reporter is speaking in front of various bodies in black bags, supposedly in Germany. Suddenly, one of the "cadavers" starts to move and tries to get up. The post's caption suggests that the scene is "theater" and that it has to do with "media terrorism" regarding Covid-19. According to the post's text, the report says that there are many people dying in the country from the Omicron variant.
	111 reactions 31 comments 235 shares		X The information is FALSE. The content of the recording does not say anything regarding the number of deaths from Covid-19 or the Omicron variant in Germany. The recording shows a...

8. References

Aguado-Guadalupe, Guadalupe; Bernaola-Serrano, Itziar (2020). "Verificación en la infodemia de la Covid-19. El caso *Newtral*". *Revista latina de comunicación social*, n. 78, pp. 289-308. <https://doi.org/10.4185/RLCS-2020-1478>

Aleixandre-Benavent, Rafael; Castelló-Cogollos, Lourdes; Valderrama-Zurián, Juan-Carlos (2020). "Información y comunicación durante los primeros meses de Covid-19. Infodemia, desinformación y papel de los profesionales de la información". *Profesional de la información*, v. 29, n. 4, e290408. <https://doi.org/10.3145/epi.2020.jul.08>

Almansa-Martínez, Ana; Fernández-Torres, María-Jesús; Rodríguez-Fernández, Leticia (2022). "Desinformación en España un año después de la Covid-19. Análisis de las verificaciones de *Newtral* y *Maldita*". *Revista latina de comunicación social*, n. 80, pp. 183-200. <https://doi.org/10.4185/RLCS-2022-1538>

- Andréu-Abela, Jaime** (2000). *Las técnicas de análisis de contenido: una revisión actualizada*. Sevilla: Fundación Centro de Estudios Andaluces.
<https://abacoenred.com/wp-content/uploads/2019/02/Las-técnicas-de-análisis-de-contenido-una-revisión-actualizada.pdf>
- Argiñano, José-Luis; Goikoetxea-Bilbao, Udane** (2021). "El rol de fact-checkers de los influencers españoles de la alimentación en *Instagram*". *Comunicación y medios*, v. 30, n. 44, pp. 14-27.
<https://doi.org/10.5354/0719-1529.2021.64567>
- Azer, Jaylan; Blasco-Arcas, Lorena; Harrigan, Paul** (2021). "#Covid-19: forms and drivers of social media users' engagement behavior toward a global crisis". *Journal of business research*, v. 135, pp. 99-111.
<https://doi.org/10.1016/j.jbusres.2021.06.030>
- Bustos-Díaz, Javier; Ruiz-del-Olmo, Francisco-Javier** (2020). "Comunicar en tiempos de crisis en las redes sociales. Estrategias de verificación e intermediación informativa en los casos de *Facebook*, *Instagram* y *Twitter* durante la Covid-19". *Hipertext.net*, n. 21, pp. 115-125.
<https://doi.org/10.31009/hipertext.net.2020.i21.10>
- Capilla, Pablo** (2019). "¿De qué hablamos cuando hablamos de posverdad? Análisis del término en siete diarios de calidad". *Profesional de la información*, v. 28, n. 3, e280309.
<https://doi.org/10.3145/epi.2019.may.09>
- Casares-Corrales, Alfredo** (2021). *La hora del periodismo constructivo: El poder transformador de la información orientada al futuro y a las soluciones*. Pamplona: Eunsa. ISBN: 978 84 313 3562 5
- Caspari, Gino** (2022). "Instagram as a tool for archaeological science communication". *Digital applications in archaeology and cultural heritage*, v. 24, e00219.
<https://doi.org/10.1016/j.daach.2022.e00219>
- Cassany, Roger; Cortiñas, Sergi; Elduque, Albert** (2018). "Communicating science: the profile of science journalists in Spain [Comunicar la ciencia: El perfil del periodista científico en España]". *Comunicar*, v. 26, n. 55, pp. 9-17.
<https://doi.org/10.3916/C55-2018-01>
- Costa-Sánchez, Carmen; López-García, Xosé** (2020). "Communication and coronavirus crisis in Spain. First lessons". *Profesional de la información*, v. 29, n. 3, e290304.
<https://doi.org/10.3145/epi.2020.may.04>
- Costa-Sánchez, Carmen; Tüñez-López, Miguel** (2019). "Contenidos audiovisuales en social media. Análisis comparativo de *Facebook* y *YouTube*". *Fonseca, journal of communication*, n. 19, pp. 223-236.
<https://doi.org/10.14201/fjc201919223236>
- Dafonte-Gómez, Alberto; Corbacho-Valencia, Juan-Manuel; García-Mirón, Silvia** (2021). "El fact-checking en Iberoamérica: evolución reciente y mapa de situación". En: Sotelo-González, Joaquín; González-García, Joaquín (eds.). *Digital media. El papel de las redes sociales en el ecosistema educacional en tiempo de Covid-19*. España: McGraw-Hill, pp. 877-889. ISBN: 978 84 486 2586 3
<https://dialnet.unirioja.es/servlet/libro?codigo=793917>
- Dafonte-Gómez, Alberto; Míguez-González, María-Isabel; Ramahí-García, Diana** (2022). "Fact-checkers on social networks: analysis of their presence and content distribution channels". *Communication & society*, v. 35, n. 3, pp. 73-89.
<https://doi.org/10.15581/003.35.3.73-89>
- Díaz-Moreno, Naira** (2019). "Caracterizando controversias sociocientíficas en la prensa escrita. Una herramienta para el desarrollo de la alfabetización científica". *Revista eureka sobre enseñanza y divulgación de las ciencias*, v. 16, n. 1, 1102.
https://doi.org/10.25267/Rev_Eureka_ensen_divulg_cienc.2019.v16.i1.1102
- Dunwoody, Sharon** (2020). "Science journalism and pandemic uncertainty". *Media and communication*, v. 8, n. 2, pp. 471-474.
<https://doi.org/10.17645/mac.v8i2.3224>
- Espinoza-Portilla, Elizabeth; Mazuelos-Cardoza, César** (2020). "Desinformación sobre temas de salud en las redes sociales". *Revista cubana de información en ciencias de la salud*, v. 31, n. 2, e1498.
http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S2307-21132020000200002
- Fernández-Barrero, Ángeles; López-Redondo, Isaac** (2022). "Verification processes in fake news era. Some examples about Covid-19". *Ámbitos. Revista internacional de comunicación*, n. 57, pp. 124-137.
<https://doi.org/10.12795/Ambitos.2022.i57.07>
- Ferreras-Rodríguez, Eva-María** (2022). "Más allá del fact-checking: organizaciones contra la desinformación. Identificación y análisis de proyectos internacionales". *Hipertext.net*, n. 24, pp. 41-54.
<https://doi.org/10.31009/hipertext.net.2022.i24.04>

García-Avilés, José A.; Arias-Robles, Félix; De-Lara-González, Alicia; Carvajal, Miguel; Valero-Pastor, José-María; Mondéjar, Dámaso (2022). "How Covid-19 is revamping journalism: newsroom practices and innovations in a crisis context". *Journalism practice*, Online first.
<https://doi.org/10.1080/17512786.2022.2139744>

García-Crespo, Oswaldo; Ramahí-García, Diana; Dafonte-Gómez, Alberto (2021). "Fact-checkers Iberoamericanos en YouTube. Análisis de presencia y actividad". En: Blanco-Pérez, Manuel (ed.). *El progreso de la comunicación en la era de los prosumidores*. Madrid: Dykinson S. L., pp. 83-105. ISBN: 978 84 1377 644 6
<https://www.dykinson.com/libros/el-progreso-de-la-comunicacion-en-la-era-de-los-prosumidores/9788413776446>

García-Galera, María-del-Carmen; Del-Hoyo-Hurtado, Mercedes; Blanco-Alfonso, Ignacio (2020). "Disinformation and communicative intent: a proposal for fake news classification in professional journalistic environments". *Revista mediterránea de comunicación*, v. 11, n. 2, pp. 105-118.
<https://doi.org/10.14198/MEDCOM2020.11.2.16>

García-Marín, David (2020). "Infodemia global. Desórdenes informativos, narrativas fake y fact-checking en la crisis de la Covid-19". *Profesional de la información*, v. 29, n. 4, e290411.
<https://doi.org/10.3145/epi.2020.jul.11>

García-Marín, David; Merino-Ortego, Marta (2022). "Desinformación anticientífica sobre la Covid-19 difundida en Twitter en Hispanoamérica". *Cuadernos.info*, n. 52, pp. 24-46.
<https://doi.org/10.7764/cdi.52.42795>

González-Pedraz, Cristina; Campos-Domínguez, Eva (2017). "Práctica profesional del periodista científico: revisión bibliográfica de las disfunciones derivadas del entorno digital". *Revista mediterránea de comunicación*, v. 8, n. 2, pp. 225-240.
<https://doi.org/10.14198/medcom2017.8.2.14>

Grassau, Daniela; Porath, William; Ortega, Constanza (2021). *La crisis de la industria de los medios y la precarización del empleo del periodista. Informe final proyecto PLU190009. Resultados trabajo de campo 2020/2021*.
<https://mapademedios.cl/publicaciones>

Greene-González, María-Francisca; Cerda-Diez, María-Fernanda; Ortiz-Leiva, Germán (2022). "Prácticas periodísticas en tiempos de pandemia de coronavirus. Un estudio comparado entre Chile y Colombia". *Revista de comunicación*, v. 21, n. 1, pp. 195-213.
<https://doi.org/10.26441/rc21.1-2022-a10>

Guallar, Javier; Codina, Lluís; Freixa, Pere; Pérez-Montoro, Mario (2020). "Desinformación, bulos, curación y verificación. Revisión de estudios en Iberoamérica 2017-2020". *Telos: revista de estudios interdisciplinarios en ciencias sociales*, v. 22, n. 3, pp. 595-613.
<https://doi.org/10.36390/telos223.09>

Gurler, Deniz; Buyukceran, Ismail (2022). "Assessment of the medical reliability of videos on social media: detailed analysis of the quality and usability of four social media platforms (Facebook, Instagram, Twitter, and YouTube)". *Healthcare*, v. 10, n. 10, 1836.
<https://doi.org/10.3390/healthcare10101836>

Gutiérrez-Coba, Liliana M.; Coba-Gutiérrez, Patricia; Gómez-Díaz, Javier-Andrés (2020). "Las noticias falsas y desinformación sobre el Covid-19: análisis comparativo de seis países iberoamericanos". *Revista latina de comunicación social*, n. 78, pp. 237-264.
<https://doi.org/10.4185/RLCS-2020-1476>

Habibi, Sarah A.; Salim, Lidya (2021). "Static vs. dynamic methods of delivery for science communication: a critical analysis of user engagement with science on social media". *PLoS one*, v. 16, n. 3, e0248507.
<https://doi.org/10.1371/journal.pone.0248507>

Herrero, Esperanza; Herrera-Damas, Susana (2021). "El fact-checker en español alrededor del mundo: perfil, similitudes y diferencias entre verificadores hispanohablantes". *Revista de comunicación de la SEECI*, n. 54, pp. 49-77.
<https://doi.org/10.15198/seeci.2021.54.e725>

Herrero-Diz, Paula; Pérez-Escolar, Marta (2022). "Análisis de los bulos sobre Covid-19 desmentidos por *Maldita y Colombiacheck*: efectos de la infodemia sobre el comportamiento de la sociedad". *Palabra clave*, v. 25, n. 1, e2517.
<https://doi.org/10.5294/pacla.2022.25.1.7>

Herrero-Diz, Paula; Pérez-Escolar, Marta; Varona-Aramburu, David (2022). "Competencias de verificación de contenidos: una propuesta para los estudios de Comunicación". *Revista de comunicación*, v. 21, n. 1, pp. 231-249.
<https://doi.org/10.26441/rc21.1-2022-a12>

- Hoyos-Simbaña, Diego-Paúl; Lara-Aguilar, Jhonatan-Andrés; Mila-Maldonado, Juan-Arturo** (2022). "Desinformación y fact-checking en Ecuador. Análisis del rol informativo de la cuenta de *Instagram* de @ecuadorverifica en torno a las elecciones presidenciales del 2021". *Cuadernos del Centro de Estudios de Diseño y Comunicación*, n. 161, pp. 73-90.
<https://doi.org/10.18682/cdc.vi161.6979>
- Hu, Xiao-Juan** (2022). "Research on Chinese journalists' scientific literacy". *Frontiers in communication*, v. 7, 850118.
<https://doi.org/10.3389/fcomm.2022.850118>
- Jarreau, Paige B.; Dahmen, Nicole S.; Jones, Ember** (2019). "Instagram and the science museum: a missed opportunity for public engagement". *Journal of science communication*, v. 18, n. 2, A06.
<https://doi.org/10.22323/2.18020206>
- Jiang, Hongyu; Guo, Ao; Ma, Jianhua** (2020). "Automatic prediction and insertion of multiple emojis in social media text". In: *2020 International conferences on internet of things (iThings) and IEEE green computing and communications (GreenCom) and IEEE cyber, physical and social computing (CPSCoM) and IEEE smart data (SmartData) and IEEE congress on cybermatics (Cybermatics)*, pp. 505-512.
<https://doi.org/10.1109/iThings-GreenCom-CPSCoM-SmartData-Cybermatics50389.2020.00092>
- León, Bienvenido; López-Goñi, Ignacio; Salaverría, Ramón** (2022). "The Covid-19 catastrophe: a science communication mess?". *Church, communication and culture*, v. 7, n. 1, pp. 6-22.
<https://doi.org/10.1080/23753234.2022.2031236>
- León, Bienvenido; Martínez-Costa, María-Pilar; Salaverría, Ramón; López-Goñi, Ignacio** (2022). "Health and science-related disinformation on Covid-19: a content analysis of hoaxes identified by fact-checkers in Spain". *PLoS one*, v. 17, n. 4, e0265995.
<https://doi.org/10.1371/journal.pone.0265995>
- Lobato-Martínez, Miguel; Monjas-Eleta, María; Gómez-García, Salvador** (2022). "Situación y perspectivas del periodismo científico en España. Investigación prospectiva a través del método Delphi". *Estudos em comunicação*, n. 34, pp. 66-80.
<https://ojs.labcom-ifp.ubi.pt/index.php/ec/article/view/1056>
- López-García, Xosé; Costa-Sánchez, Carmen; Vizoso, Ángel** (2021). "Journalistic fact-checking of information in pandemic: stakeholders, hoaxes, and strategies to fight disinformation during the Covid-19 crisis in Spain". *International journal of environmental research and public health*, v. 18, n. 3, 1227.
<https://doi.org/10.3390/ijerph18031227>
- López-Pan, Fernando; Rodríguez-Rodríguez, Jorge-Miguel** (2020). "El fact checking en España. Plataformas, prácticas y rasgos distintivos". *Estudios sobre el mensaje periodístico*, v. 26, n. 3, pp. 1045-1065.
<https://doi.org/10.5209/ESMP.65246>
- López-Pujalte, Cristina; Nuño-Moral, María-Victoria** (2020). "La 'infodemia' en la crisis del coronavirus: Análisis de desinformaciones en España y Latinoamérica". *Revista española de documentación científica*, v. 43, n. 3, e274.
<https://doi.org/10.3989/redc.2020.3.1807>
- Maiden, Neil; Zachos, Konstantinos; Franks, Suzanne; Wells, Rebecca; Stallard, Samantha** (2020). "Designing digital content to support science journalism". In: *NordiCHI '20: Proceedings of the 11th Nordic conference on human-computer interaction: shaping experiences, shaping society*.
<https://doi.org/10.1145/3419249.3420124>
- Maldita** (2021). "Por qué no debemos utilizar 'noticias falsas' 'información falsa' o 'fake news' para referirnos a los bulos y las desinformaciones". *Maldita.es*, 11 febrero.
<https://maldita.es/newsletter-flipboard/20210211/por-que-no-utilizar-noticias-falsas-fake-news>
- Malik, Aqdas; Khan, M. Laeeq; Quan-Haase, Anabel** (2021). "Public health agencies outreach through *Instagram* during the Covid-19 pandemic: crisis and emergency risk communication perspective". *International journal of disaster risk reduction*, v. 61, 102346.
<https://doi.org/10.1016/j.ijdrr.2021.102346>
- Maniou, Theodora A.; Papa, Venetia** (2023). "The dissemination of science news in social media platforms during the Covid-19 crisis: characteristics and selection criteria". *Communication and society*, v. 36, n. 1, pp. 35-46.
<https://doi.org/10.15581/003.36.1.35-46>
- Marta-Lazo, Carmen; Rodríguez-Rodríguez, Jorge-Miguel; Peñalva, Sheila** (2020). "Competencias digitales en periodismo. Revisión sistemática de la literatura científica sobre nuevos perfiles profesionales del periodista". *Revista latina de comunicación social*, n. 75, pp. 53-68.
<https://doi.org/10.4185/RLCS-2020-1416>
- Martín-Neira, Juan-Ignacio; Trillo-Domínguez, Magdalena; Olivera-Lobo, María-Dolores** (2022). "La divulgación científica en *Instagram*: usos y estrategias desde la praxis chilena". *Cuadernos.info*, n. 53, pp. 229-252.
<https://doi.org/10.7764/cdi.53.42515>

- Martín-Neira, Juan-Ignacio; Trillo-Domínguez, Magdalena; Olvera-Lobo, María-Dolores** (2023a). “Comunicación científica tras la crisis del Covid-19: estrategias de publicación en *TikTok* en el tablero transmedia”. *Revista latina de comunicación social*, n. 81, pp. 109-132.
<https://doi.org/10.4185/RLCS-2023-1841>
- Martín-Neira, Juan-Ignacio; Trillo-Domínguez, Magdalena; Olvera-Lobo, María-Dolores** (2023b). “De la televisión a *TikTok*: nuevos formatos audiovisuales para comunicar ciencia”. *Comunicación y sociedad*, v. 20.
<https://doi.org/10.32870/cys.v2023.8441>
- Martínez-Rolán, Xabier; Dafonte-Gómez, Alberto** (2022). “Covid, hoaxes and fact checkers. Evolution of denyings in Ibero-America 2020-2022”. *Visual review. International visual culture review*, v. 10, n. 1.
<https://doi.org/10.37467/revvisual.v9.3556>
- Masip, Pere; Aran-Ramspott, Sue; Ruiz-Caballero, Carlos; Suau, Jaume; Almenar, Ester; Puertas-Graell, David** (2020). “Consumo informativo y cobertura mediática durante el confinamiento por el Covid-19: sobreinformación, sesgo ideológico y sensacionalismo”. *Profesional de la información*, v. 29, n. 3, e290312.
<https://doi.org/10.3145/epi.2020.may.12>
- Masip, Pere; Ferrer-Sapena, Antonia** (2021). “Más allá de las fake news. Anatomía de la desinformación”. *BiD: textos universitaris de biblioteconomia i documentació*, n. 46.
<https://doi.org/10.1344/BiD2020.46.08>
- Massarani, Luisa; Fernandes-Neves, Luiz-Felipe; Entradas, Marta; Lougheed, Tim; Bauer, Martin W.** (2021) “Perceptions of the impact of the Covid-19 pandemic on the work of science journalists: global perspectives”. *Journal of science communication*, v. 20, n. 7, A06.
<https://doi.org/10.22323/2.20070206>
- Mellado, Claudia; Cárcamo-Ulloa, Luis; Alfaro, Amaranta; Inai, Darla; Isbej, José** (2021). “Fuentes informativas en tiempos de Covid-19: cómo los medios en Chile narraron la pandemia a través de sus redes sociales”. *Profesional de la información*, v. 30, n. 4, e300421.
<https://doi.org/10.3145/epi.2021.jul.21>
- Míguez-González, María-Isabel; Abuín-Penas, Javier; Pérez-Seoane, Jesús** (2021). “¿Cómo utilizan los fact-checkers las redes sociales para combatir la desinformación? Análisis de la actividad de los fact-checkers iberoamericanos en *Instagram*”. En: Blanco-Pérez, Manuel (ed.). *El progreso de la comunicación en la era de los prosumidores*. Madrid: Dykinson SL, pp. 15-39. ISBN: 978 84 1377 644 6
<https://www.dykinson.com/libros/el-progreso-de-la-comunicacion-en-la-era-de-los-prosumidores/9788413776446>
- Molina-Cañabate, Juan-Pedro; Sánchez-Duarte, José-Manuel; Magallón-Rosa, Raúl** (2021). “Desinformación y fact-checking durante el primer año de Covid-19 en España. El caso de Newtral”. En: Blanco-Pérez, Manuel (ed.). *El progreso de la comunicación en la era de los prosumidores*. Madrid: Dykinson SL, pp. 40-62. ISBN: 978 84 1377 644 6
<https://www.dykinson.com/libros/el-progreso-de-la-comunicacion-en-la-era-de-los-prosumidores/9788413776446>
- Montemayor-Rodríguez, Nancy; García-Jiménez, Antonio** (2021). “Percepción de los periodistas sobre la desinformación y las rutinas profesionales en la era digital”. *Revista general de información y documentación*, v. 31, n. 2, pp. 601-619.
<https://doi.org/10.5209/rgid.79460>
- Montero-Liberona, Claudia; Halpern, Daniel** (2019). “Factores que influyen en compartir noticias falsas de salud online”. *Profesional de la información*, v. 28, n. 3, e280317.
<https://doi.org/10.3145/epi.2019.may.17>
- Ndiaye, Aïda** (2021). “Juntos contra la información errónea sobre el Covid-19: una nueva campaña en colaboración con la OMS”. *Facebook*, 10 marzo.
<https://www.facebook.com/formedia/blog/together-against-covid-19-misinformation-a-new-campaign-in-partnership-with-the-who>
- Newman, Nic; Fletcher, Richard; Robertson, Craig T.; Eddy, Kirsten; Nielsen, Rasmus-Kleis** (2022). *Digital news report 2022*.
<https://reutersinstitute.politics.ox.ac.uk/es/digital-news-report/2022>
- Palomo, Bella; Sedano-Amundarain, Jon A.** (2018). “*WhatsApp* como herramienta de verificación de fake news. El caso de *B de Bulo*”. *Revista latina de comunicación social*, n. 73, pp. 1384-1397.
<https://doi.org/10.4185/RLCS-2018-1312>
- Parra-Valcarce, David; Onieva-Mallero, Charo** (2021). “Análisis del impacto de las redes sociales sobre el tráfico web de los cibermedios nativos digitales españoles”. *Fonseca, journal of communication*, n. 22, pp. 99-117.
<https://doi.org/10.14201/fjc-v22-22696>

Post, Senja; Bienzeisler, Nils; Lohöfener, Mareike (2021). "A desire for authoritative science? How citizens' informational needs and epistemic beliefs shaped their views of science, news, and policymaking in the Covid-19 pandemic". *Public understanding of science*, v. 30, n. 5, pp. 496-514.
<https://doi.org/10.1177/096366252111005334>

Poynter (2022). *IFCN Code of Principles*.
<https://ifcncodeofprinciples.poynter.org/signatories>

Rodríguez, Marcelo; Giri, Leandro (2021). "Desafíos teóricos cruciales para la comunicación pública de la ciencia y la tecnología post pandemia en Iberoamérica". *Revista iberoamericana de ciencia, tecnología y sociedad*, v. 16, pp. 25-39.
<http://www.revistacts.net/contenido/numero-numero-especial-2021/desafios-teoricos-cruciales-para-la-comunicacion-publica-de-la-ciencia-y-la-tecnologia-post-pandemia-en-iberoamerica>

Rodríguez-Fernández, Leticia (2019). "Desinformación: retos profesionales para el sector de la comunicación". *Profesional de la información*, v. 28, n. 3, e280306.
<https://doi.org/10.3145/epi.2019.may.06>

Rodríguez-Ferrándiz, Raúl (2019). "Posverdad y fake news en comunicación política: breve genealogía". *Profesional de la información*, v. 28, n. 3, e280314.
<https://doi.org/10.3145/epi.2019.may.14>

Rodríguez-Pérez, Carlos (2020). "Una reflexión sobre la epistemología del fact-checking journalism: retos y dilemas". *Revista de comunicación*, v. 19, n. 1, pp. 243-258.
<https://doi.org/10.26441/RC19.1-2020-A14>

Rodríguez-Pérez, Carlos; Seibt, Taís; Magallón-Rosa, Raúl; Paniagua-Rojano, Francisco-Javier; Chacón-Peinado, Sonia (2022). "Purposes, principles, and difficulties of fact-checking in Ibero-America: Journalists' perceptions". *Journalism practice*, Online first.
<https://doi.org/10.1080/17512786.2022.2124434>

Sádaba, Charo; Salaverría, Ramón (2023). "Combatir la desinformación con alfabetización mediática: análisis de las tendencias en la Unión Europea". *Revista latina de comunicación social*, n. 81, pp. 17-33.
<https://doi.org/10.4185/rllcs-2023-1552>

Salaverría, Ramón; Buslón, Nataly; López-Pan, Fernando; León, Bienvenido; López-Goñi, Ignacio; Erviti, María-Carmen (2020). "Desinformación en tiempos de pandemia: tipología de los bulos sobre la Covid-19". *Profesional de la información*, v. 29, n. 3, e290315.
<https://doi.org/10.3145/epi.2020.may.15>

Salvat, Guiomar (2021). "El lugar del periodismo ciudadano desde la credibilidad y la confianza". *Estudios sobre el mensaje periodístico*, v. 27, n. 2, pp. 639-648.
<https://doi.org/10.5209/esmp.71039>

Sánchez-González, María; Sánchez-Gonzales, Hada M.; Martos, Javier (2022). "Innovación editorial en redes sociales de los verificadores hispanos de la #CoronavirusFactCheck Alliance: contenidos y visión de sus responsables". *Revista latina de comunicación social*, n. 80, pp. 135-161.
<https://doi.org/10.4185/rllcs-2022-1535>

Sidorenko-Bautista, Pavel; Alonso-López, Nadia; Giacomelli, Fabio (2021). "Fact-checking in TikTok. Communication and narrative forms to combat misinformation". *Revista latina de comunicación social*, n. 79, pp. 87-113.
<https://doi.org/10.4185/RLCS-2021-1522>

Sidorenko-Bautista, Pavel; Cabezuelo-Lorenzo, Francisco; Herranz-de-la-Casa, José-María (2021). "Instagram como herramienta digital para la comunicación y divulgación científica: el caso mexicano de @pictoline". *Chasqui. Revista latinoamericana de comunicación*, n. 147, pp. 143-162.
<https://doi.org/10.16921/chasqui.v1i147.4472>

Sixto-García, José; Vázquez-Herrero, Jorge; López-García, Xosé (2022). "Journalists' self-perception of their profession in Spain: analysis of social and technological challenges". *Tripodos*, n. 52, pp. 111-128.
<https://doi.org/10.51698/tripodos.2022.52p111-128>

Swart, Jöelle; Peters, Chris; Broersma, Marcel (2018). "Shedding light on the dark social: the connective role of news and journalism in social media communities". *New media and society*, v. 20, n. 11, pp. 4329-4345.
<https://doi.org/10.1177/1461444818772063>

Torrado, Santiago (2022). "Iván Duque rechaza la despenalización del aborto y la equipara a 'una práctica anticonceptiva'". *El país*, 22 febrero.
<https://elpais.com/internacional/2022-02-22/ivan-duque-rechaza-la-despenalizacion-del-aborto-y-la-equipara-a-una-practica-anticonceptiva.html>

Varona-Aramburu, David; Sánchez-Muñoz, Gema (2016). "Las redes sociales como fuentes de información periodística: motivos para la desconfianza entre los periodistas españoles". *Profesional de la información*, v. 25, n. 5, pp. 795-802.
<https://doi.org/10.3145/epi.2016.sep.10>

Vernal-Vilicic, Teresa P.; Valderrama-Zenteno, Lorena B. (2022). "Comunicación pública de la ciencia y la tecnología en Iberoamérica". *Cuadernos.info*, n. 52.
<https://doi.org/10.7764/cdi.52.50593>

Villa-Gracia, Alberto-Daniel; Cerdán-Martínez, Víctor (2020). "Bulos durante la pandemia del Covid-19 en España: un estudio a través de *Google Trends*". *Revista latina de comunicación social*, n. 78, pp. 169-182.
<https://doi.org/10.4185/rllcs-2020-1473>

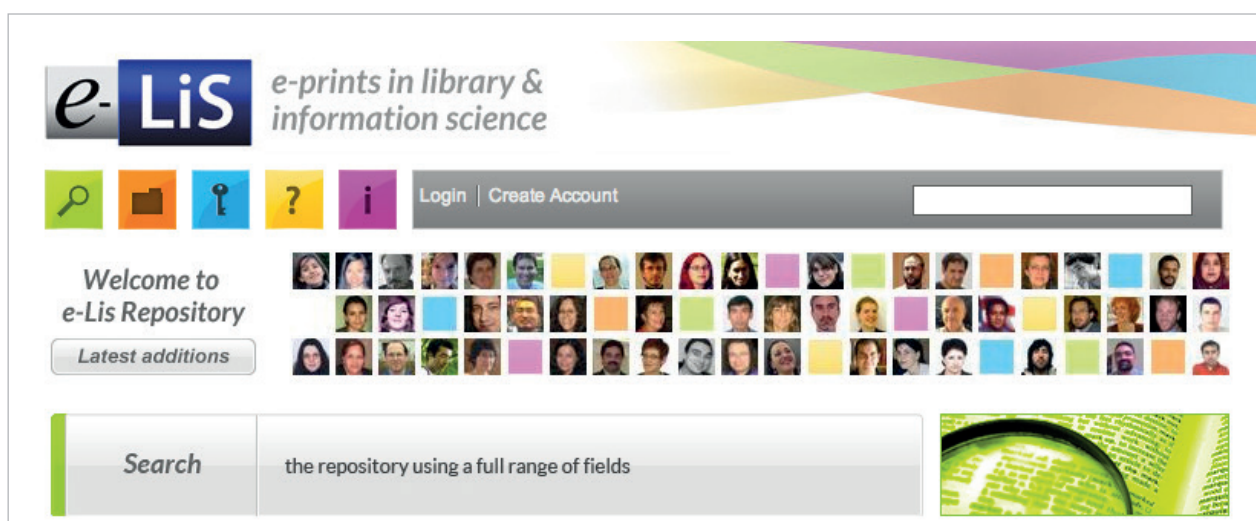
Wang, Yichen; Han, Richard; Lehman, Tamara-Silbergleit; Lv, Qin; Mishra, Shivakant (2022). "Do *Twitter* users change their behavior after exposure to misinformation? An in-depth analysis". *Social network analysis and mining*, v. 12, n. 1, 167.
<https://doi.org/10.1007/s13278-022-00992-8>

WHO (2020). *Managing the COVID-19 infodemic: Promoting healthy behaviours and mitigating the harm from misinformation and disinformation*. Organización Mundial de la Salud.
<https://www.who.int/news/item/23-09-2020-managing-the-covid-19-infodemic-promoting-healthy-behaviours-and-mitigating-the-harm-from-misinformation-and-disinformation>

Zeng, Jing; Schäfer, Mike S.; Allgaier, Joachim (2021). "Research perspectives on *TikTok* & its legacy apps | reposting 'Till Albert Einstein is *TikTok* famous': The memetic construction of science on *TikTok*". *International journal of communication*, v. 15, pp. 3216-3247.
<https://ijoc.org/index.php/ijoc/article/view/14547>

Zurita-Andión, José-Luis (2019). "El 'engagement' y las nuevas narrativas en el diseño de la comunicación digital". *Estudios sobre el mensaje periodístico*, v. 25, n. 2, pp. 1249-1261.
<https://doi.org/10.5209/esmp.64836>

Give visibility to your work by depositing it in e-LIS, the largest international repository on library & information science, and communication



<http://eprints.rclis.org>

Disinformation narratives in Spain: reach, impact and spreading patterns

Jaume Suau; David Puertas-Graell

Nota: Este artículo se puede leer en español en:
<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87312>

Recommended citation:

Suau, Jaume; Puertas-Graell, David (2023). "Disinformation narratives in Spain: reach, impact and spreading patterns". *Profesional de la información*, v. 32, n. 5, e320508.

<https://doi.org/10.3145/epi.2023.sep.08>

Manuscript received on 8th March 2023

Accepted on 07th June 2023



Jaume Suau ✉
<https://orcid.org/0000-0003-4480-4441>

Universitat Ramon Llull
Facultat de Comunicació i Relacions
Internacionals Blanquerna
Plaça Joan Coromines, s/n
08001 Barcelona, Spain
jaumesm@blanquerna.url.edu



David Puertas-Graell
<https://orcid.org/0000-0003-0484-3526>

Universitat Ramon Llull
Facultat de Comunicació i Relacions
Internacionals Blanquerna
Plaça Joan Coromines, s/n
08001 Barcelona, Spain
davidpg5@blanquerna.url.edu

Abstract

We present data from a survey conducted in Spain ($N = 1003$) in March 2022. We analysed fact-checker activity to obtain daily information regarding disinformation content encountered in the three weeks before the survey was launched. The research team analysed the material found to identify content that was related or that belonged to similar narratives. The goal was to identify the key disinformation narratives that were spreading before the survey, rather than just isolated content, to test the reach and impact of disinformation narratives, as well as spreading patterns, through survey research. Results point towards the fact that disinformation narratives were spread among a majority of respondents, with TV and social media being the main media responsible for spreading them. In addition, those that received the narratives before were more likely to believe them, indicating the disinformation narratives' potential high impact.

Keywords

Disinformation; Narratives; Reach; Impact; Surveys; Research; Spreading; Dissemination; Trust; Misleading; Information disorders; Fact checking; Fact checkers.

1. Introduction

The study of the real impact of disinformation in our societies is a complex endeavour. Foreign-sponsored misleading content and anonymous bots and social media profiles mix with domestic actors and content creators, who seem to always be one step ahead of fact-checkers and automated countermeasures (Almenar *et al.*, 2022). Election times, or relevant current issues such as covid-19 pandemic or the Russian invasion of Ukraine, trigger alarms about the vast array of content created to mislead the public. Scandals such as *Cambridge Analytica* or proven involvement of Russian operators in USA elections showed how disinformation campaigns are aimed at diminishing societal trust and reinforce pre-existing societal cleavages (Colley; Granelli; Althius, 2020; Casero-Ripolles; Doménech-Fabregat; Alonso-Muñoz, 2023). This also involves news media, who sometimes act, willingly or not, as amplifiers of such narratives (Waisbord, 2018). Consequently, news media, which have been absorbed by the attention economy (Wu, 2020) and are currently in search of successful business models (Newman *et al.*, 2022), contribute to said diffusion by utilising techniques associated with sensationalism that increase the effectiveness of misinformation disseminated (Staender *et al.*, 2021). Moreover, the role of the audience in disinformation spreading patterns needs to be also taken into account. As "unwilling crowds" (Starbird; Arif; Wilson, 2019) audiences not just passively consume misleading content, but share it, reshape it into a different format, or comment on it, contributing to its dissemination. Previous research has found, precisely, that the goal of disinformation campaigns is to move beyond like-minded audiences to reach broader publics, through multiple kinds of content and spreading strategies, continuously through time and aiming at social reproduction and remediation (Wilson; Starbird, 2020; Flore, 2020). Hence, as stated by Starbird, Arif and Wilson,

"evaluating disinformation is less about the truth value of one or more pieces of information and more about how those pieces fit together to serve a particular purpose" (Starbird; Arif; Wilson, 2019, p. 3).



Taking this into consideration, our goal is to analyse the impact of disinformation in Spain from a holistic point of view. We will focus our attention on disinformation narratives, understanding narratives as a group or collective of content that refers to the same idea or story, based on current societal polarised issues, with a moral base that plays with emotions through a certain degree of plausibility (Dawson, 2021). Previous studies have focused on the capacity of organised disinformation campaigns to increase societal polarisation and distrust (Colley; Granelli; Althuis, 2020) through the use of narratives (Weinberg; Dawson, 2021). Such narratives are designed to target specific societal groups or audiences, with content designed for them and spread through different platforms (Dawson, 2021). However, the ultimate goal of disinformation campaigns is not just to reach these specific publics, but to get to broader audiences. To do so, the common strategy is to play with emotions and polarising or divisive political topics (Bánkuty-Balogh, 2021), increasing visibility and spreadability, ensuring that common audiences become unaware collaborators in the spread of misleading content (Bastos; Mercea, 2018). Here it is key that the narratives have a certain degree of plausibility or credibility. At least, some part of true facts that can be traced or accepted, which increases their chances of being accepted by a broader audience (Starbird; Arif; Wilson, 2019).

Hence, we will not focus on specific content or stories shared in a particular platform, but rather, on how these pieces fit together to “build narratives”. Around any given narrative or core idea multiple different related content exist, originally created as part of a disinformation campaign, or remediated and spread by anonymous citizens or amplified by news media or political actors. A single piece of content can be easily forgotten. However, the underlying narrative is more likely to be remembered, and therefore, easier to be tracked and researched. More specifically, we are interested in researching the reach (understood as the spread of such content among a certain audience) and impact (the capacity of content to pass as true and therefore, mislead those who receive it) of disinformation narratives, together with their spreading patterns. In the theoretical background section we will develop these concepts more in-depth, explaining also the hypothesis that will guide our research. In the methodology section it will be introduced how disinformation narratives were identified, together with how our survey has been designed. The results and discussion sections will show our main findings and relate them with previous literature in the field.

2. Defining reach and impact of disinformation narratives

With already some years of existence, the field of disinformation research has seen some papers that attempt to summarise the main findings and trends in the field (Jankowski, 2018; Kapantai *et al.*, 2021; García-Marín; Salvat-Martinrey, 2022; Pérez-Escobar; Lilleker; Tapia-Frade, 2023). At an initial stage, the concept that dominated the field was *fake news*, rather than disinformation. However, as the concept was used in political discussion to blame non-aligned news media, it became more an example of societal struggles to hegemonise reality rather than a useful academic construct that might guide empirical research (Farkas; Schou, 2018). The field now uses the broadly accepted concept of disinformation, referring to false or misleading information that is deliberately spread through multiple platforms to deceive or manipulate public opinion. Wardle and Derakhshan (2017) include it in the typology of information disorders, based on the motives behind the creation or spread of such content. It divides between mis-, dis- and mal- information, being the first one when false information is shared without intention, the second when false content is shared intentionally and, the third, when genuine information is spread with the intention to cause harm. As Kapantai *et al.* (2021) argue, the field of disinformation studies can be classified according to two main areas of research: i) studies that conduct research on the extent or degree of dissemination of disinformation and ii) studies that are more interested in the effects or impact of disinformation content. It can be added here a third area of research, gathering studies that focus on the remedies or strategies focused on countering or mitigating the effects of disinformation. Here we will find research that focuses on how media literacy helps to identify misleading information (Amazeen; Bucy, 2019; Vraga; Bode; Tully, 2022; Sádaba; Salaverría, 2023) or the effects of fact checking institutions, with interesting results about the capacity of fact-checks to shape factual beliefs, but having minimal effects on candidates evaluations despite proven falsehoods (Nyhan *et al.*, 2020). This is an area of promising research, as different authors have proven the capacity of citizens to identify misleading content once exposed to fact checks, although partisan motivated reasoning, cognitive bias and well established political beliefs seems to be still of high relevance (Hameleers, 2020; Freiling *et al.*, 2021). Furthermore, more recently in this third area of research it can be included also research that focuses on regulatory and self-regulatory mechanisms, as national governments and EU start to adopt regulations to counter disinformation, especially in relation with online platforms (Espaliú-Berdud, 2022; Napoli, 2019). However, despite its relevance, in this research we will focus on the aforementioned two main areas (the extent or degree of disinformation dissemination and its effects or impact), due their stronger relation with our research goals.

2.1. Reach of disinformation

An analysis of previous research shows that to study the dissemination of disinformation content has been a main concern since early years. This has even given rise to what some authors have called “misinformation on misinformation” (Altay; Berriche; Acerbi, 2023). That is, how narratives about online misinformation continue to gain traction despite evidence that its prevalence and impact are overstated. Previous research has mostly focused on the propagation of disinformation on social media, with a special focus on *Twitter* and, to lesser extent, *Facebook*. These kinds of studies focus generally on one single social media, researching the dissemination of a certain content or story. Different studies have shown how false content spreads faster and deeper than truth or verified information, especially if the content on *Twitter* is related to political issues (Vosoughi; Roy; Aral, 2018). Furthermore, it seems common or ordinary citizens are

the ones responsible for the spread of false content on social media, rather than automated profiles or bots (**Brennen et al.**, 2020). Citizens tend to share content about public issues that they consider more likely to be true or that relates with pre-existing attitudes or particularly strong beliefs (**Buchanan**, 2020), preferably among like-minded peers rather than through open platforms in which anyone can see, and reply (**Das; Schroder**, 2020; **Suau**, 2015). Content directly identified (correctly or not) as untrue or of dubious nature is less likely to be shared or interacted with (**Tandoc; Lim; Ling**, 2020). Hence, disinformation has a certain participatory nature: instead of passive receivers or consumers of disinformation content, audiences have an important position in the spread of such content (**Wilson; Starbird**, 2020; **Wanless; Berk**, 2021). As **Starbird, Arif** and **Wilson** (2019) define, audiences act as “unwilling crowds”, sharing, commenting or even remaking into a different format those contents that are related with stronger emotions or beliefs.

Furthermore, even in highly polarised scenarios such as the US elections Russian-sponsored accounts on *Facebook* and *Twitter*, despite publishing great amounts of content, hardly managed to reach general audiences (**Spangher et al.** 2018; **Zannettou et al.**, 2019). **Arce-García, Said-Hung** and **Mottareale** (2022) found similar results in Spain when studying astroturfing strategies of covid-related disinformation dissemination in *Twitter*, concluding that reach of such content was low among the overall population. Astroturfing –Astroturf company sells artificial grass for sports fame avenues– is not limited to what happens online. However, it is currently used to describe situations where certain users or bots act in an organised manner. These users are generally presented anonymously on social networks, appear to have no connection with one another, and followed by a small number of individuals. All of these factors are connected in a way that give credibility to the idea that public opinion has generated a movement. This type of circumstance occurs preferably in a political context, implying the emergence of more specific concepts such as political astroturfing (**Howard**, 2006; **Walker**, 2014),

“a campaign in which participants appear to be part of a genuine grassroots movement or sentiment, while it is in fact orchestrated centrally and top down” (**Kellet et al.**, 2019, p. 1).

This common type of manipulative behaviour that creates additional problems for social media platforms and the online environment in general (**Chan**, 2022) has been applied and associated with the spread of disinformation (**Cho et al.**, 2011; **Leiser**, 2016), as well as with its psychological effects (**Zerback; Töpfl; Knöpfle**, 2021).

Studies focused on the dissemination or spread of disinformation paid attention also to closed platforms. Due its technological affordances and popularity among the population of many countries, *WhatsApp* has been considered a fertile environment for disinformation to spread (**Chagas**, 2022; **Maros et al.**, 2020). Hence, it has been the focus of most of the research, despite the fact that its information is encrypted and therefore complexifies research on what users share. **Masip et al.** (2020) highlighted the role of *WhatsApp* as spreader of disinformation during confinement, although most citizens did not report in their survey high level exposure to disinformation. **Cardoso, Sepúlveda** and **Narciso** (2022) in a research about audio misinformation spread in *WhatsApp* also during Covid-19 pandemic noticed that sharing (and producing) misinformation occurs spontaneously, in relation with strong emotional attitudes towards certain topics or events. Hence, the conceptual model of dis/mis/mal information (**Wardle**, 2018) based on the intention to mislead might prove useful to assess some content with proven authorship, but complex to apply if we take into account the participatory nature of disinformation. Regarding formats of disinformation spread on *WhatsApp*, the research of **Moreno-Castro et al.** (2021) shows that during Covid-19 pandemic in Spain audio, images and videos were similarly relevant (30% of found content each), with texts threads having some more relevance in terms of numbers (around 39%). In their research focused on two different political events in Brazil, **Resende et al.** (2019) found a predominance of images as the most common format, identifying also spreading patterns as most content on *WhatsApp* has been spotted previously on social media or other websites. Interestingly, their research also shows that disinformation content remains being shared on the platform for more than 4 days, spreading through different *WhatsApp* groups. Despite the relevance of previous research, it is still hard to assess the real dissemination of disinformation content spread through *WhatsApp* among the general population. Moreover, as disinformation becomes “participatory” and spread through social media and messaging apps it is necessary to consider the extent to which this kind of content may reach different societal groups. Recent research discusses the existence or not of the echo-chamber effect, which states that due selective exposure and attention citizens tend to be exposed, especially on social media platforms, to ideologically aligned content (**Masip; Suau; Ruiz-Caballero**, 2020; **Cardenal et al.**, 2019). Furthermore, confirming, even if at least partially, the echo-chamber theory, previous research found that fact checks are generally spread among similar audiences, hardly reaching those most likely to be convinced about misleading information (**Hameleers; Van-der-Meer**, 2019; **Shin; Thorson**, 2017).

Although social media and messaging apps focused most of the attention of studies aimed at assessing the dissemination of disinformation, relevant research has been done also on tabloid or highly polarised news media outlets. For example, **Chadwick, Vaccari** and **O’Loughlin** (2018) found that readers of tabloid press were more likely to disseminate misleading content. Other research has focused on websites labelled as fake news producers, finding that its relevance in terms of daily users or penetration in society is scarce. In the US, **Guess, Nyhan** and **Reifler** (2020) argue that sharing from websites identified as producers of fake news was relatively scarce in 2016, being more likely between citizens 65 or older and among right-wing ideology and/or pro-republican party. Studies conducted in different countries show that such spaces, even including their social media presence, gather just between 0.15 to 6% of the population (**Grinberg et al.**, 2019; **Allen et al.**, 2020; **Fletcher; Nielsen**, 2018). Particularly interesting is the findings of **Allen et al.** (2020), which indicate that in the US disinformation spread from TV is highly more relevant than in social media, although the focus

here is mostly still on hyper partisan news. It is precisely the role of traditional news media in the spread of disinformation which is still less understood and researched (**Tsfati et al.**, 2020). Few attention has been given to the possibility that traditional news media outlets may also be producers of disinformation. However, their amplifying effect has been widely researched (**Phillips**, 2018): as news media have to cover speeches and declarations of political leaders, is inevitable that news media end up contributing to the spread of disinformation and even conspiracy theories, even if they report with clear warnings, corrections and red flags (**Lewandowsky et al.**, 2012; **Walter et al.**, 2020; **Waisbord**, 2018).

At this point, it is relevant to highlight that in our country of study, Spain, TV has an important relevance as a source of news, still higher than social media, while print and radio news reach around a quarter of the population (**AIMC**, 2022; **Newman et al.**, 2022). Hence, our study of the spread of disinformation narratives needs to take into account also more traditional formats of news media consumption, such as TV, radio and press, to test the spread and impact of disinformation narratives. At the same time, it is necessary also to reflect on the lack of trust that Spanish citizens show on traditional media outlets. As shown in periodic studies (see here **Newman et al.**, 2002, as well as previous *Digital News Reports*) trust on news has been declining since 2017. Despite generalised lack of media credibility and reliability, news media outlets, or legacy media, are showing still higher levels of trust than online media or online media platforms (**Fernández-Torres; Almansa-Martínez; Chamizo-Sánchez**, 2021; **Besalú; Pont-Sorribes**, 2021). However, research in Spain has shown that this generalised lack of trust, combined with political polarisation, creates a general perception of news media as disinformation producers. **Masip, Suau and Ruiz-Caballero** (2020) has shown how Spanish citizens clearly position news media outlets according to ideological positions. Furthermore, they tend to identify those news media identified as ideologically non-affine as disinformation producers, not as legitimate journalism, a pattern that repeats in different countries and that can be attributed to political polarisation (**Van-der-Linden; Panagopoulos; Roozenbek**, 2020).

As mentioned in the introduction, reach can be understood as the capacity of certain content to spread among the general population. Or, at the level of localised reach, the capacity of disinformation content to target specific audiences or subpopulations (**Allen et al.**, 2020). Despite general concerns about the spread of disinformation, as shown in former paragraphs, studies that researched the reach of such content point precisely towards a situation of limited reach regarding the general population. Previous studies focused mostly on certain kinds of content on one social media platform. Although our approach is different, researching disinformation narratives rather than specific pieces of content, measuring reach in terms of the percentage of respondents who have heard about a certain narrative. Hence, basing on previous studies our first hypothesis will be as follows:

H1: Disinformation narratives have limited reach among the public, with higher degrees depending on the polarisation nature of the narrative or its connection to political issues.

H2: Disinformation narratives will be more spread through TV than digital media, with social media and messaging apps having also a predominant role.

However, the dissemination of disinformation narratives does not ensure their impact. Despite the growing body of literature on disinformation, we still do not fully understand the effects that the production and dissemination of such content has in society (**Zannetou et al.**, 2019). Previous research provides different considerations of the concept of impact. Research based on elections time in different countries highlighted the capacity of disinformation campaigns to shape election results, which will be a proof of their impact (**Kazeem**, 2018; **Kapantai et al.**, 2021). Other studies about public opinion stress the importance of disinformation in polarising society and contributing to extremism (**Müller; Schwarz**, 2020; **Bursztyn; González; Yanagizawa-Drott**, 2018). In Spain, a study by **Casero-Ripolles, Doménech-Fabregat and Alonso-Muñoz** (2023) based on citizens' self-reporting perceptions highlights the capacity of disinformation to shape citizens' positions and beliefs. Relevant here are other methodological approaches based on experiment design (**Richter**, 2017). Participants were shown video clips from *RT* and/or from *BBC*, to then test their opinions about the subject. Participants just exposed to *RT* content were more likely to express a negative opinion of Western policy and were even more likely to trust several fake narratives about Russian-speaking Ukrainians. As pointed out by **Starbird, Arif and Wilson** (2019) the body of literature on the impact of disinformation provides so far different results, as well as a varied selection of methodological approaches. The key aspect here is that, although it is possible that effects of disinformation content are lower than is perceived by the general public, it is plausible that these effects are in fact hard to measure, as it happened with more traditional formats of propaganda (**Bittman**, 1986). Effects are likely going beyond the impact of a certain content, being more acute when narratives are spread through time and in multiple formats. Hence, rather than individuals affected by a particular story on social media, what is more likely is that perceptions are shaped if exposed to a continuous flow of misleading information (**Katz; Lazarsfeld; Roper**, 2017; **Marwick**, 2018).

The idea that exposure may have an effect on attitudes, beliefs and political positions has been a long time present in communication studies (**Gilbert; Tafarodi; Malone**, 1993; **Nyhan**, 2020), highlighting also the potential effect of repetition and exposure through time to different formats about the same topic, disregarding how fake the content might be (**Fazio**, 2015; **Prior**, 2005; **Guess; Nyhan; Reifler**, 2020). As disinformation narratives are designed to attract attention and play with emotions and its contagion (**Kramer; Guillory; Hancock**, 2014; **Stieglitz; Dang-Xuan**, 2013) or current cleavages in society, they might be specially successful in shaping views and perceptions. As **Flore** (2020) points out, repetition increases the ease with which content is analysed and processed, which diminishes accuracy of heuristic mechanisms to assess its veracity.

Hence, we will assess the impact of disinformation narratives, measured in terms of how many respondents believe each narrative, with our third and fourth hypotheses:

H3: Those citizens that have received the narrative before are more likely to believe it than those that have never heard about it.

H4: The effect of “having received the narrative before” is similar in all narratives under study.

3. Methodology

Our methodological approach started with the identification of the most relevant disinformation narratives. We choose to study a three-weeks period, from February 14th to March 7th 2022. No elections at local, regional or national level took place during these weeks, which accomplished our goal to choose a time frame in which no particular relevant event happened, to test disinformation narratives out of elections or particular breaking events. To identify the narratives we used a methodology tested in a previous research at regional level during last Catalan elections (**Suau, Cruz and Yeste**, pending publication). This consists in studying the different material identified as disinformation by Spanish fact-checkers *Newtral*, *Maldita* and *Verificat*, based on the fact-checks published in their websites. The fourth week analysis was conducted on a daily basis, building an excel database that classified fact-checks according to key words and topics addressed. In total, 163 fact-checks have been collected. After the data collection process was concluded our goal was to cluster all content into disinformation narratives. The selection was shared with journalists from fact checker *Newtral*, so both research team and fact checkers could work on clustering the topics related to fact-checks into disinformation narratives. A team of two people from *Newtral* and two researchers worked separately to codify the collected fact-checks. Per each fact-check each team assigned gathered the following information into an excel file: i) a description of the disinformation content identified in the fact-check; ii) up to 6 keywords to resume it; iii) a statement summarising the narrative or idea that the disinformation content is referring to. The two excels were then shared between both teams, to make comments and analyse differences. This process took three days, so by March 10th both teams agreed on 6 main disinformation narratives. To choose the narratives, the following criteria were followed: i) Must be narratives that have been linked to more than one fact-check; ii) Must be narratives referred to national issues, as the survey was to be distributed in all Spain; iii) Must be narratives related to content spread through multiple platforms. The narratives were then identified with a keyword and a description sentence (see table 1). The process has to be done in a short time period, as the main intention of our survey is to test disinformation narratives that are currently happening or that have spread just some days ago.

As Table 1 shows identified narratives are related to different topics. Not surprisingly, two of them (Ns 1 & 5) are related to covid-19 pandemic. Another one, about the war in Ukraine (N2). Despite their relevance, none of these topics is divisive in the sense that represents a certain political cleavage in Spain. Conversely, the rest are related to divisive political issues in current Spanish politics. N3 is related to immigration, while N4 to economy and party positions and N6 to the Catalan case, particularly relevant during recent years. As a conclusion, we expect narratives 3, 4 and 6 to be of a more polarising nature in terms of analysis of results.

Table 1. Identified disinformation narratives (originals in Spanish) (source: authors' own elaboration)

Number	Narrative	Code
1	Poland, Sweden and Norway have filed a lawsuit against the <i>World Health Organization</i> accusing it of infecting Europe.	Covid&WHO
2	Russia is surrounded by hundreds of <i>NATO</i> bases, of neighbouring countries just China and Mongolia have none.	Russia&NATO
3	Young immigrants can collect 450 rental assistance, 250 social bonus plus 200 for being immigrants. If they are “MENAS”, up to a supplement of 1,125 euros.	MENAS
4	The labour law reform has been approved by a vote cast in error. As it is an error, it can be corrected and if the president of parliament does not respond to these requests, the government party is committing an illegality.	Labour law
5	The mRNA-based covid vaccines are scientific trials of unproven effectiveness and with side effects not yet detected	Covid vaccines
6	The Spanish language is not taught in schools in Catalonia. It is common for children to finish school with a much lower level of knowledge than in the rest of Spain	Spanish in CAT

The survey was launched on March 10th until 16th, online, with $N=1.003$, among Spanish residents between 18 and 74 years old, with a margin of error of $\pm 3,4\%$ for a confidence level of 95% and $p=q=0,5$. The survey company *Gesop* was hired to conduct the fieldwork, which was conducted with the usual stratified with uniform affixation sampling method. The strata are formed by the crossing of sex and age, at a rate of 125 interviews in each of the eight strata resulting. Within each stratum, the questionnaires have been distributed accordingly, proportional by autonomous community and size of the municipality of residence. The final distribution included 49% of men and 51% of women. Regarding age 16,95% from 18-29 years old, 28,81% from 30-44, 32% from 45-49 and 22,23% from 60-74. The survey utilised an online panel and ensured the anonymity of respondents. Personal information was not collected as part of the survey process, thereby eliminating any risks associated with the storage and handling of sensitive data. The survey design

adhered to the principles outlined in the *General Data Protection Regulation (GDPR)* to ensure compliance with data protection regulations. As such, the survey methodology prioritised the protection of participants' privacy and the confidentiality of their responses. In terms of data storage and access, the survey data was stored on secure servers provided by a reputable survey platform (*Gesop*). These servers were compliant with industry-standard security protocols to safeguard the data from unauthorised access. Only the research team had access to the survey data, and strict confidentiality measures were implemented to ensure that individual respondents could not be identified or linked to their responses. Regarding informed consent, participants were explicitly informed about the purpose of the research, the voluntary nature of their participation, and the confidentiality measures in place. They were provided with a clear and concise explanation of the survey's objectives and potential risks and benefits. Prior to participating, respondents were required to provide their consent electronically by affirmatively agreeing to a consent statement presented at the beginning of the survey. The consent statement reiterated the voluntary nature of participation and assured respondents of the confidentiality and anonymity of their responses. All methodological aspects of the research are in line with the standards of the authors' university Ethical Committee.

Our methodology allowed us to identify the most relevant disinformation narratives

The survey was structured according to the aforementioned narratives. After an initial set of socio-demographic questions (including here ideology, party identification and media engagement¹), each narrative was presented with a set of questions:

- has the respondent heard the narrative before;
- does the respondent agree with the statement (1-5);
- which media or platform have they heard it from, just one possible answer (if this was the case); and
- had they shared related content.

We conducted the statistical analysis using the *R* programming language, a widely used tool for statistical computing and data analysis. *R* provides a comprehensive suite of packages and functions for conducting various statistical procedures. Specifically, for this study, we utilised several *R* packages for data analysis, such as "tidyverse," "ggplot2," and "dplyr," among others. These packages allowed us to efficiently manipulate, clean, and visualise the survey data. Additionally, we employed various statistical procedures available in *R* to examine the research questions and test hypotheses. The Results section explains the different procedures used to test the aforementioned hypothesis.

4. Results

Our first hypothesis (H1) stated that disinformation narratives were likely going to have a limited reach among the public, with differences among them depending on the topic addressed (higher the more political or polarised). However, results show otherwise (see figure 1). Just one of the narratives shows limited reach (N1-Covid&WHO), being heard by 29% of respondents. Other two narratives (N3-MENAS and N5-CovidVaccines), related to immigration and covid-related disinformation, were spread among 60%. Despite being a high percentage, there were three other narratives that overpass 70% reach. These were N2 (Russia&NATO), N4 (Labor law) and N6 (Spanish in CAT). Hence, results showed that 5 of the 6 identified narratives were highly spread among the respondents of our survey. The next step is, then, to identify how citizens were exposed to those narratives. Hence, our first hypothesis is not verified, as all narratives but one were spread among more than 50% of respondents, showing then high levels of reach, rather than low ones. Furthermore, apart from N1, all other narratives show levels of reach between 60-70%, being just N6 close to 73%. The small differences among narrative in terms of reach does not allow us to decide if a more polarising nature of a certain narrative implies higher reach levels.

Our hypothesis 2 states that disinformation narratives will likely be more spread through television rather than digital media, with also social media playing an important role. Table 2 summarises our findings here. It needs to be highlighted that in order to structure the table we have grouped different answers into the six final categories. The category "TV" groups both "news at TV" and also "entertainment programs at TV", although this last one reported actually almost irrelevant results. Category "Social Media" includes most of the common social media (*Twitter, Facebook, Ins-*

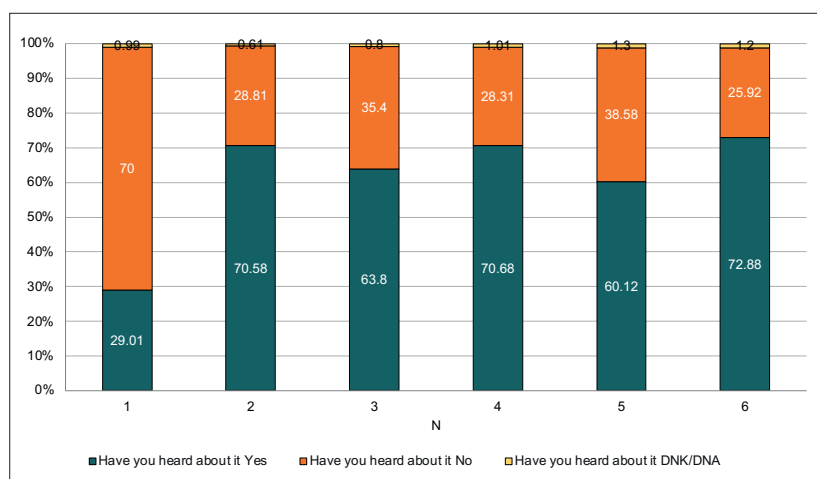


Figure 1. Reach of disinformation narratives

tagram, TikTok...) but also messaging Apps such as *WhatsApp* or *Telegram*. The category “Public talk” gathers results from different situations such as “Talking with friends or family” or “Talking at workplace”. “Radio” includes traditional radio and also podcasts (although this last one reported almost irrelevant results). Finally, the category “Digital” collects both websites of legacy media and digital native ones. The table stresses in dark green the most common channel of dissemination per each narrative, with a lighter green for the second one and an even lighter green for the third one. As it can be seen, TV is the most relevant spreading channel in three of the narratives (N2, N4 and N6) while Social Media dominates in the other three (N1, N3 and N5). While the role of Press, Public talk and Radio is scarcely relevant, Digital represents a relevant role in the spread of disinformation narratives. Although it never reaches first position, it represents between 14 and 27% of the spread of disinformation narratives. Furthermore, it is relevant here to stress that spreading patterns show great degrees of variability. For example, N4 (labour law) was heard by 60% from TV, while just 14% heard it from Digital and 13% from Social Media. This might be the fact that is related to a political debate that was happening at that time between the two main political parties in Spain. Similar results show N6 (Spanish in CAT) with also greater presence through TV and around 18-20% in Digital and Social Media). In comparison, N1 (Covid&WHO) and N3 (MENAS) show very different patterns, being spread through Social Media (36 and 41% respectively). Hence, N1 seems to follow a mostly online pattern, as is also the one in which Digital shows higher levels (27%) while is the one with lower results at TV (21%). Finally, N5 (Covid Vaccines) is the one that shows more similar results among the three most common channels: 36% from Social media, 32% from TV and 19% from Digital. Results confirm our hypothesis 2, which states that disinformation narratives were going to be more spread through TV than digital media, with social media and messaging apps having also a predominant role. Interestingly, two of the three narratives (Ns 4 & 6) that we identified as part of political debate are reaching citizens mainly through television. The other narrative that is being mostly heard from TV is N2 about war in Ukraine. This might be because the conflict had a daily presence in the news. In general, narratives 2, 4 and 6 the ones where TV plays a greater role were also the ones with higher levels of reach. For the rest of the narratives, social media plays a predominant role. These are the ones with lower levels of reach, although as stated the difference is of around just 10 points. Except for N1 which was the only narrative with a low level of reach (29%). This is the narrative that shows a very different spreading pattern: mostly online through social media and Digital, with the lowest relevance from TV (21%).

Table 2. Spreading channels of disinformation narratives

N	Press	Digital	TV	Public talk	Radio	Social Media
1	7.2993	27.0073	21.1679	0.7299	7.6642	36.1314
2	2.2760	21.0930	43.8540	0.6070	5.6150	26.5550
3	3.4880	18.6050	28.0730	3.9870	4.4850	41.3620
4	4.1787	14.1210	60.0865	0.8646	7.0605	13.6888
5	3.2990	19.4440	32.2920	2.2570	5.9030	36.8060
6	4.1310	18.0910	49.0030	2.5640	6.2680	19.9430

To complete our study, hypothesis 3 and 4 are focused on the impact of disinformation narratives. H3 proposes that those citizens that have received the narrative before are more likely to believe it than those that have never heard about it. To start, table 3 shows the general results of trust on disinformation narratives, grouped in Not believing, Neutral and Believing (the original question was a 1-5 likert scale, in which we grouped 1-2 results and 4-5 ones, being 3 the neutral one). As it can be seen, the different narratives show different levels of acceptance among survey respondents. Just one narrative (N2 Russia&NATO) is trusted by more than 50%, while other two show results close to 46% (N4 Labor Law and N6 Spanish in CAT). The other three narratives show levels of acceptance between 22 and 37%.

If we compare these results with Figure 1 we can see that those narratives with higher percentage of dissemination among survey respondents are also those that are more trusted: N2 (Russia&NATO), N4 (Labor law) and N6 (Spanish in CAT). In order to fully test this relationship and test H3, a series of bivariate chi2 tests is performed, where the dependent variable is believing, being neutral, or not believing the narrative, and the independent variable is having been exposed to it before. Cramer's V values are used to evaluate the strength of the successive bivariate associations. Table 4 summarises the results of the chi2 tests. Of the six narratives studied, 3 of them show a strong relationship between having received the narrative before and believing it, in the other three the effect is moderate-strong. Hence, having received the narrative beforehand is positively and significantly associated with

Table 3. Trust on narratives

N	Trust on narrative			
	Not believing	Neutral	Believing	DNK/DNA
1	49.55	27.43	22.13	0.89
2	17.41	29.12	52.14	1.33
3	39.58	21.33	37.98	1.11
4	29.51	22.53	46.16	1.80
5	35.49	29.81	33.89	0.81
6	36.69	16.75	45.56	1.00

an increased tendency to believe in it, regardless of the topic under discussion. This points towards a relevant effect of disinformation content on those citizens exposed to it, confirming our hypothesis 3. Finally, our last hypothesis, H4, is also confirmed. It argues that the effect of “having received the narrative before” will be similar in all narratives under study. As seen in Table 4, the Cramer’s V results show that the strength of the associations is similar, with Ns 1, 2 and 4 showing large effects, others showing medium-large effects. Although Cramer’s V

Table 4. Trust on narratives depending on received before

N	Association measures	
	Cramer’s V	p value
1	0.480 (df2)	0.000
2	0.404 (df2)	0.000
3	0.262 (df2)	0.000
4	0.423 (df2)	0.000
5	0.307 (df2)	0.000
6	0.318 (df2)	0.000

does not allow us to compare between narratives, the fact that most of them show similar results can be indicative of the relevance of the “having received the narrative before” effect, especially taking into account that the topics of each narrative are very diverse and their spread level, as seen before, shows also different patterns.

5. Discussion

Our research has presented a methodology to research the spread and impact of disinformation narratives from a holistic approach that does not rely on data from a single social media or dissemination channel. Such a methodology, like any other, has its own limitations, but allows us to evaluate disinformation not in terms of single pieces of information but at narrative level, assessing the impact that such narratives have in a certain society. It also allows us to “test” public opinion, avoiding survey design based on citizens’ own perceptions of the impact of disinformation. Former research (Vosoughi; Roy; Aral, 2018; Spangher *et al.*, 2018; Zannettou *et al.*, 2019; Arce-García; Said-Hung; Mottareale-Calvanese, 2022; Masip *et al.*, 2020; Grinberg *et al.*, 2019; Allen *et al.*, 2020; Fletcher; Nielsen, 2018) has indicated that disinformation hardly reached relevant numbers of citizens, especially among those studies that are single platform-based. Our research indicates that disinformation narratives might be more spread among public opinion than what is generally perceived in platform-based studies. The level of reach of most narratives in our research is between 60-72% of survey respondents, being just one of the narratives at low levels of reach (29%). This allows us to state that the detected narratives reached the general public, rather than being spread just among certain societal groups. Hence, just one narrative (N1) plays at the level of localised reach (Allen *et al.*, 2020; Dawson, 2021). As explained in the introduction, to reach dissemination among the general public is one of the main goals of disinformation campaigns (Polletta; Callahan, 2019). The goal of this research is not to determine the origin of each one of the spotted narratives. However, we do know that disinformation campaigns, foreign and domestic, intend to build on existing cleavages and use emotion and relate to political issues to enhance polarisation (Richter, 2017; Colley; Granelli; Althuis, 2020; Bánkuty-Balogh, 2021). Three of our spotted narratives can be identified as related to current political debates (Ns 3, 4 and 6), while other two are related to vaccination in Covid-19 pandemic (Ns 1 and 5) and the Russian invasion of Ukraine (N2). Neither covid vaccines nor Russian war in Ukraine are specially divisive topics in Spain. In general, narratives related to current political cleavages reached higher percentages of the population, confirming that disinformation spreads at higher levels if related to emotional and divisive political issues. The only exception is N2, but its higher levels of reach may be due to the constant presence of the Ukraine war in news spaces.

To analyse each narrative spreading pattern can be also interesting to consider its origin. Foreign and domestic disinformation campaigns are more likely to be spread through social media and messaging apps (Starbird; Arif; Wilson, 2019) rather than through traditional news media. The different results that can be seen in spreading patterns among spotted narratives can be interpreted in this line. However, it cannot be disregarded the participatory nature of disinformation in analysing spreading patterns. In all narratives except N4 and N6 social media (including messaging apps) showed high percentages of spread. This can indicate the presence of disinformation campaigns, but also the potential of these narratives to be spread by common citizens on social media who share content in line with their political beliefs or that addresses emotionally own beliefs and values (Buchanan, 2020; Wilson; Starbird, 2020). Interestingly, among those narratives more spread through social media we have the two related with covid vaccination (N1 and 5), plus the narrative that addresses immigration topic (N3). While vaccination has never been a divisive political issue in Spain, immigration is a relevant one. Our research shows that disinformation narratives spread predominantly through social media do have the potential to reach relevant percentages of the general population. However, we cannot identify the origin of such content, to establish if a certain narrative originated through an organised campaign, as our methodological approach does not consider this issue. Nevertheless, in line of results we can argue that single-platform studies that pointed out the limited reach of disinformation campaigns on a single platform (Spangher *et al.*, 2018; Zannettou *et al.*, 2019; Arce-García; Said-Hung; Mottareale-Calvanese, 2022) might be relevant in relation to this platform, but need to be complemented if the aim is to assess real spread or dissemination of such content at societal level. Mo-

Most of disinformation narratives reached more than 50% of survey participants

reover, the relevant results of social media as spreading pattern indicate that such mechanisms of dissemination reach relevant percentages of the overall population, which may contradict the theory of social media and messaging apps as spaces where just politically in line content is found (Masip; Suau; Ruiz-Caballero, 2020; Cardenal et al., 2019).

Study disinformation narratives allows us to assess impact beyond single pieces of content

Furthermore, confirming hypothesis 2, our results have shown that one of the main spreading patterns of disinformation in Spain are news programs on television. Generally disregarded in most disinformation studies, mostly based on online platforms, TV seems to deserve more attention. The relevant role of TV, which in Spain is one of the most common sources of news, can be of a double nature. Firstly, television programs can act as spreaders by repetition. Former research has pointed out the potential role of television as an amplifier of disinformation narratives produced by political actors (Phillips, 2018; Tsfaty et al., 2020). Our results may confirm these theories, as television is normally the first or second main spreading pattern, but at closer distance than others. Another possibility, which may cohabitate with the former one, is that television programs are not just amplifiers but also producers of disinformation. This might be the case of some narratives, mainly spread through television and with lower percentages in other formats. For example, N4 and to a lesser extent N6 are narratives strongly related to political issues, where television shows much higher percentages than the others (in N4, for example, TV represents 60% while the second spreading pattern, Digital, just 14%). Hence, the role of journalists in relation to disinformation cannot be avoided, both as amplifiers but also as producers. Our findings suggest that future research should focus on both traditional and digital media as sources of disinformation, rather than just mere unwilling or passive spreaders as is sometimes assumed. This means to also focus on their ownership in terms of transparency and possible economic or political influence that drive their reporting and editorial line. Similar research in other countries may shed some light on this, to better understand if such a pattern exists in other societal contexts.

However, in our opinion the most relevant (and worrisome) result in our research is the proof that being exposed to disinformation content plays an important role in assuming the inherent narrative, confirming H3. Furthermore, confirming H4, this effect is relevant in all narratives under study, disregarding their more or less political or emotional nature. Our findings highlight that the effect of being exposed to disinformation narratives is consistent for all topics. No differences have been found in spreading patterns, gender or age during the testing of results prior drafting the paper. Hence, the effects of being exposed to disinformation narratives are widespread, being similar for narratives that reach 29% of the respondents and also for those that passed 70%. Previous studies have shown that disinformation content can shape citizens' attitudes and positions, but were mostly focused on elections or particular breaking events (Kazeem, 2018; Silverman; Singer-Vine, 2016; Kapantai et al., 2021; Richter, 2017). As pointed out by Starbird, Arif and Wilson (2019) there is a need to better understand how disinformation narratives work through time, shaping public opinion in "normal" times. Results presented in this paper contribute to this line of research by showing how being exposed to disinformation content has an impact on citizens' assessment of what is true or not. We think that this is a relevant aspect of our methodological design, allowing us to detect an aspect of disinformation narratives that may pass as unnoticed in other research methodologies: as explained, disinformation effects are hard to measure and exist beyond the spread or impact of a certain singular piece of content, which makes them hard to measure, as happened with traditional forms of propaganda (Bittman, 1986). The adoption of narratives as a pivotal concept in the research (Weinberg; Dawson, 2021) allowed us to track impact through time and in multiple formats that may pass unnoticed in other methodological designs.

Nevertheless, our methodological design also presents limitations that future research may address to complete the whole picture. Firstly, once we are aware of the relevance of narratives we need to focus on their repetition. It might happen that citizens are particularly influenced if they are exposed to continuous flows of disinformation-related content (Katz; Lazarsfeld; Roper, 2017; Marwick, 2018). This was not considered in our methodological design, and we strongly suggest future research to address this issue. Likewise, the type of emotion that the different narratives chosen could generate in citizens has not been considered, as this is another interesting factor related with the spread of disinformation (Kramer; Guillory; Hancock, 2014; Stieglitz; Dang-Xuan, 2013). Moreover, despite that our Cramer's V values have been significant, survey research has always limitations. For example, such tests do not allow differentiating between competing causal explanations for such results. Further research may also study other possible factors that may influence the significant relationship found in our research: ideology of research participants may also shape their trust levels on the different narratives (Suau, Cruz and Yeste, pending publication). On the other hand, these particular messages might have *selection bias*, reaching a larger proportion of people already primed for believing in them. Further research would be needed in order to differentiate between these alternative explanations.

This research proves that disinformation has high impact on citizens

6. Note

1. For media engagement we include different questions related to kind of media consumption and frequency of news' consumption. These variables were not used in the analysis of results presented in this paper.

7. References

- AIMC (2022). *Estudio general de medios*. Asociación para la Investigación de Medios de Comunicación (AIMC). <https://reporting.aimc.es/index.html#/main/cockpit>
- Allen, Jennifer; Howland, Baird; Mobius, Markus; Rothschild, David; Watts, Duncan J. (2020). "Evaluating the fake news problem at the scale of the information ecosystem". *Science advances*, v. 6, n. 14, eaay3539. <https://doi.org/10.1126/sciadv.aay3539>
- Almenar, Ester; Puertas-Graell, David; Ruiz-Caballero, Carlos; Suau, Jaume; Masip, Pere (2022). "Conocimiento y uso de los verificadores de datos: ¿una herramienta para la regeneración de la autoridad periodística en un entorno polarizado?". *Anàlisi: quaderns de comunicació i cultura*, v. 66, pp. 113-129. <https://doi.org/10.5565/rev/analisi.3472>
- Altay, Sacha; Berriche, Manon; Acerbi, Alberto (2023). "Misinformation on misinformation: conceptual and methodological challenges". *Social media + society*, v. 9, n. 1. <https://doi.org/10.1177/20563051221150412>
- Amazeen, Michelle A.; Bucy, Erik P. (2019). "Conferring resistance to digital disinformation: the inoculating influence of procedural news knowledge". *Journal of broadcasting & electronic media*, v. 63, n. 3, pp. 415-432. <https://doi.org/10.1080/08838151.2019.1653101>
- Arce-García, Sergio; Said-Hung, Elías; Mottareale-Calvanese, Daria (2022). "Astroturfing as a strategy for manipulating public opinion on Twitter during the pandemic in Spain". *Profesional de la información*, v. 31, n. 3, e310310. <https://doi.org/10.3145/epi.2022.may.10>
- Bánkuty-Balogh, Lilla-Sarolta (2021). "Novel technologies and geopolitical strategies: disinformation narratives in the countries of the Visegrád group". *Politics in central Europe*, v. 17, n. 2, pp. 165-195. <https://doi.org/10.2478/pce-2021-0008>
- Bastos, Marco; Mercea, Dan (2018). "The public accountability of social platforms: lessons from a study on bots and trolls in the Brexit campaign". *Philosophical transactions of the Royal Society A. Mathematical, physical and engineering sciences*, v. 376, n. 2128, 20180003. <https://doi.org/10.1098/rsta.2018.0003>
- Besalú, Reinald; Pont-Sorribes, Carles (2021). "Credibility of digital political news in Spain: comparison between traditional media and social media". *Social sciences*, v. 10, n. 5, 170. <https://doi.org/10.3390/socsci10050170>
- Brennen, J. Scott; Simon, Felix M.; Howard, Philip N.; Nielsen, Rasmus-Kleis (2020). *Types, sources, and claims of Covid-19 misinformation*. University of Oxford. Reuters Institute. <https://reutersinstitute.politics.ox.ac.uk/types-sources-and-claims-covid-19-misinformation>
- Buchanan, Tom (2020). "Why do people spread false information online? The effects of message and viewer characteristics on self-reported likelihood of sharing social media disinformation". *Plos one*, v. 15, n. 10, e0239666. <https://doi.org/10.1371/journal.pone.0239666>
- Bursztyn, Leonardo; González, Alessandra L.; Yanagizawa-Drott, David (2018). *Misperceived social norms: female labor force participation in Saudi Arabia*. National Bureau of Economic Research. Working paper series 24736. <https://doi.org/10.3386/w24736>
- Cardenal, Ana S.; Aguilar-Paredes, Carlos; Cristancho, Camilo; Majó-Vázquez, Sílvia (2019). "Echo-chambers in online news consumption: evidence from survey and navigation data in Spain". *European journal of communication*, v. 34, n. 4, pp. 360-376. <https://doi.org/10.1177/0267323119844409>
- Cardoso, Gustavo; Sepúlveda, Rita; Narciso, Inês (2022). "WhatsApp and audio misinformation during the Covid-19 pandemic". *Profesional de la información*, v. 31, n. 3. <https://doi.org/10.3145/epi.2022.may.21>
- Casero-Ripollés, Andreu; Doménech-Fabregat, Hugo; Alonso-Muñoz, Laura (2023). "Percepciones de la ciudadanía española ante la desinformación en tiempos de la Covid-19: efectos y mecanismos de lucha contra las noticias falsas". *Icono 14*, v. 21, n. 1. <https://doi.org/10.7195/ri14.v21i1.1988>
- Chadwick, Andrew; Vaccari, Cristian; O'Loughlin, Ben (2018). "Do tabloids poison the well of social media? Explaining democratically dysfunctional news sharing". *New media & society*, v. 20, n. 11, pp. 4255-4274. <https://doi.org/10.1177/1461444818769689>
- Chagas, Viktor (2022). "Whatsapp and digital astroturfing: a social network analysis of Brazilian political discussion groups of Bolsonaro's supporters". *International journal of communication*, v. 16, n. 25, pp. 2431-2455. <https://ijoc.org/index.php/ijoc/article/view/17296>

- Chan, Jovy** (2022). "Online astroturfing: a problem beyond disinformation". *Philosophy & social criticism*, Online first. <https://doi.org/10.1177/01914537221108467>
- Cho, Charles H.; Martens, Martin L.; Kim, Hakkyun; Rodrigue, Michelle** (2011). "Astroturfing global warming: it isn't always greener on the other side of the fence". *Journal of business ethics*, 104, pp. 571-587. <https://doi.org/10.1007/s10551-011-0950-6>
- Colley, Thomas P.; Granelli, Francesca; Althuis, Jente** (2020). "Disinformation's societal impact: Britain, Covid and beyond". *Defence strategic communications*, v. 8, n. 1, pp. 89-140. <https://doi.org/10.1080/10.30966/2018.RIGA.8.3>
- Das, Anupam; Schroeder, Ralph** (2020). "Online disinformation in the run-up to the Indian 2019 election". *Information, communication & society*, v. 24, n. 12, pp. 1762-1778. <https://doi.org/10.1080/1369118X.2020.1736123>
- Dawson, Grant** (2021). "Classical realism, status, and emotions: understanding the Canada/Saudi Arabia dispute and its implications for global politics". *Global studies quarterly*, v. 1, n. 4. <https://doi.org/10.1093/isagsq/ksab027>
- Edgerly, Stephanie; Vraga, Emily K.** (2019). "News, entertainment, or both? Exploring audience perceptions of media genre in a hybrid media environment". *Journalism*, v. 20, n. 6, pp. 807-826. <https://doi.org/10.1177/1464884917730709>
- Espaliú-Berdud, Carlos** (2022). "Legal and criminal prosecution of disinformation in Spain in the context of the European Union". *Profesional de la información*, v. 31, n. 3, e310322. <https://doi.org/10.3145/epi.2022.may.22>
- Farkas, Johan; Schou, Jannick** (2018). "Fake news as a floating signifier: hegemony, antagonism and the politics of falsehood". *Javnost-the public*, v. 25, n. 3, pp. 298-314. <https://doi.org/10.1080/13183222.2018.1463047>
- Fazio, Lisa K.; Brashier, Nadia M.; Payne, B. Keith; Marsh, Elizabeth J.** (2015). "Knowledge does not protect against illusory truth". *Journal of experimental psychology general*, v. 144, n. 5, pp. 993-1002. <https://doi.org/10.1037/xge0000098>
- Fernández-Torres, María-Jesús; Almansa-Martínez, Ana; Chamizo-Sánchez, Rocío** (2021). "Infodemic and fake news in Spain during the Covid-19 pandemic". *International journal of environmental research and public health*, v. 18, n. 4, 1781. <https://doi.org/10.3390/ijerph18041781>
- Fletcher, Richard; Nielsen, Rasmus-Kleis** (2018). "Are people incidentally exposed to news on social media? A comparative analysis". *New media & society*, v. 20, n. 7, pp. 2450-2468. <https://doi.org/10.1177/1461444817724170>
- Flore, Massimo** (2020). *Understanding citizens' vulnerabilities (II): from disinformation to hostile narratives*. Luxembourg: European commission. ISBN: 978 9276143086 <https://doi.org/10.2760/271224>
- Freiling, Isabelle; Krause, Nicole M.; Scheufele, Dietram A.; Brossard, Dominique** (2023). "Believing and sharing misinformation, fact-checks, and accurate information on social media: the role of anxiety during Covid-19". *New media & society*, v. 25, n. 1, pp. 141-162. <https://doi.org/10.1177/14614448211011451>
- García-Marín, David; Salvat-Martinrey, Guiomar** (2022). "Tendencias en la producción científica sobre desinformación en España. Revisión sistematizada de la literatura (2016-2021)". *Adcomunica*, n. 23, pp. 23-50. <https://doi.org/10.6035/adcomunica.6045>
- Gilbert, Daniel T.; Tafarodi, Romin W.; Malone, Patrick S.** (1993). "You can't not believe everything you read". *Journal of personality and social psychology*, v. 65, n. 2, pp. 221-233. <https://doi.org/10.1037/0022-3514.65.2.221>
- Grinberg, Nir; Joseph, Kenneth; Friedland, Lisa; Swire-Thompson, Briony; Lazer, David** (2019). "Fake news on Twitter during the 2016 US presidential election". *Science*, v. 363, n. 6425, pp. 374-378. <https://doi.org/10.1126/science.aau2706>
- Guess, Andrew M.; Nyhan, Brendan; Reifler, Jason** (2020). "Exposure to untrustworthy websites in the 2016 US election". *Nature human behaviour*, v. 4, n. 5, pp. 472-480. <https://doi.org/10.1038/s41562-020-0833-x>

- Hameleers, Michael** (2020). "Separating truth from lies: comparing the effects of news media literacy interventions and fact-checkers in response to political misinformation in the US and Netherlands". *Information, communication & society*, v. 25, n. 1, pp. 110-126.
<https://doi.org/10.1080/1369118X.2020.1764603>
- Hameleers, Michael; Van-der-Meer, Toni G. L. A.** (2020). "Misinformation and polarization in a high-choice media environment: how effective are political fact-checkers?" *Communication research*, v. 47, n. 2, pp. 227-250.
<https://doi.org/10.1177/0093650218819671>
- Howard, Philip N.** (2006). "New media campaigns and the managed citizen". Cambridge, UK: Cambridge University Press. ISBN: 978 0511615986
<https://doi.org/10.1017/CBO9780511615986>
- Jankowski, Nicholas W.** (2018). "Researching fake news: a selective examination of empirical studies". *Javnost-The public*, v. 25, n. 1-2, pp. 248-255.
<https://doi.org/10.1080/13183222.2018.1418964>
- Kapantai, Eleni; Christopoulou, Androniki; Berberidis, Christos; Peristeras, Vassilios** (2021). "A systematic literature review on disinformation: toward a unified taxonomical framework". *New media & society*, v. 23, n. 5, pp. 1301-1326.
<https://doi.org/10.1177/1461444820959296>
- Katz, Elihu; Lazarsfeld, Paul F.; Roper, Elmo** (2017). *Personal influence: the part played by people in the flow of mass communications*. Routledge. ISBN: 978 1412805070
- Kazeem, Yomi** (2018). "Nigerian media houses are forming a coalition to combat fake news ahead of next year's elections". *Quartz Africa*, November 29.
<http://qz.com/africa/1478737/fake-news-mediacollaborate-ahead-of-nigeria-2019election>
- Keller, Franciska B.; Schoch, David; Stier, Sebastian; Yang, JungHwan** (2020). "Political astroturfing on Twitter: how to coordinate a disinformation campaign". *Political communication*, v. 37, n. 2, pp. 256-280.
<https://doi.org/10.1080/10584609.2019.1661888>
- Kramer, Adam D. I.; Guillory, Jamie E.; Hancock, Jeffrey T.** (2014). "Experimental evidence of massive-scale emotional contagion through social networks". *Proceedings of the National Academy of Sciences*, v. 111, n. 24, pp. 8788-8790.
<https://doi.org/10.1073/pnas.1320040111>
- Leiser, Mark** (2016). "Astroturfing, 'cyberturfing' and other online persuasion campaigns". *European journal of law and technology*, v. 7, n. 1.
<https://ejlt.org/index.php/ejlt/article/view/501>
- Lewandowsky, Stephan; Ecker, Ulrich K.; Seifert, Colleen M.; Schwarz, Norbert; Cook, John** (2012). "Misinformation and its correction: continued influence and successful debiasing". *Psychological science in the public interest*, v. 13, n. 3, pp. 106-131.
<https://doi.org/10.1177/1529100612451018>
- Maros, Alexandre; Almeida, Jussara; Benevenuto, Fabrício; Vasconcelos, Marisa** (2020). "Analyzing the use of audio messages in Whatsapp groups". In: *Proceedings of the web conference 2020*, pp. 3005-3011.
<https://homepages.dcc.ufmg.br/~fabricio/download/www20-maros.pdf>
- Marwick, Alice E.** (2018). "Why do people share fake news? A sociotechnical model of media effects". *Georgetown law technology review*, v. 2, n. 2, pp. 474-512.
<https://georgetownlawtechreview.org/wp-content/uploads/2018/07/2.2-Marwick-pp-474-512.pdf>
- Masip, Pere; Aran-Ramspott, Sue; Ruiz-Caballero, Carlos; Suau, Jaume; Almenar, Ester; Puertas-Graell, David** (2020). "Consumo informativo y cobertura mediática durante el confinamiento por el Covid-19: sobreinformación, sesgo ideológico y sensacionalismo". *Profesional de la información*, v. 29, n. 3, e290312.
<https://doi.org/10.3145/epi.2020.may.12>
- Masip, Pere; Suau, Jaume; Ruiz-Caballero, Carlos** (2020). "Percepciones sobre medios de comunicación y desinformación: ideología y polarización en el sistema mediático español". *Profesional de la información*, v. 29, n. 5, e290527.
<https://doi.org/10.3145/epi.2020.sep.27>
- Moreno-Castro, Carolina; Vengut-Climent, Empar; Cano-Orón, Lorena; Mendoza-Poudereux, Isabel** (2021). "Estudio exploratorio de los bulos difundidos por Whatsapp en España para prevenir o curar la Covid-19". *Gaceta sanitaria*, v. 35, n. 6, pp. 534-540.
<https://doi.org/10.1016/j.gaceta.2020.07.008>
- Müller, Karsten; Schwarz, Carlo** (2020). "Fanning the flames of hate: social media and hate crime". *Journal of the European economic association*, v. 19, n. 4, pp. 2131-2167.
<https://doi.org/10.1093/jea/jvaa045>

- Napoli, Philip M.** (2019). *Social media and the public interest: media regulation in the disinformation age*. Columbia university press. ISBN: 978 0231184540
- Newman, Nic; Fletcher, Richard; Robertson, Craig T.; Eddy, Kirsten; Nielsen, Rasmus-Kleis**(2022). *Digital news report 2022*. Oxford: Reuters institute for the study of journalism.
https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2022-06/Digital_News-Report_2022.pdf
- Nyhan, Brendan** (2020). "Facts and myths about misperceptions". *Journal of economic perspectives*, v. 34, n. 3, pp. 220-236.
<https://doi.org/10.1257/jep.34.3.220>
- Nyhan, Brendan; Porter, Ethan; Reifler, Jason; Wood, Thomas J.** (2020). "Taking fact-checks literally but not seriously? The effects of journalistic fact-checking on factual beliefs and candidate favorability". *Political behavior*, 42, pp. 939-960.
<https://doi.org/10.1007/s11109-019-09528-x>
- Pérez-Escobar, Marta; Lilleker, Darren; Tapia-Frade, Alejandro** (2023). "A systematic literature review of the phenomenon of disinformation and misinformation". *Media and communication*, v. 11, n. 2, pp. 76-87.
<https://doi.org/10.17645/mac.v11i2.6453>
- Phillips, Whitney** (2018). *The oxygen of amplification*. Data & society.
<https://datasociety.net/library/oxygen-of-amplification/>
- Polletta, Francesca; Callahan, Jessica** (2019). "Deep stories, nostalgia narratives, and fake news: storytelling in the Trump era". In: *Politics of meaning/meaning of politics: cultural sociology of the 2016 US presidential election*, pp. 55-73.
https://doi.org/10.1007/978-3-319-95945-0_4
- Prior, Markus** (2005). "News vs. entertainment: how increasing media choice widens gaps in political knowledge and turnout". *American journal of political science*, v. 49, n. 3, pp. 577-592.
<https://doi.org/10.1111/j.1540-5907.2005.00143.x>
- Resende, Gustavo; Melo, Philipe; Sousa, Hugo; Messias, Johnnatan; Vasconcelos, Marisa; Almeida, Jussara-Marques; Benevenuto, Fabrício** (2019, May). "(Mis)information dissemination in *Whatsapp*: gathering, analyzing and countermeasures". In: *The world wide web conference*, pp. 818-828.
<https://doi.org/10.1145/3308558.3313688>
- Richter, Monika L.** (2017). *The Kremlin's platform for 'useful idiots' in the west: an overview of RT's editorial strategy and evidence of impact*. European values. Kremlin watch report, 18 September.
<http://www.elespiadigital.com/images/stories/Documentos9/Overview-of-RTs-Editorial-Strategy-and-Evidence-of-Impact.pdf>
- Sádaba, Cháro; Salaverría, Ramón** (2023). "Tackling disinformation with media literacy: analysis of trends in the European Union". *Revista latina de comunicación social*, n. 81, pp. 17-32.
<https://www.doi.org/10.4185/RLCS-2023-1552>
- Shin, Jieun; Thorson, Kjerstin** (2017). "Partisan selective sharing: the biased diffusion of fact-checking messages on social media". *Journal of communication*, v. 67, n. 2, pp. 233-255.
<https://doi.org/10.1111/jcom.12284>
- Silverman, Craig; Singer-Vine, Jeremy** (2016). "Most Americans who see fake news believe it, new survey says". *Buzz-Feed news*, December 6.
<https://www.buzzfeednews.com/article/craigsilverman/fake-news-survey>
- Spangher, Alexander; Ranade, Gireeja; Nushi, Besmira; Fourney, Adam; Horvitz, Eric** (2018). "Analysis of strategy and spread of Russia-sponsored content in the US in 2017". *Arxiv*, 1810.10033.
<https://doi.org/10.48550/arXiv.1810.10033>
- Staender, Anna; Humprecht, Edda; Esser, Frank; Morosoli, Sophie; Van-Aelst, Peter** (2021). "Is sensationalist disinformation more effective? Three facilitating factors at the national, individual, and situational level". *Digital journalism*, v. 10, n. 6, pp. 976-996.
<https://doi.org/10.1080/21670811.2021.1966315>
- Starbird, Kate; Arif, Ahmer; Wilson, Tom** (2019). "Disinformation as collaborative work: surfacing the participatory nature of strategic information operations". *Proceedings of the ACM on human-computer interaction*, v. 3, n. 127.
<https://doi.org/10.1145/3359229>
- Stieglitz, Stefan; Dang-Xuan, Linh** (2013). "Emotions and information diffusion in social media—sentiment of microblogs and sharing behavior". *Journal of management information systems*, v. 29, n. 4, pp. 217-248.
<https://doi.org/10.2753/MIS0742-1222290408>
- Suau, Jaume** (2015). *Citizens and online media participation. Attitudes and motivations towards participatory journalism and other online practices in London and Barcelona*. Doctoral dissertation. Universitat Ramon Llull.
<http://hdl.handle.net/10803/289347>

- Suau, Jaume; Cruz, Irene; Yeste, Elena** (pending publication). "Assessing the impact of disinformation narratives in a polarized electoral campaign: the case of 2021 Catalan elections". Paper presentado en la *Eighth conference of the International Journal of Press/Politics*, Loughborough University, 21-23 September 2022).
- Tandoc, Edson C.; Lim, Darren; Ling, Rich** (2020). "Diffusion of disinformation: how social media users respond to fake news and why". *Journalism*, v. 21, n. 3, pp. 381-398.
<https://doi.org/10.1177/1464884919868325>
- Tsfati, Yariv; Boomgaarden, Hajo G.; Strömbäck, Jesper; Vliegenthart, Rens; Damstra, Alyt; Lindgren, Elina** (2020). "Causes and consequences of mainstream media dissemination of fake news: literature review and synthesis". *Annals of the international communication association*, v. 44, n. 2, pp. 157-173.
<https://doi.org/10.1080/23808985.2020.1759443>
- Van-der-Linden, Sander; Panagopoulos, Costas; Roozenbeek, Jon** (2020). "You are fake news: political bias in perceptions of fake news". *Media, culture & society*, v. 42, n. 3, pp. 460-470.
<https://doi.org/10.1177/0163443720906992>
- Vosoughi, Soroush; Roy, Deb; Aral, Sinan** (2018). "The spread of true and false news online". *Science*, v. 359, n. 6380, pp. 1146-1151.
<https://doi.org/10.1126/science.aap9559>
- Vraga, Emily K.; Bode, Leticia; Tully, Melissa** (2022). "Creating news literacy messages to enhance expert corrections of misinformation on Twitter". *Communication research*, v. 49, n. 2, pp. 245-267.
<https://doi.org/10.1177/0093650219898094>
- Waisbord, Silvio** (2018). "Truth is what happens to news: on journalism, fake news, and post-truth". *Journalism studies*, v. 19, n. 13, pp. 1866-1878.
<https://doi.org/10.1080/1461670X.2018.1492881>
- Walker, Edward T.** (2014). *Grassroots for hire: public affairs consultants in American democracy*. New York: Cambridge university press. ISBN: 978 1107619012
- Walter, Nathan; Cohen, Jonathan; Holbert, R. Lance; Morag, Yasmin** (2020). "Fact-checking: a meta-analysis of what works and for whom". *Political communication*, v. 37, n. 3, pp. 350-375.
<https://doi.org/10.1080/10584609.2019.1668894>
- Wanless, Alicia; Berk, Michael** (2021). "Participatory propaganda: the engagement of audiences in the spread of persuasive communications". In: Herbert, David; Fisher-Høyrem, Stefan (eds.). *Social media and social order*. De Gruyter Open Poland, pp. 111-139. ISBN: 978 8366675612
<https://doi.org/10.2478/9788366675612-009>
- Wardle, Claire** (2018). "The need for smarter definitions and practical, timely empirical research on information disorder". *Digital journalism*, v. 6, n. 8, pp. 951-963.
<https://doi.org/10.1080/21670811.2018.1502047>
- Wardle, Claire; Derakhshan, Hossein** (2017). *Information disorder: toward an interdisciplinary framework for research and policy making*. Strasbourg: Council of Europe.
<https://edoc.coe.int/en/media/7495-information-disorder-toward-an-interdisciplinary-framework-for-research-and-policy-making.html>
- Weinberg, Danna B.; Dawson, Jessica** (2021). *From anti-vaxxer moms to militia men: Influence operations, narrative weaponization, and the fracturing of American identity*. Massachusetts, The Brookings institution.
- Wilson, Tom; Starbird, Kate** (2020). "Cross-platform disinformation campaigns: lessons learned and next steps". *The Harvard Kennedy school (HKS) misinformation review*, v. 1, n. 1.
<https://doi.org/10.37016/mr-2020-002>
- Wu, Tim** (2020). "Comerciantes de atención: la lucha épica por entrar en nuestra cabeza". Capitán Swing Libros. ISBN: 978 8412064599
- Zannettou, Savvas; Sirivianos, Michael; Blackburn, Jeremy; Kourtellis, Nicolas** (2019). "The web of false information: rumors, fake news, hoaxes, clickbait, and various other shenanigans". *Journal of data and information quality*, v. 11, n. 3, pp. 1-37.
<https://doi.org/10.1145/3309699>
- Zerback, Thomas; Töpfl, Florian; Knöpfle, Maria** (2021). "The disconcerting potential of online disinformation: persuasive effects of astroturfing comments and three strategies for inoculation against them". *New media & society*, v. 23, n. 5, pp. 1080-1098.
<https://doi.org/10.1177/1461444820908530>

Crude, anonymous, partisan, sectoral and anti-elitist. Electoral disinformation in Spain (2019-2022)

David Lava-Santos; José Gamir-Ríos; Germán Llorca-Abad

Nota: Este artículo se puede leer en español en:
<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87314>

Recommended citation:

Lava-Santos, David; Gamir-Ríos, José; Llorca-Abad, Germán (2023). "Crude, anonymous, partisan, sectoral and anti-elitist. Electoral disinformation in Spain (2019-2022)". *Profesional de la información*, v. 32, n. 5, e320506.

<https://doi.org/10.3145/epi.2023.sep.06>

Manuscript received on 9th March 2023
Accepted on 3rd July 2023



David Lava-Santos ✉
<https://orcid.org/0000-0002-2352-5083>

Universitat de València
Av. Blasco Ibáñez, 32
46010 Valencia, Spain
lasanda@alumni.uv.es



José Gamir-Ríos
<https://orcid.org/0000-0001-5379-6573>

Universitat de València
Av. Blasco Ibáñez, 32
46010 Valencia, Spain
jose.gamir@uv.es



Germán Llorca-Abad
<https://orcid.org/0000-0002-0934-8420>

Escuela Politécnica Superior de Gandía
Calle Paranimf, 1
46730 Gandía, Valencia, Spain
gerlloab@upvnet.upv.es

Abstract

The aim of this research is to analyse the electoral misinformation circulating in the public space as a whole during campaign periods. The sample is made up of 481 rebuttals published by the verification media *Maldita.es*, *Newtral*, *Efe Verifica* and *Verificat* in relation to 409 pieces of misinformation during the six campaigns carried out in Spain during the political cycle that began in the general election in November 2019, which also includes the regional elections held in Galicia, the Basque Country, Catalonia, Madrid, Castilla y León, and Andalusia. The methodology consists of a content analysis of 13 variables articulated around five dimensions of study: scope of generation and/or dissemination, format, epistemological authority, prominence, theme, and discourse. The results show: (1) the predominance of the citizen space of social networks as a field for the generation and distribution of electoral misinformation; (2) the crude construction of misinformation, with a preponderance of text as a misinformation element accompanied by multimedia resources shared without alteration; (3) the pre-eminence of problematic materials disseminated anonymously and without citing sources and those actually distributed by the political class, especially the right wing; (4) the majority role of the political class, particularly the left, generally characterised in a negative manner, although the political class of the right receives a greater proportion of positive treatment; (5) the abundance of sectoral and ideological themes, with a prominent presence of attacking elites as the predominant populist discursive feature; and (6) the profusion of direct attacks on political rivals, especially related to ideological and management issues.

Keywords

Political communication; Disinformation; Misinformation; Electoral campaigns; Elections; Social media; Social networks; Political hoaxes; Political disinformation; Fact-checking; Fact checkers; Cheapfakes; Populism; Negativity; Content analysis; Spain.



Funding

This research was carried out within the framework of the “Flujos de desinformación, polarización y crisis de la intermediación mediática” project (PID2020-113574RB-I00), financed by the *Ministry of Science and Innovation* of Spain.

1. Introduction

The appeal of analysing the spread of “deliberately” false information in the public sphere (Salaverría *et al.*, 2020, p. 2) increased with Donald Trump’s rise to power in 2016 (Orbegozo-Terradillos; Morales-i-Gras; Larrondo-Ureta, 2020). Subsequent research has shown that during the US electoral campaign, 10% of the tweets contained hoaxes or conspiracy theories, and 15% presented extremely biased information (Bovet; Makse, 2019). However, if there is one time when the spread of hoaxes reached a peak, it was in the early 2020s when, in the context of the COVID-19 pandemic, misinformation on health, science, and health policy issues increased dramatically (Brennen *et al.*, 2020).

The phenomenon of disinformation has been present in politics for decades (Salaverría *et al.*, 2020). However, the consolidation of digital platforms has made it easier for citizens with ideological interests (Herrero-Diz; Pérez-Escolar; Plaza-Sánchez, 2020) to share any kind of negative information originated by elites (Chadwick; Vaccari, 2019), and to be self-sufficient in the production and distribution of potentially viral messages. The content of these messages is twisted to display misleading information and offer political actors the opportunity to open up a range of nodes through which to disseminate disinformation, fuelling populist drifts and employing increasingly crude and critically-charged language (Ernst *et al.*, 2017), which proliferates in crises (Pérez-Curiel; Velasco-Molpeceres, 2020).

Political disinformation is a social problem, as it promotes a clear polarisation in vote casting, a weakening of democratic health (Pira, 2019) and an increase in distrust towards the media and political actors as a whole (Casero-Ripollés; Doménech-Fabregat; Alonso-Muñoz, 2023). Citizens consider this to be a problem that puts the democratic system at risk. In this sense, Herrero-Diz, Pérez-Escolar and Plaza-Sánchez (2020) note that 79% of the Spanish population recognises their difficulty in differentiating between false and untrue content. Faced with the growing concern about this phenomenon, information verifiers, media and social researchers have tried to alleviate the situation, delving deeper into the search for disinformation and identifying what type of disinformation is promoted, through which channels and in what intensity (Almansa-Martínez; Fernández-Torres; Rodríguez-Fernández, 2022). However, as Casero-Ripollés, Doménech-Fabregat and Alonso-Muñoz (2023) show, people still perceive a clear lack of resources to solve the problem, and they blame the media and political actors themselves for not being forceful in eradicating disinformation in the digital sphere.

Disinformation research is becoming increasingly widespread in the social sciences, especially about the COVID-19 pandemic (García-Marín; Salvat-Martinrey, 2021). Even on occasions when the object of study is political, its centrality lies in the investigation of the narrative surrounding the coronavirus (Pérez-Curiel; Velasco-Molpeceres, 2020), with little research seeking to respond to the typology and discourse of disinformation in campaign periods (Rodríguez-Hidalgo; Herrero; Aguaded, 2021), where disinformation plays a fundamental role in the deconstruction of the political image.

The aim of this research is therefore to analyse the electoral disinformation circulating in the public space as a whole during the campaign periods. The sample is made up of 481 rebuttals published by the verifying media *Maldita.es*, *Newtral*, *Efe Verifica* and *Verificat* on 409 pieces of disinformation covering the six campaigns carried out in Spain during the political cycle that began with the general election in November 2019. Specifically, the study analyses the scope of generation and/or dissemination of disinformation, its format, its issuers and sources, its protagonists, its themes, and its discourse, with particular attention paid to populism and negativity.

2. Literature review**2.1. Background and disinformation typologies: sources, formats and spheres of dissemination of hoaxes in the contemporary public sphere**

The internet was the main access point used by citizens during the 2020 state of emergency to obtain information related to the COVID-19 pandemic (Brennen *et al.*, 2020). Despite the efforts made by governmental organisations, which invested large amounts of money on online information resources, and by social networks themselves, which introduced special information services such as the *COVID-19 Info Center* proposed by Facebook (Nielsen *et al.*, 2020), the spread of hoaxes in the digital public sphere was inevitable.

Social networks themselves were the main channel for transmitting disinformation about the coronavirus (Nielsen *et al.*, 2020). Published studies on this phenomenon (Magallón-Rosa; Sánchez-Duarte, 2021; Rodríguez-Andrés, 2018) point to the consolidation of disinformation in modern Western democratic societies not being understood without taking the formal characteristics of the media and social context into account (Pérez-Curiel; Velasco-Molpeceres, 2020). Thus, before the advent of the internet and dig-

“Within the hybrid media ecosystem misinformation circulates faster and spreads more widely”

ital platforms, disinformation circulated more slowly, as it was propagated in a local context (**Gutiérrez-Coba; Coba-Gutiérrez; Gómez-Díaz**, 2020). The epicentre of its origin and its reproduction format was reduced to fake news and oral information.

Disinformation does feed into the polarising strategy of populism

In contrast, within the hybrid media ecosystem (**Chadwick**, 2013), misinformation circulates faster and spreads more widely (**Vosoughi; Roy; Aral**, 2018). At a time of decline in journalism (**Calvo; Cano-Orón; Llorca-Abad**, 2022), when the traditional role of the media as information watchdogs is being diluted, social networks, with their innovative and user-friendly design, provide users with digital tools that allow them to create disinformative content. The lack of digital literacy required to create hoaxes (**Cerdán-Martínez; García-Guardia; Padilla-Castillo**, 2020) has favoured the implementation of *cheapfakes* (**Gamir-Ríos; Tarullo**, 2022), i.e. disinformation created in an amateurish way by users using their own mobile devices.

Despite pressure from different spheres (**Weidner; Beuk; Bal**, 2020), social networks have not implemented effective measures against disinformation and are the ideal space to disseminate hoaxes. In this scenario of information confusion, the work carried out by fact-checking platforms is increasingly necessary (**Ramon-Vegas; Mauri-Ríos; Rodríguez-Martínez**, 2020). Some studies have pointed out that these agents are key in the collection, verification and dissemination of falsehoods circulating in the public sphere (**Ramon-Vegas; Mauri-Ríos; Rodríguez-Martínez**, 2020). Others have taken a more nuanced view, finding that in periods of open crisis, as in the case of COVID-19, citizens tend to give more credibility to the traditional media, especially television (**Masip et al.**, 2020).

The social impact that the phenomenon of disinformation has had in recent years (**Calvo et al.**, 2022) has led many researchers to take an interest in studying the anatomy of hoaxes. To this end, several works published to date have studied the composition of the disinformation collected by the various verifiers, and the object of analysis ranges from the format, through the sources, the protagonists, and the degree of alteration of the hoaxes. In terms of the main format of the disinformation, the hoaxes that have circulated about COVID-19 have been mainly textual, both exclusively and combined with other formats (**Peña-Ascacibar; Bermejo-Malumbres; Zanni**, 2021; **Salaverría et al.**, 2020).

In the political sphere, **Paniagua-Rojano, Seoane-Pérez** and **Magallón-Rosa** (2020) noted the prevailing use of text on social media during the 10 November 2019 election, and **López-Martín, Gómez-Calderón** and **Córdoba-Cabús** (2023) found that the majority of political hoaxes debunked by *International Fact-Checking Network (IFCN)*-accredited verifiers between 1 January and 31 March 2022 featured plain text (32.56%), while 30.23% contained the misinformation in the text embedded into the image, and 19.77% of unretouched photographs contained purely textual falsehoods.

In regard to the credibility derived from the epistemological authority granted to the contents (**Bochenski**, 1974), several studies on the sources of hoaxes related to politics (**Paniagua-Rojano; Seoane-Pérez; Magallón-Rosa**, 2020), feminism (**Malquín-Robles; Gamir-Ríos**, 2023; **Herrero-Diz; Pérez-Escolar; Plaza-Sánchez**, 2020) or the pandemic (**Gutiérrez-Coba; Coba-Gutiérrez; Gómez-Díaz**, 2020) have corroborated the predominance of anonymous hoaxes (**Gamir-Ríos; Lava-Santos**, 2022; **Almansa-Martínez; Fernández-Torres; Rodríguez-Fernández**, 2022) over those attributed to real sources (**Salaverría et al.**, 2020). It has also been found that the producers of these materials barely alter the multimedia resources they disseminate beyond the insertion of text (**Gamir-Ríos; Tarullo**, 2022) and that the main protagonists involved in disinformation are political actors who are negatively evaluated (**Almansa-Martínez; Fernández-Torres; Rodríguez-Fernández**, 2020).

Following that which has been described in this section, we offer the following research question:

RQ1: In what sphere, who are the protagonists and their attributes, and with what format, sources and degree of alteration is political disinformation disseminated in electoral campaigns?

2.2. Us versus them: the construction of populist disinformation agendas

The overabundance of information resulting from the new communicative paradigm, in which the combination of media, political, and citizen logic is fundamental in creating a vision of social reality (**Casero-Ripollés**, 2018), causes major issues to be displaced to a position far removed from popular interest, while sensationalism, demagoguery, and emotionality occupy a privileged space (**Orbegozo-Terradillos; Morales-i-Gras; Larrondo-Ureta**, 2020). Information chaos is the product of a modern, liquid society (**Bauman**, 2002), in which objective facts have less and less influence and propagandistic narrative practices condition political news (**Bañuelos-Capistrán**, 2020).

Around this idea, **Pérez-Tornero et al.** (2018) coined the term “media neopopulism” to refer to a populism supported by the spectacular nature of politics, the erosion of the classical media, the need to bring leaders with media charisma to power and, as has been said, the hatching of emotionality to the detriment of the dissemination of logical issues (**Orbegozo-Terradillos; Morales-i-Gras; Larrondo-Ureta**, 2020). In climates where ideological polarisation is the norm (**Llorca; Fabregat-Cabrera; Ruiz-Callado**, 2021),

“the constant deepening of divisions, the theoretical rejection of any kind of negotiation, and the constant hammering of conspiracy theories” (**Waisbord**, 2020, p. 272)

are, according to **Arias-Maldonado** (2016), the basis for constructing the stereotype of the antagonistic “other”, i.e., a corrupt elite constantly confronting a pure and noble people (**Savarino**, 2006).

The construction of this antagonistic vision of reality can be seen in the communicative strategy followed by far-right activists and groupings in Germany, through the use of disinformation with an “unhygienic” theme focused on discrediting the government’s management of COVID-19, the latter being understood as a fraction of the “elite establishment” (**Vieten**, 2020, p. 12). In Europe, **Pérez-Curiel** and **Rivas-de-Roca** (2022) identified conspiracy theories, corruption and the economy as the main themes that followed a strategy framed around criticism of the elite and defence of the commons, and which contained hoaxes that came to be propagated through mainstream media.

The negative connotation that has surrounded contemporary populism is therefore evident. In his work, Laclau explained that no political action distances itself from populism, as the construction of the people and social division are the “political acts par excellence” (**Laclau**, 2005, p. 195). This assumption conditions all definitions that have tried to offer empirical knowledge about populism, being diverse, broad and ambiguous (**Groshek; Engelbert**, 2013). This phenomenon is not just therefore an ideology or political regime (**Mouffe**, 2019), but a particular form of political discursivity (**Lava-Santos**, 2023).

From the perspective of **Aalberg et al.** (2017), this conceptualisation of populism rejects any pejorative and authoritarian connotations and can occur in left-wing or right-wing groupings (**Jagers; Walgrave**, 2007). From this premise, it is interpreted on one hand that populist movements can emerge in historical periods marked by economic and political crises, and as proof of this are the extreme right-wing nativist parties such as the *National Front* (France) or the *Northern League* (Italy). On the other hand, it is also deduced that parties that were already present in political life, as well as the social audience itself, are impregnated with the identifying elements of this populist discourse (**Lava-Santos**, 2023).

While populism does not need disinformation to reach certain social sectors, disinformation does feed into the polarising strategy of populism (**Waisbord**, 2020). Deploying an “us/them” binary narrative (**Wondreys; Mudde**, 2022), disinformation actors often employ this dialectic by avoiding political debate (**Rivas-Venegas**, 2021), claiming the need to recover popular sovereignty, disowning groups and collectives considered demeaning, and repeatedly appealing to patriotic and cultural values (**Engesser; Fawzi; Larsson**, 2017; **Ernst et al.**, 2019).

Based on the cognitive bias hypothesis (**Saiz-Vélez**, 2020), according to which information on political issues affects individuals more than any other issue (**Orbegozo-Terradillos; Morales-i-Gras; Larrondo-Ureta**, 2020), several authors have sought to ascertain which populist discursive elements mark the political agenda of candidates in electoral campaigns. As recent evidence, **Lava-Santos** (2021b) observed in the electoral campaign for the 2021 Catalan Parliament election that defending the commons and the ostracism of immigrant groups are presented in issues related to health, education, or work, while criticism of the elites and the defence of popular sovereignty are framed in ideological and personal issues. Internationally, **Pérez-Curiel** and **Domínguez-García** (2021) showed that Trump’s criticism of elites is conditioned to campaign-related and general ideological issues, while patriotic values are introduced in sectoral issues.

However, the populist behaviour of political hoaxes spread in the digital sphere has not been analysed in detail. We therefore propose the following research question:

RQ2: What populist discursive elements are most frequently employed in the various disinformative political issues in electoral campaigns?

2.3. Negative disinformation campaigns: from the unofficial to the political

When the First World War ended, the Bolshevik political police first used the term ‘disinformatzia’ (**Rodríguez-Andrés**, 2018) to describe actions aimed at preventing the establishment of the communist regime in Moscow. Linked to its warlike origin, **Rodríguez-Andrés** (2018, p. 235) stresses that any disinformative element aims to “attack the one who is considered as an adversary”. In politics, the communicative practice characterised by criticising and attacking the rival acquires the name of negative campaigning. Far from being a novelty (**Greer; LaPointe**, 2004), its consolidation as an electoral strategy dates back to the 1950s, with the birth of television. Nowadays, this type of campaign has been strengthened thanks to the exceptional space for expression created by social networks (**Cabo; García-Juanatey**, 2016).

Both concepts –disinformation and negative campaigning– are closely related, especially in the consequences they can have on society (**Aguerri; Miró-Llinares**, 2023), but they also present differences. Disinformation is based on the use of lies (**Aguerri; Miró-Llinares**, 2023), its intention can be offensive (**Montes**, 2022) by discrediting the rival, but it can also have the purpose of distorting reality in favour of the disseminator of the hoax (defensive disinformation). On the contrary, negative campaigns can be based on falsehoods as part of their discursive resource, or can simply be constructed through pejorative lexis (**Aguerri; Miró-Llinares**, 2023) using truthful reasoning. In any case, criticism of the rival includes all forms of attack, regardless of its level of unfairness, dishonour, irrelevance, or manipulation (**Walter; Vliegthart**, 2010).

While it has been shown that the introduction of disinformative agendas can manipulate public opinion and condition citizens’ voting decisions (**Pira**, 2019), one of the questions that has been addressed is the effect that the dissemination of a negative campaign can have on the electorate. According to **García-Beaudoux** and **D’Adamo** (2013), using the

communicative strategy of attack can increase citizens' interest in the political debate. However, negative campaigning also presents risks for the political actors who adopt it (Joathan; Alves, 2020). García-Beaudoux and D'Adamo (2013) distinguish three;

- "boomerang effect", as the saturation of negative information produces a rebound effect that ends up hurting the attacking actor.
- "victim syndrome", in which citizens perceive an attack as unfair and generate positive feelings towards the attacked candidate.
- "double damage" effect occurs when the negative campaign affects both the attacker and the attacked.

To avoid these potential risks, political actors begin to rely on the actions of anonymous groups or individuals not linked to official political coalitions (Joathan; Alves, 2020). This gives rise to the concept of unofficial negative campaigning, in which political cyber-violence plays a fundamental role (Villar-Aguilés; Pecourt-Gracia, 2020) and its development thanks to the trolling subculture of "cyber-ghettos" (Johnson; Bichard; Zhang, 2009) and pseudo-media that disseminate "poor quality" information (Cervera, 2018, p. 9).

Another aspect to consider is the behaviour of negativity emanating from platforms such as *Telegram* (Tirado-García, 2023) or *Twitter* (Lava-Santos, 2021b). Among the main results obtained, it is seen that discrediting candidates' programme proposals, ideological components or personal traits is regularly reproduced on social networks. Even in the study of negative campaigning on Facebook during the 2016 election, Marcos-García, Alonso-Muñoz and López-Merí (2021) showed that the use of negative campaigning was based on personal, ideological and previous management criticisms of the actors concerned. Based on this premise, we present the following research question.

RQ3: Which social actors are making the attacks and what is the behaviour of these critics?

3. Methodology

3.1. Materials

This paper analyses the disinformation related to the six electoral campaigns that took place in Spain during the political cycle that began in the wake of the general election held on 10 November 2019*. It does so by studying 409 pieces of disinformation debunked in 481 publications by *Maldita.es*, *Newtral*, *Efe Verifica* and *Verificat*; the four Spanish media accredited by the *International Fact-Checking Network*. The investigation of disinformation based on disinformation published by verifying media is a well-established methodology in Spain (Almansa-Martínez; Fernández-Torres; Rodríguez-Fernández, 2022; Gamir-Ríos; Lava-Santos, 2022; Gutiérrez-Coba; Coba-Gutiérrez; Gómez-Díaz, 2020; Herre-ro-Diz; Pérez-Escolar; Plaza-Sánchez, 2020; Noain-Sánchez, 2021; Peña-Ascacibar; Bermejo-Malumbres; Zanni, 2021; Salaverría *et al.*, 2020) and internationally (Brennen *et al.*, 2020).

In addition to the rebuttals linked to the campaign that started the cycle, the materials include those referring to the campaigns for the elections to

- the *Parliament of Galicia* and the *Basque Parliament* held on 12 July 2020;
- the *Parliament of Catalonia* on 14 February 2021;
- the *Assembly of Madrid* on 4 May 2021;
- the *Parliament of Castilla y León* on 13 February 2022;
- the *Parliament of Andalusia* on 19 June 2022.

The timeframe analysed in the six cases includes the official electoral campaign periods—eight days for the general election and fifteen days for the others—, the reflection and voting days, and the day after the elections. The composition of the timeframe coincides with that usually studied in research on digital political communication in electoral campaigns (Gamir-Ríos *et al.*, 2022; Lava-Santos, 2021a).

The composition of the research corpus went through three stages. The first was the capture phase, which consisted of using the web-scraping tool *Octoparse* to download the 1,042 texts published during the periods analysed on the websites:

- *Maldito Buló* (495) and *Newtral Fakes* (137) for the study of disinformation circulating on social networks and corresponding to the citizen and media spheres;
<https://maldita.es/malditobulo>
<https://www.newtral.es/zona-verificacion/fakes>
- *Maldito Dato* (160) and *Newtral Fact-checks* (115), for those issued from the political sphere;
<https://maldita.es/malditodato>
<https://www.newtral.es/zona-verificacion/fact-check>
- *Efe Verifica*>*Verificaciones* (61) and *Verificat*>*Verificaciones* (74), for the three spheres.
<https://maldita.es/malditodato>
<https://www.newtral.es/zona-verificacion/fact-check>

Disinformation agendas can manipulate public opinion and condition the citizen vote

The reason for the inclusion of *Maldito Dato* and *Newtral* fact-checks is that the aim of the research is not limited to the analysis of disinformation disseminated on social networks, but extends to that circulating in the public sphere as a whole; therefore, the materials should not be restricted to the rebuttal of hoaxes, but should also include the verification of statements made by the political class and pieces published in the media.

The second –the narrowing phase– involved selecting only the 669 publications strictly related to electoral campaigns. To do so, we took into consideration the typology of issues developed by **Mazzoleni** (2010) based on Patterson, which distinguishes between political issues, policy issues, personal issues and campaign issues.

The third definition –the results of which are shown in Table 1– involved discarding publications that referred to more than one piece of misinformation, as this was a study of misinformation and not of verifiers' publications. This reduced the sample to 481 pieces of disinformation referring to 409 problematic materials, a figure that has been taken as the basis for calculating the percentages in the presentation of results.

Table 1. Details of materials by campaign and verifying media

	Gen19	Gal-Eus20	Cat21	Mad21	CyL22	And22	Total
<i>Efe Verifica</i>	0	1	7	5	1	2	16
<i>Maldito Bulo</i>	64	23	43	49	17	19	215
<i>Maldito Dato</i>	21	3	6	4	1	1	36
<i>Newtral Fact-checks</i>	18	6	24	18	23	25	114
<i>Newtral Fakes</i>	14	6	7	11	6	17	61
<i>Verificat</i>	4	5	9	3	5	13	39

3.2. Method

The research applies the classic methodology of content analysis (**Krippendorff**, 2004; **Neuendorf**, 2016) to the 409 pieces of misinformation that make up the sample. To this end, a codebook of thirteen variables was drawn up, as shown in Table 2, and articulated around five study dimensions: scope of dissemination, format, epistemological authority, prominence, subject matter and discourse.

Table 2. Variables and categories of the content analysis

Dimension	Variable	Reliability	Categories (and subcategories)
A. Scope	1. Scope of dissemination	$\alpha = 1$	Political / Media / Citizen
B. Format (only if 1=citizen)	2. Presence of text	$\alpha = 1$	Absent / Present
	3. Presence of multimedia resource	$\alpha = 0.914$	Absent / Present
	4. Disinformative element	$\alpha = 0.872$	Text / Multimedia resource
	5. Type of media resource (only if 3=present)	$\alpha = 0.823$	Link / Audio / Image (capture, composition, photo) / Video
	6. Degree of alteration of the media resource (only if 3=present)	$\alpha = 0.909$	Existing / Reconfigured / Fabricated
C. Epistemological authority	7. Issuing authority	$\alpha = 0.899$	Anonymous account / Troll or fake account / Left-wing political class / Right-wing political class / Institutions / Left-wing journalists or media / Right-wing journalists or media / Hyper-partisan media / Fake news media / Experts / Other people of public relevance / Digital Influencers
	8. Source	$\alpha = 0.904$	Anonymous / Fictitious / Impersonated / Real
D. Protagonism	9. Protagonist	$\alpha = 0.975$	Without public relevance or citizenship in general / Institutions (Head of state / Judiciary / Armed Forces / Electoral Authority / Health Institutions / Supranational bodies) / Political class (PAE Left / PAE Right / PANE) / Social agents (Progressive activism / Conservative activism / Trade unions/ Companies) / Citizenship (Women / LGBTQI+ / Senior citizens / Migrants / Citizens without public relevance) / Agents with public relevance (Progressive media and journalists / Conservative media and journalists / Progressive referents / Conservative referents / Experts) / Other
	10. Attribute	$\alpha = 0.84$	Negative / Neutral / Positive
E. Themes and discourse	11. Theme	$\alpha = 0.878$	Political issues / Campaign issues / Personal issues / Policy issues
	12. Populist strategy	$\alpha = 0.83$	Defence of the people / Attack on the elites / Claim for sovereignty / Marginalisation of the different / Appeal to patriotic values
	13. Type of attack	$\alpha = 0.785$	No attack / Personal / Ideological / Management / Programmatic / Background and track record / Association with denigrating groups / Electoral

The first dimension comprises only the Scope of dissemination (v1) variable that, based on the distinction of interrelated spheres in political communication proposed by **Mazzoleni** (2010), categorises misinformation according to whether it has been generated and initially disseminated by:

- a) members of the political class (political system), regardless of whether they have done so in institutional, media or social network environments;
- b) journalists or the media (media system), regardless of whether they have done so in the media products themselves or on social networks; and
- c) identified or anonymous citizens (citizen-electoral system), on social networks.

“ The composition of the corpus went through a capture stage, another delimitation stage and a final definition stage ”

The second dimension studies the format of problematic materials disseminated in the public sphere through five variables. The analysis of this dimension does not apply to disinformation corresponding to the political and media sphere, as its formal characteristics –mainly statements and texts with the appearance of news– are so homogeneous that they would distort the results.

- On one hand, the dichotomous Presence of text (v2) and Presence of multimedia resources (v3) variables analyse the inclusion of these elements in the disinformation disseminated; while the Disinforming element (v4) variable, also dichotomous, is interested in which of them is the source of the deception.
- On the other hand, and only applicable to disinformation distributed via multimedia resources, the Type of resource (v5) and Degree of alteration (v6) variables classify the multimedia elements according to their format and degree of manipulation, which distinguishes between existing when it has been shared without alterations, reconfigured when it has been altered, and fabricated when it has been created ad hoc.

The definition of the variables in this dimension is based on the cheapfakes-deepfakes spectrum proposed by **Paris and Donovan** (2019) and the complexity curve proposed by **Gamir-Ríos and Tarullo** (2022).

The third dimension analyses the epistemological authority of disinformation by combining two variables.

- On one hand, the Issuing Instance (v7) variable categorises the origin of the disinformation based on various categories established after an initial study of the research corpus and whose definition takes into account the typology of disinformation producers proposed by **Tucker et al.** (2018) in their review.
- On the other hand, the Source (v8) variable analyses the attribution of disinformation based on the typology proposed by **Salaverría et al.** (2020):
 - a) anonymous, when the source is not mentioned;
 - b) fictitious, when the identity of the source is invented;
 - c) impersonated, when the disinformation is attributed to an existing source but which was not involved in it; and
 - d) real, when the identity is correctly attributed to an existing source, although the content is false.

The fourth dimension observes the protagonist of disinformation through the combination of two variables.

- On one hand, the Protagonist (v9) variable classifies the person or entity that is the object of disinformation based on a catalogue of options drawn up after an exploratory approach to the corpus and structured into two levels; the main one distinguishes between:
 - a) no protagonist or citizens in general;
 - b) institutions;
 - c) political class, which distinguishes between non-state level parties (PANE) and state level parties (PAE);
 - d) social agents;
 - e) collectives;
 - f) agents with public relevance; g) others.
- On the other hand, the Attribute (v10) variable analyses the characterisation of the objects of disinformation on a traditional three-grade scale: negative, neutral and positive.

Finally, the fifth dimension deals with the subject matter and discourse of disinformation. Based on **Mazzoleni** (2010), the Theme (v11) variable distinguishes between ideological (political issues), sectoral (policy issues), personal (personal issues) or electoral (campaign issues). Assuming that disinformation always presents populist features, since, as **Waisbord** (2020) states, it is part of its core strategy, the Populist strategy (v12) variable takes into account the populist discursive features defined by **Engesser et al.** (2017) for political messages and differentiates between:

- a) defence of the people, understood as the citizenry as a whole;
- b) attack on political, economic or media elites;
- c) vindication of national sovereignty in the face of globalisation or the interference of external agents;
- d) ostracism of those who are different or of minorities (such as migrants or the LGBTQI+ community) and
- e) appeal to patriotic values.

Finally, the Type of attack (v13) variable combines and adapts the classifications of **García-Beaudoux** and **D'Adamo** (2013) and **Valera-Ordaz** and **López-García** (2014), and distinguishes between:

- a) no attack;
- b) personal, including private life, competence and appearance;
- c) ideological, referring to non-sectoral political issues;
- d) managerial, linked to the sectoral results of holding public office of an executive nature;
- e) programmatic, related to electoral proposals;
- f) background and track record, not linked to management;
- g) association with groups considered denigrating by broad sectors of the population; and
- h) electoral, referring to the campaign itself, to incidents that occurred during its course or to suspicions about the election results.

The coding was carried out by one of the three authors after having conducted a test with another, independently and separately, on 12.5% of the randomly selected misinformation sample (n=52). The Krippendorff's Alpha coefficients obtained, shown in Table 2 and calculated using *ReCal* software (**Freelon**, 2013), corroborate the reliability of the results.

4. Results

4.1. Analysis of the scope of dissemination

The study of the scope of dissemination of disinformation (v1), reflected in Table 3, shows that the problematic materials disproved by the four Spanish verifiers about the six electoral campaigns that constitute the object of study come mainly from the citizen system (56.5%), having been generated and disseminated on social networks or messaging applications by people with no public relevance, or anonymously. The second space with the highest concentration of misinformation is the political system (40.1%), which includes statements made by the political class, regardless of whether they have been made in electoral environments –such as rallies or debates–, institutional environments –such as parliaments or seats of government–, media environments –such as interviews–, or digital environments –such as social networks. The media space, referring to journalists and the media, both in networks and in the media products themselves, occupies a residual place in the generation and circulation of disinformation (3.4%). The correlation by the campaign is similar in the general, Galician and Basque, Catalan and Madrid elections, while in the Castilian-Leonese and Andalusian elections, disinformation promoted by the political system predominates.

Table 3. Spheres of generation and dissemination of electoral disinformation

	Gen19		Gal-Eus20		Cat21		Mad21		CyL22		And22		Total	
	n	%	n	%	n	%	n	%	n	%	n	%	n	%
Political system	37	40.2	11	26.8	37	43.5	24	29.3	24	51.1	31	50.0	164	40.1
Media system	4	4.3	3	7.3	3	3.5	1	1.2	1	2.1	2	3.2	14	3.4
Citizen system	51	55.4	27	65.9	45	52.9	57	69.5	22	46.8	29	46.8	231	56.5
Total	92	100.0	41	100.0	85	100.0	82	100.0	47	100.0	62	100.0	409	100.0

4.2. Format analysis

Analysis of the format of the 231 pieces of misinformation corresponding to the citizen sphere shows, as seen in Table 4, that practically all (98.3%) are disseminated using text (v2) and that the presence of multimedia resources (v3) is also in the majority, although in a smaller proportion (69.7%). In terms of the typology of these resources (v5), graphic-textual compositions stand out, present on 24.2% of occasions; photographs, on 22.5%; and videos, on 13%. The element on which misinformation is based (v4) is, above all, text (64.9%, n=150), while misinformation through multimedia resources is a minority (35.1%, n=81).

Table 4. Formats of electoral disinformation generated and/or disseminated at the citizen level

	Gen19		Gal-Eus20		Cat21		Mad21		CyL22		And22		Total	
	n	%	n	%	n	%	n	%	n	%	n	%	n	%
Presence of text	49	96.1	25	92.6	45	100	57	100	22	100	29	100	227	98.3
Presence of multimedia resources	35	68.6	21	77.8	30	66.7	32	56.1	19	86.4	24	82.8	161	69.7
Link	1	2.0	1	3.7	0	0.0	0	0.0	0	0.0	0	0.0	2	0.9
Audio	4	7.8	2	7.4	2	4.4	0	0.0	0	0.0	2	6.9	10	4.3
Image: screenshot	2	3.9	1	3.7	3	6.7	3	5.3	0	0.0	2	6.9	11	4.8
Image: composition	11	21.6	7	25.9	10	22.2	14	24.6	6	27.3	8	27.6	56	24.2
Image: photograph	8	15.7	4	14.8	10	22.2	11	19.3	9	40.9	10	34.5	52	22.5
Video	9	17.6	6	22.2	5	11.1	4	7.0	4	18.2	2	6.9	30	13.0
Total citizen system	51		27		45		57		22		29		231	

In turn, the observation of the degree of alteration (v6) of the disinformation originating in the civic space that has multimedia elements (n=161) reveals the predominance of pre-existing materials shared without alteration (51.6%), as opposed to those created ad hoc (40.4%) and those reconfigured (8.1%). The cross-analysis of variables 5 and 6, reflected in Table 5, shows that the most frequent combinations of multimedia resources and degree of alteration are pre-existing photographs (31.7%), fabricated graphic-textual compositions (26.7%), pre-existing videos (14.3%) and reconfigured compositions (7.5%), which are generally based on the incorporation of typographic elements into real photographs.

Table 5. Degree of alteration of the multimedia resources distributed in the citizen system according to their type

	Existing		Reconfigured		Manufactured		Total	
	n	%	n	%	n	%	n	%
Link	0	0.0	0	0.0	2	1.2	2	1.2
Audio	2	1.2	0	0.0	8	5.0	10	6.2
Image: screenshot	6	3.7	0	0.0	5	3.1	11	6.8
Image: composition	1	0.6	12	7.5	43	26.7	56	34.8
Image: photograph	51	31.7	0	0.0	1	0.6	52	32.3
Video	23	14.3	1	0.6	6	3.7	30	18.6
Total multimedia in citizen system	83	51.6	13	8.1	65	40.4	161	100.0

4.3. Analysis of epistemological authority

The most frequent sources of disinformation (v7) are anonymous accounts or accounts without public relevance, which generate half of the problematic material debunked by the verifying media (51.1%); the right-wing political class generates a quarter (25.4%); and the left-wing political class is responsible for 14.2%. In terms of sources (v8), the majority of misinformation comes from real sources (46%), although almost as much does not explicitly state its provenance (41.1%). Misrepresentation constructed by impersonating sources represents 11.7% of the sample, while those that invent them occupy a residual place (0.2%).

The cross-analysis of both variables, shown in Table 6, shows that the most frequent combinations of issuing body and source are disinformation generated by anonymous or non-publicly relevant accounts that do not make their source explicit (40.1%), those disseminated in a real way by the right-wing political class (24.7%), those disseminated in a real way by the left-wing political class (13.9%), and those disseminated by anonymous or non-publicly relevant accounts through the impersonation of their sources (8.3%).

4.4. Analysis of prominence

The people or groups most frequently involved in misinformation (v9) are the political class linked to left-wing state parties (35.9%), those linked to right-wing state parties (23.7%), migrants (9.3%), the political class linked to non-state parties (9.3%), and those linked to non-state parties (9.3%).



Image 1. Example of disinformation with alteration of the multimedia resource through fabrication.

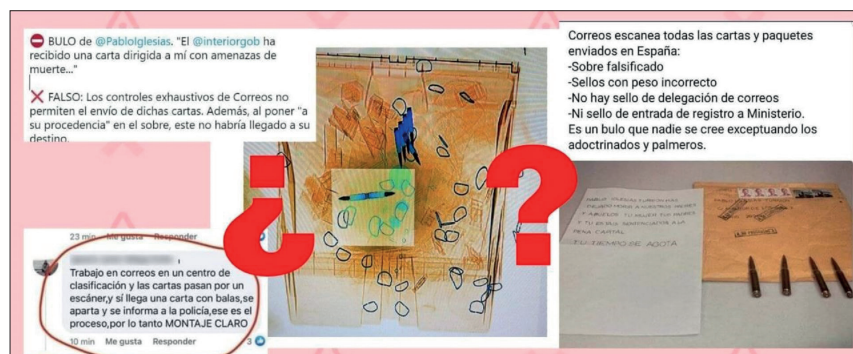


Image 2. Example of disinformation without alteration of the multimedia resource.

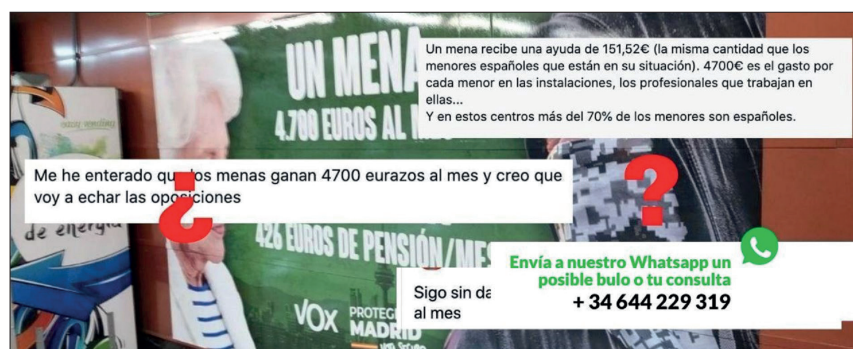


Image 3. Example of disinformation whose source is the right-wing political class.

Table 6. Source of electoral disinformation according to the issuing body

	Anonymous		Fictitious		Impersonated		Real		Total	
	n	%	n	%	n	%	n	%	n	%
Anonymous or irrelevant account	164	40.1	0	0.0	34	8.3	11	2.7	209	51.1
Troll or fake account	1	0.2	0	0.0	10	2.4	4	1.0	15	3.7
Left-wing political class	1	0.2	0	0.0	0	0.0	57	13.9	58	14.2
Right-wing political class	3	0.7	0	0.0	0	0.0	101	24.7	104	25.4
Institutions	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Journalist / Progressive media	0	0.0	0	0.0	0	0.0	1	0.2	1	0.2
Journalist / Conservative media	0	0.0	0	0.0	0	0.0	3	0.7	3	0.7
Hyper-partisan media	1	0.2	0	0.0	1	0.2	4	1.0	6	1.5
Fake-news media	2	0.5	1	0.2	3	0.7	6	1.5	12	2.9
Experts	0	0.0	0	0.0	0	0.0	1	0.2	1	0.2
Person with public relevance	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Digital influencers	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Total	172	42.1	1	0.2	48	11.7	188	46.0	409	100.0

Table 7. Protagonists of electoral disinformation according to their attributes

	Negative		Neutral		Positive		Total	
	n	%	n	%	n	%	n	%
No protagonist / general public	5	1.2	29	7.1	0	0.0	34	8.3
Institutions	35	8.6	3	0.7	0	0.0	38	9.3
Head of State	7	1.7	1	0.2	0	0.0	8	2.0
Judiciary	1	0.2	0	0.0	0	0.0	1	0.2
Armed Forces / Army	2	0.5	1	0.2	0	0.0	3	0.7
Electoral authority	16	3.9	1	0.2	0	0.0	17	4.2
Catholic Church	0	0.0	0	0.0	0	0.0	0	0.0
Health institutions	1	0.2	0	0.0	0	0.0	1	0.2
Supranational bodies	8	2.0	0	0.0	0	0.0	8	2.0
Political class	217	53.1	13	3.2	46	11.2	276	67.5
PAE Left	127	31.1	9	2.2	11	2.7	147	35.9
PAE Right	64	15.6	3	0.7	30	7.3	97	23.7
PANE	26	6.4	1	0.2	5	1.2	32	7.8
Social actors	9	2.2	1	0.2	0	0.0	10	2.4
Progressive activism	3	0.7	1	0.2	0	0.0	4	1.0
Conservative activism	0	0.0	0	0.0	0	0.0	0	0.0
Trade unionism	1	0.2	0	0.0	0	0.0	1	0.2
Entrepreneurship	5	1.2	0	0.0	0	0.0	5	1.2
Collectives	41	10.0	1	0.2	1	0.2	43	10.5
Women	1	0.2	1	0.2	0	0.0	2	0.5
LGBTQI+	0	0.0	0	0.0	0	0.0	0	0.0
Seniors	0	0.0	0	0.0	0	0.0	0	0.0
Migrants	38	9.3	0	0.0	0	0.0	38	9.3
Citizen without public relevance	2	0.5	0	0.0	1	0.2	3	0.7
Publicly relevant actors	3	0.7	0	0.0	2	0.5	5	1.2
Progressive media and journalists	2	0.5	0	0.0	0	0.0	2	0.5
Conservative media and journalists	1	0.2	0	0.0	1	0.2	2	0.5
Progressive opinion leaders	0	0.0	0	0.0	1	0.2	1	0.2
Conservative media and journalists	0	0.0	0	0.0	0	0.0	0	0.0
Experts	0	0.0	0	0.0	0	0.0	0	0.0
Other	2	0.5	1	0.2	0	0.0	3	0.7
Total	312	76.0	48	12.0	49	12	409	100.0



Image 4. Example of disinformation by the left-wing political class.

ties (7.8%) and the electoral authorities (4.2%). The attributes conveyed about these actors (v10) are mostly negative (76%), while neutral and positive characterisations are in the minority (12% in both cases).

The combined observation of both variables, shown in Table 7, shows that the most frequent protagonists and characterisations are the left-wing political class treated negatively (31.1%), the right-wing political class characterised in the same way (15.6%), migrants portrayed in a similarly adverse light (9.3%), the right-wing political class linked to positive attributes (7.3%), and the non-state political class portrayed in a detrimental light (6.4%).

All the protagonists of misinformation are characterised in a predominantly negative way, but the political class ascribed to the right or centre-right PAEs is, among the most frequent actors, the one that obtains the highest positive treatment, 30.9%, compared to 15.6% of the PANEs and 7.5% of the left or centre-left PAEs.



Image 5. Example of disinformation by the right-wing political class.

4.5. Thematic and discursive analysis

The most frequent thematic macro-category of electoral disinformation (v11) is that corresponding to political-sectoral issues (53.8%), followed by political-ideological issues (26.9%), those referring to the campaigns themselves (13.7%) and those linked personally to the candidates (5.6%). In turn, the most common populist discursive feature (v12) is the attack on the elites (51.6%), followed by the defence of the people (23.5%), the marginalisation of those who are different (18.8%) and the appeal to patriotic values (5.6%). No disinformation has been detected in which the predominant populist discursive feature is the vindication of national sovereignty in the face of globalisation or the action of external agents.

The cross-analysis of both variables, shown in Table 8, shows that the most frequent combinations of populist themes and discourse are, with very similar frequencies, the defence of the people on sectoral issues (18.3%), the attack on the elites on ideological issues (18.1%) and the attack on the elites on sectoral aspects (17.8%). They are followed by marginalising the different on sectoral issues (13.9%) and attacking elites on issues related to campaigns or their results (10.5%).

Attacking the elites is the most common populist discursive strategy in political issues (67.3%), campaign issues (76.8%) and personal issues (91.3%); in contrast, in policy issues, defence of the people predominates (34.1%), although there is also a significant presence of attacks on the elites (33.2%) and the ostracising of minority groups (25.9%).

Table 8. Themes of electoral disinformation according to the populist discursive strategy employed

	Political		Campaign		Personal		Policy		Total	
	n	%	n	%	n	%	n	%	n	%
Defence of the people	11	2.7	10	2.4	0	0.0	75	18.3	96	23.5
Attacking the elites	74	18.1	43	10.5	21	5.1	73	17.8	211	51.6
Claiming sovereignty	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
Marginalisation of the different	18	4.4	2	0.5	0	0.0	57	13.9	77	18.8
Appeal to patriotic values	7	1.7	1	0.2	2	0.5	15	3.7	25	6.1
Total	110	26.9	56	13.7	23	5.6	220	53.8	409	100.0

Finally, the presence of direct attacks on the political class (v13) is present in 46% of the disinformation analysed (n=188). Among those with such negativity, those containing ideological attacks (34.6%) and management attacks (30.3%) stand out, although those attacking through references to personal attributes (10.6%) and with the dissemination of suspicions about the course of the campaigns or election results (10.1%) are also noteworthy. As Table 9 shows, 50.5% of the attacks come from the public sphere, 45.7% from the political sphere and only 3.7% from the media. The most frequent combinations are management attacks generated from the political sphere (26.1%), ideological attacks arising from the civic (20.2%) or political (12.8%) sphere, and personal attacks circulated from the citizen system (9.6%).



Image 6. Example of disinformation with the presence of anti-elitist attacks as a populist discursive feature.

Table 9. Negativity of electoral disinformation according to its scope of dissemination

	Political		Media		Citizen		Total	
	n	%	n	%	n	%	n	%
Personal	1	0.5	1	0.5	18	9.6	20	10.6
Ideological	24	12.8	3	1.6	38	20.2	65	34.6
Management	49	26.1	0	0.0	8	4.3	57	30.3
Programme	1	0.5	3	1.6	5	2.7	9	4.8
Background and track record	5	2.7	0	0.0	7	3.7	12	6.4
Partnership with groups	4	2.1	0	0.0	2	1.1	6	3.2
Electoral	2	1.1	0	0.0	17	9.0	19	10.1
Total	86	45.7	7	3.7	95	50.5	188	100.0

5. Discussion and conclusions

The results of the study show that, although most of the hoaxes circulating during the electoral campaign are disseminated by citizens through social networks (RQ1), official political communication also has disinformative features. In line with **Herrero-Diz, Pérez-Escolar and Plaza-Sánchez (2020)**, hoaxes find their natural ecosystem in the virtual environment, so, logically, more than half of the campaign disinformation verified by fact-checkers circulates on digital platforms (**Herrero-Diz; Pérez-Escolar; Plaza-Sánchez, 2020**). The use of persuasive and propagandistic rhetoric based on fallacies (**Pérez-Curiel; Velasco-Molpeceres, 2020**) has always been present in political discourse, but content analysis has confirmed that politicians intensify the use of disinformation in their statements during electoral periods.

Continuing with RQ1, the predominance of disinformative formats containing textual elements has been corroborated in all campaigns. Likewise, the presence of multimedia elements, such as photographs or videos, which have not undergone any alteration, beyond the insertion of an accompanying text including the disinformation, is common. The preponderance of falsehoods in which the disinformation resides in the text and not in the audiovisual material confirms the findings of research by **López-Martín, Gómez-Calderón and Córdoba-Cabús (2023)**, **Gamir-Ríos and Tarullo (2022)**, **Gutiérrez-Coba, Coba-Gutiérrez and Gómez-Díaz (2020)** and **Salaverría et al. (2020)**, who point to cheapfakes as the main problem of political disinformation (**Gamir-Ríos; Tarullo, 2022**). Of course, there are sound and graphic-textual hoaxes that obey compositions with a relatively low degree of complexity (**Gamir-Ríos; Tarullo, 2022**).

In addition to the difficulty of detecting impersonated sources in the digital environment (**Salaverría et al., 2020**), there are also a number of anonymous individuals or individuals with no public relevance who support and disseminate the deception itself. Possibly, as it is a common citizen who is the source of the dialectical usurpation, they manage to capture the attention of other users, who in turn virtualise the same content, believing that the political actor is the author of the information.

The loss of rigour and the deficient work of the journalistic contrast of information has become a social and

“ The groups that most frequently lead misinformation are left-wing politicians ”

democratic problem (López-Martín; Gómez-Calderón; Córdoba-Cabús, 2023). Several studies have shown that conventional media are characterised by reproducing disinformative content that circulates in the digital environment (López-Martín; Gómez-Calderón; Córdoba-Cabús, 2023; Salaverría *et al.*, 2020). These statements contradict what is presented in this study, as only four fake news items have been verified, which seems to indicate that these channels take extreme caution with content and ensure that the material is truthful during campaign periods.

Disinformation often resides in the text, demonstrating how easily a hoax can be created and disseminated

Looking at the protagonists of the hoax and its attribute, the bulk of hoaxes about possible election rigging and hoaxes about subsidies received by immigrants is striking. Reinforcing the conclusions obtained by Gamir-Ríos, Tarullo and Ibáñez-Cuquerella (2021), producing mistrust regarding the democratic system (Almansa-Martínez; Fernández-Torres; Rodríguez-Fernández, 2020) and immigration is a communicative practice that occurs frequently in the electoral campaign, and which accentuates a populist discourse warning of a dangerous otherness that threatens the people. Furthermore, disinformation to the detriment of left-wing political actors and the benefit of the right-wing political class points to, on one hand, the fact that the main producers of disinformation are right-wing activists and politicians (Chadwick; Vaccari, 2019; Gamir-Ríos; Lava-Santos, 2022). On the other hand, the argument that disinformation in campaigning can be used to offend rivals or in self-defence is strengthened (Montes, 2022).

Focusing on RQ2, it can be observed that disinformation circulating during the campaign recurrently deals with sectoral and ideological issues, leaving campaign-related issues and personal aspects of candidates in the background. Another distinctive feature is that criticism of elites is present in all thematic categories, including personal aspects. Likewise, when disinformation disseminators present a thematic focus on issues such as health, education, labour and immigration, they construct a populist discourse in defence of the people. In turn, a populist interest in marginalising immigrant and feminist groups is observed when the hoaxes present in the verifiers have an ideological and confrontational theme, something that fits with what has been observed in previous studies (Lava-Santos, 2021a; Pérez-Curiel; Domínguez-García, 2021).

Finally, the findings from the cross-referencing of the variable's scope of dissemination and type of attack (RQ3) reveal a common axis in the criticisms made by politicians and citizens. Both social agents misinform by attacking the ideological aspects of their respective rivals. On the contrary, the communicative strategy carried out in the political sphere ratifies the systematic use of disinformation that criticises the previous management of the adversary, while the citizens collectively utters a greater personal discrediting in the codified hoaxes. These results contradict the conclusions provided by Tirado-García, (2023) and Marcos-García, Alonso-Muñoz and Casero-Ripollés (2021), who state that the negative campaigning present on the social networks of politicians usually focused on discrediting the personal characteristics of the candidates.

The limitation lies in the selection of a sample collected exclusively from the rebuttals provided publicly by the four verifiers. A large number of hoaxes spread during the campaigns analysed may have been propagated by other media, which has prevented the collection of a larger amount of disinformation that may have gone unnoticed among the fact-checkers. In this sense, the results should be considered ascribed to the study variables. One of the main limitations of disinformation research at an international level lies in the need to support learning about fact-checking techniques among academics and researchers, which, in addition to serving as a training element, would make it easier to identify hoaxes circulating in the public sphere and digital politics, within the timeframe of different campaigns.

Future research could provide a comparative study of the phenomenon in upcoming electoral campaigns. It would also be of interest to carry out a comparison between the populist discourses, themes and negativity present in the disinformation circulating in different countries,

The attack on the elites is the most common populist discursive strategy in ideological, sectoral and personal issues

which would provide an overview of the quality of their media systems. Furthermore, an analysis of the influence of electoral disinformation on citizens during the same campaigns studied in this paper would complement the results.

6. References

Aalberg, Toril; Esser, Frank; Reinemann, Carsten; Strömbäck, Jesper; De-Vreese, Claes (eds.) (2017). *Populist political communication in Europe*. Routledge. ISBN: 978 1 315623016
<https://doi.org/10.4324/9781315623016>

Aguerrí, Jesús C.; Miró-Llinares, Fernando (2023). "¿Comunicación ofensiva y de odio o desinformación programada? Un análisis de las características y discurso del evento conspiranoico #ExposeBillGates". *IDP: revista de internet, derecho y política*, n. 47.
<https://doi.org/10.7238/idp.v0i37.397192>

- Almansa-Martínez, Ana; Fernández-Torres, María-Jesús; Rodríguez-Fernández, Leticia** (2022). “Desinformación en España un año después de la Covid-19. Análisis de las verificaciones de *Newtral y Maldita*”. *Revista latina de comunicación social*, n. 80, pp. 183-200.
<https://www.doi.org/10.4185/RLCS-2022-1538>
- Arias-Maldonado, Manuel** (2016). “La digitalización de la conversación pública: redes sociales, afectividad política y democracia”. *Revista de estudios políticos*, n. 173, pp. 27-54.
<https://doi.org/10.18042/cepc/rep.173.01>
- Bañuelos-Capistrán, Jacob** (2020). “Deepfake: la imagen en tiempos de la posverdad”. *Revista panamericana de comunicación*, v. 1, n. 2, pp. 51-61.
<https://doi.org/10.21555/rpc.v0i1.2315>
- Bauman, Zygmunt** (2002). “In the lowly nowhereilles of liquid modernity: Comments on and around Agier”. *Ethnography*, v. 3, n. 3, pp. 343-349.
<https://doi.org/10.1177/146613802401092788>
- Bochenski, Józef M.** (1974). “Logic and ontology”. *Philosophy East and West*, v. 24, n. 3, pp. 275-292.
<https://doi.org/10.2307/1398384>
- Bovet, Alexandre; Makse, Hernán A.** (2019). “Influence of fake news in Twitter during the 2016 US presidential election”. *Nature communications*, v. 10, n. 1.
<https://doi.org/10.1038/s41467-018-07761-2>
- Brennen, James-Scott; Simon, Felix M.; Howard, Philip N.; Nielsen, Rasmus-Kleis** (2020). *Types, sources, and claims of Covid-19 misinformation*. University of Oxford. Reuters Institute. Oxford Internet Institute. Oxford Martin School.
- Cabo-Isasi, Alex; García-Juanatey, Ana** (2016). *El discurso del odio en las redes sociales: un estado de la cuestión*. Ajuntament de Barcelona.
https://www.injuve.es/sites/default/files/2019/02/noticias/el_discurso_del_odio_en_rrss.pdf
- Calvo, Dafne; Cano-Orón, Lorena; Llorca-Abad, Germán** (2022). “Covid-19 vaccine disinformation on YouTube: analysis of a viewing network”. *Communication & society*, v. 35, n. 2.
<https://doi.org/10.15581/003.35.2.223-238>
- Calvo, Dafne; Valera-Ordaz, Lidia; Requena-i-Mora, Marina; Llorca-Abad, Germán** (2022). “Fact-checking in Spain: Perception and trust”. *Catalan journal of communication & cultural studies*, v. 14, n. 2, pp. 287-305.
https://doi.org/10.1386/cjcs_00073_1
- Casero-Ripollés, Andreu** (2018). “Research on political information and social media: Key points and challenges for the future”. *Profesional de la información*, v. 27, n. 5, pp. 964-974.
<https://doi.org/10.3145/epi.2018.sep.01>
- Casero-Ripollés, Andreu; Doménech-Fabregat, Hugo; Alonso-Muñoz, Laura** (2023). “Percepciones de la ciudadanía española ante la desinformación en tiempos de la Covid-19: efectos y mecanismos de lucha contra las noticias falsas”. *Icono 14*, v. 21, n. 1.
<https://doi.org/10.7195/ri14.v21i1.1988>
- Cerdán-Martínez, Víctor; García-Guardia, María-Luisa; Padilla-Castillo, Graciela** (2020). “Alfabetización moral digital para la detección de deepfakes y fakes audiovisuales”. *Cuadernos de información y comunicación*, v. 25.
<https://doi.org/10.5209/ciyc.68762>
- Cervera, José** (2018). “Contra la verdad, las noticias falsas no existen”. *Cuadernos de periodistas*, n. 35.
<https://www.cuadernosdeperiodistas.com/la-verdad-las-noticias-falsas-no-existen>
- Chadwick, Andrew** (2013). *The hybrid media system: Politics and power*. Oxford: Oxford University Press. ISBN: 978 0 1993451 1 3
- Chadwick, Andrew; Vaccari, Cristian** (2019). *News sharing on UK social media: Misinformation, disinformation, and correction*. Online Civic Culture Centre (O3C); Loughborough University.
<https://www.lboro.ac.uk/research/online-civic-culture-centre/news-events/articles/o3c-1-survey-report-news-sharing-misinformation>
- Engesser, Sven; Fawzi, Nayla; Larsson, Anders-Olof** (2017). “Populist online communication: Introduction to the special issue”. *Information, communication & society*, v. 20, n. 9, pp. 1279-1292.
<https://doi.org/10.1080/1369118X.2017.1328525>
- Ernst, Nicole; Blassnig, Sina; Engesser, Sven; Büchel, Florin; Esser, Frank** (2019). “Populists prefer social media over talk shows: An analysis of populist messages and stylistic elements across six countries”. *Social media+ society*, v. 5, n. 1.
<https://doi.org/10.1177/2056305118823358>

- Freelon, Dean** (2013). “ReCal OIR: Ordinal, interval, and ratio intercoder reliability as a web service”. *International journal of internet science*, v. 8, n. 1, pp. 10-16.
https://www.ijis.net/ijis8_1/ijis8_1_freelon.pdf
- Gamir-Ríos, José; Cano-Orón, Lorena; Fenoll, Vicente; Iranzo-Cabrera, María** (2022). “Evolución de la comunicación política digital (2011-2019): ocaso de los blogs, declive de *Facebook*, generalización de *Twitter* y popularización de *Instagram*”. *Observatorio (OBS*)*, v. 16, n. 1, pp. 90-115.
<https://doi.org/10.15847/obsOBS16120221879>
- Gamir-Ríos, José; Lava-Santos, David** (2022). “Desinformación sobre historia, memoria democrática y simbología del Estado en redes sociales: Estudio de los bulos desmentidos por los *fact-checkers* españoles”. *Estudios sobre el mensaje periodístico*, v. 28, n. 4, pp. 777-788.
<https://doi.org/10.5209/esmp.82846>
- Gamir-Ríos, José; Tarullo, Raquel** (2022). “Predominio de las *cheapfakes* en redes sociales: Complejidad técnica y funciones textuales de la desinformación desmentida en Argentina durante 2020”. *adComunica*, n. 23, pp. 97-118.
<https://doi.org/10.6035/adcomunica.6299>
- Gamir-Ríos, José; Tarullo, Raquel; Ibáñez-Cuquerella, Miguel** (2021). “Multimodal disinformation about otherness on the internet: The spread of racist, xenophobic and islamophobic fake news in 2020”. *Anàlisi*, n. 64, pp. 49-64.
<https://doi.org/10.5565/rev/analisi.3398>
- García-Beaudoux, Virginia; D’Adamo, Orlando** (2013). “Propuesta de una matriz de codificación para el análisis de las campañas negativas”. *Opera*, n. 13, pp. 7-23.
<https://www.redalyc.org/articulo.oa?id=67530878002>
- García-Marín, David; Salvat-Martinrey, Guiomar** (2021). “Investigación sobre desinformación en España. Análisis de tendencias temáticas a partir de una revisión sistematizada de la literatura”. *Fonseca, journal of communication*, v. 23, pp. 199-225.
<https://doi.org/10.14201/fjc202123199225>
- Greer, Jennifer D.; LaPointe, Mark E.** (2004). “Cyber-campaigning grows up: A comparative content analysis of websites for US Senate and gubernatorial races, 1998-2000”. In: Gibson, Rachel; Römmele, Andrea; Ward, Steven. *Electronic democracy*. Routledge, pp. 132-148. ISBN: 978 0 2033579 8 9
<https://doi.org/10.4324/9780203357989>
- Groshek, Jacob; Engelbert, Jiska** (2013). “Double differentiation in a cross-national comparison of populist political movements and online media uses in the United States and The Netherlands”. *New media & society*, v. 15, n. 2, pp. 183-202.
<https://doi.org/10.1177/1461444812450685>
- Gutiérrez-Coba, Liliana; Coba-Gutiérrez, Patricia; Gómez-Díaz, Javier-Andrés** (2020). “Las noticias falsas y desinformación sobre el Covid-19: análisis comparativo de seis países iberoamericanos”. *Revista latina de comunicación social*, n. 78, pp. 237-264.
<https://www.doi.org/10.4185/RLCS-2020-1476>
- Herrero-Diz, Paula; Pérez-Escolar, Marta; Plaza-Sánchez, Juan-Francisco** (2020). “Desinformación de género: análisis de los bulos de *Maldito Feminismo*”. *Icono 14*, v. 18, n. 2, pp. 188-216.
<https://doi.org/10.7195/ri14.v18i2.1509>
- Jagers, Jan; Walgrave, Stefaan** (2007). “Populism as political communication style: An empirical study of political parties’ discourse in Belgium”. *European journal of political research*, v. 46, n. 3, pp. 319-345.
<https://doi.org/10.1111/j.1475-6765.2006.00690.x>
- Joathan, Ícaro; Alves, Marcelo** (2020). “O Twitter como ferramenta de campanha negativa não oficial: uma análise da campanha eleitoral para a Prefeitura do Rio de Janeiro em 2016”. *Galáxia (São Paulo)*, n. 43, pp. 81-98.
<https://doi.org/10.1590/1982-25532020141565>
- Johnson, Thomas J.; Bichard, Shannon L.; Zhang, Weiwu** (2009). “Communication communities or “cyberghettos?”: A path analysis model examining factors that explain selective exposure to blogs”. *Journal of computer-mediated communication*, v. 15, n. 1, pp. 60-82.
<https://doi.org/10.1111/j.1083-6101.2009.01492.x>
- Krippendorff, Klaus** (2004). *Content analysis: An introduction to its methodology*. California: SAGE. ISBN: 978 1 4129 8315 0
<https://doi.org/10.4135/9781071878781>
- Laclau, Ernesto** (2012). *La razón populista*. Argentina: Fondo de Cultura Económica. ISBN: 978 950 5576357

- Lava-Santos, David** (2021a). “El discurso populista de Vox en las elecciones generales de 2019: Análisis comparativo de las agendas temáticas en televisión y Twitter”. *Dígitos*, n. 7, pp. 37-65.
<https://doi.org/10.7203/rd.v1i7.200>
- Lava-Santos, David** (2021b). “La campaña negativa en las elecciones catalanas de 2021: Estudio del mensaje emanado por los candidatos durante los debates electorales de TVE, TV3 y La sexta”. *Sphera pública*, v. 1, n. 21, pp. 54-88.
- Lava-Santos, David** (2023). “Comunicación populista y estrategia temática en la cuenta oficial de Instagram de los candidatos españoles durante las campañas del 28A y el 10N de 2019”. *Papers. Revista de sociologia*, v. 108, n. 1, e3082-e3082.
<https://doi.org/10.5565/rev/papers.3082>
- Llorca, Elena; Fabregat-Cabrera, María-Elena; Ruiz-Callado, Raúl** (2021). “Desinformación populista en redes sociales: la tuitosfera del juicio del Procés”. *Observatorio (OBS*)*, v. 15, n. 3, pp. 124-146.
<https://doi.org/10.15847/obsOBS15320211835>
- López-Martín, Álvaro; Gómez-Calderón, Bernardo; Córdoba-Cabús, Alba** (2023). “La desinformación en auge: un análisis de los bulos sobre política española”. *Visual review. International visual culture review*, v. 10, n. 3.
<https://doi.org/10.37467/revvisual.v10.4596>
- Magallón-Rosa, Raúl; Sánchez-Duarte, José-Manuel** (2021). “Information verification during Covid-19. Comparative analysis in Southern European countries”. *Janus.net, e-journal of international relations*, v. 12, n. 1, pp. 152-165.
<https://doi.org/10.26619/1647-7251.DT21.10>
- Malquín-Robles, Andrea; Gamir-Ríos, José** (2023). “Desinformación y sexismo digital. El feminismo y su agenda como objeto de engaño en español”. *Icono14*, v. 21, n. 1.
<https://doi.org/10.7195/ri14.v21i1.1963>
- Marcos-García, Silvia; Alonso-Muñoz, Laura; Casero-Ripollés, Andreu** (2021). “The influence of the negative campaign on Facebook: The role of political actors and citizens in the use of criticism and political attack in the 2016 Spanish general elections”. *Social sciences*, v. 10, n. 10.
<https://doi.org/10.3390/socsci10100356>
- Marcos-García, Silvia; Alonso-Muñoz, Laura; López-Meri, Amparo** (2021). “Campañas electorales y Twitter: La difusión de contenidos mediáticos en el entorno digital”. *Cuadernos.info*, n. 48, pp. 27-47.
<https://doi.org/10.7764/cdi.48.1738>
- Masip, Pere; Aran-Ramspott, Sue; Ruiz-Caballero, Carlos; Suau, Jaume; Almenar, Ester; Puertas-Graell, David** (2020). “Consumo informativo y cobertura mediática durante el confinamiento por el Covid-19: sobreinformación, sesgo ideológico y sensacionalismo”. *Profesional de la información*, v. 29, n. 3.
<https://doi.org/10.3145/epi.2020.may.12>
- Mazzoleni, Gianpietro** (2010). *La comunicación política*. Madrid: Alianza. ISBN: 978 84 90975176
- Montes, Julio** (2022). “La desinformación: un arma moderna en tiempos de guerra”. *Cuadernos de periodistas*, n. 44.
<https://www.cuadernosdeperiodistas.com/la-desinformacion-un-arma-moderna-en-tiempos-de-guerra>
- Mouffe, Chantal** (2019). *Por un populismo de izquierda*. Buenos Aires: Siglo XXI Editores. ISBN: 978 987 629 883 4
- Neuendorf, Kimberly A.** (2016). *The content analysis guidebook*. Los Angeles: SAGE. ISBN: 978 1 4129279474
- Nielsen, Rasmus-Kleis; Fletcher, Richard; Newman, Nic; Brennen, J. Scott; Howard, Philip N.** (2020). *Navigating the ‘infodemic’: How people in six countries access and rate news and information about coronavirus*. Reuters Institute for the Study of Journalism.
<https://reutersinstitute.politics.ox.ac.uk/infodemic-how-people-six-countries-access-and-rate-news-and-information-about-coronavirus>
- Noain-Sánchez, Amaya** (2021). “Desinformación y Covid-19: Análisis cuantitativo a través de los bulos desmentidos en Latinoamérica y España”. *Estudios sobre el mensaje periodístico*, v. 27, n. 3, pp. 879-892.
<https://doi.org/10.5209/esmp.72874>
- Orbegozo-Terradillos, Julen; Morales-i-Gras, Jordi; Larrondo-Ureta, Ainara** (2020). “Desinformación en redes sociales: ¿compartimentos estancos o espacios dialécticos? El caso Luther King, Quim Torra y El confidencial”. *Revista mediterránea de comunicación*, v. 11, n. 2, pp. 55-69.
<https://doi.org/10.14198/MEDCOM2020.11.2.2>
- Paniagua-Rojano, Francisco; Seoane-Pérez, Francisco; Magallón-Rosa, Raúl** (2020). “Anatomía del bulo electoral: la desinformación política durante la campaña del 28-A en España”. *Revista Cidob d’afers internacionals*, v. 124, pp. 123-146.
https://www.cidob.org/es/articulos/revista_cidob_d_afers_internacionals/124/anatomia_del_bulo_electoral_la_desinformacion_politica_durante_la_campana_del_28_a_en_espana

- Paris, Britt; Donovan, Joan** (2019). *Deepfakes and cheapfakes: The manipulation of audio and visual evidence*. Data & Society Research Institute.
<https://datasociety.net/library/deepfakes-and-cheap-fakes>
- Peña-Ascacíbar, Gonzalo; Bermejo-Malumbres, Eloy; Zanni, Stefano** (2021). “Fact checking durante la Covid-19: Análisis comparativo de la verificación de contenidos falsos en España e Italia”. *Revista de comunicación*, v. 20, n. 1, pp. 197-215.
https://doi.org/10.26441/rc20.1-2021-a11_
- Pérez-Curiel, Concha; Domínguez-García, Ricardo** (2021). “Discurso político contra la democracia. Populismo, sesgo y falacia de Trump tras las elecciones de EEUU (3-N)”. *Cultura, lenguaje y representación*, n. 26, pp. 7-28.
<https://doi.org/10.6035/clr.5807>
- Pérez-Curiel, Concha; Rivas-de-Roca, Rubén** (2022). “Exploring Populism in times of crisis: An analysis of disinformation in the European context during the US elections”. *Journalism and media*, v. 3, n. 1, pp. 144-156.
<https://doi.org/10.3390/journalmedia3010012>
- Pérez-Curiel, Concha; Velasco-Molpeceres, Ana-María** (2020). “Tendencia y narrativas de *fact-checking* en Twitter. Códigos de verificación y *fake news* en los disturbios del Procés (14-O)”. *AdComunica*, n. 20, pp. 95-122.
<https://doi.org/10.6035/2174-0992.2020.20.5>
- Pérez-Tornero, José-Manuel; Tayle, Sally-Samy; Tejedor, Santiago; Pulido, Cristina** (2018). “¿Cómo afrontar las noticias falseadas mediante la alfabetización periodística: estado de la cuestión?”. *Doxa comunicación*, n. 26, pp. 211-235.
<https://doi.org/10.31921/doxacom.n26a10>
- Pira, Francesco** (2019). “Las elecciones europeas de 2019 y las *fake news*”. *Barataria*, n. 26, pp. 67-83.
<https://doi.org/10.20932/barataria.v0i26.510>
- Ramon-Vegas, Xavier; Mauri-Ríos, Marcel; Rodríguez-Martínez, Ruth** (2020). “Redes sociales y plataformas de *fact-checking* contra la desinformación sobre la Covid-19”. *Hipertext.net*, n. 21, pp. 79-92.
<https://doi.org/10.31009/hipertext.net.2020.i21.07>
- Rivas-Venegas, Miguel** (2021). “Contra las mujeres: el discurso misógino de Vox. ‘Palabras gruñido’ del nacional-populismo español”. *Investigaciones feministas*, v. 12, n. 1, pp. 67-78.
<https://doi.org/10.5209/infe.69585>
- Rodríguez-Andrés, Roberto** (2018). “Fundamentos del concepto de desinformación como práctica manipuladora en la comunicación política y las relaciones internacionales”. *Historia y comunicación social*, v. 23, n. 1, pp. 231-244.
<https://doi.org/10.5209/HICS.59843>
- Rodríguez-Hidalgo, Claudia; Herrero, Javier; Aguaded, Ignacio** (2021). “Journalistic verification against lies in the electoral processes of Ecuador and Spain”. *Universitas-XXI*, n. 34, pp. 39-60.
<https://doi.org/10.17163/uni.n34.2021.02>
- Saiz-Vélez, Jorge-Enrique** (2020). “Crisis de las democracias: una conjugación entre desinformación, desconfianza, polarización política y populismo”. *Gestión y finanzas*, v. 3, n. 5.
<https://revistas.ut.edu.co/index.php/gestionyfinanzas/article/view/2574>
- Salaverría, Ramón; Buslón, Nataly; López-Pan, Fernando; León, Bienvenido; López-Goñi, Ignacio; Erviti, María-Carmen** (2020). “Desinformación en tiempos de pandemia: Tipología de los bulos sobre la Covid-19”. *Profesional de la información*, v. 29, n. 3, e290315.
<https://doi.org/10.3145/epi.2020.may.15>
- Savarino, Franco** (2006). “Populismo: perspectivas europeas y latinoamericanas”. *Espiral (Guadalajara)*, v. 13, n. 37, pp. 77-94.
https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1665-05652006000300003
- Tirado-García, Alejandra** (2023). “The negative campaign on *Telegram*: The political use of criticism during the 2021 Community of Madrid elections”. *Social sciences*, v. 12, n. 2, 93.
<https://doi.org/10.3390/socsci12020093>
- Tucker, Joshua A.; Guess, Andrew; Barberá, Pablo; Vaccari, Cristian; Siegel, Alexandra; Sanovich, Sergey; Stukal, Denis; Nyhan, Brendan** (2018). *Social media, political polarization, and political disinformation: a review of the scientific literature*. SSRN.
- Valera-Ordaz, Lidia; López-García, Guillermo** (2014). “Agenda and frames in the websites of the People’s Party (PP) and the Spanish Social Workers’ Party (PSOE) in the 2011 cyber campaign”. *Revista latina de comunicación social*, n. 69, pp. 41-66.
<https://nuevaepoca.revistalatinacs.org/index.php/revista/article/download/963/1572>

Vieten, Ulrike M. (2020). "The 'new normal' and 'pandemic populism': The Covid-19 crisis and anti-hygienic mobilisation of the far-right". *Social sciences*, v. 9, n. 9.
<https://doi.org/10.3390/socsci9090165>

Villar-Aguilés, Alicia; Pecourt-Gracia, Juan (2021). "Antifeminismo y troleo de género en *Twitter*. Estudio de la subcultura trol a través de #STOPfeminazis". *Teknokultura*, v. 18, n. 1, pp. 33-44.
https://doi.org/10.5209/TEKN.70225_

Vosoughi, Soroush; Roy, Deb; Aral, Sinan (2018). "The spread of true and false news online". *Science*, v. 359, n. 6380, pp. 1146-1151.
<https://doi.org/10.1126/science.aap9559>

Waisbord, Silvio (2020). "¿Es válido atribuir la polarización política a la comunicación digital?: Sobre burbujas, plataformas y polarización afectiva". *Revista SAAP*, v. 14, n. 2, pp. 249-279.
<https://doi.org/10.46468/rsaap.14.2.A1>

Walter, Annemarie S.; Vliegthart, Rens (2010). "Negative campaigning across different communication channels: Different ball games?". *The international journal of press/politics*, v. 15, n. 4, pp. 441-461.
<https://doi.org/10.1177/1940161210374122>

Weidner, Kelly; Beuk, Frederik; Bal, Anjali (2020). "Fake news and the willingness to share: a schemer schema and confirmatory bias perspective". *Journal of product and brand management*, v. 29, n. 2, pp. 180-187.
<https://doi.org/10.1108/JPBM-12-2018-2155>

Wondreys, Jakub; Mudde, Cas (2022). "Victims of the pandemic?: European far-right parties and COVID-19". *Nationalities papers*, v. 50, n. 1, pp. 86-103.
<https://doi.org/10.1017/nps.2020.93>

Anuario ThinkEPI 2022

<https://thinkepi.profesionaldelainformacion.com/index.php/ThinkEPI>



How to teach the elderly to detect disinformation: a training experiment with *WhatsApp*

Charo Sádaba; Ramón Salaverría; Xavier Bringué-Sala

Nota: Este artículo se puede leer en español en:
<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87324>

Recommended citation:

Sádaba, Charo; Salaverría, Ramón; Bringué-Sala, Xavier (2023). "How to teach the elderly to detect disinformation: a training experiment with *WhatsApp*". *Profesional de la información*, v. 32, n. 5, e320504.

<https://doi.org/10.3145/epi.2023.sep.04>

Manuscript received on 18th March 2023

Accepted on 14th July 2023



Charo Sádaba ✉
<https://orcid.org/0000-0003-2596-2794>

Universidad de Navarra
Facultad de Comunicación
Campus Universitario, s/n
31080 Pamplona, Spain
csadaba@unav.es



Ramón Salaverría
<https://orcid.org/0000-0002-4188-7811>

Universidad de Navarra
Facultad de Comunicación
Campus Universitario, s/n
31080 Pamplona, Spain
rsalaver@unav.es



Xavier Bringué-Sala
<https://orcid.org/0000-0001-7943-7982>

Universidad de Navarra
Facultad de Comunicación
Campus Universitario, s/n
31080 Pamplona, Spain
jbringue@unav.es

Abstract

According to recent studies, most of the Spanish population identifies disinformation as a social problem and believes that it could endanger democracy and the stability of the country. In this context, many institutions point out the need for media literacy campaigns and initiatives that alleviate the possible harmful social effects of the phenomenon, especially among vulnerable audiences. While children and young people are the continuous target of this type of action, few so far have targeted the elderly. This article analyzes the effectiveness of a training action to increase the ability to detect false news in this age group. A 10-day course was designed, and a sample of 1,029 individuals over 50 years of age residing in Spain who are smartphone users was selected. Participants were divided into an experimental group (n=498), who were invited to take the course, and a control group (n=531). An *ex ante* and *ex post* study was carried out to determine the effects of the course on their ability to detect false news. The results reveal that those who took the course were more successful in identifying the news as true or false than the members of the control group. The results confirm the opportunity and convenience of designing media literacy actions aimed at those over 50 years of age, a social group particularly exposed to disinformation.

Keywords

Disinformation; Media literacy; Vulnerable groups; Older people; Elderly; Aged persons; Smartphones; Training activities; Digital competence; Internet; False news detection; Fake news; Critical thinking; Social networks; *WhatsApp*; Experiments.



Acknowledgements

The authors would like to thank María Zabala for her support during the research, as well as Ryan C. Moore and Jeffrey T. Hancock for allowing us to build our study on an adaptation of their questionnaire.

Funding

This study was funded by *MediaWise*, which is based at the *Poynter Institute for Media Studies*; support was provided by *Meta*.

It also received funding from *Iberifier*, a project funded by the *European Commission* in CEF-TC-2020-2 (*European Digital Media Observatory*) call, number 2020-EU-IA-0252.

1. Introduction

The information society allows citizens constant access to content from all kinds of sources, official and personal, professional and amateur, well-intentioned and ill-intentioned. The complexity of this ecosystem, where truthful information coexists with false or erroneous content, highlights the need to train users to be critical of the information they receive and share. The term *disinformation* refers to

“false or misleading content that is spread with an intention to deceive or secure economic or political gain, and which may cause public harm” (*European Commission*, 2019).

This is a phenomenon that has become commonplace in the consumption of content on social networks and often reaches public opinion. When information is abundant, errors can appear, which are commonly referred to as “misinformation” (**Burnam**, 1975), a phenomenon different from disinformation, which points to deliberate falsehoods. These lies, which spread rapidly in the highly polarized context in which we find ourselves (**Sádaba; Salaverría**, 2023), are potentially dangerous for citizens and also for social cohesion (**Brennen et al.**, 2020; **Salaverría et al.**, 2020). Although the phenomenon is not directly attributable to technology, the widespread access to the Internet and consumption of content on social networks, along with the social and political tensions of recent years, have been accompanied by an increase in false content, with diverse intentions (**Shu et al.**, 2020).

Spain does not escape from this global reality, and, according to the *I Study on disinformation in Spain* (*Uteca; Universidad de Navarra*, 2022), 95.8% of the population identifies it as a social problem. Ninety-one percent of those surveyed believe that disinformation can endanger democracy and the stability of a country, and 83.3% consider that the pandemic and the war in Ukraine have contributed to its growth. The *Uteca* study also corroborates two noteworthy elements: it identifies age as a relevant variable (**Santibáñez-Velilla; Latorre-Santibáñez; Tejada-Sánchez**, 2014), and it evidences that older people tend to continue to trust more in the traditional media they consume more frequently. Along with age, other studies point out that gender and educational level may be significant variables in the behavior towards this type of content (**Ramírez-García; González-Fernández; Sedeño-Valdellós**, 2017). It is also clear that the third-person bias is confirmed in this area (**Corbu; Oprea; Frunzaru**, 2021), since there is a tendency to think that others are more likely to be deceived than oneself (**Altay; Acerbi**, 2023; **Martínez-Costa et al.**, 2022). Added to this is the subjective factor that **Watson** (1960) calls “confirmation bias,” which implies that people tend to favor information that confirms their own beliefs or prejudices. This bias has given rise to a debate on the extent to which social networks can generate a certain “echo chamber” effect (**Iandoli; Primario; Zollo**, 2021), which in any case adds more urgency to the fight against disinformation.

2. Media literacy: part of the solution

It is easy to understand that a problem of this complexity does not have a simple solution, but requires a coordinated effort of numerous agents and at all levels (**Cucarella; Fuster**, 2022). Experts have been warning for years about the dangers of disinformation and the need for the media and journalists to work to combat it and not promote it (**Gal-dón-López**, 1994; **García-Matilla**, 1999; **Pérez-Tornero**, 2008). The emergence of *fact-checkers* can be understood as a sectoral response to this problem (**Nieminen; Rapeli**, 2019). Although *fact-checking* serves to limit the public dissemination of false content (**Carnahan; Bergan**, 2022), its effect is insufficient, and, as certain studies point out, in certain contexts it is even innocuous (**Margolin; Hannak; Weber**, 2018; **Oeldorf-Hirsch et al.**, 2020). It has been found, for example, that the ability of *fact-checkers* to disprove political disinformation among the population through fact-checking is substantially attenuated by citizens’ pre-existing beliefs, ideology, and knowledge (**Walter et al.**, 2020).

As a complement to strategies against disinformation based on information verification, the efforts of governments and institutions to address the negative consequences of the phenomenon are increasing (**Sádaba; Salaverría**, 2023; **Wardle; Derakhshan**, 2017). Within the framework of these initiatives, media literacy emerges as part of the solution, as it underscores the need to equip citizens with resources and personal skills that enable them to cope with their own biases in the first place, and with the disinformation strategies of which they may ultimately fall victim.

In the early 1990s, the concept of media literacy was defined as

“the ability of a citizen to access, analyze and produce information for specific results” (**Aufderheide**, 1993, p. 6).

Years later, a simpler definition was agreed upon:

“The active attitude and critical thinking about the messages we receive and create” (Namlr, 2007, p. 3).

In the European context, media literacy has been understood as the

“the ability to access, analyse, evaluate and create messages across a variety of contexts” (Livingstone, 2004, p. 18).

The digital scenario has added new skills required to this media competence, understood in a more integrative way and that challenges not only the formal educational stage, but also lifelong learning (Kačínová; Sádaba, 2021). Therefore, as recognized by Ferrés and Piscitelli,

“media competence has to face [...] this complexity, combining the enhancement of participatory culture with the development of critical capacity” (Ferrés; Piscitelli, 2012, p. 77).

At stake are issues that raise the need for citizens to have knowledge and skills to verify the information they consume (Redondo, 2018).

2.1. The elderly: a vulnerable public in the face of disinformation

In recent years, numerous digital and media literacy efforts have been developed targeting the younger population, a segment of the population characterized by a lower critical capacity for information and intense exposure to digital content (European Commission, 2022; Herrero-Diz; Conde-Jiménez; Reyes-de-Cózar, 2021). Moreover, the fact that it is easy to reach youth through formal education channels has made this population group an essential audience in the design of any media literacy strategy. Along with the young, the elderly also emerge as a potentially vulnerable group. Although it is true that, given their maturity and experience, they are attributed with greater critical capacity, their knowledge and skills to select and weigh the information they receive through the Internet may be more limited (Papí-Gálvez; La-Parra-Casado, 2022). If, in addition, this information reaches them through “second generation” networks, such as *WhatsApp* or *Telegram*, and given that they tend to place their trust not in the medium, but in those who send them the information (Valera-Ordaz et al., 2022), they can become not only consumers of disinformation, but also disseminators. In fact, the *Digital News Report 2021* highlighted the role of *WhatsApp* in Spain in the dissemination of false information (Amoedo-Casais et al., 2021).

Therefore, the digital competence of older people, or in many cases its absence, becomes a key factor in the fight against the harmful effects of disinformation. In addition to constituting, from a demographic point of view, a considerable part of society, older people consume a lot of information and are interested and concerned about informational issues (Brashier; Schacter, 2020). Although academic interest in media literacy for this age group is growing, as Ramírez-García, González-Fernández and Sedeño-Valdellós (2017) point out, both research and the design of actions and strategies for this audience are still limited.

This article explores to what extent a training action can improve the ability to identify misinformation among the over 50s. The evidence obtained may not only open up new avenues of research but also provide ideas for the design of plans aimed at this essential public in democratic and social life. The aim is to answer two specific questions:

1. Can a course designed for this age group, over 50, increase their ability to detect fake news?
 - 1.1. Following Ramírez-García, González-Fernández and Sedeño-Valdellós (2017), is gender a relevant variable in the possible effectiveness of this course?
2. Can this training contribute to the degree of security with which this audience consumes online information?

3. Methodology

In order to answer the research questions, we proceeded to analyze the results and impact of the course “What to do to detect false information on the Internet,” launched in Spain in April 2022 by the international media literacy project *MediaWise*, promoted by the *Poynter Institute* of the United States and the Spanish digital native medium *Newtral*, with the support of *Meta*. Aimed at people over 50 years of age, it was a free 10-day course, through *WhatsApp*, with lessons of at least 5 minutes a day. Users had to sign up by sending a message, and, from that moment on, they received a daily session that included a video with simple and practical techniques, designed to improve judgment in the use of Internet information resources and digital skills in general. The content of the videos went from general knowledge to targeting specific skills: the reality of disinformation, the main channels of dissemination of this type of content, what aspects to look for when doubting the veracity of a news item, and tips and ideas for learning how to check it for oneself (image search, use of *fact-checkers*, etc.). In addition to the video, each day they were asked a simple question about their content.

In order to test the effectiveness of this course in detecting disinformation, a two-wave, *ex ante* and *ex post* research was designed and applied to two groups: an experimental group, which received an invitation to take the course, and a control group. Two questionnaires were developed to assess the baseline situation and the possible evolution in terms of participants’ ability to identify fake news and their skills to contrast information or their general knowledge of Internet-related terms. Both groups completed two questionnaires.

Participants were selected from among men and women residing in Spain who were over 50 years of age and users of a *smartphone*. It is true that this factor, required by the fact that the course took place via *WhatsApp*, implies some limitations on the study universe: compared to 96% of users in Spain who have a smartphone, only 51% of those over 55 have one (*IAB Spain*, 2019a). Among these mobile users, *WhatsApp* (90%) and *Facebook* (66%) are the two most used social networks (*IAB Spain*, 2019b). The number of participants in the study is detailed in Table 1.

Table 1. Number of participants in the study

	Control	Experimental	Total
<i>Ex ante</i>	531	498	1,029
<i>Ex post</i>	448	200	648

For the design of the questionnaires used, we adapted the one created by **Moore and Hancock** (2022), once permission had been obtained. The questionnaire adapted to Spain was structured in six blocks. Block 1 asked participants to evaluate the truthfulness or falsehood of some selected headlines, using a seven-item Likert scale (from “clearly false” to “clearly true”). They were presented with a total of six headlines, three with false or inaccurate information and three true news items. Among the false (or inaccurate), one was congruent with right-wing ideological bias, one was congruent with left-wing ideological bias, and the third was neutral (see Table 2). The 12 headlines included in the two waves were extracted from media and news websites; in the case of false or inaccurate headlines, they were chosen from those that had been reviewed by a *fact-checker*. In their selection, only news items from the last month before each wave were taken into account, in order to potentially seek greater recall among the sample subjects. Block 1 also asked whether the participants had searched for information about the news before giving their opinion on the veracity of each headline. Table 1 shows the headlines used in the questionnaires of both phases.

Table 2. News headlines used in the questionnaires (*ex ante* and *ex post*)

	True news	False news
Congruent with a left-wing bias	News 4 A high-ranking Ayuso official spread via WhatsApp that his brother had received 283,000 euros from Priviet Sportive <i>The Community of Madrid would have confirmed, by mistake, that Tomás Díaz Ayuso received 283,000 in 2020 from Priviet Sportive.</i>	News 1 Moreno Bonilla allocates more resources to private healthcare while he lays off 8,000 public healthcare workers <i>Thirty-eight percent of the reinforcement healthcare workers who have been dismissed in our country have been dismissed in Andalusia.</i>
	News 10 Ayuso’s toll on Vox: no research on therapies to cure homosexuality <i>Since the arrival of the Councilor for Family, Youth and Social Policy, there has been no ex-officio action against possible cases of LGTBphobia in the Community of Madrid.</i>	News 7 The IMF is against the tax cuts proposed by the PP <i>A frontal blow to Feijóo.</i>
	News 5 Iglesias and Monedero go on the attack against the EU for vetoing Putin’s propaganda organs <i>Iglesias said that Russia Today and Sputnik report in favor of the Russian government just as Mediaset and Atresmedia report in favor of their owners.</i>	News 2 Social Communists defend Putin <i>The social-communist government is making an appalling fool of itself over Russia’s invasion of Ukraine.</i>
Congruent with right-wing bias	News 11 Education eliminates Philosophy as an elective and introduces Civic and Ethical Values <i>The contents of Philosophy will be included in the compulsory fourth-grade subject known as Education in Civic and Ethical Values.</i>	News 8 Socialist Soria City Council finances a workshop to draw vulvas on cloth bags <i>Funded by part of the 20 billion euros that Sanchez gave to the Minister of Equality.</i>
	News 6 73-year-old woman arrested in Toledo for marijuana cultivation <i>Seized 552 plants in advanced flowering stage with a total weight of 92.2 kilograms.</i>	News 3 Thousands of deer become infected with Covid and generate a new coronavirus <i>Interactions between deer and humans, or other animals, are now a cause for concern.</i>
Neutral content	News 12 A man’s true identity discovered thanks to Covid vaccine <i>In the US city of Newport.</i>	News 9 BBC reveals that the Kramatorsk massacre was the work of Ukraine, not Russia <i>The missile’s serial number belongs to the Ukrainian military.</i>

The questionnaire presented the headline decontextualized and asked the participant for his or her opinion about its veracity and then asked if he or she had used any means to contrast the information before giving his or her answer. There was no time limit for answering.

Block 2 of the questionnaire asked about the technological skills of the participants. Block 3 asked about the level of trust in the information received through all media. Block 4 inquired about the media and technology consumption habits, as well as the political preferences of the participants. Block 5 asked for demographic information. Finally, Block 6 explained the inaccuracies or falsehoods included in the headlines of Block 1.

The research was conducted between the end of April and the end of June 2022 and was carried out by *YouGov*. Specifically, the first wave was conducted between the end of April and mid-May 2022 and the second wave from the second week of June 2022 until the end of June 2022. The research design, fieldwork contracting, and data analysis were the sole responsibility of the research team. The project was approved by the *Research Ethics Committee* of the *University of Navarra*.

4. Preliminary results

4.1. Description of the sample

Of the 1,029 participants in the first phase of the study (*ex ante*), 47.9% were male and 52.1% were female; a single participant chose “other” in the definition of gender. The ages of the study participants can be seen in Table 3.

Of the 498 participants in the experimental group, 190 attended at least one of the 10 sessions of the course. For the purposes of data analysis, a subsample of 87 of these 190 cases was taken into consideration, made up of those who had attended five or more sessions. This decision, while undoubtedly limiting the predictive potential of the data, is consistent with the objective of observing results clearly linked to the course.

Among the 87 participants in the experimental group subsample, there are slightly more men (51.7%) than women (48.3%). Likewise, the difference between the percentage of men and women is somewhat more pronounced than that of the total sample. In the subsample of the experimental group, there is also a slightly higher percentage of people between 60 and 69 years of age in relation to the total sample.

Among the participants, the most frequent level of education was school graduation or vocational training. A total of 28% of the total sample and 36.7% of the experimental group had undergraduate or postgraduate university studies (see Table 5). The distribution of the total sample and the experimental group by Autonomous Community of residence can be seen in Table 6.

Table 5. Educational level of participants

Education level	Total sample		Experimental group	
	<i>n</i>	%	<i>n₂₅</i>	%
No education	23	2.2	1	1.1
School graduate	400	38.9	25	28.7
Professional training	318	30.9	29	33.3
University degree, 2 years	87	8.5	11	12.6
University degree, 4 years	143	13.9	12	13.8
Postgraduate	58	5.6	9	10.3
TOTAL	1,029	100	87	100

The political positions of the participants were distributed as follows: in the total sample, 42% declared that they belonged to left-wing ideological groups, 27.7% to the right, 18.8% said they belonged to the center, and 11.6% chose the “other” category. The distribution in the experimental group was similar, with 41.4% on the left, 26.4% on the right, 13.8% in the center, and 18.4% in the “other” category.

In terms of Internet use, 83% of both research groups reported going online several times a day, with no gender differences in this regard. The rest did so around once a day or several times a week. Table 7 shows the time spent using social networks. There are no major differences between the two groups, nor between men and women: more than half of the participants said they used the networks less than 30 minutes a day, and 20% used them more than an hour a day.

Table 3. Ages of participants in the total *ex ante* sample (n=1,029)

Age	Participants	%
50-59 years	335	32.6
60-69 years	448	43.5
>70 years	246	23.9

Table 4. Ages of participants in the experimental subsample (*n₂₅*=87)

Age	Participants	%
50-59 years	33	37.9
60-69 years	41	47.1
>70 years	13	14.9

Table 6. Place of residence of participants

Region	Total sample		Experimental group	
	n (1,029)	%	n≥5 (87)	%
Andalusia	180	17.5	15	17.2
Aragon	26	2.5	2	2.3
Principality of Asturias	39	3.8	5	5.7
Balearic Islands	23	2.2		
Canary Islands	31	3.0	3	3.4
Cantabria	15	1.5		
Castilla-La Mancha	35	3.4	2	2.3
Castilla y León	45	4.4	3	3.4
Catalonia	212	20.6	16	18.4
Valencian Community	113	11.0	10	11.5
Extremadura	26	2.5		
Galicia	36	3.5	5	5.7
La Rioja	4	0.4		
Community of Madrid	168	16.3	22	25.3
Region of Murcia	22	2.1	2	2.3
Community of Navarra	15	1.5	1	1.1
Basque Country	37	3.6	1	1.1
Ceuta/Melilla	2	0.2		

Table 7. Daily access to social networks during the last week (%)

Use times	Total sample		Experimental group	
	Man	Woman	Man	Woman
Less than 10 minutes	13.5	12.0	13.8	11.5
Between 10 and 30 minutes	15.9	14.4	16.1	14.9
Between 31 and 60 minutes	14.4	9.3	10.3	16.1
Between 1 and 2 hours	5.5	7.4	6.9	2.3
Between 2 and 3 hours	2.2	3.6	3.4	1.1
More than 3 hours	1.5	3.3	1.1	2.3

4.2. Course effectiveness

4.2.1. Ability to identify the veracity or falsity of news stories

Table 8 shows the percentages of success in identifying the veracity or falsehood of the news presented in waves 1 and 2 by both the experimental group ($n_{\geq 5}=87$) and the control group.

Table 8. Level of accuracy in detecting false and true news

	Wave	Holder	Political bias	(a) % Experimental subsample ($n_{\geq 5}=87$)	(b) Control group ($n=448$)	(a)-(b)
News 1	1	False	Left	41.3	34.2	7.1
News 2	1	False	Right	53.1	47.8	5.3
News 3	1	False	Neutral	87.5	83.8	3.6
News 4	1	True	Left	36.7	30.9	5.8
News 5	1	True	Right	49.8	45.0	4.8
News 6	1	True	Neutral	57.7	50.0	7.7
News 7	2	False	Left	33.0	37.1	-4.1
News 8	2	False	Right	59.1	60.0	-1.8
News 9	2	False	Neutral	66.6	60.2	6.3
News 10	2	True	Left	73.6	57.6	16.0
News 11	2	True	Right	40.5	26.0	14.4
News 12	2	True	Neutral	27.9	27.0	7.0

It can be seen that having attended at least five of the course sessions has a positive impact on the participants' ability to identify whether the headlines are true. The starting point of the experimental subsample is already more positive than that of the control group before the course, perhaps due to a greater prior interest in accessing reliable information, which led the participants to follow the course sessions and participate in the survey. Apart from this circumstance, what is certain is that the participants of the experimental subsample, after at least five sessions of the course, showed an improvement in the level of accuracy in identifying false or true news.

However, there are differences in the ability to recognize true and false news, respectively. The study shows that, once they have received the training, the degree of accuracy of the members of the experimental subsample improves when it comes to identifying as true the information that is indeed true. In contrast, their ability to identify false information as erroneous is only slightly improved or even slightly diminished. This difference in the ability to identify one or the other is probably related to the fact that older adults find it easier to confirm the veracity of true news from reliable sources, a habit cultivated throughout the course. In contrast, the skills needed to identify false information as erroneous require greater effort and expertise, which have not been sufficiently developed by the members of the subsample. In this sense, the knowledge and use of information verification organizations, which are much lower in this age group than the use of traditional information sources, limit their ability to identify false information.

Table 9 shows the values of subjective certainty perceived by the participants about the validity of their own judgment. In other words, it indicates how confident the participants are that they are correct about the veracity or falsity of the news. Those who have taken the course show a higher perception of confidence in assessing the falsity of the news than those in the total sample. The members of the experimental subsample reveal, in fact, greater subjective certainty of having been able to detect false information, even if they were actually confused. This overconfidence effect could be due to the fact that, in the absence of contrasted information, they consider that their training enables them to better detect falsehoods. On the other hand, they have more doubts about their own judgment in the news that they get right compared to the control group, the true ones. This phenomenon could be a double-edged sword, as it could cause the user to relax due to overconfidence in his or her ability.

Table 9. Degree of certainty in the detection of false and real news

	Wave	Holder	Political bias	(a) % Experimental subsample (n ₂₅ =87)	(b) Control group (n=448)	(a)-(b)
News 1	1	False	Left	23.1	21.6	1.6
News 2	1	False	Right	29.7	33.8	-4.1
News 3	1	False	Neutral	35.3	37.9	-2.7
News 4	1	True	Left	22.8	23.6	-0.8
News 5	1	True	Right	24.3	20.5	3.9
News 6	1	True	Neutral	17.1	16.4	0.7
News 7	2	False	Left	25.3	15.0	10.3
News 8	2	False	Right	35.7	17.8	17.9
News 9	2	False	Neutral	31.8	17.9	13.9
News 10	2	True	Left	28.7	21.8	6.9
News 11	2	True	Right	25.0	16.9	8.1
News 12	2	True	Neutral	27.3	10.3	17.0

4.2.2. Accuracy about the veracity or falsity of the news

Incidence of the gender variable

From the descriptive point of view, the judgment of the participants regarding the veracity or falsity of the news is analyzed, interpreting that the news are deemed false when they are scored between 1 and 3, and true when they are scored between 5 and 7, within a range from "1 = clearly false" to "7 = clearly true." In the case of certainty about the correctness of the assessment of the truthfulness or falsity of the news item, the analysis focuses on the mean number of responses to the question "How certain are you of your opinion?", taking as data the response to the option "5 = totally certain," within a scale whose range is "1 = not at all certain, 5 = totally certain." In this section, the results of this analysis are considered according to the gender of the participants. The descriptive analysis is complemented with the application of Student's t-test for independent samples, taking gender (male or female) as the independent variable and the dependent variables of the evaluation of the truthfulness or falsehood of each news item, and the certainty of correctness of the previous evaluation.

As can be seen in Table 10, for both waves and groups, the statistical analysis of the comparison of means indicates a low incidence of the gender variable in the correctness of the participants regarding the veracity or falsity of the news.

In this sense, the only cases that indicate the presence of this influence belong to the experimental group, one in the first wave of the study and two in the second. It is noteworthy that in these three cases, women outperformed men in terms of success.

From a descriptive point of view and according to the political orientation of the headline, women in the experimental group have a higher hit rate in all those headlines with a “right-wing” political bias, and men outperform women in three of the four headlines with a “left-wing” orientation. It is noteworthy that the headlines with a neutral political orientation in the first wave obtain the highest hit rates, for both men and women, compared to those news items with some type of political bias.

In the control group, there is little difference in the hit rate between men and women. The absolute mean difference between the two is 3.6 points. Men are slightly more successful than women in all the news items presented during the first wave, and women are more successful than men in three of the six news items of the second wave, with a tie between both genders in two other news items of this wave.

The differences within the experimental group, from the descriptive point of view, are somewhat more pronounced and varied. The absolute mean of the differences in the hit rate by gender reaches 11.9 points. In general terms, women score more correctly in both waves. In two of the three cases where there is a notable difference between men and women, the statistical analysis of the comparison of means indicates that this may be influenced by the gender variable. Finally, there is no clear pattern of accuracy by gender depending on whether the news is true or false, from both a descriptive and a statistical point of view.

Table 10. Accuracy about the veracity or falsehood of the news, according to the gender variable

	Wave	Holder	Political bias	Experimental group				Total sample			
				Total	Man	Woman	p	Total	Man	Woman	p
News 1	1	False	Left	41,3	42.2	40.5	0.353	34.2	36.9	31.6	0.594
News 2	1	False	Right	53.1	46.7	59.5	0.613	47.8	50.8	44.8	0.633
News 3	1	False	Neutral	87.5	84.4	90.5	0.645	83.8	85.6	82.1	0.271
News 4	1	True	Left	36.7	37.8	35.7	0.318	30.9	33.1	28.8	0.645
News 5	1	True	Right	49.8	40.0	59.5	0.016	45.0	46.2	43.9	0.409
News 6	1	True	Neutral	57.7	51.1	64.3	0.474	50.0	50.4	49.5	0.902
News 7	2	False	Left	33.0	42.2	23.8	0.362	37.1	36.0	38.2	0.295
News 8	2	False	Right	59.1	44.4	73.8	0.014	60.9	57.6	64.2	0.279
News 9	2	False	Neutral	66.6	68.9	64.3	0.611	60.2	60.6	59.9	0.417
News 10	2	True	Left	73.6	73.3	73.8	0.953	57.6	53.8	61.3	0.867
News 11	2	True	Right	40.5	33.3	47.6	0.116	26.0	27.5	24.5	0.313
News 12	2	True	Neutral	27.9	17.8	38.1	0.016	20.9	21.6	20.3	0.397

Safety on the hit valuation

Table 11 reflects the results of those participants who affirm total certainty about the accuracy of their assessment of the truthfulness or falsity of the news presented in the research. In this aspect, the statistical analysis of mean comparison indicates an almost null incidence of the gender variable during the first wave. However, during the second wave, this incidence is present in half of the news items evaluated in the control group and is almost non-existent in the case of the experimental group. This may indicate two realities: by repeating the same pattern of analysis and evaluation of news items in the same group, the statistical incidence of the gender variable increases. On the other hand, the fact of having participated in the training course significantly attenuates the influence of gender on the assessment of the correctness of the true or false nature of the news.

By groups, from the descriptive data, the control group offers fewer differences in the perception of security, the absolute mean difference between both genders being 5.1 points. In all the news items evaluated, men slightly outperformed women in the “totally safe” option. The experimental group offers more variety in this respect: the differences in absolute certainty on the correctness vary more; in absolute terms, it is 11.9 points between the two groups. During the first wave, greater confidence is equally distributed between the two genders, and men are more confident of being correct in the second wave.

Table 11. "I am totally sure" about the veracity or falsehood of the news, according to the gender variable

	Wave	Holder	Political bias	Experimental group				Total sample			
				Total	Man	Woman	<i>p</i>	Total	Man	Woman	<i>p</i>
News 1	1	False	Left	23,1	20.0	26.2	0.249	21.6	23.3	19.8	0.771
News 2	1	False	Right	29.7	35.6	23.8	0.310	33.8	36.4	31.1	0.056
News 3	1	False	Neutral	35.3	46.7	23.8	0.747	37.9	39.0	36.8	0.831
News 4	1	True	Left	22.8	28.9	16.7	0.018	23.6	24.6	22.6	0.262
News 5	1	True	Right	24.3	20.0	28.6	0.726	20.5	22.5	18.4	0.537
News 6	1	True	Neutral	17.1	22.2	11.9	0.237	16.4	18.6	14.2	0.239
News 7	2	False	Left	25.3	26.7	23.8	0.685	15.0	18.2	11.8	0.011
News 8	2	False	Right	35.7	33.3	38.1	0.312	17.8	19.9	15.6	0.107
News 9	2	False	Neutral	31.8	42.2	21.4	0.247	17.9	22.0	13.7	0.011
News 10	2	True	Left	28.7	31.1	26.2	0.699	21.8	23.3	20.3	0.935
News 11	2	True	Right	25.0	33.3	16.7	0.994	16.9	22.0	11.8	0.003
News 12	2	True	Neutral	27.3	37.8	16.7	0.410	10.3	14.0	6.6	<.001

5. Discussion and conclusions

As a starting point, it is apparent that the ability to discern whether a news item is true or false in the digital environment cannot be presupposed in the case of this age group, which is therefore particularly vulnerable (Papí-Gálvez; La-Parrá-Casado, 2022; Santibáñez-Velilla; Latorre-Santibáñez; Tejada-Sánchez, 2014). That said, and responding to the first research question of this article, it is possible to affirm that the realization of the course has proved effective in increasing the ability of participants to identify the truthfulness/falseness of the information presented to them. In relation to the incidence of the gender variable, and in line with what was pointed out by Ramírez-García, González-Fernández and Sedeño-Valdellós (2017), sex did mark some differences in this ability in the *ex ante* analysis, but the course has mitigated such differences in a way to the point of making them disappear in the sample in the *ex post* analysis.

Regarding the second question, where we asked whether greater knowledge may imply greater security, the results show that the perception of being better prepared to deal with disinformation may have a rebound effect that is particularly evident in the face of false information. Detecting falsehood in the news requires advanced technical skills, which are generally lacking in those over 50. On the contrary, confirming that information is true is easier for them, thanks to the traditional media that they regularly consume. (Brashier; Schacter, 2020). On the other hand, the increase in training to verify the veracity or falsity of a news item significantly reduces the effect that gender can have on the perception of being right in the judgment.

It is therefore possible to conclude that training actions on media literacy can have a positive effect on the population over 50. The previous interest that these people have in information makes them victims of hoaxes more frequently and, also, unwitting collaborators of their dissemination, as they do not have specific knowledge to verify what they hear or see. It cannot be forgotten that it is advisable to work with an increased concept of media competence (Kačínová; Sádaba, 2021), which also includes digital reality and its specific competences, as well as other possible areas.

This research seeks to provide ideas to better understand how older people face the challenge of disinformation and to offer useful clues for the design of effective media literacy strategies for this sector of the population. It becomes clear that it is necessary to seek and design strategies and activities that are adapted to their reality and that are easily accessible. Public administrations and third-sector organizations must be alert and aware of the vulnerability of this public and make firmer and more creative decisions to reach them.

6. Limitations of the study and next steps

Although it provides encouraging and positive data on the effectiveness of a training action to improve the defense capacity of the over-50 age group in the face of disinformation, the study has some limitations. It would be worth analyzing in depth the incidence of some sociological traits on the ability to distinguish between true and false news, as well as on the effectiveness of a training action. It could be explored, for example, whether factors such as political ideology, level of education, or frequency of Internet connection affect the ability to identify false content. The incidence of age, especially after 65 years of age, is also an element that could receive special attention. Another limitation, which points to the need to find creative and truly accessible ways to reach this age group, has to do with the platform (*WhatsApp*) selected for this course: while the content has proven capable of facilitating positive results in those who have completed a significant number of its sessions, it has been difficult to get a large sample of users who have completed it in its entirety.

It is possible that other formats and platforms, more accessible and understandable for this age group, may need to be considered in order to disseminate this valuable content. In this sense, it is true that the action itself had the limitations of the platform used, which made it difficult to offer support or reference material that could complement the videos.

7. References

- Altay, Sacha; Acerbi, Alberto** (2023). "People believe misinformation is a threat because they assume others are gullible". *New media & society*. Online first.
<https://doi.org/10.1177/14614448231153379>
- Amoedo-Casais, Avelino; Vara-Miguel, Alfonso; Negrodo, Samuel; Moreno-Moreno, Elsa; Kaufmann-Argueta, Jürg** (2021). *Digital news report España 2021*. Pamplona: Servicio de Publicaciones de la Universidad de Navarra.
<https://doi.org/10.15581/019.003>
- Aufderheide, Patricia** (1993). *Media literacy. A report of the national leadership conference on media literacy*. Queenstown, Maryland: The Aspen Institute. ISBN: 0 89843 137 9
<https://eric.ed.gov/?id=ED365294>
- Brashier, Nadia M.; Schacter, Daniel L.** (2020). "Aging in an era of fake news". *Current directions in psychological science*, v. 29, n. 3, pp. 316-323.
<https://psycnet.apa.org/doi/10.1177/0963721420915872>
- Brennen, J. Scott; Simon, Felix M.; Howard, Philip N.; Nielsen, Rasmus-Kleis** (2020). *Types, sources, and claims of Covid-19 misinformation*. Reuters Institute for the Study of Journalism Factsheet.
<https://reutersinstitute.politics.ox.ac.uk/types-sources-and-claims-covid-19-misinformation>
- Burnam, Tom** (1975). *The dictionary of misinformation*. New York, NY: Thomas Y. Crowell. ISBN: 978 0 690001471
- Carnahan, Dustin; Bergan, Daniel E.** (2022). "Correcting the misinformed: The effectiveness of fact-checking messages in changing false beliefs". *Political communication*, v. 39, n. 2, pp. 166-183.
<https://doi.org/10.1080/10584609.2021.1963358>
- Corbu, Nicoleta; Oprea, Denisa-Adriana; Frunzaru, Valeriu** (2021). "Romanian adolescents, fake news, and the third-person effect: A cross-sectional study". *Journal of children and media*, v. 16, n. 3, pp. 387-405.
<https://doi.org/10.1080/17482798.2021.1992460>
- Cucarella, Lluís; Fuster, Pau** (2022). *Informe. Alfabetización mediática: contexto actual, legislación, casos de éxito, herramientas y recursos, y percepción y propuestas de especialistas y profesores*. Laboratorio de Periodismo. Fundación Luca de Tena.
<https://laboratoriodeperiodismo.org/wp-content/uploads/2023/02/informe-alfabetizacion-mediatica.pdf>
- European Commission** (2019). *Tackling online disinformation*.
<https://ec.europa.eu/digital-single-market/en/tackling-online-disinformation>
- European Commission** (2022) *Guidelines for teachers and educators on tackling disinformation and promoting digital literacy through education and training*. Directorate General for Education, Youth, Sport and Culture.
<https://data.europa.eu/doi/10.2766/28248>
- Ferrés, Joan; Piscitelli, Alejandro** (2012). "Media competence. Articulated proposal of dimensions and indicators" [La competencia mediática: propuesta articulada de dimensiones e indicadores]. *Comunicar*, v. XIX, n. 38, pp. 75-82.
<https://doi.org/10.3916/C38-2012-02-08>
- Galdón-López, Gabriel** (1994). *Desinformación: método, aspectos y soluciones*. Pamplona: Eunsa. ISBN: 978 84 31312947
- García-Matilla, Agustín** (1999). "Escuela, televisión y valores democráticos". *Comunicar*, n. 13, pp. 107-110.
<https://doi.org/10.3916/C13-1999-16>
- Herrero-Diz, Paula; Conde-Jiménez, Jesús; Reyes-de-Cózar, Salvador** (2021). "Spanish adolescents and fake news: level of awareness and credibility of information (Los adolescentes españoles frente a las fake news: nivel de conciencia y credibilidad de la información)". *Culture and education*, v. 33, n. 1.
<https://doi.org/10.1080/11356405.2020.1859739>
- IAB Spain** (2019a). *Estudio anual de mobile & connected devices 2019*.
<https://iabspain.es/estudio/estudio-anual-de-mobile-connected-devices>
- IAB Spain** (2019b). *Estudio anual de redes sociales 2019*.
<https://iabspain.es/estudio/estudio-anual-de-redes-sociales-2019>
- Iandoli, Luca; Primario, Simonetta; Zollo, Giuseppe** (2021). "The impact of group polarization on the quality of online debate in social media: A systematic literature review". *Technological forecasting and social change*, v. 170, 120924.
<https://doi.org/10.1016/j.techfore.2021.120924>

- Kačínová, Viera; Sádaba, Charo** (2021). "Conceptualización de la competencia mediática como una 'competencia aumentada'". *Revista latina de comunicación social*, n. 80, pp. 21-38.
<https://doi.org/10.4185/RLCS-2022-1514>
- Livingstone, Sonia** (2004). "What is media literacy?". *Intermedia*, v. 32, n. 3, pp. 18-20.
<http://eprints.lse.ac.uk/id/eprint/1027>
- Margolin, Drew B.; Hannak, Aniko; Weber, Ingmar** (2018). "Political fact-checking on *Twitter*: When do corrections have an effect?". *Political communication*, v. 35, n. 2, pp. 196-219.
<https://doi.org/10.1080/10584609.2017.1334018>
- Martínez-Costa, María-Pilar; López-Pan, Fernando; Buslón, Nataly; Salaverría, Ramón** (2022). "Nobody-fools-me perception: Influence of age and education on overconfidence about spotting disinformation". *Journalism practice*, Online first.
<https://doi.org/10.1080/17512786.2022.2135128>
- Ministerio de la Presidencia, Relaciones con las Cortes y Memoria Democrática (2022). *Lucha contra las campañas de desinformación en el ámbito de la seguridad nacional. Propuestas de la sociedad civil*. Madrid: Departamento de Seguridad Nacional; Ministerio de la Presidencia, Relaciones con las Cortes y Memoria Democrática.
<https://www.dsn.gob.es/es/documento/lucha-contra-campañas-desinformación-ámbito-seguridad-nacional-propuestas-sociedad-civil>
- Moore, Ryan C.; Hancock, Jeffrey T.** (2022). "A digital media literacy intervention for older adults improves resilience to fake news". *Scientific reports*, n. 12, 6008.
<https://doi.org/10.1038/s41598-022-08437-0>
- Namle (2007). *Core principles of media literacy education in the United States*. National Association for Media Literacy Education.
<https://namle.net/wp-content/uploads/2020/09/Namle-Core-Principles-of-MLE-in-the-United-States.pdf>
- Nieminen, Sakari; Rapeli, Lauri** (2019). "Fighting misperceptions and doubting journalists' objectivity: A review of fact-checking literature". *Political studies review*, v. 17, n. 3, pp. 296-309.
<https://doi.org/10.1177/1478929918786852>
- Oeldorf-Hirsch, Anne; Schmierbach, Mike; Appelman, Alyssa; Boyle, Michael P.** (2020). "The ineffectiveness of fact-checking labels on news memes and articles". *Mass communication and society*, v. 23, n. 5, pp. 682-704.
<https://doi.org/10.1080/15205436.2020.1733613>
- Papí-Gálvez, Natalia; La-Parra-Casado, Daniel** (2022). *Informe 2022. Cátedra de brecha digital generacional. Las personas mayores en la era de la digitalización en la Comunidad Valenciana (datos 2021)*. Alicante: Cátedra de Brecha Digital Generacional.
<http://hdl.handle.net/10045/122914>
- Pérez-Tornero, José-Manuel** (2008). "La sociedad multipantallas: retos para la alfabetización mediática". *Comunicar*, v. 16, n. 31, pp. 15-25.
<https://doi.org/10.3916/c31-2008-01-002>
- Ramírez-García, Antonia; González-Fernández, Natalia; Sedeño-Valdellós, Ana-María** (2017). "La competencia mediática en la población mayor. Diagnóstico de la realidad española". *Signo y pensamiento*, v. 36, n. 70, pp. 94-111.
<https://doi.org/10.11144/Javeriana.syp36-70.cmpm>
- Redondo, Myriam** (2018). *Verificación digital. Manual contra bulos y desinformación internacional*. Barcelona: UOC. ISBN: 978 84 91801313
- Sádaba, Charo; Salaverría, Ramón** (2023). "Combatir la desinformación con alfabetización mediática: análisis de las tendencias en la Unión Europea". *Revista latina de comunicación social*, n. 81, pp. 17-33.
<https://doi.org/10.4185/RLCS-2023-1552>
- Salaverría, Ramón; Buslón, Nataly; López-Pan, Fernando; León, Bienvenido; López-Goñi, Ignacio; Erviti, María-Carmen** (2020). "Desinformación en tiempos de pandemia: tipología de los bulos sobre la Covid-19". *Profesional de la información*, v. 29, n. 3, e290315.
<https://doi.org/10.3145/epi.2020.may.15>
- Santibáñez-Velilla, Josefina; Latorre-Santibáñez, Carmen; Tejada-Sánchez, Soraya** (2014). "Evaluación de la competencia mediática estética en personas mayores". *Revista didáctica, innovación y multimedia*, n. 30.
<https://dimglobal.net/revistaDIM30/revista30ARcompetenciamayores.htm>
- Shu, Kai; Bhattacharjee, Amrita; Alatawi, Faisal; Nazer, Tahora H.; Ding, Kaize; Karami, Mansooreh; Liu, Huan** (2020). "Combating disinformation in a social media age". *WIREs data mining knowledge and discovery*, n. 10, e1385.
<https://doi.org/10.1002/widm.1385>

Uteca; UNAV (2022). *I Estudio sobre la desinformación en España*. Uteca; Universidad de Navarra.
<https://dadun.unav.edu/handle/10171/63643>

Valera-Ordaz, Lidia; Requena-i-Mora, Marina; Calvo, Dafne; López-García, Guillermo (2022). "Unraveling disinformation: Notions and discourses from the Spanish population". [Desenredando la desinformación: Nociones y discursos de la población española]. *Comunicar*, v. XXX, n. 72, pp. 21-32.
<https://doi.org/10.3916/C72-2022-02>

Walter, Nathan; Cohen, Jonathan; Holbert, R. Lance; Morag, Yasmin (2020). "Fact-checking: A meta-analysis of what works and for whom". *Political communication*, v. 37, n. 3, pp. 350-375.
<https://doi.org/10.1080/10584609.2019.1668894>

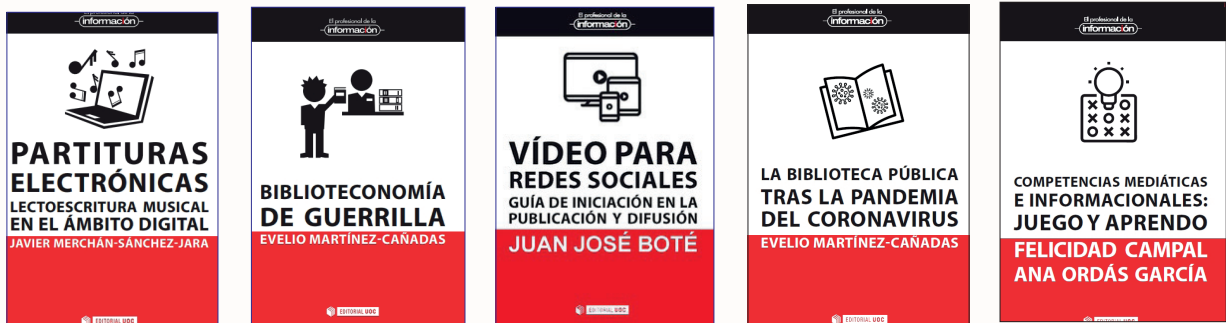
Wardle, Claire; Derakhshan, Hossein (2017). *Information disorder: Toward an interdisciplinary framework for research and policy making*. Strasbourg: Council of Europe.
<https://edoc.coe.int/en/media/7495-information-disorder-toward-an-interdisciplinary-framework-for-research-and-policy-making.html>

Watson, Peter C. (1960). "On the failure to eliminate hypotheses in a conceptual task". *Quarterly journal of experimental psychology*, v. 12, n. 3, pp. 129-140.
<https://doi.org/10.1080/17470216008416717>

Pocket book collection

Profesional de la información (Editorial UOC)

Last published titles



More information:

<https://www.profesionaldelainformacion.com/producto/libros-epi-uoc>

Fact-checking in war: Types of hoaxes and trends from a year of disinformation in the Russo-Ukrainian war

Raúl Magallón-Rosa; Carolina Fernández-Castrillo; Miriam Garriga

Nota: Este artículo se puede leer en español en:

<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87315>

Recommended citation:

Magallón-Rosa, Raúl; Fernández-Castrillo, Carolina; Garriga, Miriam (2023). "Fact-checking in war: Types of hoaxes and trends from a year of disinformation in the Russo-Ukrainian war". *Profesional de la información*, v. 32, n. 5, e320520.

<https://doi.org/10.3145/epi.2023.sep.20>

Article received on March 10th 2023

Approved on September 14th 2023



Raúl Magallón-Rosa ✉
<https://orcid.org/0000-0002-2236-7802>

Universidad Carlos III de Madrid
Departamento de Periodismo y Comunicación
Audiovisual
Madrid, 133
28903 Getafe (Madrid), Spain
raul.magallon@uc3m.es



Carolina Fernández-Castrillo
<https://orcid.org/0000-0001-6108-6440>

Universidad Carlos III de Madrid
Departamento de Periodismo y Comunicación
Audiovisual
Madrid, 133
28903 Getafe (Madrid), Spain
carolfer@hum.uc3m.es



Miriam Garriga
<https://orcid.org/0009-0004-4797-7959>

Universidad Carlos III de Madrid
Departamento de Periodismo y Comunicación
Audiovisual
Madrid, 133
28903 Getafe (Madrid), Spain
mgarriga@pa.uc3m.es

Abstract

This study explores the verification of the contents related to the coverage of the Russo-Ukrainian war during the first year of the conflict. We address the analysis of false information collected from the *EDMO* database by the Spanish fact-checking organizations: *AFP Factual* and *Comprovem*, *EFE Verifica*, *RTVE Verifica*, *Maldita.es*, *Newtral* and *Verificat*. Based on the results obtained, a typology has been established to identify the style, format and content of the misinformation under study. In this way, we follow the main trends in the manipulation dynamics that shaped the media coverage of the latest war in European territory. In total, up to 307 verifications by verifiers working in the Spanish context are analyzed. Unlike the hoaxes related to the origin of the coronavirus, in which the preferred format was the text message, in the invasion of Ukraine visual evidence has prevailed in the media coverage of the conflict during the initial phase of the conflict. We will see which are the social networks in which the greatest traffic and viralization of false news is detected, exposing users to manipulative content to a greater extent. Likewise, the potential implementation of a transnational network to combat disinformation in war contexts will be assessed. We will pay special attention to the important role of Spanish fact-checkers both at the European level and in identifying false information, avoiding its dissemination in Latin America. And, finally, we will detect the new challenges that war fact-checking faces, as a result of the evolution of falsification strategies in the construction of the collective narrative about the Russo-Ukrainian war in the post-truth era.

Keywords

Media literacy; Hoaxes; News coverage; Disinformation; Warlike fact-checking; Fake news; Fact-checkers; Verification; Infociation; War journalism; Propaganda; Wars; Ukraine; Russia.



Funding

This article is a result of *Iberifier (Iberian Media Research & Fact Checking)*, a project that has received funding from the *European Commission*, agreement CEF-TC-2020-2 (*European Digital Media Observatory*) with reference 2020-EU-IA-0252.

1. Introduction

The aim of this research is to analyse the role of fact-checking organisations in the journalistic coverage of Russia's invasion of Ukraine and the verifications made during the first year of the war. This study explores the challenges that this new scenario poses for the field of verification and reporting in armed conflicts, focusing on the coverage of the Russian-Ukrainian war.

We address the analysis of the verifications collected in the database of the *European Digital Media Observatory (EDMO)* by the Spanish fact-checking organisations¹. In total, up to 307 verifications carried out by the participating institutions are analysed: *AFP Factual* (including *AFP Comprovem*), *EFE Verifica*, *RTVE Verifica*, *Maldita.es*, *Newtral* and *Verificat*. In addition to the distribution of the contents verified by these organisations, the narratives are analysed according to themes, months, countries involved, etc.

Unlike the hoaxes related to the origin of the coronavirus, where text messages were the preferred format (**Molina-Cañabate; Magallón-Rosa, 2023**), in the invasion of Ukraine the relevance of visual evidence prevailed in media coverage of the conflict during the initial phase of the fighting (**Fernández-Castrillo; Ramos-Vera, 2023a; 2023b**), becoming the focus of attention, monitoring and verification by fact-checking organizations.

Furthermore, the evolution in the development of increasingly diversified strategies for the propagation of hoaxes leads us to evaluate the role of decontextualization and the part played, in this respect, by certain actors on social platforms as possible predominant elements in the current information war.

Referred to as the First World War of the 21st century (**Martino, 2022**) because of the level of global involvement by groups of internet users who have helped in the verification of information through *Open-Source Intelligence (OSINT)* tools for the collection, analysis and linking of data for the dissemination of public information, but also because of the consequences for international geopolitics. It is one of the most recent examples of hybrid warfare in which the use of platforms, media and social networks is added to military, political, technological, economic, and psychological strategies to destabilise the adversary and influence international public opinion (**Carmack, 2022; Petrocchi, 2022; Susska; Chernii; Sukharevska, 2022**).

This is therefore an ideal context for the development of new methods of manipulation and the viralization of hoaxes, which makes it necessary to review the current state of information and content verification strategies, networks and actions at the international level. Based on the data collected, we will evaluate Spain's situation in the global fight against disinformation in war contexts –taking the Russian-Ukrainian conflict as an object of study–, as well as the limits and possibilities of setting up a transnational fact-checking network in war contexts.

Finally, a series of recommendations are put forward in the form of a decalogue concerning how to respond to disinformation in war conflicts and to try to help improve media coverage in this type of situation.

2. State of play**2.1. Disinformation and the presence of social platforms in war journalism**

In the case of war conflicts where the interest of international public opinion is shared and is reflected in the daily media coverage, social platforms have become a fundamental source and channel for the circulation of content. They are intermediaries that allow journalists and media to access a wide variety of content that would otherwise be very difficult to obtain in situ in such a short time, such as images and testimonies of some of the victims (**Sacco; Bossio, 2015**).

As an example, the integration of testimonies of Albanian Internet users through their blogs in the coverage of the Kosovo War served to denounce human rights violations by the Serbian authorities in the late 1990s (**Carreras-Álvarez; Román-San-Miguel, 2011**) and, among the main achievements of the Arab Spring, is the convergence between traditional and social media, both for the organisation and mobilisation of citizens and to vindicate and give visibility to the events (**Soengas-Pérez, 2013; AlSayyad; Guvenec, 2015; Ortiz-Galindo, 2016; Soengas-Pérez; Assif, 2017**).

The construction of the collective narrative on war conflicts is no longer the exclusive monopoly of the media due to the influence of the dynamics introduced by the growing presence of User-Generated Content (UGC) (**Fernández-Castrillo, 2014**) but also by the propaganda strategies of the countries involved.

“ This study explores the challenges that this new scenario poses for the field of verification and reporting in armed conflicts, focusing on the coverage of the Russian-Ukrainian war ”

This situation is particularly significant in the case of crisis and disaster coverage in general –and war in particular–, due to the limitations of access to the main locations and sources of information (**Micó et al.**, 2008; **Sánchez-Gonzales; Martos-Moreno**, 2020).

In this context, the proliferation of news content makes it difficult to identify the origin and authorship of the information, reinforcing the role of the image as the axis of a collective narrative that is illustrated rather than narrated (**Rodríguez-Tranche**, 2019). Photography as a supposed guarantee of maximum objectivity gives greater strength to the manipulation of the image (**Caballo-Méndez; De-Santiago-Mateos**, 2021), leading to a lack of contextualisation, which is one of the main threats when it comes to interpreting distant realities with a broad impact, as in the object of study in question.

It is also worth noting the important evolution of the role of the journalist from gatekeeper to gatewatcher (**Bruns**, 2003), especially evident in war contexts in which the journalist’s work as a “content curator” is enhanced (**Loo-Vázquez et al.**, 2016), where in addition to filtering and blocking information, he or she must focus on the selection and dissemination of other people’s content.

In the case of the Russian invasion of Ukraine, social networks have not only become relevant for the distribution and viralization of certain content, but also for the lack of access to the platforms themselves within the territories in conflict. In March 2022, Russia’s communications regulatory agency, *Roskomnadzor*, blocked access to *Facebook* and *Twitter* throughout the country in the face of the restriction by these social networks –and also by the European Union– on the dissemination of news from media outlets such as *Russia Today* and the agencies *Sputnik* and *RIA Novosti* (**Milmo**, 2022; **Kemp**, 2023; *Statcounter*, 2023)².

A recent study on the presence of social networks in the coverage of the Russian-Ukrainian war by two of the main digital media in the countries involved –*Ukrayinska Pravda* and *Rossiyskaya Gazeta*– highlights the need for

“greater transparency in the selection of UGC published in the media, especially in the case of screenshots, a type of content that is increasingly present based on appropriationism and the selection of fragments from user-created videos” (**Fernández-Castrillo; Ramos-Vera**, 2023a).

Therefore, as **Himma-Kadakas** and **Ojamets** (2022) point out, the development of advanced verification skills through social networks and the alliance of *OSINT* organisations with fact-checking organizations are two new factors determining the coverage of war journalism today.

2.2. Prevailing narratives and disinformation in the Russian-Ukrainian war

In the most recent imaginary of the Russo-Ukrainian conflict we find series of narratives that, according to **Lazarenko** (2019), are summarised in a Russian strategy centred on the incorporation of Crimea, Russia’s fraternal link with Ukraine and the protection of Russian speakers in the diaspora due to Russophobia (**Karpchuk; Yuskiv**, 2021), versus a Ukrainian narrative centred on the heroic struggle against invading imperialism and Russian-funded and supported separatists.

These narratives sometimes lead to sub-narratives about the “neo-Nazi insurgency in Kiev” to justify Russia’s position in the conflict and contribute to extending the battlefield to the struggle for hegemony between imaginaries. From this perspective, the main Russian disinformation narratives detected in the wake of the Euromaidan protests in 2013-2014 have focused on portraying Ukraine as a ‘Nazi state’, a ‘failed state’ or a ‘non-state’.

In this context, it is worth mentioning **Zawadzki’s** (2022) study of pro-Russian, anti-Ukrainian and anti-US accounts on *Twitter*, *Facebook* and *Vkontakte* –the Russian version of *Facebook*, created in 2006–. According to the results, some of the monitored accounts were deleted and many of the remaining accounts were classified as bots.

For his part, **Ibitowa** (2022) attempted to document the fact that media coverage is always tainted by attempted propaganda manipulation by the parties to the conflict. As an example, he notes that the killing of two people on 15 November 2022 in the Polish village of Przewodow –six kilometres from the Ukrainian border– was initially attributed to a missile attack by the Russian army, but eventually both Warsaw and Washington were forced to exonerate Moscow (**Ibitowa**, 2022).

In this evolution of tactics and narrative, **Aral** (2023) observes that Russian disinformation has shifted from the third decade of the 21st century from bots to the support of real profiles to create credible narratives and cascade them.

Temporally, in this evolution of the Russo-Ukrainian war, the first phase focused on discrediting countries supporting Ukraine, while later Russia’s disinformation strategy focused on curbing public debate on possible *NATO* expansion by polarising and viralizing content aimed at immigrant populations in neighbouring countries (**Aral**, 2023; **Yarova**, 2023).

By way of example, the Czech Republic has strengthened its strategic line to combat disinformation and the actions carried out by the *Centre Against Terrorism and Hybrid Threats* (*Centrum proti terorismu a hybridním hrozbám*, *CTHH*) have been joined by the *Prague Security Studies Institute* (*PSSI*), the *Czech Elves* (*Čeští elfové*), *Manipulátoři.cz*, *Demagog.cz* or the *NELEŽ* project (**Cabada**, 2022).

“The main Russian disinformation narratives detected in the wake of the Euromaidan protests in 2013-2014 have focused on portraying Ukraine as a ‘Nazi state’, a ‘failed state’ or a ‘non-state’”

Regarding the evolution of narratives in the last year of war, for Viktoriia Romaniuk, deputy director of the Ukrainian fact-checking organisation *StopFake*,

“Russian disinformation shifted its focus from the initial narratives to new ones: falsehoods related to arms deliveries to Ukraine, military support from Western countries, the food and energy crisis, and false stories discrediting refugees from Ukraine” (*Maldita.es*, 2023a).

“The first phase focused on discrediting countries supporting Ukraine, while later Russia’s disinformation strategy focused on curbing public debate on possible *NATO* expansion by polarising and viralizing content aimed at immigrant populations in neighbouring countries”

In this sense, some narratives of the invasion of Ukraine still resemble those of the post-Soviet strategy (**Vorster**, 2022), and are based on information intoxication, enemy exhaustion, inoculation of mistrust towards leaders, intensification of class divisions, incrimination of the enemy, or propagation of threats, among others (**Stancu**, 2019).

2.3. EU reports on disinformation during the conflict

In June 2018, the *European Council* mandated the then *High Representative of the Union for Foreign Affairs and Security Policy*, Federica Mogherini, and the *European Commission* to draft a project on coordination among Member States against disinformation. The *Action Plan* was presented in December of the same year and endorsed by the *European Council*.

Among the most prominent measures was an *Early Warning System (EWS)* to alert European partners in the event of a disinformation campaign and the implementation of common measures if necessary.

For its part, the *European External Action Service (EEAS)* has a specific task force on *Foreign Information Manipulation and Interference (FIMI)*, which addresses the problem of disinformation and possible foreign interference.

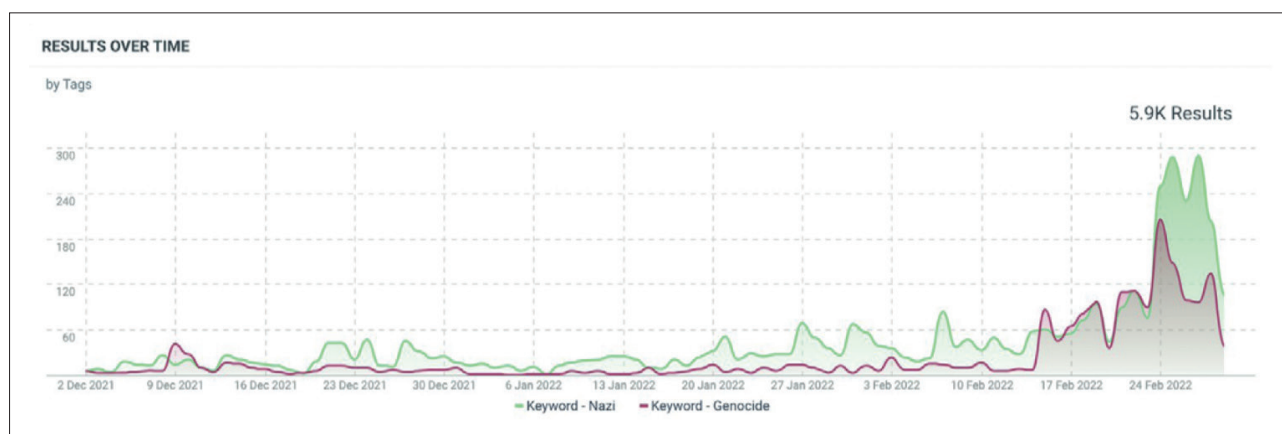
According to the 2022 *Activity Report*,

“Russia’s use of *FIMI* in its war against Ukraine has provoked an unprecedented reaction from the European Union. Sanctions against media outlets such as *RT* and *Sputnik*, which are directly controlled by the Kremlin and used as instruments of war propaganda, have demonstrated a strong determination to impose the costs of *FIMI*’s activities’ (*EEAS*, 2023, p. 3).

By 2023, this unit had the largest budget in its history with €14.6 million, which allowed it to strengthen cooperation with its international partners and boost projects on analysing and creating resilience strategies that have played a key role in documenting the actions carried out by the Kremlin since 2015.

FIMI has detected two levels of action regarding disinformation actions by the Russian state: one domestic and the other international. The first consists of manipulation and disinformation practices ranging from censorship to the banning of independent media on Russian territory in order to prevent any kind of domestic opposition to the war and targeted media coverage, including entertainment content. While the international one aims to undermine support for Ukraine, sowing doubts about who the aggressor is and exploiting sensitive issues for the EU such as migration and refugees, the rising cost of living and gas, among others (*EEAS*, 2023).

Some of the main disinformation narratives are aimed at the *Eastern Partnership (EaP)*, an initiative launched in 2009 to strengthen relations between the EU, its member states and its six eastern neighbours: Ukraine, Georgia, the Republic of Moldova, Armenia, Azerbaijan, and Belarus. Among the main objectives is to inoculate the fear of possible interference and the feeling that the citizens of these countries have no choice over their decisions (*EEAS*, 2021).



Graph 1. Presence of the terms “Nazi” and “genocide” in Russian state-controlled media coverage of Ukraine. Source: *EEAS*, 2023.

In this line, and from the *Ukraine War Resource Hub* –launched by the *EU Disinfo Lab*–, they have identified some of the main organisations and projects dedicated to combating hoaxes during the Russian-Ukrainian war, among which the following stand out: *Detector Media*, *EDMO*, *EUvsDisinfo*, *Facta.news*, *Fake fact-checks*, *ISD*, *NewsGuard Disinformation Tracking Center*, *The Washington Post’s fact-checking section*, *UkraineFacts*, etc. (*EU Disinfo Lab*, 2022).

The *Ukraine Facts* project, meanwhile, has seen fact-checking organizations from all over the world debunk more than 2,800 hoaxes about the war in Ukraine. This initiative takes over from the database created by the *International Fact-Checking Network (IFCN)* at the beginning of the pandemic as a result of learning from other collaborative fact-checking proposals such as *Verificado* (Mexico and Uruguay), *Comprova* (Brazil), *Reverso* (Argentina) and *Cross-Check* (Sánchez-Duarte; Magallón-Rosa, 2020).

2.3.1. What is EDMO?

Established in June 2020 as an EU-funded body, the *European Digital Media Observatory (EDMO)* is a partnership under the leadership of the *European University Institute* in Florence (Italy). *EDMO* brings together fact-checkers, media literacy experts and academic researchers with the aim of understanding and analysing misinformation, in collaboration with journalistic organisations, online platforms and industry professionals.

It currently has fourteen national or multinational centres (hubs or observatories) covering the 27 Member States of the European Union, as well as Norway. Each hub constitutes a network of organisations active in one or more Member States, to provide specific knowledge of local environments with the aim of strengthening the detection and analysis of disinformation campaigns (*EDMO*, 2023).

One year after the invasion, the *EDMO* database had a total of 2074 verifications. Of these, as many as 166 referred to Ukrainian President Zelensky in the headline, compared to 84 for Putin and six for Spain as a country.

EDMO’s Ukraine section (2022) had initially identified these five main narratives during the first months of the war:

1. The prejudice that Ukraine is a Nazi country.
2. The idea that President Zelensky is not fit to lead and is a Nazi.
3. Western media publish “fake news”.
4. The exaggerated economic consequences of the war for Europe.
5. Ukrainian refugees are violent and ungrateful and are treated better than EU citizens.

In addition, *EDMO* alerted international experts to the fact that disinformation detection practices have focused on English-speaking countries in the Western world, leaving Central and Eastern Europe unattended. According to *EDMO Ukraine*, disinformation is cross-border and cross-linguistic, so it is important to support preventive pre-bunking actions to avoid the spread of hoaxes in other EU languages (*EDMO*, 2022).

3. Objectives and research questions

The main objectives of this study are to collect the verifications of Spanish fact-checkers and to study the behaviour and patterns of disinformation dissemination during the first year of the Russian invasion of Ukraine.

Secondary objectives include the possibility of analysing the distribution of hoaxes by typology, subject, month, country involved, etc.

Based on these objectives, the following hypotheses were proposed:

- H1. President Zelensky-related hoaxes are one of the most potent narratives of the Russian disinformation strategy.
- H2. The first weeks of war account for a very significant part of the content verified by *fact-checking* organisations.
- H3. Hoaxes are adapted to local contexts in order to maintain the strategic tension over the war.

In turn, the research questions that were asked were:

- Q1. What is the distribution by month of the verifications carried out by the fact-checking organisations? What reading can be made of the results?
- Q2. Which verification organisations uploaded the most items to the *EDMO* database, and is there a pattern?
- Q3. How many checks were repeated between the different fact-checkers?
- Q4. What types of verifications were carried out?
- Q5. Which were the countries referred to in the verification or hoax?
- Q6. What were the main social networks or platforms that fact-checking organisations identified in their checks?

4. Methodology

The analysis and study of this type of collaborative fact-checking projects, which are becoming more and more frequent, raises the need to establish a combination of quantitative methodologies –on the number of verifications, format, time distribution, etc.– and qualitative ones, such as narrative and discursive strategies and the choice of the design of the database they share.

From this perspective, and as we pointed out in the introduction, the fact-checking organisations selected were six: *AFP Comprovem* (in Catalan), which is part of *AFP Factual*, *EFE Verifica*, *RTVE Verifica*, *Maldita.es*, *Newtral* and *Verificat*.

“ The verification work of fact-checking organisations is determined by news cycles, but also by the intensity of the coverage of the Russian-Ukrainian war itself ”

For its part, the selected sample was chosen from the database made public by the *EDMO*, which can be downloaded and updated over the course of a year. The dates of the analysis range from 24 February 2022 to 23 February 2023³.

From a methodological point of view, the first selection was to eliminate those pieces of information that had more than one update in the database itself. Afterwards, hoaxes repeated by the different fact-checking organisations were located –repetition was defined as the identical duplication of an unequivocal hoax–, either between different journalistic brands or between the same brand when it published it in more than one different article. We also detected publications that included groups of hoaxes, which were suppressed when calculating certain frequencies, because they generated distortions as they could not be considered –within the given unit of analysis– as an individual hoax⁴.

In other words, for the content analysis, two units of analysis were selected: the verification, explanation and contextualisation articles carried out by the fact-checkers and the unequivocal hoaxes identified by the journalistic organisations. In the first case, the sample was extended to 307 publications of which, after eliminating duplications, updates, and compilations, 211 verifications remained.

To draw up the distribution by geographical regions, the hoaxes were first coded by the countries involved, and on that basis, they were grouped into geographical regions, states directly involved in the conflict and the nation of origin of the verification. In this way, the following classification was established, in which the following territories could be detected:

1. Ukraine.
2. Russia.
3. Europe: Poland, United Kingdom, Serbia, Netherlands, Italy, Germany, Romania, Finland, Belarus, Czech Republic, and the Vatican. Includes hoaxes involving the European Union as a body.
4. Spain.
5. North and Central America: USA, Mexico, and Canada.
6. Middle East and North Africa: Qatar, Israel, and Afghanistan.
7. Asia and Australia: Japan, Australia, China, and North Korea.

On the other hand, it should be pointed out that the terms selected when classifying hoaxes by theme have been taken based on the hoax itself, and not its subsequent verification. In other words, if a falsehood has been spread about an attack with military vehicles, and it is subsequently corroborated that the offensive has been perpetrated solely with weapons, it will be classified under the theme of Military vehicles.

Regarding the typology of verifications, we start from the work of **Wardle** and **Derakhshan** (2017) who identify the following types of *dis/mis/malinformation*:

- Satire or parody: not intended to cause harm but with the potential to mislead.
- Misleading content: misleading use of information for the purpose of framing an issue or an individual.
- Impostor content: impersonation of content from original sources.
- False connection: headlines, images or subtitles do not match the content.
- False context: original content is shared with false contextual information.
- Manipulated content: the original information or images are manipulated with the intention to mislead.
- Fabricated content: the new content is 100% fake and is designed to mislead and cause harm.

In addition, a new one was added: the explanatory post. This is content whose veracity could not be corroborated by the fact-checkers, so they decided to provide contextualisation.

These explanatory narratives are part of the fact-checking organisations' literacy strategy that goes beyond traditional fact-checking itself. It was decided to choose this categorisation for further indexing by types, as part of the concept of misinformation includes *misinformation*, but also intentional malpractice (*disinformation* and *malinformation*).

Regarding the formats (audio, video, text, etc.) and themes, beyond the frequency of categorisation, we were also interested in analysing whether these were linked to current events or to the way the war itself unfolds.

Finally, we focused on checking whether there was an internationalisation of the hoaxes verified by Spanish organisations and whether this allowed us to identify possible patterns.

5. Analysis of results

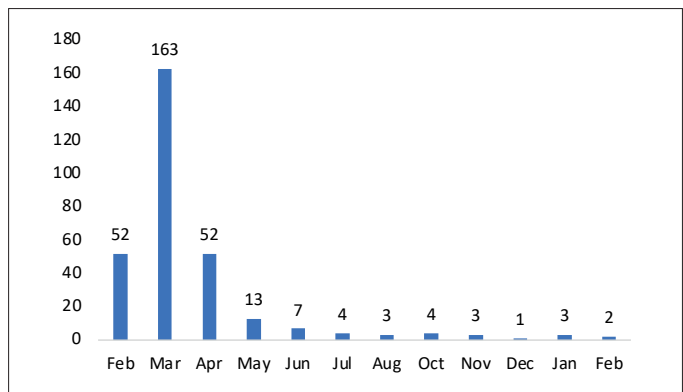
5.1. Spanish fact-checkers and their pieces on misinformation analysis

The verification work of fact-checking organisations is determined by news cycles, but also by the intensity of the coverage of the Russian-Ukrainian war itself. In order to measure how the chronological issue is related to the verification cycles, it was decided to examine two variables: the number of verifications shared monthly by each organisation and their

repetition –either by the same media outlet when different hoaxes are grouped together or when it shares coverage with other media outlets–. It was also decided to analyse the distribution of the verifications identified and uploaded to the *EDMO* database by the different journalistic projects to analyse possible dynamics and differences.

5.1.1. Month-by-month analysis

The month of March (Graph 2) is not only the month with the highest number of verification articles over the year analysed, but also the one in which the topic of the Ukrainian War is most present directly –with a total of 160 verifications.



Graph 2. Number of verifications analysed by Spanish fact-checkers distributed by month. Source: *EDMO*.

Although the percentage of misinformation in February is lower, it must be considered very high if we consider that the war had only been going on for six days and that all the claims examined during this period were about the start of the conflict. From April onwards, hoaxes about the Russo-Ukrainian war decrease, coinciding with a lower level of public interest in it (see Figure 1). In September, a turning point is perceived, with fact-checking organizations not contradicting any information.⁵

This does not mean that disinformation has stopped circulating on social networks and platforms, but rather that the content that has circulated in Spain or that fact-checking organisations have received for consultation has been significantly reduced.

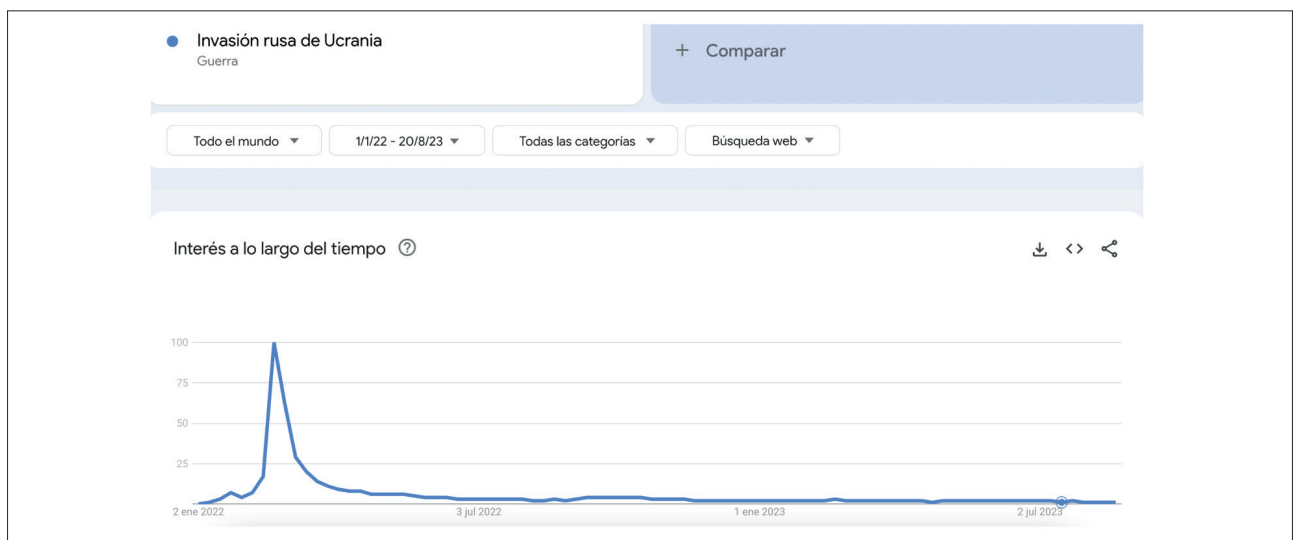
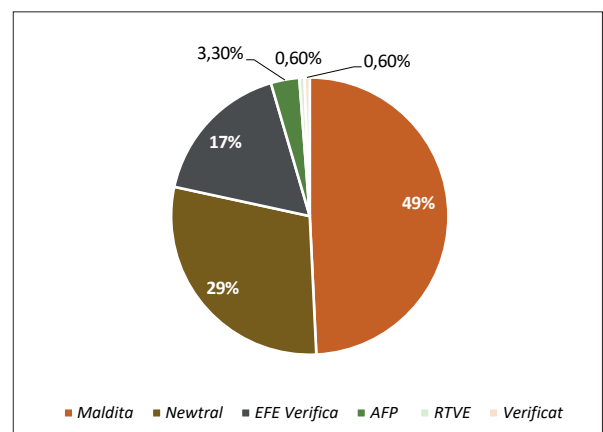


Figure 1. Search volume for *Google's* indexed term on the Russo-Ukrainian war, corresponding to “Russian invasion of Ukraine”. Source: *Google Trends*

5.1.2. Fact-checker analysis and repetitions of hoaxes

The total number of verifications identified and uploaded to the *EDMO* database has been uneven among the different verifiers. The organisation that analysed and verified the most disinformation on the conflict was *Maldita.es*, while the second most verified was *Newtral*, followed by *EFE Verifica*. In this regard, it should be mentioned that *Maldita.es* –as the driving force behind *Ukraine Facts*– is the media outlet that has focused most of its work on disproving hoaxes on this issue. Its verifications accounted for almost half of the total number of pieces carried out by journalistic organisations.



Graph 3. Percentage of hoaxes verified by each fact-checking organisation between February 2022 and February 2023.

On the other hand, we highlight that more than half of the publications categorised in the *EDMO database* are repeated, up to 164 of the initial 307 (53.42%). In this regard, we must point out that journalistic criteria, consultations on

the part of citizens themselves with the different communication channels of fact-checking organisations and the relevance of dubious content are three fundamental reasons to explain a possible coincidence in the verification of dubious or false content.

Maldita.es is the organisation that disseminated the most unequivocal verifications (70 out of 149), nearly half of the non-repeated hoaxes were shared by this medium. On the other hand, *AFP* (*Comproven* and *Factual*) repeated proportionally more hoaxes denied by other fact-checkers –80%, although they only analysed a total of 10.

In this respect, the possible competition between the organisations *Newtral* and *Maldita.es* also stands out, as the two fact-checkers with the most resources and the most relevance, which means that more than half of their hoaxes are also duplicated –61.8% and 53.02% respectively.

On the other hand, the media with the fewest repeated verifications are *RTVE Verifica* (it only carries out two analyses, but both are unequivocal), and *EFE Verifica*, which, out of 55 news items, 61.82% had not been analysed or verified by other fact-checkers.

5.1.3. Clusters of hoaxes

There are 18 verification and explanatory articles written by fact-checkers that not only analyse a single piece of disinformation, but also compile a collection of several hoaxes on the same subject in a single publication. These are characterised by bringing together disinformation belonging to the same period or associated with the same individual, such as the attack on the Mariupol hospital or the alleged drug addiction of the Ukrainian president. In this respect, *Maldita.es* is the media outlet that has used this method most frequently, with a percentage of 66.66%. *EFE*, *Newtral* and *RTVE Verifica* carry out this action in isolation, while *APF* and *Verificat* do not use this narrative and discursive technique.

5.2. Characteristics of hoaxes analysed by fact-checkers

This section analyses the verification of individual unambiguous hoaxes, that is, hoaxes that are not repeated among the different fact-checking organisations, nor are several hoaxes grouped together in a single piece of information. The aim is to study the type of format, geographic distribution, subject matter, types, etc. This is why the sample is reduced in number, from 307 to 211 checks.

5.2.1. Analysis by type of format in which misinformation is shared

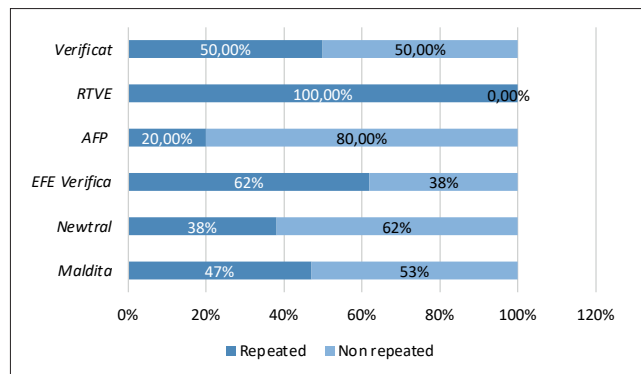
The combination of text with image and video is the most common among the verifications analysed, occurring 83.41% of the time. When there is only one format, the issuer or distributor of disinformation tends to share it through text (53%), followed by video (26%) and image (21%). The isolated use of audio is infrequent, occurring only once and not in combination with any other format.

In this sense, it is relevant that more than half of the text-only hoaxes (51.61%) correspond to the categories of fabricated content and explanatory content. In other words, almost all the fabricated content was transmitted through text –11 of the 13 total hoaxes⁶.

5.2.2. Frequency of occurrence of specific regions in the hoaxes

In most cases, the hoaxes do not refer to a single region, but to several. The country most affected by the disinformation, predictably, was Ukraine.

Russia is the second-most mentioned country in the checks, almost half of those referring to Ukraine. For its part, the Europe category not only groups together countries that specifically encompass the continent, but it has also been considered relevant to include those hoaxes that name the European Union as an institution. In this sense, part of the disinformation involves countries bordering Ukraine –Poland, Romania and Belarus– and EU or neighbouring countries –the United Kingdom, Serbia, the Netherlands, Italy, Germany, Finland and the Czech Republic. The North American presence is mostly justified by the United States, although to a lesser extent the presence of *NATO* member Canada, and Mexico can also be observed.



Graph 4. Percentage of hoaxes repeated by each fact-checker in the period of one year.

On the other hand, the media with the fewest repeated verifications are *RTVE Verifica* (it only carries out two analyses, but both are unequivocal), and *EFE Verifica*, which, out of 55 news items, 61.82% had not been analysed or verified by other fact-checkers.

On the other hand, the media with the fewest repeated verifications are *RTVE Verifica* (it only carries out two analyses, but both are unequivocal), and *EFE Verifica*, which, out of 55 news items, 61.82% had not been analysed or verified by other fact-checkers.

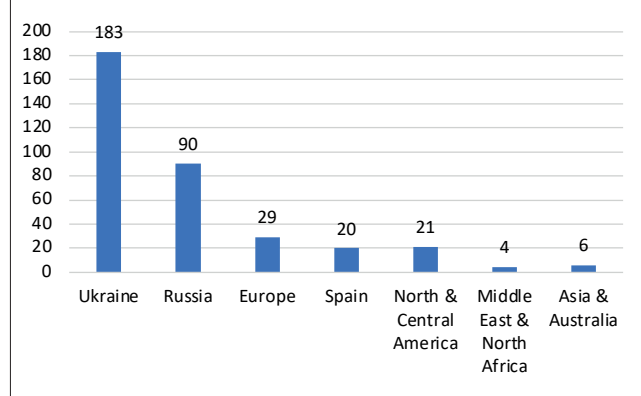
There are 18 verification and explanatory articles written by fact-checkers that not only analyse a single piece of disinformation, but also compile a collection of several hoaxes on the same subject in a single publication. These are characterised by bringing together disinformation belonging to the same period or associated with the same individual, such as the attack on the Mariupol hospital or the alleged drug addiction of the Ukrainian president. In this respect, *Maldita.es* is the media outlet that has used this method most frequently, with a percentage of 66.66%. *EFE*, *Newtral* and *RTVE Verifica* carry out this action in isolation, while *APF* and *Verificat* do not use this narrative and discursive technique.

This section analyses the verification of individual unambiguous hoaxes, that is, hoaxes that are not repeated among the different fact-checking organisations, nor are several hoaxes grouped together in a single piece of information. The aim is to study the type of format, geographic distribution, subject matter, types, etc. This is why the sample is reduced in number, from 307 to 211 checks.

The combination of text with image and video is the most common among the verifications analysed, occurring 83.41% of the time. When there is only one format, the issuer or distributor of disinformation tends to share it through text (53%), followed by video (26%) and image (21%). The isolated use of audio is infrequent, occurring only once and not in combination with any other format.

In this sense, it is relevant that more than half of the text-only hoaxes (51.61%) correspond to the categories of fabricated content and explanatory content. In other words, almost all the fabricated content was transmitted through text –11 of the 13 total hoaxes⁶.

In most cases, the hoaxes do not refer to a single region, but to several. The country most affected by the disinformation, predictably, was Ukraine.



Graph 5. Number of hoaxes referring to each geographical region.

Finally, it is necessary to highlight the number of hoaxes involving Spain as a country, up to twenty. In this regard, it should be remembered that the verifications have been published and distributed by Spanish fact-checkers, who in many cases receive requests for analysis from their fellow citizens.

It could be assumed that Spanish citizens are exposed to more misinformation about their country than about any other foreign nation. However, this figure does not exceed 10% of the total hoaxes analysed. This indicator could mean a greater importation of hoaxes and less creation of narratives developed specifically for Spain.

5.2.3. Frequency of occurrence of each topic in hoaxes

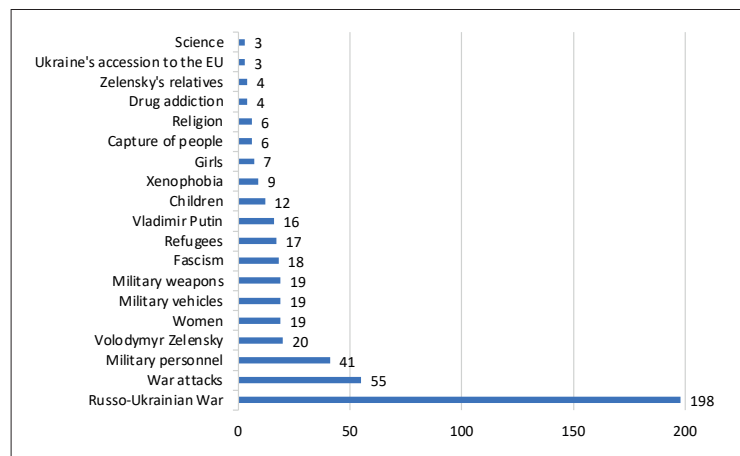
The analysis of the themes was carried out by means of twenty categories, which were selected based on the criterion of relevance to the research objectives. It was considered essential to include in the study some themes that did not have a result of more than 10% of the total, in order to complete the description qualitatively.

In accordance with the expected perception, the Russo-Ukrainian war is the theme that appeared most often in the total number of hoaxes, either directly or indirectly –contextually.

On the other hand, the Ukrainian president is the person most affected by disinformation. His Russian counterpart Vladimir Putin has 20% fewer verifications referring to him.

The Ukrainian president is accused of multiple issues: alleged drug addiction –20% of the hoaxes related to Zelensky deal with this issue–, sympathy with Nazi ideology, following beliefs such as Satanism or allegedly fleeing Ukraine.

In this vein, it is interesting to see how the narrative related to Zelensky's flight relies on different narrative strategies, such as the chroma-key recordings or a meeting that was supposed to take place in Kiev but allegedly took place in Poland. The same is true of narratives linking the Ukrainian president to cocaine use or his affinity with fascism, both of which are attributed to heterogeneous aspects. Some examples are: the fictitious presence of cocaine in his office through manipulated images, his theoretical statements about cocaine use on television through false captions, his photographs with manipulated T-shirts with swastikas, and so on.



Graph 6. Number of hoaxes referring to each subject.

On the other hand, the main characteristic of inaccurate stories involving the Russian president is that he is often not the direct protagonist. In other words, third parties refer to him and alleged statements of support for Putin are used. This is the case of tweets supposedly written by political representatives Pablo Echenique or Adrián Barbón.

On the other hand, and as can be seen in Graph 6, the number of verifications that point to women stands out, usually being presented in the dichotomy of victims or heroines. The victims are presented as vulnerable either because of their age –girls or elderly women– or because they are pregnant.

At the same time, a remarkable percentage of hoaxes about refugees stand out, ones which are characterised by content that may incite xenophobia (47%) either because people fleeing war are racially degraded when crossing the border into Poland, or because they are accused of false acts of uncivil behaviour in countries where they have been granted asylum.

On the occasions when reference is made to the capture of individuals, disinformation broadcasters spread rumours about arrests, arrest warrants or prisoners of war.

Finally, although there is not a great deal of repetition in the theme of Ukraine's accession to the European Union, it is essential to mention that there are as many as three narratives that attempt to link Ukraine's EU membership with its participation in the war.

5.2.4. Typology

As noted in the methodology, and based on the typology of **Wardle** and **Derakhshan** (2017), it should be noted that the most frequent type of misinformation is false context, in which text is combined with an image or video.

More than half of hoaxes follow this structure, whereby potential issuers prefer to start from what they consider to be "visual evidence", which accompanied by text would serve as a way of justifying the veracity of the claims.

In this regard, it should be pointed out that, through the classification of disinformation used, a hoax can go from being considered fabricated to being considered false context, only by incorporating a real, non-manipulated image or video.

The categories of impostor, manipulated, fabricated and explanatory content ranged from 6.2 to 10%. Specifically, and in reference to impostor content, 70% of the hoaxes grouped in this category have supplanted the identity of renowned foreign media outlets –CNN, *Al Jazeera*, *Time*, *BBC* and *Excelsior*. CNN was the most affected media brand.

From this perspective, and beyond the original typology presented by Wardle and Derakhshan (2017), it is essential to highlight the explanatory content. With this type of information, news organisations have found a method and a narrative format to contextualise disinformation that, due to its characteristics, cannot be disproved with evidence, but whose narratives are similar to certain hoaxes.

On the other hand, misleading content only constitutes 3.8% of all the verifications analysed, and it has been noted that it always attempts to negatively frame an individual, whether a natural person (50%) or a legal personality (50%) understood as a State or a political community of law such as the European Union.

Finally, it should be noted that the false connection is the second-least used type of disinformation –the last is satire–, and in all cases the issuer decided to incorporate false subtitles to spread disinformation about the presidents of different countries. The most disadvantaged representative when it came to modifying his original words in another language through text was the president of Russia, Vladimir Putin (60%). Volodymyr Zelensky and Joe Biden were also victims of the fake connection, but to a lesser extent.

5.2.5. Distribution channel

The distribution channel is no longer so relevant in the analysis of disinformation circulation because strategies are increasingly diversified. However, the preference for certain networks or platforms allows us to analyse the routes of access to disinformation.

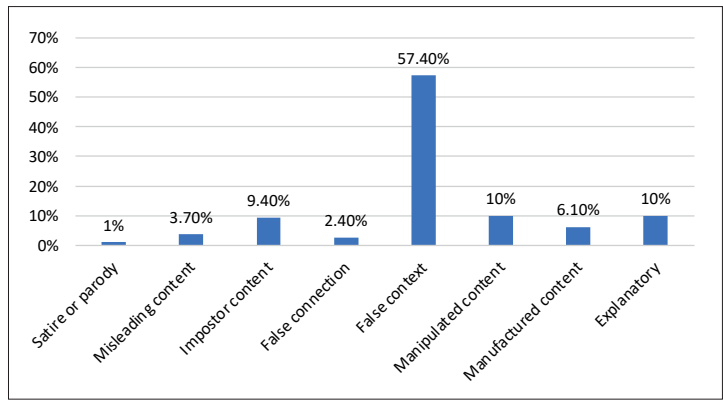
The distribution of hoaxes has been detected through three types of platforms: social networks, instant messaging applications and websites. The results shown in Graph 8 are consistent with the perception of distrust of Spanish citizens towards social networks (Amoedo et al., 2021). It should be noted that, in addition to the prominent role that TikTok has acquired, the Social networks category refers only to those contents in which the verifiers did not name a specific social network, but rather indicated it generically.

5.2.6. Results provided by fact-checking organisations

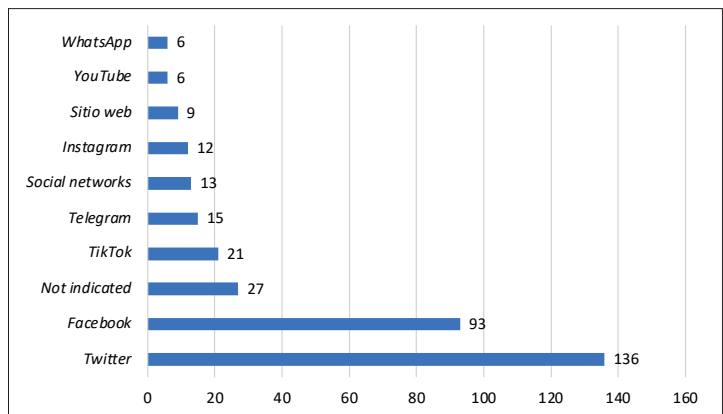
Finally, it should be noted that a part of our results coincides with those offered by the fact-checking organisations themselves on the anniversary of the first year of war in Ukraine.

One example is *Newtral*, which analysed up to 144 hoaxes about the war (*Newtral*, 2023). In March alone, *Newtral* disproved 53 of the 144 disinformation hoaxes about the invasion, which was the month in which they verified the most false content related to the conflict (Pereñeló, 2023). After March, April and May 2022, January 2023 was the month with the most verifications uploaded to the *EDMO* database.

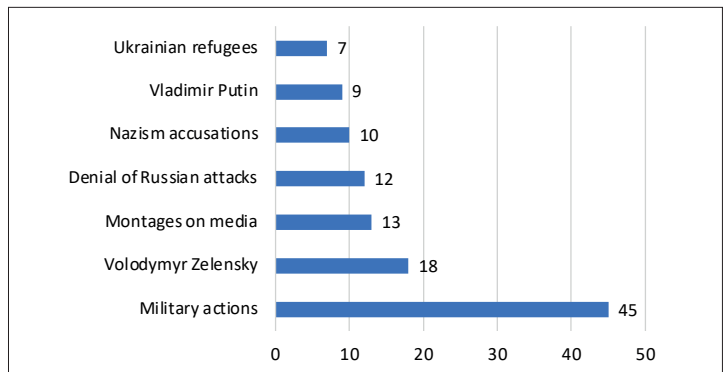
For its part, the *Ukraine Facts database* developed by the *IFCN* and led by *Maldita.es* had, after the first year of war, categorised up to 357 verifications in Spain –the same as in India, making it the country with the second highest number of total verifications after Ukraine (703).



Graph 7. Percentage of hoaxes according to type of disinformation.



Graph 8. Distribution channel of hoaxes according to the platform used.



Graph 9. Main disinformation narratives on Russia and Ukraine. Source: *Newtral*.

Of these, more than 170 were verified by the journalistic organisation *Maldita.es*. According to the organisation's data, the ten most frequently verified misinformation contain images and videos that were taken elsewhere on dates prior to the war, fake tweets, dissemination of stills from video games, or attempts to discredit real images through claims that they were staged (*Maldita.es*, 2023b).

“ The hoaxes related to President Zelensky are the most powerful narratives of the Russian disinformation strategy ”

6. Discussion and Conclusions

In relation to the first two hypotheses, it is confirmed that hoaxes related to President Zelensky are one of the most powerful narratives of the Russian disinformation strategy. Similarly, the data also indicate that the first weeks of the war accounted for a very significant part of the content verified by fact-checking organisations.

Also highlighted is how hoaxes are internationalised to adapt to local contexts depending on different news cycles and the involvement of the countries mentioned.

On the other hand, and as we have seen in previous research, the role of Spanish fact-checking organisations (mainly *Maldita.es* and *Newtral*) and their relevance in verifying dubious content is very relevant, even in a context of global disinformation such as the Russian invasion of Ukraine.

Their relevance is significant not only because of the number of verifications, but also because they often anticipate hoaxes that end up circulating in Latin American countries. From this perspective, we could point out that more hoaxes were identified in Spain than in other countries, but also that the role and prominence of these fact-checking organisations means that less content is left unverified than in other countries.

We also consider as very positive the attitude of the EU institutions towards open data on such an important issue as misinformation. The collaboration between fact-checking organisations, but also the fact that researchers can access data quickly, not only allows for a much better understanding of the phenomenon, its consequences, and limitations, but also for the development of more precise responses and strategies to identify both the narratives and their viralization cycles. As we have seen, the development of verification skills and the emergence of *OSINT* organisations that have played an important role in verifying images and videos of the Russian invasion of Ukraine are two new factors shaping the coverage of war journalism today.

The need to standardise the fields of data collection, verification processes and methodologies, and prior training, as well as to agree on formats and designs for how verified information is presented, is also apparent. In this sense, we can consider that the repetition of almost half of the hoaxes can be interpreted as a good sign, since a greater number of publications on the same narrative also implies a greater propagation of refutations. However, it is necessary to further analyse the effect of repeated messages on the reception of potential consumers who are informed by the different fact-checking organisations during the war processes.

In conclusion, and based on the results obtained, a series of recommendations are put forward to tackle disinformation and improve news coverage of today's wars:

1. The need for correspondents to be trained in image and video verification also becomes apparent, as they are the preferred witnesses of conflicts. Their role as intermediaries and journalistic brand for the media organisation they represent gives them a double credibility to tell what is happening but also to disprove possible misinformation.
2. From the point of view of the internal organisation of collaborative *fact-checking* projects, it is essential to establish beforehand the common criteria of the database to be shared by all organisations. In order to carry out subsequent analyses, it is essential that all media can carry out their documentation work with as much standardisation as possible.
3. In addition to recovering the importance of the distribution channel in internal categorisation, it is important to establish the time variable of verification. From this perspective, the date and time of identification of the possible misinformation as an internal category, as well as the date of publication of the verification, are information with a relevant subsequent use. An analysis of the time elapsed between the two periods can be crucial to establish early warning mechanisms.
4. On the other hand, indicating the specific social network or platform through which the manipulated information has gone viral can help to get a clearer idea of the target audience of the disinformation and the level of reliability of the platform itself in its detection mechanisms.
5. The differentiation between the potential perpetrator and the distributor of hoaxes through social networks and platforms by fact-checking organisations can provide a better understanding of the possible objectives and intentions behind disinformation and, also, identify the coordination of certain campaigns.
6. It would be advisable for fact-checking organisations to broaden their pool of experts to reduce a possible gender gap in the contextualisation and verification of conflict-related narratives.
7. We believe it is appropriate for fact-checking organisations that have an internal database to categorise verifications to explore new narratives that help to better understand the complexity of war conflicts and better explain the disinformation strategies deployed by the different actors involved.

8. Fact-checking organisations involved in collaborative initiatives should establish standardised dissemination metrics to distinguish new disinformation from already verified content that repeatedly reappears depending on the news cycle. The aim should be to reduce the viralization of already verified hoaxes as much as possible.

Also highlighted is how hoaxes are internationalised to adapt to local contexts depending on different news cycles and the involvement of the countries mentioned

9. It is recommended that verifications that affect global narratives or refer to different countries be published in English. The internationalisation of hoaxes is one of the most frequent characteristics of conflicts, crises, and global events.

10. It is essential to strengthen partnerships with fact-checking organisations in neighbouring countries. Such countries, in addition to being much more directly affected by the consequences of war, are limited by the fact that global monitoring of disinformation in more minority languages is much smaller.

7. Notes

1. The full *EDMO* database can be found at this link:

<https://edmo.eu/war-in-ukraine-the-fact-checked-disinformation-detected-in-the-eu>

2. In January 2023, Russia was home to 106 million active users of social networks and Ukraine 26.7 million. According to data collected in February 2023, the most used social networks in Russia are *VK* (75.3%); *WhatsApp* (71.5%); *Telegram* (64.4%); *Odnoklassniki* (43.5%); *TikTok* (42.6%); *Viber* (34.7%); *Instagram* (24%); *Pinterest* (13%); *Skype* (11.7%); *Discord* (8.2%); *Facebook* (7%) and *Twitter* (5.7%), among others (**Kemp**, 2023). While the most popular social platforms in Ukraine are *Facebook* (42.67%); *Twitter* (13.39%); *Instagram* (13.29%); and *YouTube* (8.54%) (*Statcounter*, 2023).

3. The possible bias of the selected sample is determined by the verifications uploaded to *EDMO* by the journalistic organisations themselves. In any case, we considered it appropriate to use this database because it allowed us to compare, through a common categorisation, the verification dynamics among Spanish organisations.

4. However, it was decided to count clusters when the unit of analysis was defined as newspaper articles disseminated by any fact-checking organisation in the *EDMO* network. This is the case for the frequency of hoaxes repeated among themselves, which refers not only to articles that show sets of hoaxes, but also to those that also expose them individually.

5. This may be because the last month of the summer was marked by a swift Ukrainian counter-offensive towards the Russian side around Kharkov.

6. This trend can be linked to the fact that it is easier to produce information from scratch using text than generating an image or a video, which not only requires more effort, but also specific tools.

8. References

AlSayyad, Nezar; Guvenc, Muna (2015). "Virtual uprisings: on the interaction of new social media, traditional media coverage and urban space during the 'Arab Spring'". *Urban studies*, v. 52, n. 11, pp. 2018–2034.

<https://doi.org/10.1177/0042098013505881>

Amoedo-Casais, Avelino; Vara-Miguel, Alfonso; Negrodo-Bruna, Samuel; Moreno-Moreno, Elsa; Kaufmann-Argueta, Jürg (2021). *Digital News Report España 2021*. Pamplona: Servicio de Publicaciones de la Universidad de Navarra.

<https://doi.org/10.15581/019.003>

Aral, Sinan (2023). "The changing face of Russia's information war against Ukraine and other democratic countries: lessons and recommendations. Interview with Professor Sinan Aral". *Security and defence quarterly*, v. 41, n. 1, pp. 88-96.

<https://doi.org/10.35467/sdq/156264>

Bruns, Axel (2003). "Gatewatching, not gatekeeping: collaborative online news". *Media international Australia*, v. 107, n. 1, pp. 31-44.

<https://doi.org/10.1177/1329878X0310700106>

Cabada, Ladislav (2022). "Russian aggression against Ukraine as the accelerator in the systemic struggle against disinformation in Czechia". *Applied cybersecurity & internet governance*, v. 1, n. 1, pp. 178-193.

<https://doi.org/10.5604/01.3001.0016.0916>

Caballo-Méndez, Daniel; De-Santiago-Mateos, Miguel-Ángel (2021). "Fotoperiodismo y 'fake news'. Orígenes y tipos de manipulación de la imagen: el caso de la Agencia EFE". *Cuadernos de documentación multimedia*, v. 32, e80836.

<https://doi.org/10.5209/cdmu.80836>

Carmack, Dustin (2022). "US must implement lessons on 'hybrid' conflict from Ukraine war". *The heritage foundation*, n. 3704.

<https://www.heritage.org/node/25023651/print-display>

Carreras-Álvarez, María-del-Valle; Román-San-Miguel, Aránzazu (2011). "Redes sociales y herramientas 2.0 en la revuelta de Egipto. El poder de derribar a un régimen". En: Verón-Lassa, José-Juan; Sabés-Turmo, Fernando (eds.). *La investigación en periodismo digital. Algunos trabajos desde el ámbito universitario*. Zaragoza: Asociación de la Prensa de Aragón, pp. 336-349. ISBN: 978 8 487 17544 2

- EDMO (2022). *EDMO 2022 Annual conference*. EDMO.
https://edmo.eu/wp-content/uploads/2022/09/2022-AC_Report_updated-Jan-2023_PDF.pdf
- EDMO (2023). *EDMO Hubs*. EDMO.
<https://edmo.eu/edmo-hubs>
- EEAS (2021). *Disinformation narratives about the Eastern partnership*. EEAS, December.
https://www.eeas.europa.eu/sites/default/files/eeas_-_factsheet_eap_en_v10.pdf
- EEAS (2023). *2022 Report on EEAS activities to counter FIMI*. EEAS, 2 February.
https://www.eeas.europa.eu/sites/default/files/documents/EEAS-AnnualReport-WEB_v3.4.pdf
- EU Disinfo Lab (2022). *Narratives and hoaxes. Ukraine war resource hub*.
<https://www.disinfo.eu/ukraine-hub>
- Fernández-Castrillo, Carolina** (2014). "Prácticas transmedia en la era del prosumidor: hacia una definición del contenido generado por el usuario (CGU)". *CIC. Cuadernos de información y comunicación*, v. 19, pp. 53-67.
https://doi.org/10.5209/rev_CIYC.2014.v19.43903
- Fernández-Castrillo, Carolina; Ramos-Vera, Celia** (2023a). "El fotoperiodismo en la era de Telegram: la invasión Rusia de Ucrania en Rossiyskaya Gazeta y Ukrayinska Pravda". En: Bonales-Daimiel, Gema; Sierra-Sánchez, Javier (eds.). *Desafíos y retos de las redes sociales en el ecosistema de la comunicación*, v. 1. Madrid: McGraw-Hill/Interamericana de España, pp. 451-468.
- Fernández-Castrillo, Carolina; Ramos-Vera, Celia** (2023b). "Social web and photojournalism: user-generated content of the Russo-Ukrainian war". [Web social y fotoperiodismo: Contenido generado por el usuario en la guerra ruso-ucrainiana]. *Comunicar*, n. 77, pp. 85-96.
<https://doi.org/10.3916/C77-2023-07>
- Himma-Kadakas, Marju; Ojamets, Indrek** (2022). "Debunking false information: investigating journalists' fact-checking skills". *Digital journalism*, v. 10, n. 5, pp. 866-887.
<https://doi.org/10.1080/21670811.2022.2043173>
- Ibitowa, Philippe** (2022). "Information warfare, propaganda and the manipulation of opinion in times of conflict: the case of the Russia-Ukraine war". *International journal for multidisciplinary research*, v. 4, n. 6.
<https://doi.org/10.36948/ijfmr.2022.v04i06.1121>
- Karpchuk, Nataliia; Yuskiv, Bohdan** (2021). "Dominating concepts of Russian Federation propaganda against Ukraine (Content and collocation analyses of Russia Today)". *Politologija*, v. 102, n. 2, pp. 116-152.
<https://doi.org/10.15388/Polit.2021.102.4>
- Kemp, Simon** (2023). "Digital 2023: the Russian Federation". *Datareportal*.
<https://datareportal.com/reports/digital-2023-russian-federation>
- Lazarenko, Valeria** (2019). "Conflict in Ukraine: multiplicity of narratives about the war and displacement". *European politics and society*, v. 20, n. 5, pp. 550-566.
<https://doi.org/10.1080/23745118.2018.1552108>
- Loo-Vázquez, José-Roberto; Gámez-Paz, Daniela; Lamarque-Vega, Evelyn; Domínguez-Quevedo, Lilieth; Haber-Guerra, Yamile** (2016). "Del gatekeeper al content curator: cambiar algo para que no cambie nada". *Comunicología cubana*, v. 20, n. 92, pp. 1149-1171.
<https://www.revistarazonypalabra.org/index.php/ryp/article/view/293>
- Maldita.es* (2023a). "Verificar una guerra cuando las bombas caen cerca de casa: cómo los 'fact-checkers' ucranianos se enfrentan a los bulos de la invasión rusa". *Maldita.es*, 24 febrero.
<https://maldita.es/malditobulo/20230224/bulos-invasion-ucrania-stopfake>
- Maldita.es* (2023b). "#UkraineFacts: Las 10 desinformaciones más veces desmentidas por los verificadores un año después del inicio de la guerra". *Maldita.es*, 23 febrero.
<https://maldita.es/malditobulo/20230223/ukrainefacts-10-desinformaciones-mas-desmentidas>
- Martino, John** (2022). "Z: perception as weaponry in the Russian/Ukraine conflict". In: Morelock, Jeremiah; Ziotti-Narita, Felipe (eds.). *The return of history. The Russo-Ukrainian war and the global spectacle*. Ribeirão Preto: Centro Universitário Barão de Mauá, pp. 125-146.
<https://doi.org/10.56344/lv2-c5>
- Micó, Josep-Lluís; Canavilhas, João; Masip, Pere; Ruiz, Carles** (2008). "La ética en el ejercicio del periodismo: credibilidad y autorregulación en la era del periodismo en internet". *Estudos em comunicação*, v. 4, pp. 15-39.
<http://www.ec.ubi.pt/ec/04/pdf/02-Jose-Lluís-Mico-La-etica-en-el-ejercicio-del-periodismo.pdf>

- Milmo, Dan** (2022). "Russia blocks access to Facebook and Twitter". *The Guardian*, 4 March.
<https://www.theguardian.com/world/2022/mar/04/russia-completely-blocks-access-to-facebook-and-twitter>
- Molina-Cañabate, Juan-Pedro; Magallón-Rosa, Raúl** (2023). "Tipologías de desinformación y autoridad percibida en las narrativas sobre el origen del coronavirus en España". *RAE-IC, Revista de la Asociación Española de Investigación de la Comunicación*, v. 10, n. 19, pp. 184-208.
<https://doi.org/10.24137/raeic.10.19.9>
- Newtral** (2023). *Un año de la invasión rusa en Ucrania: datos, análisis y desinformación*. Newtral, 24 febrero.
<https://www.newtral.es/invasion-rusa-ucrania-especial/20230224>
- Ortiz-Galindo, Rocío** (2016). "Social cybermovements: a review of the concept and theoretical framework". *Communication & society*, v. 29, n. 4, pp. 165-182.
<https://doi.org/10.15581/003.29.4.sp.165-183>
- Perelló, Blanca** (2023). "Un año de guerra y más de 100 bulos sobre Rusia y Ucrania". *Newtral*, 23 febrero.
<https://www.newtral.es/bulos-guerra-rusia-ucrania/20230223>
- Petrocchi, Maurizio** (2022). "The use of psychology as strategy in hybrid warfare". *The international scientific journal of humanities*, v. 1, n. 1.
<https://doi.org/10.55804/TSU-ti-1/Petrocchi>
- Rodríguez-Tranche, Rafael** (2019). *La máscara sobre la realidad. La información en la era digital*. Alianza.
- Sacco, Vittoria; Bossio, Diana** (2015). "Using social media in the news reportage of war & conflict: opportunities and challenges". *The journal of media innovations*, v. 2, n. 1, pp. 59-76.
<https://doi.org/10.5617/jmi.v2i1.898>
- Sánchez-Duarte, José-Manuel; Magallón-Rosa, Raúl** (2020). "Infodemia y COVID- 19. Evolución y viralización de informaciones falsas en España". *Revista española de comunicación en salud*, suplemento 1, pp. 31-41.
<https://doi.org/10.20318/recs.2020.5417>
- Sánchez-Gonzales, Hada; Martos-Moreno, Javier** (2020). "Telegram como herramienta para periodistas: percepción y uso". *Revista de comunicación*, v. 19, n. 2, pp. 245-261.
<https://doi.org/10.26441/RC19.2-2020-A14>
- Soengas-Pérez, Xosé** (2013). "El papel de Internet y de las redes sociales en las revueltas árabes: una alternativa a la censura de la prensa oficial". *Comunicar*, v. 41, n. 21, pp. 147-155.
<https://doi.org/10.3916/C41-2013-14>
- Soengas-Pérez, Xosé; Assif, Mohamed** (2017). "Cyberactivism in the process of political and social change in Arab Countries". *Comunicar*, v. 53, pp. 49-57.
<https://doi.org/10.3916/C53-2017-05>
- Stancu, Mihai-Cristian** (2019). "Information war. Case study: the role of Russia Today for coverage of the mission conducted by Russia in East Ukraine". *Bulletin of Carol I National Defence University*, v. 8, n. 4, pp. 45-53.
<https://doi.org/10.12753/2284-9378-19-07>
- Statcounter** (2023). *Social media stats in Ukraine - February*. Statcounter global stats.
<https://gs.statcounter.com/social-media-stats/all/ukraine>
- Susska, Olga; Chernii, Liudmyla; Sukharevska, Hanna** (2022). "Media as a tool of manipulative technology of Russian infoaggression in the Ukrainian media space". *Journal of interdisciplinary research*, v. 12, n. 01-25, pp. 228-233.
http://www.magnanimitas.cz/ADALTA/120125/papers/A_41.pdf
- Vorster, Oli-Rose** (2022). "The soviet information machine: The USSR's influence on modern Russian media practices & disinformation campaigns". *London School of Economics, Undergraduate Political Review*, v. 4, n. 2, pp. 106-112.
<https://upr.lse.ac.uk/articles/abstract/49>
- Wardle, Claire; Derakhshan, Hossein** (2017). *Information disorder: toward an interdisciplinary framework for research and policy-making*. Strasbourg: Council of Europe.
<https://edoc.coe.int/en/media/7495-information-disorder-toward-an-interdisciplinary-framework-for-research-and-policy-making.html>
- Yarova, Alla** (2023). "Thematic patterns of Russian disinformation". *Baltic journal of legal and social sciences*, v. 4, pp. 158-165.
<https://doi.org/10.30525/2592-8813-2022-4-19>
- Zawadzki, Tomasz** (2022). "Examples of Russian information war activity at the beginning of Ukrainian crisis". *International conference: Knowledge-based organization*, v. 28, n. 1, pp. 146-150.
<https://doi.org/10.2478/kbo-2022-0023>

Teachers in a world of information: Detecting false information

Kamil Kopecký; Dominik Voráč; René Szotkowski; Veronika Krejčí; Karolína Mackenzie; Magdalena Ramos-Navas-Parejo

Nota: Este artículo se puede leer en español en:

<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87136>

Recommended citation:

Kopecký, Kamil; Voráč, Dominik; Szotkowski, René; Krejčí, Veronika; Mackenzie, Karolína; Ramos-Navas-Parejo, Magdalena (2023). "Teachers in a world of information: Detecting false information". *Profesional de la información*, v. 32, n. 5, e320501.

<https://doi.org/10.3145/epi.2023.sep.01>

Manuscript received on 12th October 2022

Accepted on 29th March 2023



Kamil Kopecký

<https://orcid.org/0000-0003-2860-946X>

Univerzita Palackého v Olomouci
Faculty of Education
Žižkovo nám. 5
77140 Olomouc, Czech Republic
kamil.kopecky@upol.cz



Dominik Voráč

<https://orcid.org/0000-0002-3970-0130>

Univerzita Palackého v Olomouci
Faculty of Education
Žižkovo nám. 5
77140 Olomouc, Czech Republic
dominik.vorac01@upol.cz



René Szotkowski

<https://orcid.org/0000-0002-1261-6481>

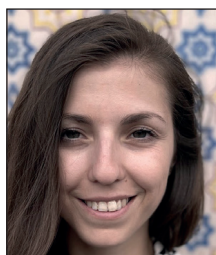
Univerzita Palackého v Olomouci
Faculty of Education
Žižkovo nám. 5
77140 Olomouc, Czech Republic
rene.szotkowski@upol.cz



Veronika Krejčí

<https://orcid.org/0000-0002-2130-7613>

Univerzita Palackého v Olomouci
Faculty of Education
Žižkovo nám. 5
77140 Olomouc, Czech Republic
veronika.krejci@upol.cz



Karolína Mackenzie

<https://orcid.org/0000-0002-9652-3230>

Univerzita Palackého v Olomouci
Faculty of Education
Žižkovo nám. 5
77140 Olomouc, Czech Republic
karolina.mackenzie01@upol.cz



Magdalena Ramos-Navas-Parejo ✉

<https://orcid.org/0000-0001-9477-6325>

Universidad de Granada
Facultad de Ciencias de la Educación y el
Deporte de Melilla
Ctra. Alfonso XIII, s/n
52005 Melilla, Spain
magdalena@ugr.es

Abstract

The media have advanced rapidly in recent years, bringing with them new challenges, such as the need to acquire adequate training to be successful in today's global world. Since anyone can participate in public media communication, misinformation is an element that shapes today's society. This article focuses on the extent to which Czech teachers believe fake news and conspiracy theories. We conducted an online survey with a total of 2,155 teachers from all regions of the Czech Republic. The teachers assessed a total of 34 statements divided into three thematic groups: statements related to the European Union, statements related to the Covid-19 disease pandemic and statements focusing on well-known historical, cultural and social phenomena. More than half (61.14%) of Czech teachers were able to correctly judge the truthfulness of statements. Nevertheless, 15.59% of Czech teachers admit that they are not able to assess the statements in a relevant way, and only 10.41% believed some of the fake news or conspiracy theories.

Keywords

Media literacy; Disinformation; Misinformation; Hoaxes; Fake news; Conspiracy theories; Fact-checking; False information; Teachers; Educators; Czech Republic; Surveys; Primary schools; Secondary schools; European Union; Covid-19.



Funding

Funded by O2 Czech Republic (contract research).

1. Introduction

The twenty-first century is often referred to in public as the information age or the post-factual age. There is a reason for these labels –with the development of the media and the advent of the Internet, the amount of information that people are surrounded by has increased many times over, and naturally, with the amount of information, the amount of information that is partially or completely false has also increased. As early as 1964, debates raged about whether there was an ‘information explosion’ (Havrilesky, 1972) but the volume of information circulating is also increasing on the internet itself –in 2000 there were more than 17 million websites available on the internet, in 2018 the number of websites was more than 1.6 billion. However, more information does not make people make better decisions, but can negatively influence them (Zheng *et al.*, 2020). An example from around the world is the experience of the COVID-19 disease pandemic, where at least several hundred people died due to misinformation (Islam *et al.*, 2020). Thus, in addition to the pandemic itself, the *World Health Organization* has focused on the so-called ‘infodemic’ manifested by the large amount of fake news about COVID-19 in order to prevent unnecessary loss of life caused by misinformation (WHO, n.d.).

The large amount of information increases the demands on the teachers in terms of information and media literacy. In the Czech Republic, the case of a primary school teacher who, at the time of the Russian invasion of Ukraine, told her pupils, for example, that the capital Kiev was not facing any attacks and that they should not trust the public *Czech Television* because it belongs to a media group tied to the billionaire Soros has been widely discussed (Ciroková, 2022). The potential problems of critical reception of media messages were highlighted by research from Slovakia, which is a neighbouring country of the Czech Republic and until 1993 formed one state with the Czech Republic. For example, research entitled *Basic skills for the Hoax century from the perspective of teachers: claim vs. reality* by Focus Agency and the *Centre for Environmental and Ethical Education Pyskyřice*, conducted in 2019 with the support of the *British Council*, produced interesting results (Goda, 2019): up to 57% of teachers out of 523 respondents would recommend one of the Slovak conspiracy websites as a verified source of information to their students. This could be due, for example, to the fact that teachers are not regular readers of all the media mentioned in the survey, ignored the “I don’t know” option and recommended even those media that they do not actually know.

Other interesting surveys on teachers and misinformation include the *Focus Agency* survey (Modrako, 2020) commissioned by the Slovak political party *Spolu* and the *European People’s Party (EPP)* in 2020 among second grade primary school teachers in Slovakia, which showed that some Slovak teachers are subject to misinformation. For example, of a sample of 518 Slovak teachers, almost a third said that the vaccination against COVID-19 was only a preparation for chipping people, or that COVID-19 was just a “fluke”. Already in this research it was clear that teachers –like the rest of the population– are influenced by misinformation and disinformation.

2. Methodology**2.1. Research identification**

The research *Czech Teacher in the World of Media* was conducted by the *Centre for Prevention of Risky Virtual Communication* of the *Faculty of Education at Palacký University in Olomouc* and *O2 Czech Republic*. The research was funded by *O2 Czech Republic* and *The Media Literacy of Teachers and Students as Key Issues of Education in the 21st Century* project of the *Faculty of Education at Palacký University in Olomouc*, and a research group from the *University of Granada*, with which researchers from the Czech Republic have long cooperated, also collaborated on partial analyses.

The research included four thematic parts, in this text we focus only on the part dedicated to the perception and critical assessment of false information in the online environment with the working title *Czech teacher in the world of information*.

2.2. Research procedure

The research is primarily quantitative (with some qualitative elements). The basic research instrument was an anonymous online questionnaire distributed to all regions of the Czech Republic (using the Czech Republic’s own database of schools). The data collection took place from 1 February 2021 to 15 June 2021. In the following weeks, the evaluation of the partial results and their interpretation took place. Selected items were also subjected to relational analysis using t-tests. *Statistica* statistical software was used for detailed data analysis.

The research included four thematic parts, in this text we focus only on the part dedicated to the perception and critical assessment of false information in the online environment with the working title *Czech teacher in the world of information*.

2.3. Population and sample

A total of 2155 teachers from all over the Czech Republic aged 21-82 participated in the research. The average age was 47.21 years (median 47, mode 56). 77.82% were women and 22.18% were men. The characteristics of the research population (age, education, gender) correspond to the official statistics of the *Ministry of Education, Youth and Sports* of the Czech Republic.

Most of the teachers who participated in the research have a pedagogical education at the master's level (78.10%), university education without pedagogical education with additional pedagogical studies was confirmed by 8.72% of the respondents.

64.17% of the respondents were primary school teachers (22.69% were primary school teachers of the first level, 41.48% were primary school teachers of the second level), 24.73% were secondary school teachers. Most of the respondents were from the Moravian-Silesian Region (13.64%), Olomouc Region (10.72%), the Capital City of Prague (9.33%) and the South Moravian Region (8.68%).

The normality of the distribution of the research sample was tested based on the age of the respondents. According to the presented histogram of the monitored variable supplemented with the Shapiro-Wilk test for testing normality $SW-W = 0.9826$; $p = 0.0000$ indicates a normal distribution of the research sample.

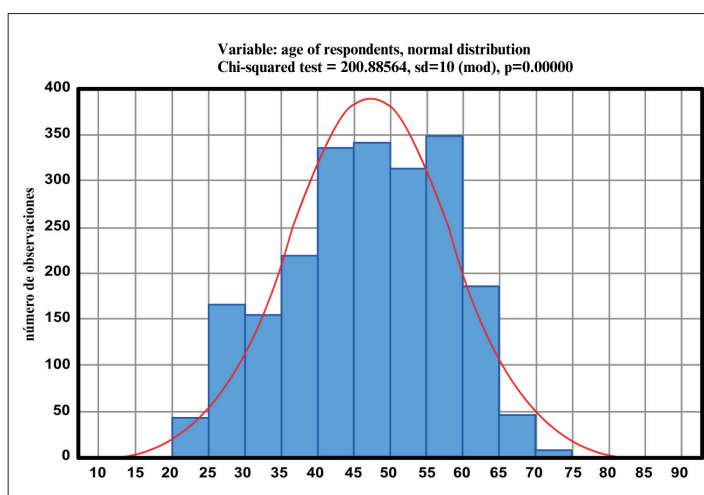


Chart 1. Characteristics of the research population.

2.4. Research tool

A special questionnaire was developed and divided into 5 parts:

The initial part (7 questions) was devoted to demographic data (gender, age, region, educational attainment) and information related to the teacher's experience (what type of school he/she teaches at, what subjects he/she teaches, how long he/she has been teaching, etc.).

The second part (13 questions) was devoted to teachers' active use of media (newspapers, magazines, television, internet portals, etc.). In this part we asked which media teachers regularly follow, which ones they consider trustworthy, which ones they would recommend to their students, whether they can indicate which media belong to the public media, etc.

The third part (3 thematic groups of questions) of the questionnaire focused on the recognition of different types of true and false information. This part was divided into 3 areas,

- information about the European Union,
- information related to the COVID-19 pandemic and
- information related to politics, technology and history.

This part of the questionnaire included a total of 45 statements (some of which were then used in the calculation of the final results).

All statements included in this section were publicly known and widespread, the battery of questions contained both true and false (disinformation, conspiracy) information. Teachers were given the opportunity to indicate whether they agreed, disagreed, or did not know the statement (a scale was used to measure this). In the following sections of this text we work with the results of this part.

The fourth part included 3 questions focused on the distinction between public and private media, the recognition of verified and unverified profiles on social networks and spam.

The fifth part was aimed at teachers in relation to media education and its teaching in school settings and contained a total of 14 items.

The validity and reliability of the questions included in the questionnaire were verified with a test group of teachers (21 persons) and subsequently corrected to make the questions unambiguous and fully understandable. The questionnaire was then extended to all schools in the Czech Republic through its own online system *Researcher 1.0*.

2.5. Research objectives

Several research objectives/questions were formulated in the framework of our research, which were further operationalized. Specifically, we pursued the following:

1. What type of media do teachers actively use (in their personal lives, in teaching, etc.)?
2. Which media do teachers find credible?
3. Whether they can name which media belong to the so-called public media (and distinguish them from e.g. commercial media).

4. How is media education implemented in their school environment?
5. Which media do they actively use in their teaching?
6. Are Czech teachers able to distinguish true news from false news (disinformation, misinformation, conspiracy)?

In this text, we focus only on research objective number 6 –teachers' ability to distinguish which information in the public space is true and verified from information that is false and that can be considered misinformation or misinformation. The teachers involved in the research were given a didactic test containing a battery of statements, which were then assessed by the teachers with regard to their truthfulness. The results were then evaluated and converted into a final score. The battery of questions and the teachers' success in detecting truthfulness is presented in the following sections of this text. Thus, the primary research questions for this part of the study were:

1. Are Czech school teachers able to distinguish true from false information?
2. What is their success rate in recognizing true and false statements?
3. How many teachers admit that they are unable to decide whether a statement is true or false at a given moment?
4. Do teachers know the technical ways of tagging official profiles in social networking environments?

3. Results

In our research, we focused on whether Czech teachers are able to assess the truthfulness of selected statements that include both true and verified information and false content (including various forms of hoaxes, misinformation, conspiracies, etc.). Teachers were given the opportunity to indicate whether they considered the statement to be true or false, whether they were unsure of the answer and simply did not know, or whether they were unable to critically evaluate the information (e.g. they did not have enough information about it, did not understand the topic, etc.).

Teachers were presented with a battery of 34 statements that were divided into 3 main topic areas:

- A. Information and misinformation about the European Union,
- B. Information and misinformation about the COVID-19 pandemic
- C. Other information and misinformation from the world.

The assessed statements included highly frequent disinformation, but also satirical news mapped by both domestic and foreign fact-checking servers (*Snopes, Manipulators, Hoax.cz, etc.*), which have significantly affected social networks in the Czech Republic (especially *Facebook*) in the last 5 years. A large part of the statements has been refuted by official governmental and supranational institutions, in the case of COVID-19 e.g. the *Ministry of Health* of the Czech Republic or *WHO*, in the case of misinformation about the European Union directly by the EU institutions (detailed analyses of the statements are freely available on the websites of these organisations)¹.

Each thematic area was evaluated separately. It was mapped how many teachers identified the statement correctly, how many incorrectly, but also how many teachers admitted that they were not able to critically evaluate and assess the statement.

3.1. True and false information about the European Union

The topic of the European Union is a very sensitive one in the Czech Republic and trust in the European Union in the Czech Republic fluctuates quite a lot (according to *Eurobarometer* data, it was only 39% in 2019-2020, rising to 48% in 2021) (*Evropská komise, 2021*). The Czech Republic's membership in the European Union is then associated with a lot of false or distorted information, hoaxes and misinformation that are disseminated through the internet and which are gradually being denied by the EU representation in the Czech Republic. And it is from these media disseminated reports that we selected a set of 9 statements which were then assessed by the respondents of the research –teachers. A large part of the statements in our research are also regularly explained by the *European Commission* Representation in the Czech Republic through specialised websites (*Evropská komise, 2023*).

1) The European Union wants to ban the letter “ř”

A specific feature of the Czech language is the letter and vowel “ř”, which is not used in any other European country. However, the news that the European Union wants to ban this specific letter from the Czechs –precisely because it is not used anywhere else and is therefore unnecessary– has spread through the internet. However, this is a satirical report, not true information (**Nutil, 2016**) This report has nevertheless managed to create a wave of hatred against the European Union in the online environment.

The message was correctly identified by 80.79% of the teachers surveyed and was considered untrue. 9.33% stated that they did not know and could not judge the statement for its truthfulness. Only 2% of the teachers identified the report as true information.

2) The European Union has banned us from producing rum

This statement belongs to the typical hoaxes and misinformation that appear in the Czech media space in connection with the European Union regulations. The European Union regulates only the name –the word rum should be used to describe a drink made from sugar cane distillate (Czech rum was made from alcohol and rum essence). After joining the European Union, the product started to be referred to as Tuzemský or Tuzemák. Another part of this myth is the issue of flavourings used in the production of rum, which contain carcinogenic substances, to which the European Union has drawn attention –here the Czech Republic has been granted an exemption. Czech Tuzemák has therefore not disappeared from the market (*Evropská komise, 2023*).

64.5% of the teachers correctly assessed the statement about the ban on rum production, while 10.86% of the teachers considered this statement to be true. 10.02% of the teachers said that they could not judge the statement because they did not have enough information about it.

3) The European Union has decreed that no curved bananas may be sold

The European Union does not regulate the curvature of bananas, it only regulates what parameters the goods should have in each quality class. 76.47% of the teachers assessed this statement correctly, 11% said they could not assess the statement. Only 2.88% of the teachers were incorrect in confirming the truth of the statement.

4) Following an European Union's decision, products labelled as "spreadable butter" had to be renamed

This is true information, again related to the regulation of names –the European Union has never banned the production of spreadable butter but has demanded the unification of names– the term butter can be used on the common European market to describe products that contain 80% milk fat, which Czech spreadable butter did not meet.

56.29% of teachers correctly identified the statement as true. 21.67% rated the statement as false (the high error rate may be due to respondents confusing the regulation of the name of the product with its prohibition), and 8% said they could not judge the statement.

5) A European Union regulation required that doughnuts be packaged in microtene bags to 'eliminate the risk of contamination and to preserve their health and quality during the period of sale'

This is one of the myths associated with the European Union –the well-known "doughnut decree" was created by the *Ministry of Health* of the Czech Republic in 2002 and was not based on any European regulations, its validity ended a few years later. However, it still survives as a myth.

The statement was correctly assessed by 54.29% of teachers and evaluated as false. 10.95% of teachers then rated the myth as true. 12.85% could not decide and 16.75% said they could not judge the statement. The increased number of respondents who stated that they could not judge the statement indicates that they were not familiar with it.

6) As a result of the Czech Republic's accession to the EU, legislation has been introduced that de facto bans classic Czech slaughterhouses

This is again an untrue statement, the organisation of domestic slaughterhouses is regulated directly by the laws of the Czech Republic and the accession to the EU and subsequent membership has not led to any change in the legal regulation of domestic slaughter of pigs (*Zastoupení Evropské komise v ČR, 2016b*).

51.04% of respondents correctly identified the statement as untrue, 13.09% as true (i.e. incorrect). 18.61% again stated that they could not assess the statement (this is probably again a less widespread myth).

7) The European Union has ordered the phasing out of conventional incandescent light bulbs from the market

Indeed, the European Union has ordered the gradual withdrawal of conventional incandescent bulbs from the market, following an agreement between national government representatives and members of the *European Parliament* (including those from the Czech Republic) (*Zastoupení Evropské komise v ČR, 2016a*). 53.13% of teachers correctly identified the statement as true. 14.76% were wrong and identified it as false, 14% said they could not judge the statement.

8) The refugee quotas introduced by the European Union oblige us to accept a certain number of refugees from war-affected countries (Syria, Eritrea, Iraq)

It is a more complex statement and a more complex statement: although the refugee quotas oblige us to accept a certain number of refugees, the Czech Republic has refused to accept these quotas and has not respected this regulation. Thus, the statement can be seen from different angles –on the one hand, there is the EU's commitment, but on the other hand, there is the Czech Government's refusal to be bound. This has also led to different ways of evaluating the statement.

36.06% of the teachers evaluated the statement as true, 33.78% as false, 14.66% could not decide and 11.09% could not evaluate the statement. This statement and its evaluation are not included in the final score.

9) The Czech Republic gives more to the EU than it receives from it

The Czech Republic is a net beneficiary –it is one of the Member States that receive significantly more from the EU budget than it puts into it (*Irozhlas, 2021*) This myth is often associated with other issues such as the diversion of funds out of the Czech Republic (e.g. in the case of corporate profits) or EU subsidy fraud– but this is not primarily a problem of the EU, but of the Czech Republic's own legislation and its own subsidy recipients (and their potential fraud).

59.58% of teachers correctly identified the statement as false, only 8.26% agreed with the statement. Again, 11.09% stated that they could not assess the statement.

Interesting results are also provided by the Student's T-Test, in which we looked at whether there were differences in the responses of women and men. For example, in the case of the statement that the Czech Republic gives more to the EU than it receives from it, men performed significantly better ($\alpha = 0,05$) than women. See the following graph for more details.

3.2. Overall results (educators' assessment of the truth of claims)

On the basis of the analysis of the above statements about the EU, it can be said that 62% of the respondents from among Czech teachers can correctly determine their truth (average values of 8 statements). 10.57% of the respondent teachers determined the truth of the statements incorrectly (average values of 8 statements).

Given that 13.02% of teachers indicated that they could not judge the statement at this time, it is possible that the final score could improve but also worsen. This is because one of the characteristics of media literacy is the ability to admit that I do not have enough relevant information at a given moment to judge a given statement.

3.3. Teachers and the coronavirus pandemic

The COVID-19 pandemic has swept the global media space and brought a great deal of information and misinformation related to the origin of the disease, its treatment, testing, preventive measures, vaccination, etc. That is why we focused the second thematic area on this topic and let the educators evaluate a total of 19 statements.

In the analysis, we chose a variety of statements that comment on the characteristics of the virus (COVID-19 is just a common flu), its origin (COVID-19 was developed and spread by Bill Gates' team), its spread (COVID-19 is spread by suspicious planes, packages from Aliexpress), its goals (COVID-19 is used to reduce the population, it is a biological weapon, COVID-19 is spread by 5G networks), etc. For most of the questions, teachers correctly identified the misinformation or conspiracy associated with COVID-19, as well as being able to separate statements that are true from conspiracies and speculations for which there is no evidence.

Among the statements assessed, we also included statements that are more difficult to verify and which we did not include in the final assessment of the whole section. These were the following statements:

1) The numbers of deaths with COVID-19 do not match the reality presented in the media

28.12% of the teachers rated this statement as true, 27.70% rated it as false, 16.29% of the teachers could not decide and another 23.94% could not evaluate the statement.

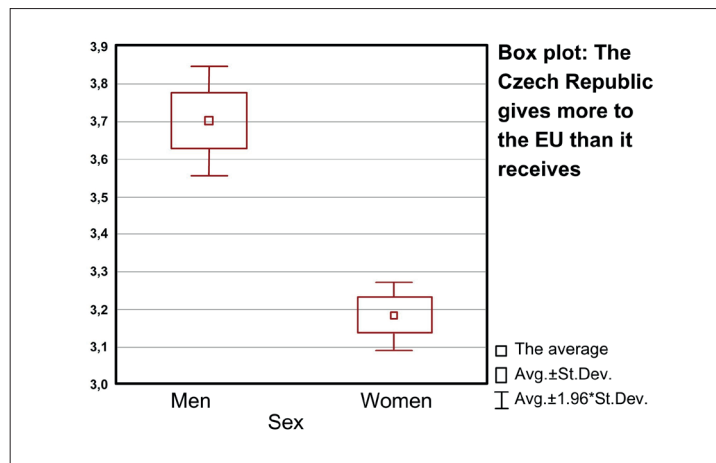


Chart 2. The Czech Republic gives more to the EU than it receives. Methodological note: We coded the questionnaire items that were scaled as follows: 5 = strongly disagree, 1 = strongly agree. Similarly, we observed whether the response of men or women was true or false. Thus, in this case, men who express disagreement with the statement score better.

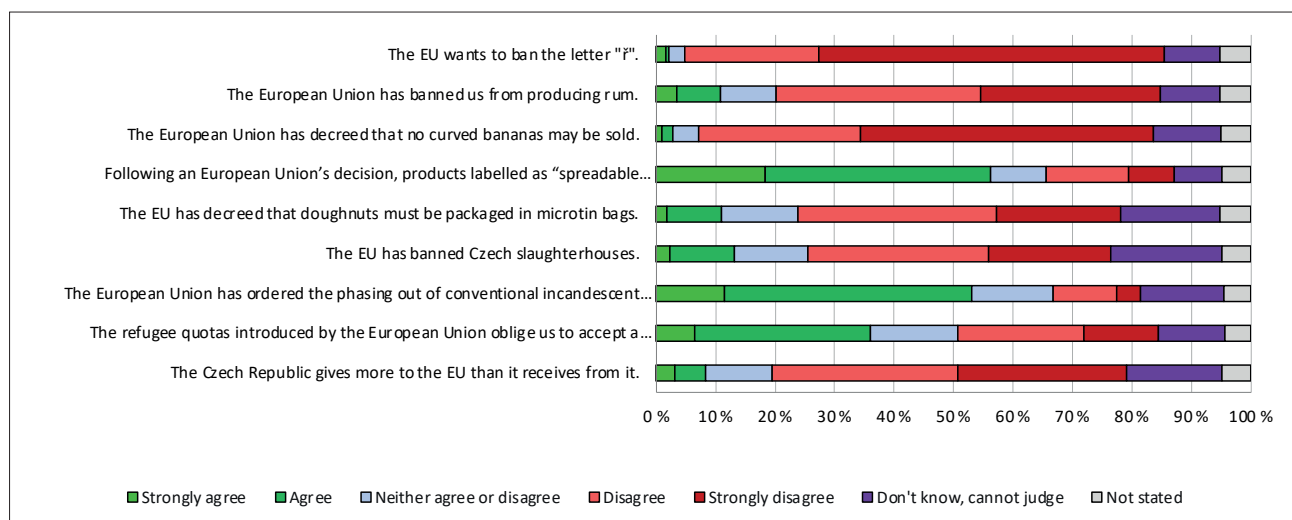


Chart 3. Teachers and information about the European Union (teachers' assessment of the truth of the statements), n=2155.

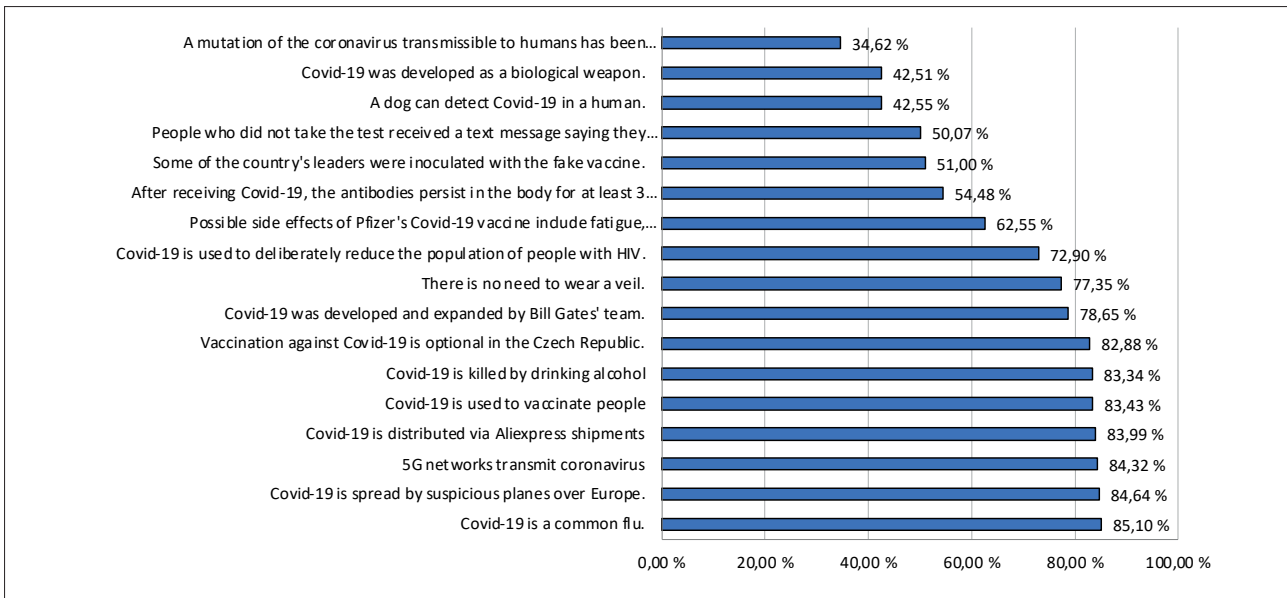


Chart 4. Educators and COVID-19 (correct assessment of statements), n=2155.

The ambivalence of the respondents is logical, as a number of persons died in a COVID positive state, but the disease itself may not have been the primary cause of death (it only accelerated other health problems). Data presented on the official website of the *Ministry of Health* of the Czech Republic (www.mzcr.cz) can be defined as deaths of persons with COVID-19 (Daňková; Zvolský; Dušek, 2020). However, it is pure speculation whether the media are lying to us and purposefully distorting the information.

2) Measures introduced by the State against the spread of the coronavirus restrict freedom and democracy

29.88% of the teachers marked this statement as true, 41.76% of the teachers marked it as false, 21.81 could not decide. Many institutions are seeking an answer to the question of whether the measures introduced by the State against the spread of coronavirus restrict freedom and democracy, and the *European Parliament* has also drawn attention to the problem (Grogan, 2022). It is clear that the COVID-19 pandemic has affected people's lives and that some human rights and freedoms have had to be curtailed. It must be said that the *Charter of Fundamental Rights and Freedoms* itself provides for the restriction of rights when they are measures necessary in a democratic society for the protection of the rights and freedoms of others, for the protection of public order, health, morals, property or the security of the State. The question is whether the rights have been restricted adequately to the situation.

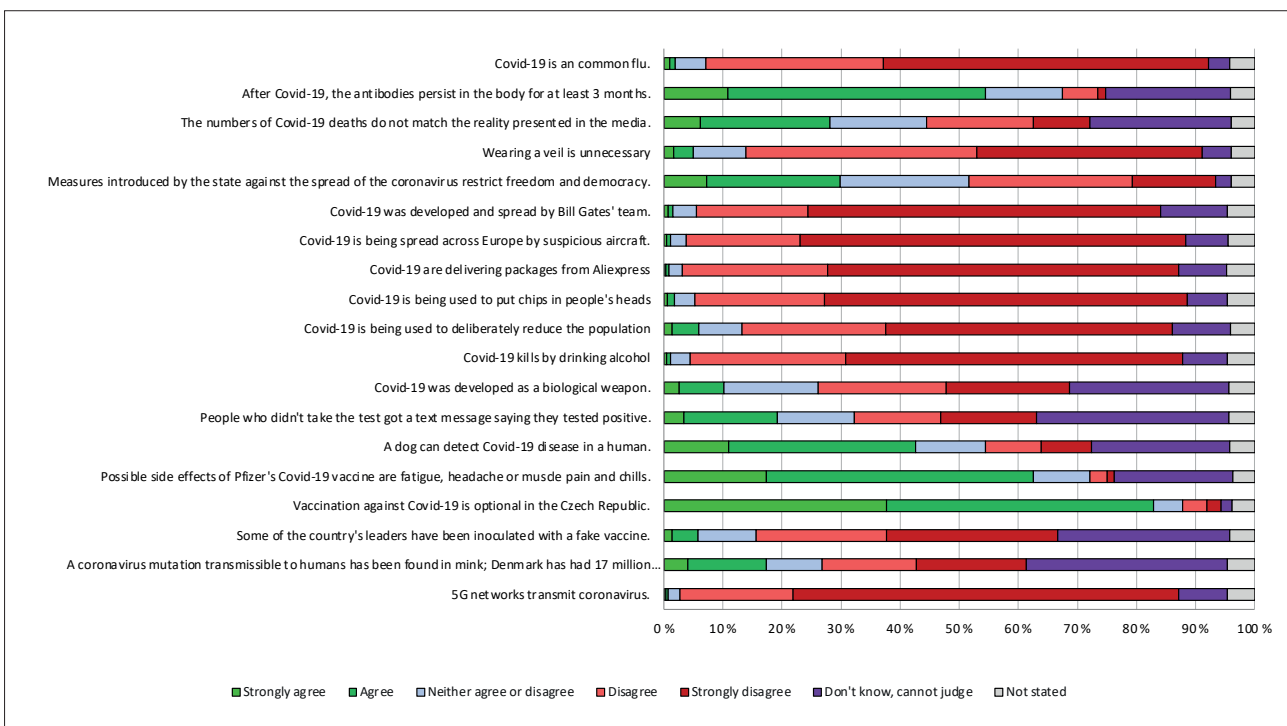


Chart 5. Overall results (teachers' assessment of the truth of the statements), n=2155.

3) People who didn't take the test got a text message saying they tested positive

The information that people who didn't take the test received a text message saying that they were positive (Jadrný, 2020) took a very interesting turn. This message was indeed originally a hoax and alarmist message, however, it was mimicked by online perpetrators and misused for so-called phishing (fraudulent solicitation of personal information). The original alarm message was thus misused to commit fraudulent activity.

This message was evaluated as true by 19.16% of respondents and false by 30.90%, however, due to the ambiguous conclusion caused by the publicized phishing cases, we evaluate both types of answers as acceptable and do not include this question in the final score.

Summary

Based on the analysis of the above statements related to the COVID-19 pandemic, 67.91% of the teacher respondents were able to correctly identify the truth of the statements (average of 17 statements). 4.56% of the teacher respondents were able to incorrectly identify the truth of the statements (average of 17 statements).

Given that 15.09% of teachers indicated that they could not judge the statement at this time, it is possible that the final score could improve but also worsen. This is because one of the characteristics of media literacy is the ability to admit that I do not have enough relevant information at a given moment to judge a given statement. However, the question is whether I choose a relevant source of information when verifying the information.

3.4. Teachers and information from the world around us

In this section, we presented educators with a total of 17 statements that included several different types of information to assess the truthfulness of the statements. On the one hand, educators had to assess statements that can be classified as conspiracy (chemtrails, *NWO*, etc.), statements that fall into the realm of hoaxes and misinformation, and historical myths that have persisted to the present day and are commonly believed (e.g., the height of Napoleon, Viking helmets).

We have also included several real events from history, for which knowledge of historical realities is necessary to judge (e.g. photographs of the dead, tea in British army tanks, nomination of totalitarian dictators for the Nobel Peace Prize, etc.).

We will now look at some statements and their assessment:

A. Conspiratorial statements

The Americans didn't land on the Moon in 1969, it's a giant hoax

74.8% of the teachers correctly identified that this is not a true statement and that Americans actually landed on the moon. Conspiracy theorists, on the other hand, claim that the moon landing was fictitious and all was made in film studios, which conspiracy theorists say can be determined from official film footage and photo documentation. However, all of the conspiracists' arguments have already been repeatedly refuted (Godwin, 2019; Link, 2019).

Aircraft deliberately disperse chemicals that serve to reduce the population

The so-called chemtrails theory is one of the very popular and widespread conspiracy theories (Llanes-Álvarez *et al.*, 2016; Xiao; Cheshire; Bruckman, 2021), yet 80.7% of teachers correctly identified that it is not a true statement. Only 1.16% identified the statement as true.

The attack on the World Trade Center in the USA on 11 September 2001 was planned and executed by US intelligence services

64.13% found this statement to be false, but this does not mean that the rest of the teachers believe in the conspiracy theory about the attack on the World Trade Center (Sardarizadeh, 2021) –only 4.59% of the teachers interviewed believe this theory.

The world is controlled by a select group of people who want to establish a world order (the so-called *NWO*)

Conspiracies about the so-called establishment of a world order by a group of powerful people (Barkun, 2012) are a common part of chain letters and various websites. However, 57.12% of the educators rated this statement as untrue, while 7.01% of the educators surveyed believed in conspiracies about the *NWO*. 18.75% of the educators also stated that they were unable to judge the statement.

The 2020 US presidential election was purposefully rigged in favour of Biden

We included this statement in our set in the context of the massive wave of disinformation and conspiracies spread by ex-President Trump and others in the wake of the US election (Sardarizadeh, 2021; Tollefson, 2021). Therefore, we were interested in how Czech educators would react to this statement. 64.87% said that the statement was untrue and that they did not believe in conspiracies related to the manipulation of the US elections. This conspiracy was supported by only 4.92 % of teachers. 17.68 % of teachers again remained cautious and said that they were unable to judge the statement.

B. Hoaxes and other disinformation (hoax classics)

We've also included several notorious hoaxes in our set of statements:

If you put an egg between two mobile phones calling each other, it will boil within an hour

This is a classic hoax that is more than 20 years known (*Hoaxes.org*, 2006). Still only 58.79% of educators rated the statement as false. Only 1.62% believed the hoax. However, more than a quarter of the educators (28.26%) also stated that they could not evaluate the statement and were unsure whether cooking an egg this way could actually happen.

If we have to withdraw money from an ATM in an emergency situation (someone is blackmailing me), we can just enter the PIN backwards and it calls the police

This is a well-known worldwide hoax (*Biryukov*, 2015), which was correctly detected by 59.49%. Only 2.8% of the interviewed teachers caught on to this statement. However, a large number of educators (26.50%) again stated that they could not judge the statement and did not have enough information to judge it.

C. Historical myths that have persisted to the present day

In this section, we have included statements that are historical errors or propaganda of the time, but which have persisted to the present day. Teachers have often made mistakes in this section.

Napoleon had a complex about his height, which was below average for that time. Hence the term “Napoleon complex” in psychology

This is one of the most famous and widespread myths, as Napoleon was in fact taller than average and the myth of his short height was created by a miscalculation of English and French measurements (different sizes of inches) and also by contemporary propaganda (*Mclivenna*, n.d.).

65.94% of the teachers rated the statement as true and believed that Napoleon would indeed be small in stature. Only 13.32% of teachers rated the statement as false. Thus, it can be said that this myth is still quite successful after two hundred years.

Vikings had horned helmets

Another myth is that Vikings had horned helmets (which we also see in many fairy tales, films, advertisements, etc.). In reality, Vikings did not have horned helmets (they would have been difficult to fight with, easy to rip off the head, etc.) and this myth came about thanks to Richard Wagner’s opera *The ring of the Nibelungs* in the 19th century (*Gritton*, 2017).

52.95% of the teachers agree that the Vikings did indeed have horned helmets, so they misjudged the statement. 21.21% of the teachers said that the statement was false, while 12.58% of the teachers said that they were unable to evaluate the statement.

D. Historical truths

We have also included several historical truths with the statements, but historical knowledge is necessary to assess them. Therefore, we also do not include them in the final assessment of the success of detecting false statements.

In 19th century England, it was common to photograph deceased relatives in poses as if they were alive (for example, seated)

This is true (*Bell*, 2016). 21.25% of educators correctly identified. 10.90% then rated the statement as false and 50.39% said they could not evaluate the statement. Here our assumption that knowledge of the topic + historical reality is necessary for this type of statement is evident.

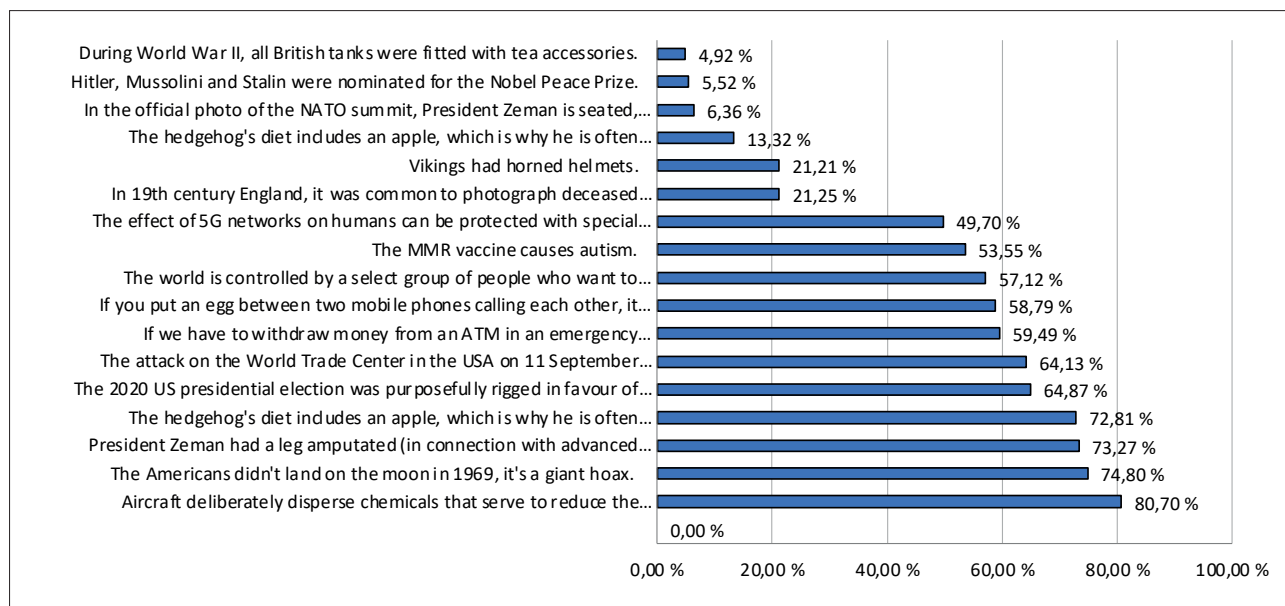


Chart 6. Educators and information from the world (correct assessment of statements), n=2155.

During World War II, all English tanks contained tea brewing accessories

Again, this is a true statement, (O'Brien, 2019) which only 4.92% of teachers correctly identified, while 17.45% identified the statement as false. Importantly, 58.93% of educators responded that they could not evaluate the statement.

Hitler, Mussolini and Stalin were nominated for the Nobel Peace Prize

This is again a historical truth, the above dictators were indeed nominated for the Nobel Peace Prize (Mendes, 2019). This statement was correctly identified by 5.52% of the teachers, but 72.16% of the teachers identified it as untrue, while 12.85% of the respondents could not judge the statement.

4. Summary

On the basis of the analysis of the above statements, it can be said that 53.53% of the respondent teachers can correctly determine the truth of the statements (average values of 14 statements). 16.79% of the respondent teachers determined the truth of the statements incorrectly (average values of 14 statements). We excluded from the evaluation 4 statements that require specific historical knowledge and for which the number of teachers who could not judge the statement was above 45%.

Given that 18.67% of educators indicated that they could not assess the statement at this time, it is possible that the final score could improve but also worsen. This is because one of the characteristics of media literacy is the ability to admit that I do not have enough relevant information at a given moment to judge a given statement. However, the question is whether I choose a relevant source of information when verifying the information.

General summary

Based on the results of verifying the truthfulness of 38 statements (divided into 3 thematic groups, not including statements for which more than 45% of teachers were unable to assess the statement for various reasons) that contained false information (hoaxes, conspiracies, misinformation, but also true information), we can say that on average:

61.14% of Czech teachers are able to correctly judge the truthfulness of statements (even 66.78% of teachers in the area of EU-related misinformation).

15.59% of Czech teachers admit that they are not able to assess the statements in a relevant way (with respect to ignorance of facts, ignorance of the issues, etc.). This is a rather positive finding - educators are aware of their limits with regard to knowledge and are able to work on themselves.

10.41% of Czech teachers misjudge the truth of the statements made, believing false content, misinformation, conspiracies, hoaxes and myths.

5. Discussion

5.1. Revision of teacher training as a necessity

The global transformation of human society and the ever-increasing importance of digital technologies and the internet in people's lives bring with them increased demands for knowledge and skills to navigate this world. At the moment we are experiencing an information pandemic (infodemic), characterised by an extreme increase in online content, which, however, does not undergo any process of checking, reliability or validity.

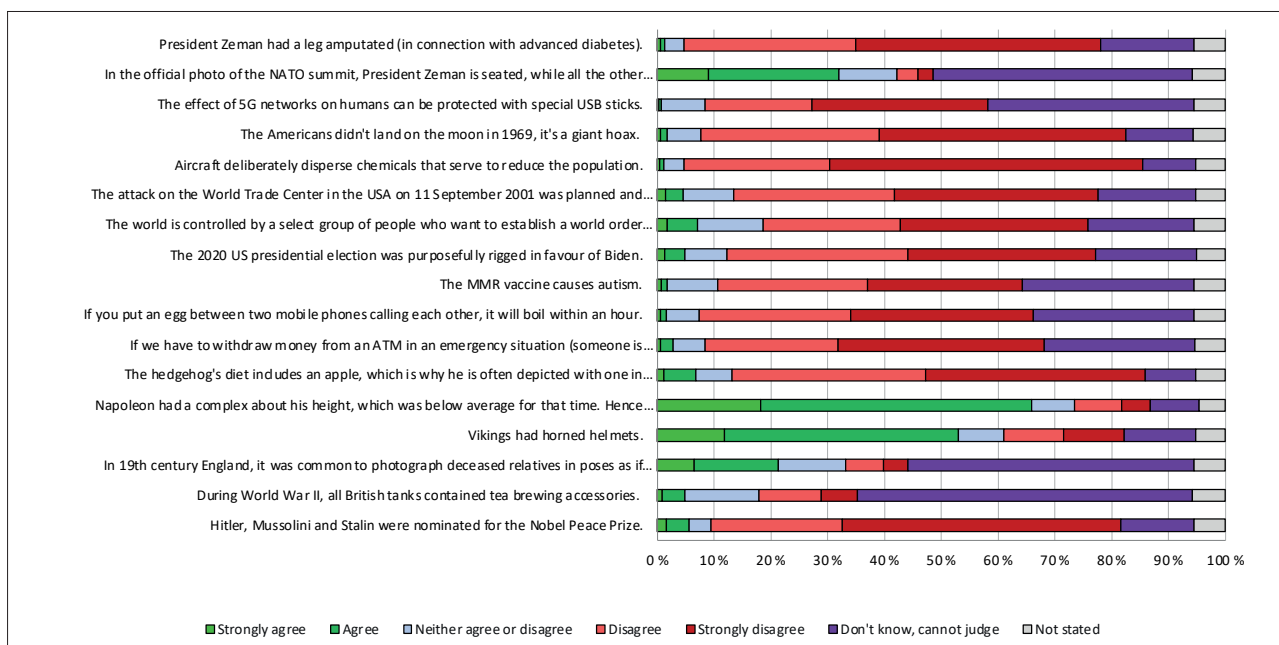


Chart 7. Overall results (teachers' assessment of the truth of the statements), n=2155.

It is imperative that the education system responds to these changed conditions and includes topics related to online content validation in the educational curriculum already at primary school level. This is happening to a very limited extent, and education systems very often rely on external providers of these forms of education (companies, the non-profit sector, etc.) rather than integrating these issues into their own curricula –especially in the form of compulsory subjects. Research Czech teacher in the world of media (Kopecký *et al.*, 2021) investigated, among other things, whether Czech teachers consider media education and teaching related to verifying information to be important, and 91% of teachers actually stated that media education is important –and this already in primary school (35% of teachers stated that they want topics related to misinformation and identifying relevant sources of information to be taught in their school). At the same time, 21% of teachers said they did not agree that the number of hours devoted to media education should be increased at the expense of other subjects.

A major problem is the formal teacher training (implemented through university education at pedagogical, philosophical, natural science and other faculties) in the Czech Republic, which often responds to the changed social situation with great delay and inconsistency, and there are great differences between faculties preparing future teachers. Media education appears very little in teacher training and in most cases takes the form of only optional electives, which often do not pass through all teacher training programmes.

5.2. Selected problems related to the implementation of media education in the Czech school environment

The lack of teacher training in the field of media education is not the only problem encountered in the implementation of teaching in practice –in the environment of primary and secondary schools in the Czech Republic. There are several problems:

Lack of a fixed time allocation for teaching media education

The time space for the implementation of media education in the environment of Czech schools comes from the so-called available hours, the use of which is decided by the schools themselves. It can easily happen that media education will not be taught at all, or its share will be minimal.

Lack of responsibility for the implementation of media education

One of the key problems facing media education is the lack of responsibility for the implementation of media education among specific teachers with specific qualifications (e.g. teachers of civic education, mother tongue, etc.). In other words, the responsibility for the implementation of media education is shared, which leads to the fact that it is very often implemented in practice in a very limited form.

Fear of a swelling educational curriculum

Other problems related to the implementation of media education in schools include teachers' concerns about the ever-increasing curriculum (syllabus). The results of our research show that teachers are concerned that more intensive implementation of media education in education would lead to an increase in educational content and thus bring more teaching load.

Parents disagree with the implementation of media education

In recent years, many untruths (and in some cases outright misinformation) have been spread in the public space about media education, leading many parents to have distorted ideas about the implementation of media education at school. Among the numerous misinformation were, for example, statements that media education is only “political training”, that it is “brainwashing” (Erhart; Kropáčková, 2018), that the goal of media education is to present the “only right truth”, etc. Unfortunately, these statements are very often spread by top representatives of Czech politics, often very active disseminators of disinformation content in the online environment. However, the school must not resign itself to this situation and patiently remind the importance of media education and its implementation.

The very nature of media education as a dynamic field

Media education is a very dynamic field. This dynamism is due to the rapid advances in the field of information and digital technologies, which in turn give rise to new phenomena that curriculum documents and system changes cannot quickly reflect. These are, in particular, the so-called ‘new media’ that have emerged from the digitisation of traditional media. For example, the *Framework Curriculum for Primary Education* does not make specific reference to phenomena such as disinformation, social networks, algorithms or personalised advertising. This rigidity of curriculum documents is partly compensated by projects whose outputs are also online, so that they can be updated according to modern trends, but the basic curriculum document does not force teachers to respond to these changes in society. This creates a discrepancy between what surrounds people in real life and how this world is reflected in school. For example, Jirák, Šťastná and Zezulková writes:

“In the future, therefore, the basic principles of media education will need to be rethought towards an understanding of the principles of new media and their constitutive role in shaping society and its activities” (Jirák; Šťastná; Zezulková, 2018).

Fragmentation of educational materials

Related to the previous problem is the fact that teachers themselves do not know where to look for new materials. There are several resources for media education, the most dominant of which is probably the One World in Schools project,

but these may not suit everyone. These online resources do not even have a clear plan for how to teach media education throughout the year and tend to focus on specific topics. Another problem may be the lack of materials for certain subjects (e.g. mathematics, geography, etc.).

Media education with a focus on education cannot currently be studied in the Czech Republic either as part of a bachelor's or master's degree at a university. Czech universities do offer a media studies programme, but it does not focus on education. The only field of study that deals with media education in education is the doctoral programme *Literacy and Media Education in Education* at the *Faculty of Education of Palacký University in Olomouc*, which, however, is a combination of didactics of literature and media education, so it is not a "pure" form of studying media education.

6. Limits and perspectives of the study

Although the data collection through the electronic questionnaire was targeted and the instrument was distributed directly to specific teachers, it is difficult to check the identity of individual respondents and to ask complex and challenging questions using this instrument. Here it would be appropriate to follow up with e.g., qualitative research.

Due to the size and number of items in the questionnaire, it was not possible to give an exhaustive explanation of the individual items of the research instrument, which could have caused a meaning discrepancy between the researcher's and respondent's perceptions.

The research instrument does not reflect the whole spectrum of the media world, the media space, only its subparts. Similarly, it does not map teachers' media literacy as a whole, but only its subparts.

7. References

- Barkun, Michael** (2012). "New world order conspiracies: the new world order and the illuminati". In: Barkun, Michael. *A culture of conspiracy apocalyptic visions in contemporary America*. California Scholarship Online, pp. 39-64. ISBN: 978 0 520238053
<https://doi.org/10.1525/california/9780520238053.003.0003>
- Bell, Bethan** (2016). "Taken from life: the unsettling art of death photography". *BBC news*, 5 June.
<https://www.bbc.com/news/uk-england-36389581>
- Biryukov, Vladislav** (2015). "Fact or fiction? Reversing your PIN can call the police in an emergency". *Kaspersky official blog*, November 16.
<https://www.kaspersky.com/blog/fact-or-fiction-inverted-pin-code/10559>
- Ciroková, Kristýna** (2022). "Potrestejte žáka, co mě nahrál, žádá učitelka, která šířila dezinformace". *Seznam zprávy*.
<https://www.seznamzpravy.cz/clanek/domaci-kauzy-potrestejte-zaka-co-me-nahral-zada-ucitelka-ktera-sirila-dezinformace-198997>
- Daňková, Šárka; Zvolský, Miroslav; Dušek, Ladislav** (2020). *Mortalitní data, kódování a vymezení umrtí na Covid-19*. Strana 2 z 11.
<https://onemocneni-aktualne.mzcr.cz/doc/zemreli-COVID-19-metodicky-rozbor.pdf>
- Erhart, Michal; Kropáčková, Renata** (2018). "Neziskovky jako terč? Podle Volného vymývají dětem mozky, pomáhají tam, kde stát zaspal, oponuje Gajdůšková". *Irozhlas*.
https://www.irozhlas.cz/zpravy-domov/lubomir-volny-spd-alena-gajduskova-cssd_1808081346_haf
- Evropská komise* (2021). "Eurobarometr 2021: důvěra česků v EU je nejvyšší od roku 2013". *Business INFO*. Cz.
<https://www.businessinfo.cz/clanky/eurobarometr-2021-duvera-cechu-v-eu-je-nejvyssi-od-roku-2013>
- Evropská komise* (2023). *Evropský rok dovedností 2023*.
https://ec.europa.eu/czech-republic/news/euomyty_cs
- Goda, Jakub** (2019). "Hlavné správy a Zem a vek patria podľa prieskumu medzi najdôveryhodnejšie médiá medzi učiteľmi". *Rubriky*.
<https://dennikn.sk/blog/1571691/hlavne-spravy-a-zem-a-vek-patria-podla-prieskumu-focusu-medzi-najdoveryhodnejsie-media-medzi-ucitelmi/?ref=in>
- Godwin, Richard** (2019). "One giant... lie? Why so many people still think the Moon landings were faked". *The Guardian*, 10 July.
<https://www.theguardian.com/science/2019/jul/10/one-giant-lie-why-so-many-people-still-think-the-moon-landings-were-faked>
- Gritton, Jim** (2017). "Wagner and the trope of the horned helmet".
https://www.researchgate.net/publication/335219184_Wagner_and_the_Trope_of_the_Horned_Helmet
- Grogan, Joelle** (2022). "The impact of Covid-19 measures on democracy, the rule of law and fundamental rights". *Think tank*.
[https://www.europarl.europa.eu/thinktank/en/document/IPOL_STU\(2022\)734010](https://www.europarl.europa.eu/thinktank/en/document/IPOL_STU(2022)734010)
- Havrilesky, Thomas** (1972). "The information explosion, technological innovativeness and the competitive ethic". *Land economics*, v. 48, n. 4, pp. 347-356.
<https://doi.org/10.2307/3145311>

- Hoaxes.org* (2006). "How to cook an egg with two cell phones". *Hoaxes.org*, February 7.
http://hoaxes.org/weblog/comments/how_to_cook_an_egg_with_two_cell_phones
- Irozhlas* (2021). "Česko z Rozpočtu EU získalo o 38,7 miliardy více, než do něj odvedlo. Zůstalo tak čistým příjemcem peněz". *Irozhlas*, 29, července.
https://www.irozhlas.cz/ekonomika/eu-rozpocet-ministerstvo-financi-alena-schillerova-strukturalni-fondy_2107291544_ere
- Islam, Md-Saiful; Sarkar, Tonmoy; Khan, Sazzad-Hossain; Kamal, Abu-Hena-Mostofa; Hasan, S. M. Murshid; Kabir, Alamgir; Yeasmin, Dalia; Islam, Mohammad-Ariful; Amin-Chowdhury, Kamal-Ibne; Anwar, Kazi-Selim; Chughtai, Abrar-Ahmad; Seale, Holly** (2020). "Covid-19 related infodemic and its impact on public health: A global social media analysis". *The American journal of tropical medicine and hygiene*, v. 103, n. 4, pp. 1621-1629.
<https://doi.org/10.4269/ajtmh.20-0812>
- Jadrný, Petr** (2020). "Nešli na test, přesto dostali SMS, že jsou pozitivní. „je to absolutní nesmysl,“ řekla o hoaxu jágrová". *Irozhlas*, 1. Října.
https://www.irozhlas.cz/zpravy-domov/hoax-koronaviru-v-cesku-sms-chyb_2010010746_pj
- Jiráček, Jan; Šťastná Lucie, Zezulková, Markéta** (2018). *Mediální výchova jako průřezové téma*. Praha.
https://www.npi.cz/images/podkladov%C3%A11_studie/medialni_vychova.pdf
- Kopecký, Kamil; Szotkowski, René; Voráč, Dominik; Mikulcov, Klára; Krejčí, Veronika** (2021). *Czech teachers in the world of media*. Olomouc: Palacký University in Olomouc.
- Link, Devon** (2019). "Fact check: Moon landing conspiracy theory misrepresents footprint". *USA Today. News*.
<https://eu.usatoday.com/story/news/factcheck/2021/09/17/fact-check-moon-landing-conspiracy-theory-misrepresents-footprint/8380230002>
- Llanes-Álvarez, Carlos; San-Roman-Uría, Alberto; Nunes-Nancabu Paulo; Ruiz-Gippini, María; López-Landeiro, Patricia; Franco-Martín, Manuel-Ángel** (2016). "Chemtrails: an overview of the phenomenon". *European psychiatry*, v. 33, n. S1, pp. 450-450.
<https://doi.org/10.1016/j.eurpsy.2016.01.1635>
- Mclivenna, Una** (2021). "Was Napoleon short? Origins of the 'Napoleon complex'". *History.Com*.
<https://www.history.com/news/napoleon-complex-short>
- Mendes, Claudia** (2019). "The irony - Nobel Peace Prize nominees included Hitler, Stalin and Mussolini". *War history online*, February 24.
<https://www.warhistoryonline.com/instant-articles/nobel-peace-prize-nominees.html>
- Modrako, Alicia** (2020). "Exkluzívny prieskum medzi učiteľmi: Nedôverujú vláde a mnohí si myslia, že sa chystá čipovať ľudí".
- Nutil, Petr** (2016). "Hoax: EU chce zrušit písmeno Ř". *Manipulátoři.Cz*.
<https://manipulatori.cz/hoax-eu-chce-zrusit-pismeno-r>
- O'Brien, Joseph** (2019). "Tiger tanks & tea: The British cup of char in war". *War history online*, 10 March.
<https://www.warhistoryonline.com/world-war-ii/tiger-tanks-tea-the-british-cup.html>
- Sardarizadeh, Shayan** (2021). "11 September 2001: The conspiracy theories still spreading after 20 years". *BBC news*, 10 September.
<https://www.bbc.com/news/58469600>
- Tollefson, Jeff** (2021). "Tracking QAnon: How Trump turned conspiracy theory research upside down". *Nature*, v. 590, n. 7845, pp. 192-193.
<https://doi.org/10.1038/D41586-021-00257-Y>
- WHO* (2023). *Infodemic*. World Health Organization.
https://www.who.int/health-topics/infodemic#tab=tab_1
- Xiao, Sijia; Cheshire, Coye; Bruckman, Amy** (2021). "Sensemaking and the chemtrail conspiracy on the Internet: insights from believers and ex-believers". *Proceedings of the ACM on human computer interaction*, v. 5, n. 454.
<https://doi.org/10.1145/3479598>
- Zastoupení Evropské komise v ČR* (2016a) "Euromýtus: Brusel nám zakázal žárovky". *Euroactiv*, 29 August.
<https://euractiv.cz/section/energeticka-ucinnost/news/euromytus-brusel-nam-zakazal-zarovky>
- Zastoupení Evropské komise v ČR* (2016b). "Euromýtus: konec zabijaček v českých - euractiv.Cz". *Euroactiv*, 9 December.
<https://euractiv.cz/section/cr-v-evropske-unii/news/euromytus-konec-zabijacek-v-cechach>
- Zheng, Min; Marsh, Jesseca K.; Nickerson, Jeffrey; Kleinberg, Samantha** (2020) "How causal information affects decisions". *Cognitive research: Principles and implications*, n. 5.
<https://doi.org/10.1186/s41235-020-0206-z>

Disintermediation and disinformation as a political strategy: use of AI to analyse fake news as Trump's rhetorical resource on *Twitter*

Alba Diez-Gracia; Pilar Sánchez-García; Javier Martín-Román

Nota: Este artículo se puede leer en español en:
<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87347>

Recommended citation:

Diez-Gracia, Alba; Sánchez-García, Pilar; Martín-Román, Javier (2023). "Disintermediation and disinformation as a political strategy: use of AI to analyze fake news as Trump's rhetorical resource on *Twitter*". *Profesional de la información*, v. 32, n. 5, e320523.

<https://doi.org/10.3145/epi.2023.sep.23>

Manuscript received on 19th April 2023
Accepted on 27th September 2023



Alba Diez-Gracia ✉
<https://orcid.org/0000-0001-9262-3771>

Universidad de Valladolid
Facultad de Filosofía y Letras
Plaza del Campus Universitario, s/n
47011 Valladolid, Spain
alba.diez.gracia@estudiantes.uva.es



Pilar Sánchez-García
<https://orcid.org/0000-0002-6223-182X>

Universidad de Valladolid
Facultad de Filosofía y Letras
Plaza del Campus Universitario, s/n
47011 Valladolid, Spain
pilar.sanchez@uva.es



Javier Martín-Román
<https://orcid.org/0000-0003-0217-9879>

Universidad de Valladolid
Parque Científico
Campus Valladolid - Miguel Delibes
Paseo Belén, 69-75
47011 Valladolid, Spain
javiermartinroman@outlook.com

Abstract

The communicative effects of disintermediation caused by social media promote the expansion of personalist and emotional political discourses that reach the audience directly and evade the traditional journalistic filter. This phenomenon leads to new political communication tactics, but also exposes citizens to potentially fraudulent, contaminated or polarised content. In this context, framed in post-truth, the term 'fake news' gains relevance as a way of referring to disinformation and as a political and performative argument that can be weaponised. This research aims to analyse such use in the discourse of the former president Donald Trump during his presidential term (2017-2021), focussing on *Twitter* as the main platform in his political communication strategy online. To analyse this, we resort to a methodological triangulation of content, discourse, and sentiment analysis, with the latter combining both lexicon and artificial intelligence (AI) techniques through machine learning on the basis of deep learning and natural language processing, which is applied to his messages published with the term 'fake news' ($N = 768$). The analysis of the sample, provided here in an open dataset, employs self-developed software that allows each unit of analysis to be filtered and coded around its predominant themes, sentiments, and words. The main results confirm that Trump's attribution of 'fake news' focusses on three main topics: the media (53%), politics (40%) and his cabinet (33%). It also shows how the former president resorts to a personalist agenda, focussed on the defence of his proposals and his team (80%) by delegitimizing his opponents and the press, with a negative tone (72%) loaded with derogatory terms, confirming a weaponised strategy of the term 'fake news' as a political argument of disinformation and disintermediation.



Keywords

Disinformation; Disintermediation; Fake news; Political communication; Political strategy; Political personalization; Artificial Intelligence; AI; Social networks; Discourse analysis; Sentiment analysis; *Twitter*; Donald Trump; Deep learning; Machine learning; Natural language processing.

Funding

This article has been funded by the national R&D&I project “Politainment in the face of media fragmentation: disintermediation, engagement and polarisation” (Poldespol), of the *University of Valladolid* (Ref. PID2020-114193RB-100), *Ministry of Economy and Competitiveness* of Spain (2021-2023).

1. Introduction

The crystallisation of social media implies a qualitative leap in online communication through applications that prioritise interpersonal contact (Van-Dijck, 2016), as well as massive, multidirectional interaction and participation (Beer; Burrows, 2007; Pérez-Salazar, 2011; Pfister, 2011). Under these digital structures, those surfing the internet become potential prosumers who fulfil the simultaneous roles of producers and consumers (Ritzer; Dean; Jurgenson, 2012). Thus, they can select, distribute, and curate the information on different platforms (Hernández-Serrano *et al.*, 2017; Timoshenko; Hauser, 2019) without the need for specialised knowledge (Muswede, 2022).

The appearance of the user as a participant in the flow of information (Pavlicková; Kleut, 2016), and the competition between their own content and that offered by the media, whose distribution is measured, in part, by the automation of the algorithms, the search engines and the viral flow (Lokot; Diakopoulos, 2015), counter the intermediary role of the journalists as the traditional gatekeepers, since the

“power to inform and transmit reality no longer belongs solely to the media” (López-Jiménez, 2014, p. 84).

These mechanisms and forms of participation contribute to the disintermediation of mass communication in the digital environment, in which the people

“reject the representative mediation and prefer the self-representation of their particular, subjective reality” (Mencarelli, 2021, n.p.).

Faced with this change in the communicative process, the sources of information are also forced to adapt their position: politicians, enterprises and groups with differing power and social influence become prosumers (Berrocal *et al.*, 2014). They modify their communicative strategies and areas of diffusion through these networks, in order to launch their messages directly to the potential users, voters or consumers. It is a strategy that takes advantage of micro-targeting as a form of aiming at audience niches through the exhaustive study of their habits or preferences, so as to reach them in the most direct and personalised way possible (Barbu, 2014). It is a practice embedded within the marketing techniques in which artificial intelligence (AI) plays a fundamental role, and which allows massive quantities of data to be processed (Kotras, 2020) in order to personalise the content (Milan, 2015; Sánchez-García *et al.*, 2023). This gives rise to such paradoxes as the personalisation of the masses through decisions taken by machines (Ritzer, 2015) that are also capable of dealing with the public discourse (Gillespie, 2010). This mechanisation affects equally both the automation of political information (Bradshaw *et al.*, 2020) and its influence on public opinion (Murthy *et al.*, 2016); this is due to its effect on journalistic objectivity (Carlson, 2018; Wu, 2019) and its relationship with fake news (Zimmer *et al.*, 2019).

In this context, there occurs the transition from mediated politics (Castells, 2009), which needed the media as intermediaries with public opinion, to disintermediated politics, in which leaders and entities transmit their message directly to their audience through social media. Thus, the persuasive skills of the politicians are now being developed through a connected, telematic media flow (Muswede, 2022) that provides electronic devices with political goals (Grossman, 1995) and gives the Internet direct democracy (Morris, 2001, p. 1033). This is an environment that can favour interactivity, participation, and cooperation (Coleman, 2005; Fernández-Castrillo, 2014); yet it can also bring dangers, such as the encouragement of ‘popular politics’ (Berrocal *et al.*, 2022) which treats leaders as celebrities; the opportunistic use of citizens as viral instruments (García-Orosa, 2021, p. 3); or the impulse towards a new Caesarism (Rubio-Fabián, 2019, p. 46) that can pose a threat to democracy (Borgesius *et al.*, 2018).

2. Disinformation and fake news in political discourse: the case of Donald Trump

Disintermediated political communication provides an opportunity for the persuasive strategies of leaders, but it also puts reception at risk, as the content is prone to manipulation and may reach public opinion “contaminated by the wishes of the person producing it” (Enguix, 2020, p. 26). The selective presentation of social media foments intrinsic effects of this environment, such as filter bubbles (Pariser, 2011) or echo chambers (Cinelli *et al.*, 2021), both intensified by the use of such mass disinformation tools as the creation of social bots to automatically disseminate fraudulent contents (Boshmaf *et al.*, 2011) and to

“potentially manipulate discussions in social media (...), creating false narratives that take root in a large percentage of the population” (Rossetti; Zaman, 2023, n.p.).

It is a context that leads to the polarisation of both politics and the people, and which fosters an environment that favours sensationalism, polarisation and disinformation via these platforms, as opposed to professional journalism in crisis (**Cano-Orón et al.**, 2021a), as has been studied in such countries as Germany (**Papakyriakopoulos et al.**,

2018), Russia (**Ribero et al.**, 2019), Spain (**Cano-Orón et al.**, 2021b), India, or the United Kingdom (**Cheng**, 2019), as well as the electoral processes in the UK and the USA (**Fincham**, 2019). More particularly, the figure of the USA's ex-president, Donald Trump, has generated numerous works of research (**Raynauld; Turcotte**, 2018; **Singh; Wijegunawardhana**, 2018). His direct communication through *Twitter* (now X) has been a watershed in political communication, using it beyond the institutional, unlike his predecessor Obama, stating his political opinions and judgements without intermediaries. It was a propagandistic and rhetorical wager that gave him an advantage over his rivals (**Muswede**, 2022; **Das-Sarma**, 2016), which he also endowed with a controversial viral and showy resonance (**Gómez-García et al.**, 2019).

“The unique combination of the power provided by his position as a world leader and the extremely colloquial tone of his publications meant his tweets had a great «success». At the end of 2020 he had accumulated almost 80 million followers and he had published over 50,000 tweets before his *Twitter* account was permanently suspended in January of 2021” (**Machus et al.**, 2022, n.p.).

As part of this political strategy based on direct communication, the accusations of informative falsehood against his opponents, including politicians, spokespersons, and the media, represent a discursive constant associated with the context of post-truth favoured by the increase in fake news (**Journell**, 2017; **Lorenzo; Manfredi**, 2019; **Bleakley**, 2018). A series of attacks used reiteratively by Trump and reproduced by like-minded sources as figures of authority “exalted” his public and generated an effect of illusory truth that distorted their perception (**Froehlich**, 2020).

Although there are several concepts to refer to different types of false or manipulated information (**Carmi et al.**, 2020), “the term ‘fake news’ refers to a whole range of information types, from honest errors of little impact and satirical content to high impact manipulative techniques and malicious inventions” (**Kapantai et al.**, 2020, p. 5).

It is a term that is overloaded (**Wardle; Derekshan**, 2017) and inadequate for describing the complexity of the phenomenon of disinformation in an era already earmarked as “the era of fake news” (**Albright**, 2017) and which, nevertheless, acquired

“worldwide relevance in 2016, during the presidential elections in the USA (...). The term was widely used (or abused) to characterise almost any content that entered into conflict with the points of view or the programme of a particular party” (**Kapantai et al.**, 2020, p. 2).

The concept, associated with other previously existing concepts, such as post-truth, not only took on relevance because of its meaning, but also because of its instrumentalisation or performative impact (**Farkas; Schou**, 2020, p. 6) as a political weapon. Donald Trump intentionally focused part of his discourse on this, starting

“a war of rhetoric against the established media, labelling them as the fake news media” (**Farkas; Schou**, 2020, p. 6).

In this light, the above studies not only consider the interested propagation of fake news or the use of bots to support his discourse (**Rossetti; Zaman**, 2023), but also the Trumpist narrative of painting as fake the information he considered to be detrimental by means of a discursive deflection (**Ross; Ribers**, 2018). This practice is considered to be an instrumentalisation and has recently been investigated in the American context with respect to

- its reception by the public opinion (**Tong et al.**, 2020);
- the influence of biased informative sources in the success of this strategy (**Froehlich**, 2020; **Meirick; Franklyn**, 2022);
- the satirical content or memes (**Smith**, 2019);
- its use to discredit the established media (**Rossini et al.**, 2021; **Happer; Hoskins; Merrin**, 2018); and
- the effects of its dissemination on the audience's informative perception (**Tamul et al.**, 2019).

Despite the academic attention dedicated to the effects of disinformation,

“there has been less research into how the political elites and the experts use fake news as a weapon to discredit the media” (**Rossini et al.**, 2021, p. 676).

Keeping in mind the above-described context, this research offers a dual analysis, thematic agenda, and emotions, of the rhetorical and political use of the term fake news by the ex-president of the USA, Donald Trump, in his discourse published in social media during his complete term of office (2017-2021). In particular, the object of the study focuses on *Twitter* as the principal platform used in his disintermediated discursive strategy (**Machus; Mestel; Theissen**, 2022) with two main aims:

- a) to analyse the thematic agenda, the policies, and the players (the media, cabinet, and opposition) that the ex-president refers to when using the term fake news;
- b) to establish the emotional level or the polarity (positive, negative, or neutral) of his discourse based on experimentation with artificial intelligence (AI) tools using machine learning systems.

Disintermediated politics can promote participation and direct democracy, but also the instrumentalization of technologies, personalism or spectacularization

Both of them allow us to perceive a panoramic view of the instrumentalisation of the term fake news as part of the disintermediated discursive strategy and the disinformation practiced by the ex-president of the USA. At the same time, they also allow us to test technological AI tools that facilitate an automation process in the case of the digital discourse analysis.

“The absence of filters in online political communication exposes citizens to potentially fraudulent, contaminated or polarized content”

Starting from these objectives, this research proposes three main hypotheses:

- H1. Donald Trump instrumentalises the term “fake news” with a rhetorical use.
- H2. The ex-president of the USA uses the label “fake news” as an ambivalent argument to discredit the sources that oppose his discursive line with a triple strategy of attack-defence-imposition of ideas.
- H3. The messages in which Trump has recourse to the term “fake” against information from the established media are characterised by a certain polarity aimed at discrediting the said media so the people will no longer trust them.

3. Materials and methods

This research uses a triangular methodology, based on content analysis, of the discourse and emotions, centred on two axes: thematic agenda (E1) and emotional level or polarity (E2). The sample is made up of the tweets published by the ex-president with the term “fake news” during his mandate (2017-2021).

In a first phase, the sample is obtained from the tweets in “The Trump Archive”:

<https://www.thetrumparchive.com>

It is an open-access platform that brings together all the messages emitted by the politician from the very start of the records (December 10th 2016) until the account was suspended on January 8th 2021 (Courty, 2021), and which has already been used in previous works of research (Quealy, 2021; Magallón, 2018; Meeks, 2019). Starting from this open database, a first filtration is carried out based on tweets from when he was sworn in on January 20th 2017 until the elimination of his profile that fit the search term (N=970). The sample was then further filtered with the elimination of retweets, deleted tweets and those that could not be visualised or codified due to a lack of context, leaving a final sample (N=768) downloaded from The Trump Archive in JSON (see Annex 1) format which allows them to be codified using a “software” of our own elaboration that facilitates the collection and automated data analysis in the empirical work.

In a second phase, this same filtered sample is submitted to an analysis of emotions through the techniques of Lexicon and Deep Learning. Figure 1 summarises the technical procedure of each of the proposed axes, and which will be considered individually in greater detail in the following subsections.

The design work and testing of the tool is distributed according to the areas of specialisation of each author of this research work: the programmer developed and implemented the application according to the needs marked out by the researchers, who chose the labels and variables and carried out the testing and codification.

3.1. Thematic agenda: content and discourse analysis (Axis 1)

The thematic analysis of Axis 1 was carried out through a content and discourse analysis, two methodological techniques that allow rationales to be established by combining categories (Piñuel-Raigada, 2002) and textual structures to be analysed (Van-Dijk, 1990). The aim is to identify in which themes Donald Trump uses the term “fake news” as a political argument in his *Twitter* discourse. These were selected on the basis of a prior sample analysis, detecting the most repeated ones and grouping them by similarity (Márquez-Domínguez *et al.*, 2017). This allows them to be fitted into four main categories and 52 multiple choice variables (see Annex 1), since a tweet can be about one or more topics.

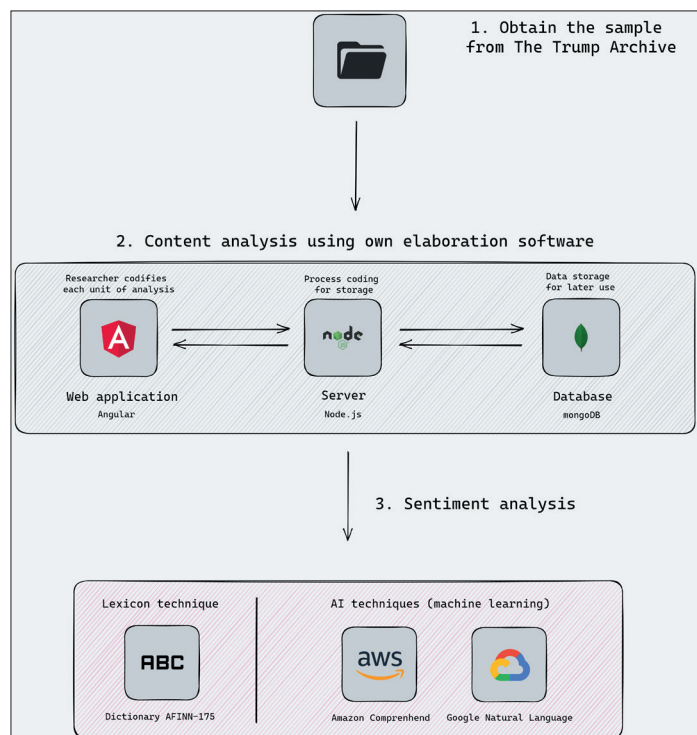


Figure 1. Technical procedure to analyse the content, discourse and emotion using own elaboration software and AI techniques

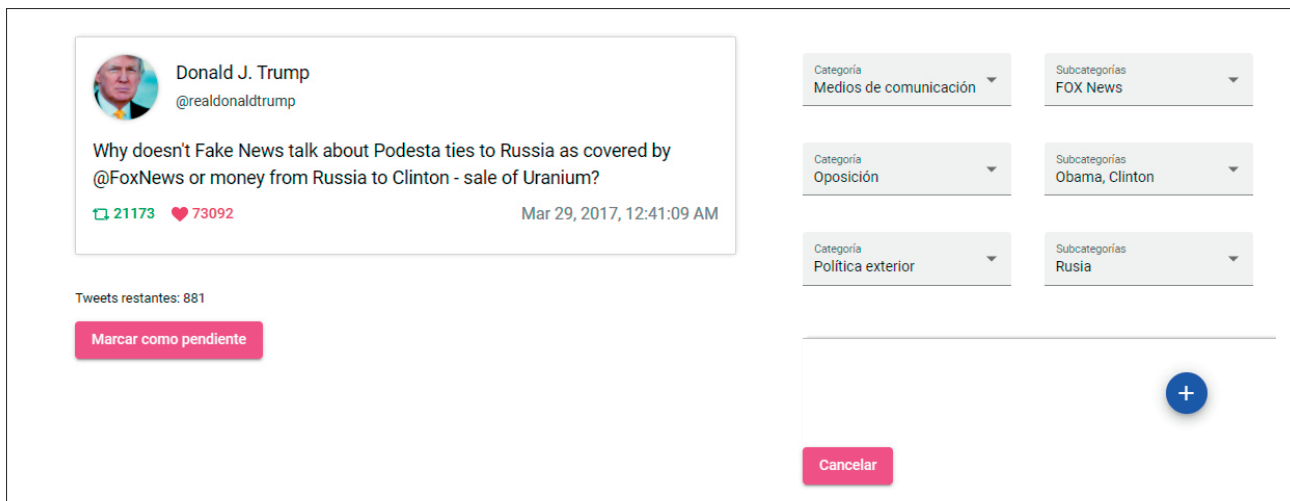


Figure 2. Visualisation of the own elaboration APP to analyse the thematic agenda of the tweets of Donald Trump concerning fake news

- C1. Politics. This includes aspects of foreign and domestic policies, and topics that can especially mobilise the republican electorate, such as arms control, economic development, or immigration (Galán-García, 2017; Cabezuelo; Manfredi, 2019). Similarly, it includes relevant political or meta-political themes that marked the agenda of the analysed period (Galán-García, 2017), such as the Black Lives Matter protests in 2020, the pardons, the impeachment, or the different electoral campaigns.
- C2. Cabinet. This includes mentions of the Trump Administration: his team, presidency, family, party, or himself.
- C3. Relation with the opposition. This brings together the major representatives of the *Democrat Party* during the previous mandate and the electoral campaign of 2020. It also includes the term “RINO” (Republicans in name only), a word frequently used by Trump to refer to those republicans who opposed him.
- C4. Relation with the media. Prominent as one of the most studied themes in previous exploratory studies (Cabezuelo; Manfredi, 2019; Márquez-Domínguez *et al.*, 2017), here expanded to differentiate between the main written, audio-visual media and the social media referred to by Donald Trump.

This analysis is made operational through a web application of our own elaboration based on the Angular framework to visualise and categorise the tweets (Figure 2), obtained from The Trump Archive; and in the administrator of the *mongoDB* database to consult, analyse and add information. This allows us to save the codification and interrelate the final results on the basis of consultations in the database. During the said coding process, a first pre-test of the intercoders was done following the criteria of Krippendorf (2004) concerning a subsample of 20% (N=153). This managed to achieve an average agreement of 93.5%. This manual phase of the content analysis was complemented by an automated sentiment analysis.

3.2. Polarity: sentiment analysis with AI: using Lexicon and Deep Learning (Axis 2)

This second phase offers a double automated sentiment analysis, “a popular method used to analyse discourse through the identification of the valence in the data from the text” (Misiejuk *et al.*, 2021, p. 376). It is a semantically oriented tool that helps to understand “public opinion concerning controversial questions in a more accurate, complete and accessible way” (Abdulla *et al.*, 2013, s.p.). It has also already been applied to the study of political discourse in *Twitter* (Kaur *et al.*, 2021). In the case of Trump, this technique has been used in previous studies applied to concrete aspects of his discourse, such as the economy (Machus; Mestel; Theissen, 2022; Colonescu, 2018), the general elections (Chandra; Saini, 2021; Xia; Yue; Liu, 2021), or domestic policy matters such as the management of the coronavirus pandemic (Dwianto; Nurmandi; Salahudin, 2021).

An analysis is used that is focused on the emotional level through a polarity qualified as positive-negative-neutral. We do so through the combination of two AI methods: Lexicon through the use of a dictionary; and a pre-trained machine learning model based on Deep Learning and Natural Language Processing (NLP). The combination of these two techniques achieves “more reliable and valid” results (Wahleed *et al.*, 2021, p. 100).

For the first type of analysis focused on the polarity, Lexicon is used as an unsupervised technique in which, in this case, a dictionary is used; that is to say, a list of words with an already assigned numerical polarity. In this case, the *Node.js* and the *Natural code* library are used. The latter includes the dictionary *Afinn-165*, which “assigns values to emotions ranging from -5 to 5; this allows a more precise quantification of the emotional content of the words than other similar lexicons” (Colonescu, 2018, p. 379).

The automated analysis commences with the filtering and cleaning of the sample (N=768). Hyperlinks and special characters (punctuation signs, hyphens, symbols, etc.) are eliminated, verbal formulas with apostrophes are replaced in or-

der to maintain the structure of the data (e.g., “you’re” becomes “you are”), and words that do not provide any information, such as conjunctions, determiners or numerals are also eliminated, thus increasing the accuracy of the analysis as much as possible.

“ Fake news arises as a concept associated with misinformation, but is also used in a post-truth context as an argument or discursive deviation ”

Having once achieved the definitive sample, the analysis of emotions divides each unit of analysis or tweet into words, which are then assigned a value individually, generating a mean polarity (Abdulla *et al.*, 2013). In order to classify the results, the research divides the range of values into Very Positive, Positive, Neutral, Negative or Very Negative.⁴

The second procedure combines *Google Natural Language* and *Amazon Comprehend*, both commercial tools that analyse emotions based on Deep learning and NLP, with two of the best scores in accuracy from the available tools (Ermakova; Henke; Fabian, 2021). The technique is unsupervised and aimed at the polarity; it is based on an intelligent, trained model, which allows a better accuracy and performance (Dang; Moreno-García; De-la-Prieta, 2020). To do so, we used the sample of tweets already refined using Lexicon. A programmed analytical instruction was executed for each of the individual tweets in the libraries available on each platform, obtaining the result with the polarity data, which was then dealt with using a summation algorithm to group the scores and thus obtain the relative percentage of each polarity within the sample.

4. Results

The analysis of content, discourse and emotions illustrates how and to which topics Donald Trump directed the concept of fake news as a political argument in his online discourse, as well as an approximation of the level of emotion or polarity. The following subsections deal with the results of the thematic content analysis and the data from the sentiment analysis, in that order.

4.1. The fake news agenda of Donald Trump

The thematic agenda in the ex-president’s discourse on fake news (Figure 3) shows a preponderance of governmental policies and actions both within the country and internationally (46.87%), as well as specifically concerning the *Trump Administration* (32.9%). Beyond the general accusation of fake news, which the ex-president used to defend the said policies, he takes direct aim at the media in over half his messages (52.6%). The last great block is taken up by his political opposition, principally the *Democrat Party* (20.7%), even though he tends to label the media as “the real opposition party.”

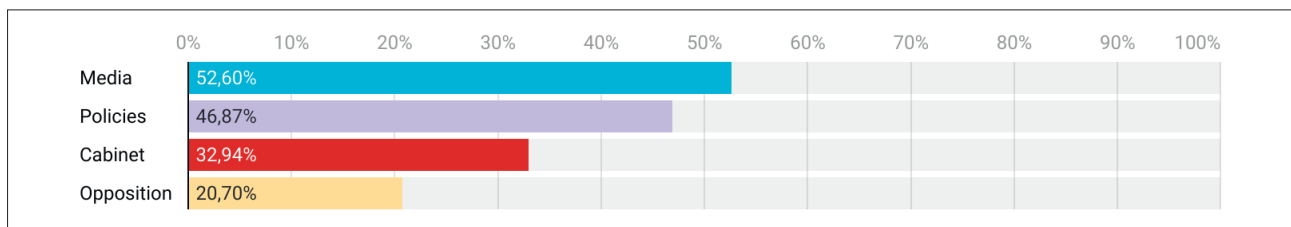


Figure 3. Donald Trump’s thematic agenda concerning fake news during his presidency (2017-2021)

4.2. Political self-referral (C1) and supporting his cabinet (C2)

The detailed results of the first category of analysis concerning the contents that refer to Trump’s policies and cabinet reveal a personalist agenda, in which his own figure as president stands out, occupying 24.7% of the mentions (Figure 4). Self-referral in his discourse mainly concerns defending himself through accusations aimed at the media and the opposition, which he characterises as haters. Worth noting is a high level of personalism, with such phrases as “your favourite president”, “a true champion of civil rights” or “your all-time favourite duly elected President, me!” This result is further confirmed through the analysis of the messages concerning the electoral campaign (16.3%), where he aggrandises his own figure using positive polling results, to which he refers as the “the polls that matter”, as opposed to the “fake news suppression polls” that do not show him as the winner. This praise of his own person can also be seen in other aspects of his political administration, when he attributes to himself the good handling of foreign policy (22.5%): “there has never been a president who has been tougher (but fair) on China or Russia”; positive numbers concerning the economy and employment (6.38%): “my Administration and I built the greatest economy in history (...) saved millions of lives”; or for his handling of immigration (3.77%).

“ The discursive tactic of the former president reveals a use of his institutional figure for the sake of political personalization and polarization in the absence of filters or journalistic verification ”

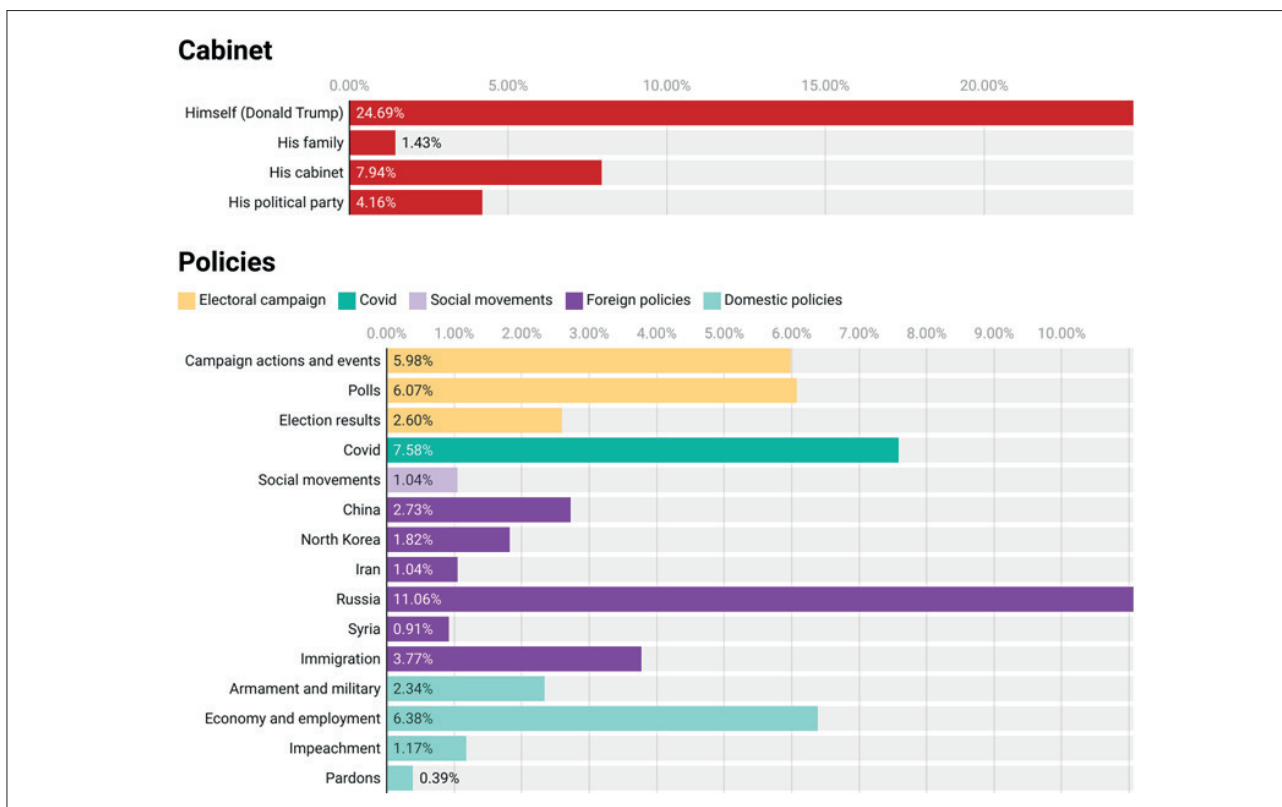


Figure 4. Details of Donald Trump's thematic fake news agenda with respect to his policies and cabinet

4.3. Relation with the opposition: politicians (C2) and discrediting of the media (C3)

The ex-president relates his discourse on fake news to the political opposition, but also to the media that are not like-minded (Figure 5). In the first case (20.7%), he refers to the opposition in mainly a generalised way, alluding to the *Democrat Party* as a whole (13.2%). As for particular candidates, Trump focuses on Joe Biden (4.6%) as his main adversary in the following General Elections of 2020. To a lesser extent and during the first half of his term of office, he also refers to Obama (2.2%) and Hillary Clinton (2%) as leaders of the previous opposition. Trump dedicates derogatory expressions and epithets to all of them: “democrats of the radical left who know nothing,” “crooked Hillary Clinton,” “slow Joe,” “Obamagate,” “illegal democrat witch hunt.”

On the other hand, his allusions to the media, either in general (19.5%) or aimed at particular targets (33.1%) are present in more than half the messages published by Trump with respect to fake news (52.6%). He uses the term indistinctly to refer to false news and the media itself, which he calls “the real opposition party” and directly links them to the *Democrat Party* by using such expressions as “the “fake news” media,” defining them as their “partner in crime,” “lap dog,” “their vehicle,” or the “illegal democrat/fa news media partnership.” In the same way, he criticises the use of unrevealed sources and encourages his audience not to believe such news (Figure 6): “The most often used phrase in the Lamestream Media, by far, is ‘sources say’, or ‘officials who spoke on the condition of anonymity’ (...) which allows Fake News to make up a phony quote from a person who doesn’t even exist”.

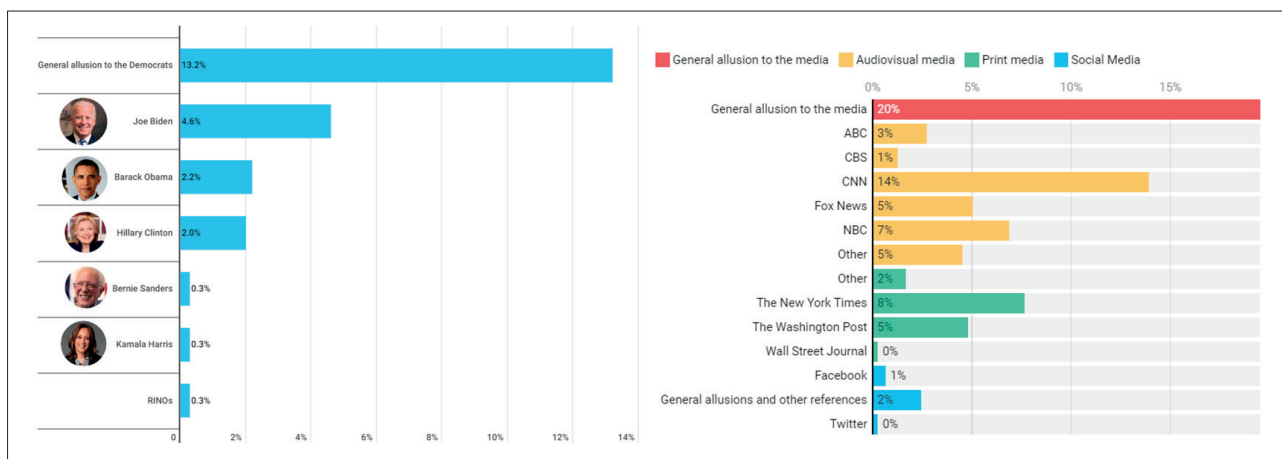


Figure 5. Trump's accusations of fake news against the political opposition and the media

We have detected more references to the audiovisual media than to the written media and the social media. In the first case, Trump focuses on *CNN* (13.9%), and in particular, the ex-presenter Chris Cuomo. Similarly, he considers *NBC* (6.9%), *ABC* (2.7%) and *CBS* (1.3%) to be opposition media. His relationship with *Fox News* (5.1%), however, changes over time.

In the case of the written media, he focuses on *The New York Times* (7.7%) and *The Washington Post* (4.8%). Allusions to social media (3.8%) indicate the relevance he gives them in spreading his discourse outside of the traditional media: "I use social media not because I like it, but because it is the only way to fight against a VERY dishonest and unfair 'press,' now often referred to as the Fake News Media;" "Fake News Media coverage of me is negative, with numerous forced retractions of false stories. Hence my use of social media, the only way to get the truth out."

This position changes when *Twitter* blocks profiles or like-minded messages: "they and the fake news, working together, want to silence the truth;" "*Twitter* is interfering in the 2020 Presidential Election".

4.4. Emotions and polarity in the discourse of Donald Trump

The results of the sentiment analysis allow us to qualify the polarity of Trump's messages in his digital narrative. The combination of the two methods (dictionaries and machine learning) and the three tools (*Afinn*, *GNL* and *Amazon Comprehend*) allow us to obtain detailed comparisons of the contents of his discourse. Figure 7 offers two types of data: on the left, the complete results of each tool in accordance with the particular variables offered by each one; and on the right, the results set out under three variables (negative, positive, or neutral), in which the values of Very Negative to Negative and Very Positive to Positive from *GNL* and *Afinn* are grouped so as to be able to compare them directly with *Amazon Comprehend*, which adds its own variable of Mixed.

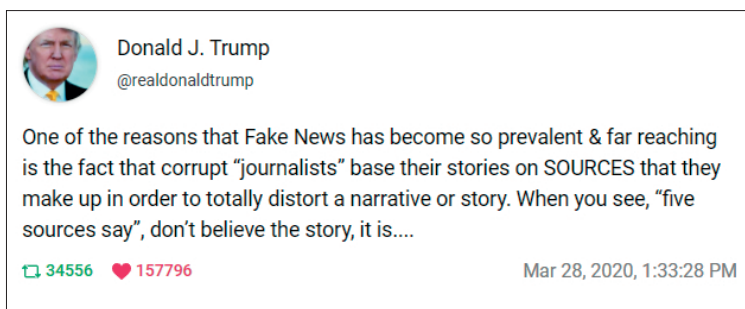


Figure 6. Tweet from Donald Trump accusing journalists of being "corrupt" and of falsifying their sources.

Source: *The Trump Archive*

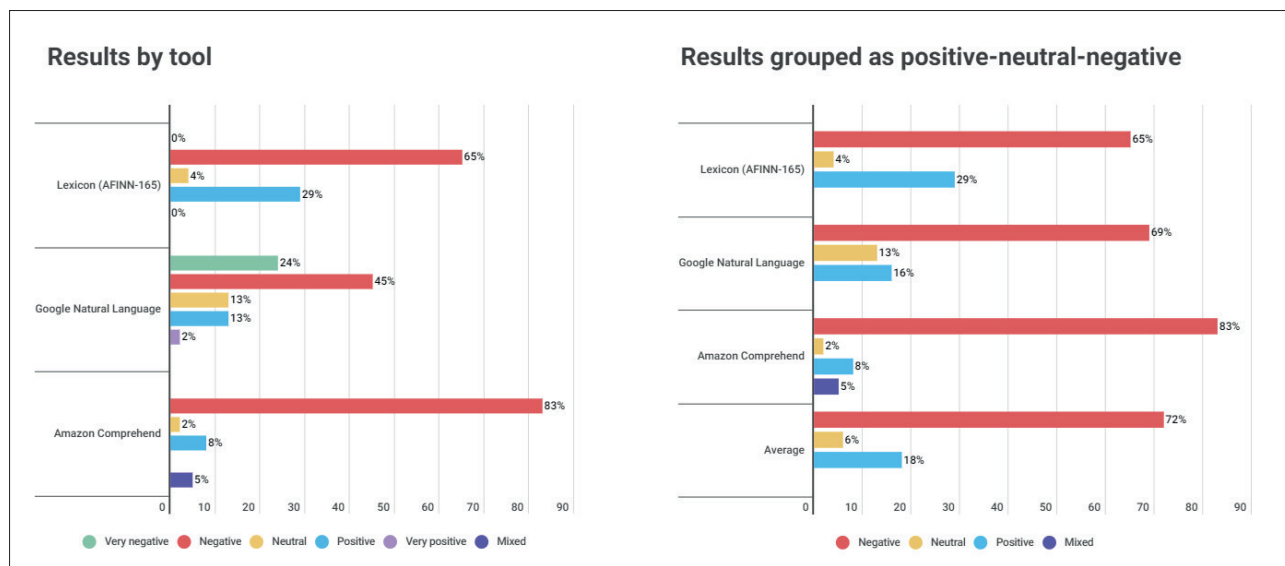


Figure 7. Analysis of emotional polarity analysis and emotional load in the fake news discourse of Donald Trump

As a whole, we can observe a mainly negative polarity in all three services, oscillating between 65% and 84%; a minority of positive between 26% and 8%; and a residual of neutrality, with 14% being the highest. These results demonstrate Trump's hostile tone which, combined with the previous thematic analysis, he uses both to attack others and to defend himself. It is a discourse with a negative tone that comes out to his global audience, directly, without filters.

The interrelation between the emotions analysis and the codified thematic categories in the content analysis reflects concrete disparities with respect to the negativity expressed by Donald Trump in his digital discourse. The figure is lower when referring to his policies (with an average polarity index of -0.0878) and cabinet (-0.0594), but increases his negative emotional load in the case of his political opposition (-0.116) and, in particular, the media (-0.1776).

6. Note

1. The dataset is available on:
<https://doi.org/10.5281/zenodo.8414751>

7. References

- Albright, Jonathan** (2017). "Welcome to the Era of Fake News". *Media and communication*, v. 5, n. 2, pp. 87-89.
<https://doi.org/10.17645/mac.v5i2.977>
- Barbu, Oana** (2014). "Advertising, microtargeting and social media". *Procedia-social and behavioral sciences*, v. 163, pp. 44-49.
<https://doi.org/10.1016/j.sbspro.2014.12.284>
- Beer, David; Burrows, Roger** (2007). "Sociology and, of and in web 2.0: some initial considerations". *Sociological research online*, v. 12, n. 5.
<https://doi.org/10.5153/sro.1560>
- Berrocal-Gonzalo, Salomé; Campos-Domínguez, Eva; Redondo-García, Marta** (2014). "Media prosumers in political communication: politainment on YouTube". *Comunicar*, v. 43, n. XXII, pp. 65-72.
<https://doi.org/10.3916/C43-2014-06>
- Berrocal-Gonzalo, Salomé; Quevedo, Raquel; García-Beaudoux, Virginia** (2022). "Pop online politics: new strategies and leaderships for new audiences". *Index comunicación*, v. 12, n. 1, pp. 13-19.
<http://hdl.handle.net/10115/18489>
- Bleakley, Paul** (2018). "Situationism and the recuperation of an ideology in the era of Trump, fake news and post-truth politics". *Capital & class*, v. 42, n. 3, pp. 419-434.
<https://doi.org/10.1177/0309816818759231>
- Borgesius, Frederik-Zuiderveen; Möller, Judith; Kruikemeier, Sanne; Fathaigh, Ronan Ó.; Irion, Kristina; Dobber, Tom; Bodó, Balázs; De-Vreese, Claes H.** (2018). "Online political microtargeting: Promises and threats for democracy". *Utrecht law review*, v. 14, n. 1, pp. 82-96.
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3128787
- Boshmaf, Yazan; Muskukhov, Ildar; Beznosov, Konstantin; Ripeanu, Matei** (2011). "The socialbot network: when bots socialize for fame and money". *Twenty-seventh annual computer security applications conference* [conference session], December 5th-9th. Orlando, USA.
<https://bit.ly/3YV2nGQ>
- Bradshaw, Samantha; Howard, Philip; Kollanyi, Bence; Neudert, Lisa-Maria** (2020). "Sourcing and automation of political news and information over social media in the United States, 2016-2018". *Political communication*, v. 37, n. 2, pp. 173-193.
<https://doi.org/10.1080/10584609.2019.1663322>
- Cabezuelo-Lorenzo, Francisco; Manfredi, Juan-Luis** (2019). "Post-truth, fake-news and political agenda in Trump's speech on Twitter". *Historia y comunicación social*, v. 24, n. 2, pp. 471.
<https://doi.org/10.5209/hics.66291>
- Cano-Orón, Lorena; Calvo, Dafne; Llorca-Abad, Germán; Mestre-Pérez, Rosanna** (2021b). "Media crisis and disinformation: the participation of digital newspapers in the dissemination of a denialist hoax". *Profesional de la información*, v. 30, n. 4.
<https://doi.org/10.3145/epi.2021.jul.12>
- Cano-Orón, Lorena; Calvo, Dafne; López-García, Guillermo; Baviera, Tomás** (2021a). "Disinformation in Facebook ads in the 2019 Spanish general election campaigns". *Media & communication*, v. 9, n. 1, pp. 217-228.
<https://doi.org/10.17645/mac.v9i1.3335>
- Carlson, Matt** (2018). "Automating judgment? Algorithmic judgment, news knowledge, and journalistic professionalism". *New media & society*, v. 20, n. 5, pp. 1755-1772.
<https://doi.org/10.1177/1461444817706684>
- Carmi, Elinor; Yates, Simeon J.; Lockley, Eleanor; Pawluczuk, Alicja** (2020). "Data citizenship: rethinking data literacy in the age of disinformation, misinformation and malinformation". *Internet policy review*, v. 9, n. 1.
<https://www.econstor.eu/handle/10419/218938>
- Castells, Manuel** (2009). *Comunicación y poder*. Alianza Editorial. ISBN: 978 84 2068499 4
- Chandra, Rohitash; Saini, Ritij** (2021). "Biden vs Trump: modelling US general elections using BERT language model". *IEEE access*, v. 9.
<https://doi.org/10.1109/ACCESS.2021.3111035>
- Cheng, Eviane** (2019). "Immigrant, nationalist and proud: A Twitter analysis of Indian diaspora supporters for Brexit and Trump". *Media and communication*, v. 7, n. 1.
<https://doi.org/10.17645/mac.v7i1.1629>

- Cinelli, Matteo; De-Francisci-Morales, Gianmarco; Galeazzi, Alessandro; Quattrociocchi, Walter; Starnini, Michele** (2021). "The echo chamber effect on social media". *Proceedings of the National Academy of Sciences*, v. 118, n. 9.
<https://doi.org/10.1073/pnas.2023301118>
- Coleman, Stephen** (2005). "New mediation and direct representation: reconceptualising representation in the digital age". *New media & society*, v. 7, n. 2, pp. 177-198.
<https://doi.org/10.1177/1461444805050745>
- Colonescu, Constantin** (2018). "The effects of Donald Trump's tweets on US financial and foreign exchange markets". *Athens journal of business & economics*, v. 4, n. 4, pp. 375-388.
<https://doi.org/10.30958/ajbe.4-4-2>
- Courty, Audrey** (2021). "Despite being permanently banned, Trump's prolific Twitter record lives on". *The conversation*, 13th January.
<https://theconversation.com/despite-being-permanently-banned-trumps-prolific-twitter-record-lives-on-152969>
- Dang, Nhan-Cach; Moreno-García, María; De-La-Prieta, Fernando** (2020). "Sentiment analysis based on deep learning. A comparative study". *Electronics*, v. 9, n. 3.
<https://doi.org/10.3390/electronics9030483>
- Das-Sarma, Matthew** (2016). "Tweeting 2016: how social media is shaping the presidential election". *Inquiries*, v. 8, n. 9.
<http://www.inquiriesjournal.com/articles/1454/tweeting-2016-how-social-media-is-shaping-the-presidential-election>
- Dwianto, Rahmad-Agus; Nurmandi, Achmad; Salahudin, Salahudin** (2021). "The sentiments analysis of Donald Trump and Jokowi's twitters on covid-19 policy dissemination". *Webology*, v. 18, n. 1, pp. 389-405.
<https://doi.org/10.14704/WEB/V18I1/WEB18096>
- Enguix-Oliver, Salvador** (2020). "Efectos de la digitalización en la relación entre periodistas y políticos: desintermediación y pactos comunicativos". *Pasajes*, v. 59, pp. 19-30.
<https://roderic.uv.es/bitstream/handle/10550/76566/7550103.pdf>
- Ermakova, Tatiana; Henke, Max; Fabian, Benjamin** (2021). "Commercial sentiment analysis solutions: a comparative study". *Webist 2021*, pp. 103-114.
<https://doi.org/10.5220/0010709400003058>
- Farkas, Johan; Schou, Jannick** (2020). *Post-truth, fake news and democracy: mapping the politics of falsehood*. Routledge. ISBN: 978 0 36732217 5
- Fernández-Castrillo, Carolina** (2014). "Prácticas transmedia en la era del prosumidor: hacia una definición del contenido generado por el usuario (CGU)". *Cuadernos de información y comunicación*, n. 19, pp. 53-67.
https://doi.org/10.5209/rev_CIYC.2014.v19.43903
- Fincham, Kelly** (2019). "Exploring political journalism homophily on Twitter: a comparative analysis of US and UK elections in 2016 and 2017". *Media and communication*, v. 7, n. 1, pp. 213-224.
<https://doi.org/10.17645/mac.v7i1.1765>
- Froehlich, Thomas J.** (2020). "A disinformation-misinformation ecology: the case of Trump". In: Visnovský, Ján; Radosinká, Jana (eds.). *Fake news is bad news. Hoaxes, half-truths and the nature of today's journalism* (pp. 13-58). IntechOpen. ISBN: 978 1 83962 423
- García-Orosa, Berta** (2021). "Disinformation, social media, bots, and astroturfing: the fourth wave of digital democracy". *Profesional de la información*, v. 30, n. 6.
<https://doi.org/10.3145/epi.2021.nov.03>
- Gillespie, Tarleton** (2010). "The politics of 'platforms'". *New media & society*, v. 12, n. 3, pp. 347-364.
<https://doi.org/10.1177/1461444809342738>
- Gómez-García, Salvador; Gil-Torres, Alicia; Carrillo-Vera, José-Agustín; Navarro-Sierra, Nuria** (2019). "Constructing Donald Trump: mobile apps in the political discourse about the president of the United States". *Comunicar*, v. 27, n. 59.
<https://doi.org/10.3916/C59-2019-05>
- Grossman, Lawrence** (1995). *The electronic republic: reshaping democracy in the information age*. Viking. ISBN: 978 0 670 86129 3
- Happer, Catherine; Hoskins, Andrew; Merrin, William** (2018). "Weaponizing reality: an introduction to Trump's war on the media". In: Hoskins, Andrew; Merrin, William (eds.). *Trump's media war* (pp. 3-22). Palgrave Macmillan. ISBN: 978 3 319 94069 4
- Hernández-Serrano, María-José; Renés-Arellano, Paula; Graham, Gary; Greenhill, Anita** (2017). "From prosumer to prodesigner: participatory news consumption". *Comunicar*, v. 25, n. 50, pp. 77-88.
<https://doi.org/10.3916/C50-2017-07>

- Journell, Wayne** (2017). "Fake news, alternative facts, and Trump: Teaching social studies in a post-truth era". *Social studies journal*, v. 37, n. 1, pp. 8-21.
<https://pcssonline.org/wp-content/uploads/2017/06/2017-Spring-PCSS-Journal.pdf>
- Kapantai, Eleni; Christopoulou, Androniki; Berberidis, Christos; Peristeras, Vassilios** (2020). "A systematic literature review on disinformation: toward a unified taxonomical framework". *New media & society*, v. 23, n. 5, pp. 1301-1326.
<https://doi.org/10.1177/1461444820959296>
- Kaur, Manpreet; Verma, Rajesh; Otoo, Frank-Nana-Kweku** (2021). "Emotions in leader's crisis communication: Twitter sentiment analysis during Covid-19 outbreak". *Journal of human behaviour in the social environment*, v. 31, n. 1-4, pp. 362-372.
<https://doi.org/10.1080/10911359.2020.1829239>
- Kotras, Baptiste** (2020). "Mass personalization: Predictive marketing algorithms and the reshaping of consumer knowledge". *Big data & society*, v. 7, n. 2.
<https://doi.org/10.1177/2053951720951581>
- Krippendorff, Klaus** (2004). *Content analysis: an introduction to its methodology*. SAGE.
- Lokot, Tetyana; Diakopoulos, Nicholas** (2015). "News bots". *Digital journalism*, v. 4, n. 6, pp. 682-699.
<https://doi.org/10.1080/21670811.2015.1081822>
- López-Jiménez, Gloria** (2014). "El proceso de desintermediación comunicativa". *Revista internacional del mundo económico y del derecho*, v. VII, pp. 69-91.
<https://bit.ly/2I2Rhuy>
- Machus, Tobias; Mestel, Roland; Theissen, Erik** (2022). "Heroes, just for one day: the impact of Donald Trump's tweets on stock prices". *Journal of behavioral and experimental finance*, v. 33.
<https://doi.org/10.1016/j.jbef.2021.100594>
- Magallón-Rosa, Raúl** (2018). "La biblioteca digital sobre Donald Trump. Fact-checking frente a fake news". *Estudios sobre el mensaje periodístico*, v. 24, n. 1, pp. 273-282.
<https://doi.org/10.5209/ESMP.59949>
- Meeks, Lindsey** (2019). "Defining the enemy: how Donald Trump frames the news media". *Journalism & mass communication quarterly*, v. 97, n. 1, pp. 211-234.
<https://doi.org/10.1177/1077699019857676>
- Meirick, Patrick C.; Franklyn, Amanda E.** (2022). "Seeing and believing pro-Trump fake news: the interacting roles of online news sources, partisanship, and education". *International journal of communication*, v. 16, pp. 3379-3401.
<https://ijoc.org/index.php/ijoc/article/view/18579>
- Mencarelli, Alberto** (2021). "Parliaments facing the virtual challenge: a conceptual approach for new models of representation". *Parliamentary affairs*, n. October.
<https://doi.org/10.1093/pa/gsab052>
- Milan, Stefania** (2015). "When algorithms shape collective action: Social media and the dynamics of cloud protesting". *Socialmedia+society*, v. 1, n. 2.
<https://doi.org/10.1177/2056305115622481>
- Misiejuk, Kamila; Scianna, Jennifer; Kaliisa, Rogers; Vachuska, Karl; Williamson-Schaffer, David** (2021). "Incorporating sentiment analysis with epistemic network analysis to enhance discourse analysis of twitter data". In: Ruis, A. R.; Lee, S. B. (eds.). *Advances in quantitative ethnography* (pp. 375-389). Springer. ISBN: 978 3 0303323 10
- Morris, Dick** (2001). "Direct democracy and the internet". *Loyola of Los Angeles law review*, v. 34, n. 3, pp. 1033-1053.
<https://digitalcommons.lmu.edu/cgi/viewcontent.cgi?referer=httpsredir=1&article=2280&context=llr>
- Murthy, Dhiraj; Powell, Alison B.; Tinati, Rramine; Anstead, Nick; Carr, Leslie; Halford, Susan J.; Weal, Mark** (2016). "Automation, algorithms, and politics. Bots and political influence: A sociotechnical investigation of social network capital". *IJOC*, n. 10, v. 2016, pp. 4952-4971.
<https://ijoc.org/index.php/ijoc/article/view/6271>
- Muswede, Thabiso** (2022). "Impact of digital media on political campaigns. Towards disintermediated political communication in Africa". *African journal of political science*, v. 10, n. 1, pp. 41-50.
<https://doi.org/10.36615/ajpsrasp.v10i1.1152>
- Papakyriakopoulos, Orestis; Hegelich, Simon; Shahrezaye, Morteza; Medina-Serrano, Juan-Carlos** (2018). "Social media and microtargeting: Political data processing and the consequences for Germany". *Big data & society*, v. 5, n. 2.
<https://doi.org/10.1177/2053951718811844>

- Pariser, Eli** (2011). *The filter bubble: what the Internet is hiding from you*. Penguin. ISBN: 978 1 59420300 8
- Pérez-Salazar, Gabriel** (2011). "La web 2.0 y la sociedad de la información". *Revista mexicana de ciencias políticas y sociales*, v. 56, n. 212, pp. 57-68.
https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0185-19182011000200004
- Pfister, Damien-Smith** (2011). "Networked expertise in the era of many-to-many communication: on Wikipedia and invention". *Social epistemology*, v. 25, n. 3, pp. 217-231.
<https://doi.org/10.1080/02691728.2011.578306>
- Piñuel-Raigada, José-Luis** (2002). "Epistemología, metodología y técnicas del análisis de contenido". *Estudios de socio-lingüística*, v. 3, 42 pp.
<https://cutt.ly/4wbptUNp>
- Quealy, Kevin** (2021). "The complete list of Trump's Twitter insults (2015-2021)". *The New York Times*, 19th January.
<https://nyti.ms/3Y3XDgR>
- Raynauld, Vincent; Turcotte, André** (2018). "Different strokes for different folks: implications of voter micro-targeting and appeal in the age of Donald Trump". In: Gillies, J. (ed.). *Political marketing in the 2016 US presidential election* (pp. 11-28). Springer. ISBN: 978 3 31 959345 6
- Ribeiro, Felipe N.; Saha, Koustuv; Babaei, Mahmoudreza; Henrique, Lucas; Messias, Johnnatan; Benevenuto, Fabricio; Gaga, Oana; Gummadi, Krishna P.; Redmiles, Elissa M.** (2019). "On microtargeting socially divisive ads: A case study of Russia-linked ad campaigns on Facebook". *Proceedings of FAT*19*, pp. 140-149.
<https://doi.org/10.1145/3287560.3287580>
- Ritzer, George** (2015). "Automating prosumption: The decline of the prosumer and the rise of the prosuming machines". *Journal of consumer culture*, v. 15, n. 3, pp. 407-424.
<https://doi.org/10.1177/1469540514553717>
- Ritzer, George; Dean, Paul; Jurgenson, Nathan** (2012). "The coming of age of the prosumer". *American behavioural scientist*, v. 56, n. 4, pp. 379-398.
<https://doi.org/10.1177/0002764211429368>
- Ross, Andrew S.; Rivers, Damian J.** (2018). "Discursive deflection: Accusation of 'fake news' and the spread of mis- and disinformation in the tweets of President Trump". *Social media+society*, v. 4, n. 2.
<https://doi.org/10.1177/2056305118776010>
- Rossetti, Michael; Zaman, Tauhid** (2023). "Bots, disinformation, and the first impeachment of U.S. President Donald Trump". *Plos one*, v. 18, n. 5.
<https://doi.org/10.1371/journal.pone.0283971>
- Rossini, Patricia; Stromer-Gallery, Jennifer; Korsunskaja, Ania** (2021). "More than 'fake news'? The media as malicious gatekeeper and a bully in the discourse of candidates in the 2020 U.S. presidential election". *Journal of language and politics*, v. 20, n. 5, pp. 676-695.
<https://doi.org/10.1075/jlp.21033.ros>
- Rubio-Fabián, Roberto** (2019). "La peligrosa desintermediación política". *La prensa gráfica*, v. 104, n. 6, pp. 46.
<http://www.repo.funde.org/id/eprint/1552>
- Sadia, Azeema; Khan, F.; Bashir, Fatima** (2018). "An overview of lexicon-based approach for sentiment analysis". *3rd International electrical engineering conference*, 5 pp.
https://ieec.neduet.edu.pk/2018/Papers_2018/15.pdf
- Sánchez-García, Pilar; Merayo-Álvarez, Noemí; Calvo-Barbero, Carla; Díez-Gracia, Alba** (2023). "Spanish technological development or artificial intelligence applied to journalism: Companies and tools for documentation, production and distribution of information". *Profesional de la información*, v. 32, n. 2.
<https://doi.org/10.3145/epi.2023.mar.08>
- Singh, Narishia; Wijegunawardhana, Anuradha** (2018). "Impact of big data and political microtargeting on Donald Trump's 2016 presidential campaign". *International research conference*, pp. 28-34.
<http://ir.kdu.ac.lk/handle/345/2485>
- Smith, Christopher A.** (2019). "Weaponized iconoclasm in Internet memes featuring the expression 'fake news'". *Discourse & communication*, v. 13, n. 3.
<https://doi.org/10.1177/1750481319835639>
- Tamul, Daniel J.; Ivory, Adrienne-Holz; Hotter, Jessica; Wolf, Jordan** (2019). "All the president's tweets: effects of exposure to Trump's 'fake news'. Accusations on perceptions of journalists, news stories, and issue evaluation". *Mass communication and society*, v. 23, n. 3.
<https://doi.org/10.1080/15205436.2019.1652760>

- Timoshenko, Aartem; Hauser, John** (2019). "Identifying customer needs from user-generated content". *Marketing science*, v. 38, n. 1.
<https://doi.org/10.1287/mksc.2018.1123>
- Tong, Chau; Gill, Hyungjin; Li, Jianing; Valenzuela, Sebastián; Rojas, Hernando** (2020). "Fake news is anything they say!" – Conceptualization and weaponization of fake news among the American public". *Mass communication and society*, v. 23, n. 5, pp. 755-778.
<https://doi.org/10.1080/15205436.2020.1789661>
- Van-Dijck, José** (2016). *La cultura de la conectividad: una historia crítica de las redes sociales*. Siglo Veintiuno. ISBN: 978 987 629 669 4
- Van-Dijk, Teun** (1990). *La noticia como discurso. Comprensión, estructura y producción de la información*. Paidós. ISBN: 84 7509 622 0
- Waheed, Anum; Salam, Abdus; Bangash, Javed-Iqbal; Bangash, Maria** (2021). "Lexicon and learn-based sentiment analysis for web spam detection". *IEEESEM*, v. 9, n. 3.
https://www.ieeesem.com/researchpaper/Lexicon_and_Learn_Based_Sentiment_Analysis_for_Web_Spam_Detection.pdf
- Wardle, Claire; Derekshan, Hossein** (2017) *Information disorder: toward an interdisciplinary framework for research and policy making*. Council of Europe report DGI09, 31st October. Brussels: Council of Europe.
<https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-researc/168076277c>
- Wu, Yanfang** (2019). "Is automated journalistic writing less biased? An experimental test of auto-written and human-written news stories". *Journalism practice*, v. 1, n. 21.
<https://doi.org/10.1080/17512786.2019.1682940>
- Xia, Ethan; Yue, Han; Liu, Hongfu** (2021). "Tweets sentiment analysis of the 2020 U.S. presidential election". *Proceedings of WWW'21*.
<https://dl.acm.org/doi/abs/10.1145/3442442.3452322>
- Zimmer, Franziska; Scheibe, Katrin; Stock, Mechtild; Stock, Wolfgang G.** (2019). "Fake news in social media: Bad algorithms or biased users?". *JISTaP*, v. 7, n. 2, pp. 40-53.
<https://doi.org/10.1633/JISTaP.2019.7.2.4>

8. Annexes

Annex 1

Coding sheet for the thematic agenda of Trump's tweets about 'fake news'

Categories	Themes	Sub-themes	
C1. Politics	Domestic policies	Economy and employment	
		Impeachment	
		Armament and military	
		Pardons	
		Immigration	
		Others	
	Foreign policies and international relations	Russia	
		China	
		North Korea	
		Iran	
		Syria	
		Others	
	Social movements and protests		
	Electoral campaign	Surveys	
		Election results	
		Campaign actions or events	
Others			
Coronavirus			

Categories	Themes	Sub-themes
C2. Cabinet	His cabinet	
	His party	
	His family	
	Himself	
C3. Political opposition	<i>Democrat Party</i>	Barack Obama
		Hillary Clinton
		Joe Biden
		Bernie Sanders
		Kamala Harris
		Others
	'RINO' (Republican in name only)	
	Others	
C4. Media and social networks	General allusion to the media	
	Print media	<i>The New York Times</i>
		<i>The Wall Street Journal</i>
		<i>The Washington Post</i>
		Others
	Audio-visual media	CNN
		Fox News
		ABC
		CBS
		NBC
		Others
	Social networks and technology companies	Facebook
		Twitter
		General references and others

Annex 2

Examples of AFINN-165 dictionary values applied to dictionary sentiment analysis

Polarity	Value	Example
Very positive	2.6–5	admire, breath-taking, heavenly
Positive	0.1–2.5	likable, meaningful, motivated
Negative	-0.1 – -2.5	totalitarian, aggression, coerce
Very negative	-2.6 – -5	fraud, torture, hell



SEDIC Rediseñando lo que somos para renovar lo que hacemos

<https://www.sedic.es>
<https://intranetsedic.es>
<https://formacionsedic.online>
 c/Gargantilla 13, local 24 Madrid 28005
 +34 639 186 570 | +34 91 593 40 59
 sedic@sedic.es

 <https://twitter.com/SEDIC20>
 <https://www.facebook.com/AsociacionSEDIC>
 <https://www.linkedin.com/company/sedic/>
 <https://www.instagram.com/sedicasociacion/>

Using video for fact-checking on *Facebook*. Analysis of the trend and reach of Ibero-American production (2016-2021)

Alberto Dafonte-Gómez; Oswaldo García-Crespo; Diana Ramahí-García

Nota: Este artículo se puede leer en español en:
<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87300>

Recommended citation:

Dafonte-Gómez, Alberto; García-Crespo, Oswaldo; Ramahí-García, Diana (2023). "Using video for fact-checking on *Facebook*. Analysis of the trend and reach of Ibero-American production (2016-2021)". *Profesional de la información*, v. 32, n. 5, e320511.

<https://doi.org/10.3145/epi.2023.sep.11>

Manuscript received on 3rd March 2023
Accepted on 24th July 2023



Alberto Dafonte-Gómez
<https://orcid.org/0000-0003-1076-4445>

Universidade de Vigo
Facultad de Comunicación
Campus A Xunqueira, s/n
36005 Pontevedra, Spain
albertodafonte@uvigo.gal



Oswaldo García-Crespo
<https://orcid.org/0000-0002-8886-9266>

Universidade de Vigo
Facultad de Comunicación
Campus A Xunqueira, s/n
36005 Pontevedra, Spain
oswaldogarcia@uvigo.gal



Diana Ramahí-García ✉
<https://orcid.org/0000-0002-8990-5337>

Universidade de Vigo
Facultad de Comunicación
Campus A Xunqueira, s/n
36005 Pontevedra, Spain
dramahi@uvigo.gal

Abstract

Social networks –and *Facebook* in particular– have become an important element of the information diet for millions of people around the world. By using them, the traditional media lose control of the distribution channel for their content, whose reach now depends, firstly, on the relevance criteria established by the networks themselves and, secondly, on the interactions generated by the audience with each publication. Very often research on journalism has tackled the issue of reach and how efficient fact-checkers are. To find answer to the sociodemographic features of their audiences or the characteristics of their posts are explored. However, factors such as the influence of the algorithms which choose the content users are shown on the social networks is not often dealt with. This article aims to contribute in both areas. Firstly, it offers a broad perspective on the publications of Ibero-American fact-checkers on *Facebook* between 2016 and 2021, focuses on the evolution of video production (n=9075) and on the views and engagement achieved by this format with respect to the rest, and relates them to changes in the *News Feed* algorithm. Secondly, it proposes a content analysis to identify formal and thematic elements in the most popular videos in the same period (n=414) and relates them to previous research. Our results show significant similarities in popular videos, but also changes in video production, a generalized decrease in the ratio of views and a drop in the interaction rate more accentuated than in all the publications of the period. Although the focus of this research does not allow us to make direct causal inferences, the trends identified coincide with the changes in the *Facebook News Feed* algorithm that were made public in those years.

Keywords

Fact-checkers; Fact-checking; Algorithms; *Facebook*; Video; Gatekeepers; Social media; Ibero-America; Disinformation; Fake news; Verification; Digital communication.



Funding

This study is part of the research project “Digital narratives against disinformation, a study of networks, themes and formats in Ibero-American fact-checkers / Narrativas digitales contra la desinformación, estudio de redes, temas y formatos en los fact-checkers iberoamericanos”, funded by the *State Research Agency (Agencia Estatal de Investigación)* (Ref. PID2019-108035RB-I00/AEI/ 10.13039/501100011033). Alberto Dafonte-Gómez is member of *Unesco Chair on Transformative Education: Science, Communication and Society* at the *University of Vigo*.

1. Introduction

In recent years the information industry has undergone a change towards more digital, social and mobile consumption. In this new scenario in which smartphones have become the main device for accessing the Internet (Ericsson, 2022) and the news, social networks are increasingly the main source of information for a large amount of people, (Kemp, 2022), especially the youngest users (Newman et al., 2022), both for this purpose, and incidentally, by recommendations and posts from contacts (Fletcher; Nielsen, 2017).

Although conventional media is still highly significant, especially in terms of trustworthiness, the democratising impact of the social media has turned these networks into a space in which they compete for the attention of audiences with alternative sources, such as politicians and influencers (Neuman et al., 2021). In this setting, organisations specialised in fact-checking have become a key part of the news ecosystem and today have reached unprecedented prominence (Guallar et al., 2020). The presence of fact-checkers on social networks often enables their contents to be spread organically (Robertson; Mourão; Thorson, 2020) by interacting with users, which multiply their reach (Margolin; Hannak; Weber, 2017) in the same areas where disinformation circulates (Humprecht, 2020). Moreover, their credibility is raised by means of metrics (Tandoc, 2021). Another point is it also drives traffic to their websites and helps them remain financially viable (Valenzuela; Correa; Gil de Zúñiga, 2017). However, the inclination Facebook users have to identify themselves with profiles which validate their beliefs, as well as the confirmation effect –when the most polarised audiences are exposed to verified contents (Nyhan; Reifler, 2010)– paints a complex picture in which the fact-checkers themselves cannot be certain whether their contents will reach the consumers who have been exposed to the news whose truthfulness has been checked (Ardèvol-Abreu; Delponti; Rodríguez-Wangüemert, 2020; Shawcross, 2016).

Currently, there is no research corpus broad enough to determine the factors which influence the results of the contents dissemination strategy of fact-checkers (Shin; Thorson, 2017). However, references can often be found which stress how important the posting format in general is (Amazeen et al., 2018; Ecker, 2019; Shawcross, 2016), and the video in particular (Elizabeth, 2016; Mantzarlis, 2016; Young, 2018). Nevertheless, there is an external factor which has been given little attention until now, which we believe is interesting: the effect the algorithm (which acts as a gatekeeper for deciding what contents are visible on the platform they are hosted on) has on reach and interaction levels.

2. Facebook as a gatekeeper: chronology of the algorithm criteria for videos

Facebook is still the social network with the greatest number of users worldwide (2.74 billion) (Kemp, 2022), but more significantly, regarding this research, it is the favourite platform for news (Newman et al., 2022) in the global average for users. It is also considered to be the main channel worldwide for spreading disinformation (Newman et al., 2021), since the design of the platform itself encourages the spread of controversial content which is rewarded in terms of users reached and time of use (Horwitz; Seetharaman, 2020). According to data from Dafonte-Gómez, Míguez-González and Ramahí-García (2022), Facebook is also the digital channel with the second most fact-checking initiatives worldwide (72.1%), just behind Twitter (75%).

In recent years different researchers have stressed the role Facebook –and other social networks and search engines– have as gatekeepers (Bozdog, 2013; Powers, 2017; West, 2017), and as intermediaries between the audience and the media, which traditionally has played this role, directly. The networks –and more specifically, their automated algorithms– not only affect what each consumer sees on them, but also determine journalism routines and editorial decisions (Madrigal; Meyer, 2018) in the rat race for visibility, whose aim is to adapt the contents to the criteria which the platform algorithm rewards especially at any given time (Grygiel; Lysak, 2021; Peterson-Salahuddin; Diakopoulos, 2020). DeVito (2017) summarised nine values –from the most to least relevant– which the automated contents filter in the News Feed revolve around: friendships, explicitly expressed user interests, previous participation of the user, implicitly expressed user preferences, age of the post, platform priorities, relationship of the page preferences expressed negatively and content quality.

In June 2014 Facebook announced a change in its algorithm with which native videos (those imported to the platform itself, rather than by links from third parties) would take priority over other types of content, such as links or photos (Welch; Zhang, 2014). This way, videos are viewed without users leaving the platform and in a more favourable setting for Facebook concerning copyright, monetisation and controlling viewer consumption metrics (Tandoc; Maitra, 2018). In the same year the time users spend viewing videos was added as a significant factor, regardless of their reactions. Since then, the video format has become an essential part of Facebook’s corporate strategy (Honan, 2016), and not only determines the content strategy for the media for distributing contents on the social network, but also creates a new way of communicating with its

audiences and interacting with them (Kite *et al.*, 2016; Rein; Venturini, 2018). In 2015, for example, *Facebook* began looking at how to activate sound, viewing with the full screen and in high definition as interesting options apart from reactions to the videos (Wang; Zhuo, 2015). In 2016 *Facebook* extended the alternative reactions to *Like* and the *Live Video*. Two new items were added for consideration in the *News Feed* settings, which, for videos, led to the prioritisation of live broadcasts (Kant; Xu, 2016). This demonstrated the platform was pushing for this format (Meese; Hurcombe, 2020) and, thus, any content that was not aligned with this strategy would lose viewing opportunities. In 2017 the new reactions supplementary to the classic *like* and which had been added in 2016 were prioritised. Moreover, with videos, the rate of finalisation and percent completion (the percentage the consumer views out of the entire duration of each video) were considered to be significant indicators, weighed in terms of their duration— a metric which replaced total viewing time for each video and which had been in operation since 2014 (Bapna; Park, 2017). In the same year they started to show those videos which consumers searched for and frequently revisited on a certain page (Smith, 2017).

Although *Facebook* announced consecutively in 2015 and 2016 that it would prioritise posts from friends in the *News Feed* (Backstrom, 2016), in 2018 a new update went on to give greater priority to posts from contacts in the *News Feed*, so that the organic scope of contents from the pages of institutions and companies began to decline gradually. This obviously had effects on its business model. According to Newberry (2022) the organic scope of a *Facebook* post in 2018 was 7.7% but by the end of 2020 it had fallen to 5.2% and the interaction ratio was at around 0.25% (and around 0.08% for pages with over 100,000 followers).

The way to make the contents from pages more relevant was now either via making paid campaigns, or generating interactions between followers which enabled the contents to be shared organically between contacts and thereby gain greater relevance for the algorithm.

With respect to the video in particular, in 2018, those contents which could capture user attention in under a minute were prioritised in the *News Feed*, and a minimum length of 3 minutes was recommended for monetising them by means of advertising (*Facebook*, 2018), although, the view was that despite video being an important part of the contents ecosystem, consuming them is essentially a passive experience in terms of interacting with the platform (Vogelstein, 2018 citing Adam Mosseri).

In 2019 this principle of capturing user attention in under 1 minute was maintained, but the minimum duration of three minutes (providing that high completed percentages were obtained) became a priority criterion for the *News Feed* (*Facebook*, 2019). That year the trend on *Facebook* was to give greater relevance to posts from close contacts with whom there was a high amount of interaction from users and videos which created organic interaction and conversation among friends and other users; it also maintained the “loyalty” criterion when consuming videos from the same page on a weekly basis in order to give them preference in the *News Feed* (Miller, 2019). In 2021, the year in which the company was renamed *META*, the push for video format became even more decided and constituted one of the main drivers for the social network (*Facebook*, 2021).

In the third quarter of 2022 10.6% of the *News Feed* from American users came from the pages they followed and 16.6% from groups they were part of. Therefore, although posts from friends were the main content –both for original posts (29.7%) and shared ones (19.6%)–, up to 15.2% of what users saw in their *News Feed* came from sources they did not follow, just from recommendations given by the algorithm (*Facebook*, 2023).

Therefore, in this respect, software for automated contents filtering on the social networks is crucial for examining the changes in the information ecosystem model (Epstein; Robertson, 2015; Gillespie, 2014), but constitutes a variable which is difficult to access and understand.

In light of this, it seems pertinent to observe the results of the contents distribution strategies of the fact-checkers and how their users behave on *Facebook*, in so far as they enable us to gain an insight into how a media scenario determined by automatic content filtering works. Moreover, we can observe the capacity fake news has to spread on social networks and how effective fact-checkers are at countering them. We have chosen the geographical scope of the research as Ibero-America: Latin American countries, Spain and Portugal. This is justified by the cultural and linguistic links which unite the countries in this category; the existence of large amounts of collaboration, precisely as a result of cultural affinity and needs. Very often, they face the same problems and need to provide alternative and complimentary outlooks to the academic fields to the dominant global north (Mitchelstein; Boczkowski, 2021), since although these initiatives are relevant, (Ryan, 2022), they have not been studied enough (Molina-Cabañate; Magallón-Rosa, 2021).

3. Research questions

In relation to the foregoing, and as there are no data to take as a starting point, we have created the following research questions:

RQ1. How has video production from Ibero-American fact-checkers evolved on *Facebook*?

RQ2. Has video production from Ibero-American fact-checkers been optimised according to the preferred format and duration criteria established by *Facebook* for showing contents on the *News Feed*?

Since user interaction and duration and video format are all the main criteria *Facebook* has for increasing the chances a post will have of appearing in the *News Feed*:

RQ3. What features of the videos (formal, topical and expressive) and types of interaction can influence the ratios which define how successful *Facebook* content is?

Since the *Facebook* algorithm criteria for making contents more visible has changed over the years, we then posed the question:

RQ4. Have the changes in the settings criteria for the *Facebook News Feed* had a negative effect on the number of views and interactions obtained by the videos posted by Ibero-American fact-checkers?

4. Methodology

The sample is made up of the Ibero-American initiatives which in January 2022 were part of the *International Fact-Checking Network (IFCN)* from the *Poynter Institute*, either as signees or as institutes pending renovation.

The study combines quantitative and qualitative perspectives.

Regarding the quantitative approach, *Crowdtangle* was used to capture the metadata from all the posts made on *Facebook* by fact-checkers from the sample between 2016 and 2021, and obtained a database containing 200,005 entries, out of which 9075 were videos (*Crowdtangle Team*, 2022).

Each registry showed the post dates, followers (likes on the page) when posting (link, live video complete, live video scheduled, native video, photo, status, video and *YouTube* –for non native videos outside *YouTube*–), interactions (likes, comments, shares, love, wow, ha ha, sad angry, care), duration and views (just for native video, live video complete), and interaction ratio, R_i (which is the result of dividing the sum of shares, comments and the 7 possible reactions from a post between the number of likes on a page at the time of posting multiplied by 100) for each registry. Additionally, for the videos posted, we calculated the viewing ratio, R_v ; shares ratio, R_s ; and the comments ratio, R_c , per follower, in a similar way to the calculation made by *Crowdtangle* for the interaction ratio, and the interaction ratio per view, R_{ipv} (number of interactions between the number of views by 100).

Once the database was reviewed we saw there was a high number of posts in 2016 in which there were no data on the followers for the website at the time of posting. This meant we could not have representative data on the viewing and interaction ratios. Therefore, we used the 2016 data to plot the general trend in the number of posts, video category and durations in the broadest way possible and to keep to the 2017-2021 period for all that concerned with calculating ratios and, as a result, the make up of the posts whose contents we wanted to analyse.

With these data, by means of the different metrics (due to their robustness against the average for a set of data with notable deviations), and percentages we analysed their year-on-year trend throughout the period, and studied the correlations between variables with the R statistics programme.

For the qualitative analysis we chose a sample of 5% of the videos with the greatest interaction ratio for each year for every fact-checker between 2017 and 2021, including all the possible categories. This came to 414 videos for the analysis, out of which 391 corresponded to formats with comprehensive metrics.

This relationship underwent a content analysis from a coding sheet. This composition was fundamentally based on already validated coding such as that by **Míguez-González** and **Dafonte-Gómez** (2022) or **García-Marín** and **Salvat-Martinrey** (2022). For the other cases inductive coding was used, given the specific nature of the material and the aim of the research, no sources were found with classifications which fit the contents analysed and enabled deductive coding. After the three researchers initially encoded on an individual basis, they reached a consensus on their findings and this was reflected in a final categorisation.

For each video the resulting analysis template was applied, in which topical and expressive aspects were taken into consideration: content (science, culture, disinformation, ecology, economics, education, gender-LGTBI, politics-laws, racism-xenophobia, health, events, miscellaneous), intention (literacy, self-promotion, rebuttal, information, positive verification, others), scope (national, international), aspect ratio (horizontal, vertical, square), number of shots (various, sole shot) and presence or absence of: graphics, subtitles, visual identification elements from the fact-checker, voice over, presenters, experts, and level of comprehension with no sound which may be total, partial or zero/very low.

The three researchers encoded independently. The differences in classification for the category variables were then resolved by agreement.

Having collected the data, descriptive and inferential statistical analysis were carried out with the R statistical software.

“ In recent years different researchers have stressed the role *Facebook* –and other social networks and search engines– have as gatekeepers, and as intermediaries between the audience and the media, which traditionally has played this role, directly ”

5. Results

5.1. Trend in video production 2016-21

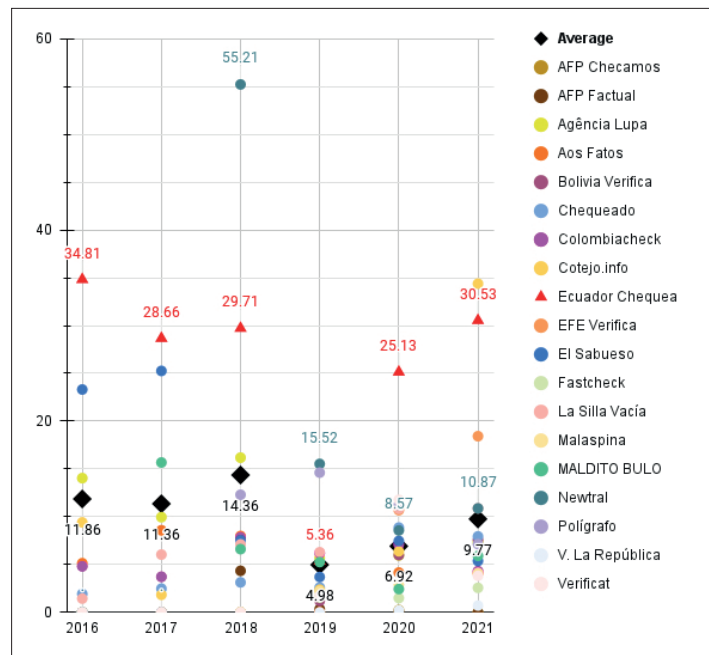
Table 1. Annual trend in posts and videos

	2016	2017	2018	2019	2020	2021
Fact-checkers	8	10	12	17	19	19
Total amount of posts	7757	12248	14634	22469	30051	36359
Average number of posts per fact-checker	969.62	1224.8	1219.5	1321.7	1581.63	1913.63
Average growth in posts	-	26.32%	-0.43%	8.38%	19.77%	20.99%
Total number of videos	284	711	1148	1463	2363	3106
Average number of videos per fact-checker	35.5	71.1	95.66	86.06	124.37	163.47
Average growth in videos	-	100.28%	34.54%	-10.04%	44.52%	31.44%
Video/posts ratio	3.66%	5.81%	7.84%	6.51%	7.86%	8.54%

As shown in table 1, throughout the study period, the number of fact-checkers belonging to the *IFCN* in Ibero-America and on *Facebook* rose notably, from 8 registries in 2016 to 19 in 2021. This increase came with a 368.72% rise in the overall number of posts between 2016 and 2021, although we must focus our analysis on the average per fact-checker.

Between 2016 and 2021 the average number of posts per fact-checker went from 969.62 to 1913.63, which shows a 97.36% rise over 6 years. Especially outstanding in this period is the rise in 2016 and 2017 (at 26.32%, it was the greatest year-on-year one).

As for the videos posted, we went from an average of 35.5 per fact-checker in 2016 to 163.47 in 2021. This shows a jump of 360.47% in the average number of videos posted over the 6 years studied. After a 100.28% growth between 2016 and 2017 and a 10% fall in 2019 vis-à-vis 2018, in 2020 and 2021 video posts increased by 44.52% and 31.44% respectively.



Graph 1. Percentage of videos broadcast by each fact-checker with respect to the total number of posts and regarding the annual average (n = 9075)

Although the percentage of videos posted shows a lower volume (under 10% in all the years studied) in relation to the total number of posts, the reality is there was a slow but sustained growth between 2016 (3.66%) and 2021 (8.54%). Therefore, the growth rate for videos, –above 30% except in 2019– surpassed that for the posts as a whole.

On analysing the level of video posts from each fact-checker, we saw that the push for this format varied widely. In absolute terms, just three fact-checkers reached 1000 videos posted within the 6-year window analysed: *Newtral* (2178, having started in 2018); *La Silla Vacía* (2042) and *Ecuador Chequea* (1015). However, the figures for video posting from each fact-checker must be put into context with respect to their own capacity to produce contents. If we establish the average percentage of videos for each fact-checker for each year analysed in terms of their total posts and we take that as the comparative reference point (graph 1) we see that *Ecuador Chequea* is the checker which had the strongest push for video within its contents plan. Also remarkable in this respect was *Newtral* whose figures for this fact-checker were above the average for the 4 years in which videos were posted, surpassing *Agência Lupa*, *La Silla Vacía*, *Polígrafo* and *EFE Verifica* –despite only recently starting– and *El Sabueso* in 2016 and 2017.

It must be pointed out that *Cotejo* (Venezuela) posts on an irregular basis and this affected the results obtained for this fact-checker, since it did not post anything between 25/08/2020 and 12/10/2021.

5.2. Trend in video formats posted 2016-21

The most frequent type of video on *Facebook* is the native one, as shown in graph 2. This is the clearly dominant category over the 6 years of this study, oscillating between a minimum percentage of 64.79% in 2016 and a maximum of 89.08% which the category reached in 2020. The lowest percentage of native video in 2016 coincided with the highest amount of them on *YouTube* (12.32%, which duplicated its second best registry in 2017) and with the second highest video percentage (14.44%) within the series studied.

Although in terms of percentages live videos did not stand out regarding the total number of videos produced, there was a notable rise in them broadcast in the period analysed, especially between 2016, 2017 and 2018 with 24, 73 and 164 videos respectively. In 2020 the number of live videos surpassed 180; and in 2021 they reached 275 with an ever higher amount of *Live Video Scheduled*.

Despite the growth in production of *Live Video* between 2016 and 2021 was around 1045.83% this observation must be qualified by the fact that this rise largely came

from the activity of a small group of initiatives which seemingly pushed especially for the format. Between 2016 and 2017 still only slightly over half the fact-checkers published *Live Video* (62.5% and 60% respectively) but in 2018 this figure jumped to 83.3%. Thus, at the same time, at 14.3% the peak value for *Live Video* was reached with respect to the total number of videos posted for the whole series analysed. From then onwards, the proportion of fact-checkers which posted *Live Video* dropped continuously: 52.9% in 2019, 42.1% in 2020 and 36.8% in 2021. *La Silla Vacía* is the most consistent of them all within the period studied and accounts for most live videos posted throughout the years: it produced 37.5% of all *Live Videos* in 2016, 67% in 2017, 38.4% in 2018, 54.7% in 2019, 53.8% in 2020 and 68% in 2021.

As for video links outside *YouTube*, these were strongly conditioned by one fact-checker, *Polígrafo*, which used the website *Sapo* to host its external videos and in 2019 accounted for 79% of the 269 videos posted in this category (the rest were from *Newtral*), while in 2020 it was the only fact-checker which used it.

Adding *YouTube* videos is also a minority activity and there was an average of 42.16 videos per year between 2016 and 2020. No specific trend was seen apart from the sharp percentage fall between 2016 and 2017 (from 12.32% to 5.39%).

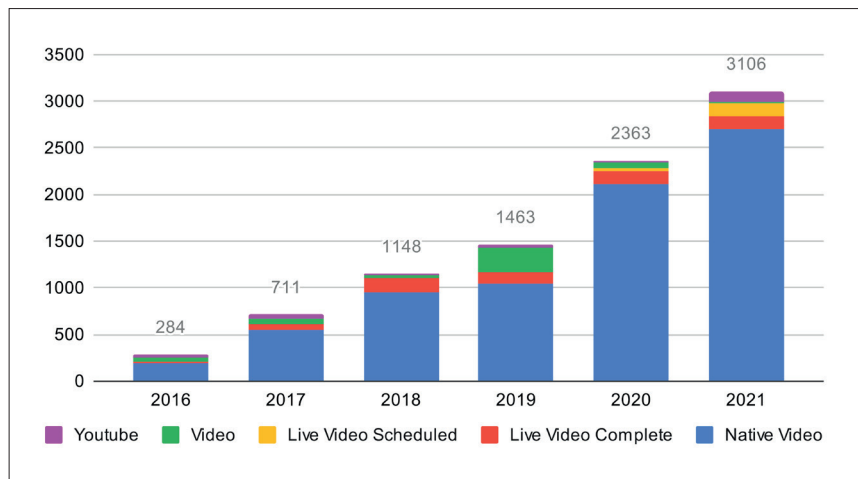
5.3. Trend in video duration

As we mentioned, *Facebook* just provides the duration of native or direct videos. These type of videos come to a total of 8369 within the study period and we have all the data for 8187 of them. This comes to 90.13% of the total number of videos posted under any format. In this group, 77.87% of the videos posted last under 3 minutes while 48.55% last less than 1 minute, a proportion which annually goes from a minimum of 37.82% in 2018 to a maximum of 58.78% in 2021.

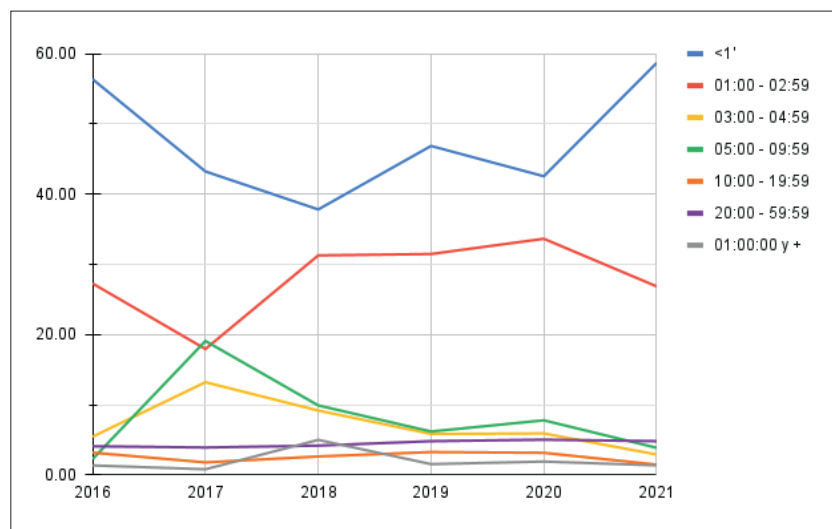
On graph 3, it was observed that those lasting under 1 minute generally had an upward trend in the last years: after a very sharp drop between 2016 and 2018 in which it went from 56.36% to 37.82%, and reached 58.68% in 2021. The period in which there was a fall in video production of under a minute was in 2017 with remarkable growth in those whose duration varied between 05:00 and 09:59 minutos and videos between 03:00 and 04:59. In 2018 these three video categories fell to the benefit of those whose duration was between 10:00 and 19:59, which kept rising until 2020.

Videos over 3 minutes long, have gradually become less prominent over the years and have gone from representing 38.8% of the total amount of video produced in 2017 to 14.5% in 2021, with a slight upturn in 2020.

While short videos are transversal and typical to most verifiers, with those lasting 20 minutes or more there are highly specific formats such as webinars, video conferences and chat sessions in live video format or retransmissions from a *Twitch* channel as native videos which just a small number of fact-checkers implement.



Graph 2. Trend in video posts according to categories (n= 9075)



Graph 3. Trend in video production according to duration (n=8369)

5.4. Viewing and interaction ratios 2017-2021

To calculate the correlations in which different ratios intervened we restricted the analysis to 2017-2021 and once again to native videos and comprehensive live ones. This gave us a sample of 7967 videos, out of which there were valid data for 7327.

Once normal data distribution was ruled out with the Kolmogorov-Smirnov test with the Lilliefors significance correction, Spearman's correlation coefficient was chosen between the different types of interaction, video duration and interaction ratios (Ri), views (Rv) and interaction per view (Ripv), with the following results.

Table 2. Spearman's correlation coefficient between types of interaction, Ri, Rv, Ripv (n=7327)

	Shares	Coms	Likes	Angry	Care	Haha	Love	Sad	Wow	Length	Views	Inter	Ri	Rv	Ripv
Shares	1.000	0.598*	0.827*	0.489*	0.283*	0.402*	0.596*	0.378*	0.496*	0.376*	0.721*	0.883*	0.588*	0.582*	0.328*
Coms	0.598*	1.000	0.663*	0.518*	0.272*	0.585*	0.543*	0.387*	0.473*	0.274*	0.632*	0.760*	0.450*	0.456*	0.247*
Likes	0.827*	0.663*	1.000	0.480*	0.291*	0.460*	0.682*	0.338*	0.500*	0.424*	0.821*	0.958*	0.576*	0.616*	0.308*
Angry	0.489*	0.518*	0.480*	1.000	0.231*	0.436*	0.348*	0.487*	0.496*	0.233*	0.487*	0.541*	0.342*	0.389*	0.128*
Care	0.283*	0.272*	0.291*	0.231*	1.000	0.242*	0.276*	0.202*	0.249*	0.179*	0.242*	0.301*	0.236*	0.170*	0.158*
Haha	0.402*	0.585*	0.460*	0.436*	0.242*	1.000	0.357*	0.297*	0.407*	0.155*	0.435*	0.539*	0.314*	0.298*	0.216*
Love	0.596*	0.543*	0.682*	0.348*	0.276*	0.357*	1.000	0.217*	0.380*	0.385*	0.594*	0.682*	0.412*	0.475*	0.188*
Sad	0.378*	0.387*	0.338*	0.487*	0.202*	0.297*	0.217*	1.000	0.440*	0.156*	0.374*	0.405*	0.267*	0.295*	0.097*
Wow	0.496*	0.473*	0.500*	0.496*	0.249*	0.407*	0.380*	0.440*	1.000	0.229*	0.487*	0.534*	0.342*	0.408*	0.117*
Length	0.376*	0.274*	0.424*	0.233*	0.179*	0.155*	0.385*	0.156*	0.229*	1.000	0.463*	0.408*	0.219*	0.383*	-0.028
Views	0.721*	0.632*	0.821*	0.487*	0.242*	0.435*	0.594*	0.374*	0.487*	0.463*	1.000	0.835*	0.459*	0.738*	-0.117*
Inter	0.883*	0.760*	0.958*	0.541*	0.301*	0.539*	0.682*	0.405*	0.534*	0.408*	0.835*	1.000	0.619*	0.637*	0.349*
Ri	0.588*	0.450*	0.576*	0.342*	0.236*	0.314*	0.412*	0.267*	0.342*	0.219*	0.459*	0.619*	1.000	0.629*	0.387*
Rv	0.582*	0.456*	0.616*	0.389*	0.170*	0.298*	0.475*	0.295*	0.408*	0.383*	0.738*	0.637*	0.629*	1.000	-0.049*
Ripv	0.328*	0.247*	0.308*	0.128*	0.158*	0.216*	0.188*	0.097*	0.117*	-0.028	-0.117*	0.349*	0.387*	-0.049*	1.000

*p-value <.0001

Exploring the relationships between interactions and views, we saw there was a slight positive correlation, (0.629) between the interaction ratios (Ri) and views (Rv). Also, if we study the relationship between views and interactions – without applying the matrix from the number of followers from the Rv and Ri variables– there was a strong correlation between both variables (0.835).

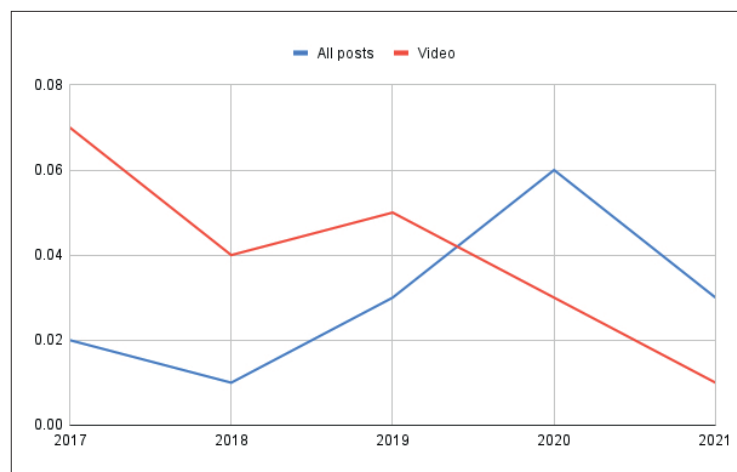
The interaction ratio (Ri) shows the highest positive correlations with shares (0.588) and likes (0,576). Moreover, the same happens with the ratio of views (0.582 for shares and 0.616 for likes).

With the interactions per view ratio (Ripv), the greatest correlations, albeit slight ones, also occur with shares (0.328) and likes (0.308), although the significant and negative (slight) correlation between the number of views and the Ripv (-0,117) is also salient.

Lastly, the duration variable shows positive and significant –albeit slight– results with respect to Ri and Rv. Therefore, it seems a longer duration does not have a negative effect on the views or interactions ratio, although the result with respect to Ripv is not significant.

To show the annual trend for the 3 ratios used as variables we chose to calculate the median due to its greater robustness as opposed to the average and to compare the comprehensive study and the sample.

In graph 4 we see the interaction ratio for the whole sample shows a rising trend, with very little variation between 2017 and 2021, whereas with videos there was a clear falling trend over the same period. Moreover, it is



Graph 4. Trend in the interaction ratio per follower. Total number of posts versus videos

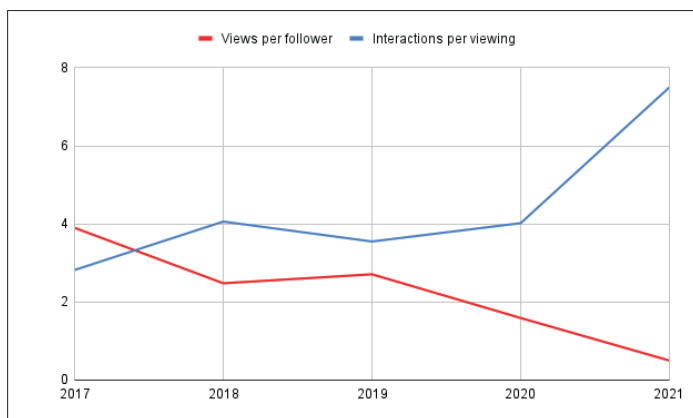
remarkable that between 2017 and 2019 videos obtained an interaction ratio median that was higher than that for the posts as a whole, but that trend was reversed in 2020 and 2021.

The ratio of views per follower for the videos shows a downward trend coupled with that for interactions per follower we saw in graph 3. However, the ratio of interactions by the number of views shows a clear positive trend in the study period (graph 5).

5.5. Analysis of the most successful videos

In the sample of videos with the highest interaction ratio per fact-checker we see that native videos are the most significant category, and in an even higher proportion than in the group of videos posted as a whole. Conversely, the videos hosted by *YouTube* are far less abundant in the most successful group than what they represent in terms of the total number of posts (table 3).

As for duration, the data show meanwhile that videos under one minute are the main group within the whole group of videos posted for all the years analysed (all). As for the most successful ones (top) those between 1:00 and 1:59 are the main category in 2019 and 2020, and show a marked trend since 2017. In the global comparison 2017-2021, videos under a minute are less prominent in the group of the most successful ones and videos with higher durations surpass them (table 4).



Graph 5. Trend in the views ratio per follower and interactions per view (n=7327)

Table 3. Trend in video production according to duration. Comparison between total and video sample

	All (2017-21)		Top (2017-21)	
	n	%	n	%
Native video	7340	83.50	365	88.16
Live video complete	627	7.13	26	6.28
Live video scheduled	185	2.10	3	0.72
Video	421	4.79	16	3.86
<i>YouTube</i>	218	2.48	4	0.97
Total	8791	100	414	100

Table 4. Trend in video production according to duration. Comparison between the total and the sample of videos with the highest interaction ratio from 2017-2021

	2017		2018		2019		2020		2021		2017-21	
	%All	%Top	%All	%Top	%All	%Top	%All	%Top	%All	%Top	%All	%Top
<1 min	43.23	50.00	37.82	29.27	46.86	22.81	42.55	24.14	58.68	35.10	48.34	30.30
01:00 - 02:59	17.94	14.29	31.27	24.39	31.47	43.86	33.64	36.21	26.86	32.45	29.37	33.33
03:00 - 04:59	13.21	7.14	9.18	24.39	5.85	12.28	5.92	13.79	2.92	6.62	5.85	11.36
05:00 - 09:59	19.09	21.43	9.91	12.20	6.19	10.53	7.79	12.93	3.87	12.58	7.38	12.88
10:00 - 19:59	1.79	0.00	2.64	7.32	3.27	10.53	3.16	3.45	1.48	5.96	2.40	5.56
20:00 - 59:59	3.92	3.57	4.18	2.44	4.82	0.00	5.03	7.76	4.82	7.28	4.72	5.81
01:00:00 y +	0.82	3.57	5.00	0.00	1.55	0.00	1.91	1.72	1.37	0.00	1.37	0.76

In the sample analysed 78.99% of the videos refer to national issues, the horizontal aspect ratio predominates (66.91%) over square (28.02%) or vertical (5.07%), and also those with various shots (79.47%) as opposed to those with a single shot.

As seen in table 5 there is a slight predominance of videos with no presenter, but experts frequently appear in order to make the messages more believable.

While visual identity elements and graphics are highly normalised in the sample and appear in over 90% of videos in both cases, subtitles (35.02%) and voiceover (23.43%) are lesser used resources. With videos which use voiceover the use of subtitles rises to 62.89%, although in the sample as a whole, 33.57% are incomprehensible without sound; in 23.43% of cases this is partial and in 43% they are completely incomprehensible without sound. 60.11% of the videos which are totally understandable without sound lack it originally (there is no presenter, nor expert, nor voiceover). However, we can show that 65.47% with virtually no or very little comprehension are those which have a presenter and/or an on-screen expert who are not subtitled.

Table 5. Presence of different elements in the videos

	Presenter	Expert	Identity	Graphics	Subtitles	Voiceover
Yes	43.00	82.13	90.58	91.55	35.02	23.43
No	57.00	17.87	9.42	8.45	64.98	76.57

Table 6. Classification of the most successful posts according to intention. Data for views and interactions

Type of post	Volume (n=414)	Ri (n=391)	Rc (n=391)	Rs (n=391)	Rv (n=391)	Ripv (n=391)
Literacy	13.04%	0.4%	0.03%	0.15%	7.55%	5.66%
Self-promotion	12.32%	0.5%	0.03%	0.09%	10.12%	4.59%
Rebuttals	22.71%	0.37%	0.04%	0.12%	9.56%	4.75%
Information	39.61%	1.06%	0.09%	0.34%	27.36%	3.88%
Verification	6.04%	0.37%	0.07%	0.06%	4.79%	4.57%
Others	6.28%	0.20%	0.03%	0.03%	4.25%	4.72%

In table 6 we see that the main part of the sample is concentrated in the information (39.61%) and rebuttals (22.71%) categories, with a very low proportion of positive verifications (6.04%). As for the analysis of interactions per category, it can be seen that information is that which creates a higher ratio of views, interactions, comments and shares per number of followers, although the ratio of interactions per view remains below the average. Neither rebuttals nor verification have salient figures, although the former category surpasses the latter in terms of the ratio of comments, shares and views per number of followers.

Table 7. Classification of the most successful posts according to topic. Data for views and interactions

Type of post	Volume (n=414)	Ri (n=391)	Rc (n=391)	Rs (n=391)	Rv (n=391)	Ripv (n=391)
Science	0.48%	1.66%	0.13%	0.60%	146.89%	4.96%
Culture	0.97%	1.19%	0.03%	0.62%	18.40%	6.19%
Disinformation	11.84%	0.50%	0.03%	0.10%	10.32%	4.65%
Ecology	2.17%	0.26%	0.04%	0.10%	4.48%	4.47%
Economy	3.62%	0.68%	0.05%	0.32%	10.92%	3.48%
Education	1.69%	0.22%	0.05%	0.12%	4.31%	5.75%
Gender-LGTBI	2.17%	0.27%	0.01%	0.04%	7.74%	3.74%
Politics and law	41.06%	0.97%	0.08%	0.27%	22.67%	4.14%
Racism and xenophobia	2.42%	1.04%	0.08%	0.27%	29.38%	3.73%
Health	24.15%	0.27%	0.03%	0.09%	5.75%	4.95%
Events	3.14%	3.24%	0.15%	1.73%	73.61%	3.81%
Miscellaneous	6.28%	0.34%	0.01%	0.12%	10.28%	5.09%

The main part of the sample is concentrated in the categories of politics and law (41.06%) and health (24.15%). The analysis of interactions and views shows the events category is the one which has the highest interaction ratio (3.24%) and is especially marked in the views ratio (73.61%) and for shares per follower (1.72%). If we omit the science and culture categories –given the small number of cases conditions the results– we can see that both politics and law, and racism and xenophobia stand out in terms of views per follower (22.67 and 29.38% respectively). In both cases this corresponds to interaction ratios of around 1%. The posts about gender and group LGBTI have the lowest rates of interactions per view and comments or shares per number of followers, although the views per follower figures are higher than those from other categories such as health.

Table 8. Trend in the median for the views ratio, interaction ratio and interaction per view ratio (%)

	2017		2018		2019		2020		2021	
Interaction ratio (total number of posts)	0.02		0.01		0.03		0.06		0.03	
	All	Top	All	Top	All	Top	All	Top	All	Top
Interaction ratio (video)	0.07	0.45	0.04	1.91	0.05	1.13	0.03	0.45	0.01	0.19
Views ratio (video)	3.90	20.72	2.48	58.35	2.71	25.40	1.59	11.98	0.50	3.25
Interaction per view ratio (video)	2.82	3.38	4.06	0.007	3.55	3.00	4.02	3.92	7.50	6.57

The videos in the sample achieve an interaction ratio which is far higher than that for the videos as a whole, but if we link the interactions per view, instead of doing this with the followers, we can see that, apart from in 2017, the results are lower than for the videos as a whole.

“Ibero-American fact-checkers made ever more posts on Facebook and the video is an ever more significant type of content”

6. Discussion and conclusions

With respect to RQ1, with the data obtained in the analysis we can state that Ibero-American fact-checkers made ever more posts on Facebook and the video is an ever more significant type of content. In the period between 2016 and 2021 there was a clear rise in the average number of posts made by every fact-checker (a 97.36% rise) and greater still was the average number of videos posted (+360.47%). Videos go from accounting for 3.66% of the whole content posted in 2016 to 8.54% in 2021, albeit this trend was not distributed evenly between the fact-checkers in the sample and there were marked differences in the drive for video.

Most of the videos posted are native (90.13%), which is in keeping with the Facebook policy for rewarding this type of videos in the News Feed as opposed to alternative external ones. If we focus on the annual trend there is a clear tendency to adapt to the algorithm: in 2016 26.76% of video posts were still “outside Facebook” (they were neither native nor live video) in 2021 this figure fell to under 4.5%. We can add to this the small amount of videos hosted by YouTube among which are those which had the highest interaction ratios. Another example of adapting to the algorithm as the data imply is the Live Video format: production in absolute terms grew year-on-year, especially from 2016 (the time at which they were first prioritised in the News Feed algorithm) to 2018. Despite this, although Live Video was considered as a priority format in the News Feed from 2016 and drove certain growth in producing contents with this format, generally speaking, it was not fully consolidated into the everyday work of the fact-checkers.

As for duration, the trend in the study period was clearly favourable to briefer video formats, and this was in keeping with the features of online video consumption (Dafonte-Gómez; Míguez-González; Corbacho-Valencia, 2020). This occurred even given that, since 2019 Facebook has prioritized videos over 3 minutes long. Not even any special effort to produce videos that reach at least one minute by the fact-checkers is observed. Videos under one minute account for 48.55% of the whole period and peak in 2021 (58.68%). At this point, however, we should remark on the slight fall in production of these kind of videos and a rise in the percentage of almost all the other groups of longer duration in 2020, after updating criteria on the most favourable duration for appearing in the News Feed and within the context of a 44.52% rise with respect to 2019. This may indicate that video production from the fact-checkers was to an extent influenced by the criteria recommended by Facebook. In relation to this, the sample of videos with the best interaction ratios, the presence of pieces over one-minute-long are far more frequent than in the total amount of videos: specifically, the category between one minute and 2:59 is the most abundant from 2019. This may be linked to the algorithm favouring these videos by giving them more visibility, among other possible factors.

Therefore, as for RQ2 and how Facebook criteria influenced the News Feed settings, we could see that prioritising native videos had an obvious and permanent effect on the contents the fact-checkers produced. Meanwhile, prioritising live video did not lead to them adopting this generally, nor long-term. The policies regarding the optimal duration for the videos coincided with irregular and temporary variations in the most frequent durations. Although there was a clear predominance of contents of under one minute (not recommended) in terms of production, it is true that among the videos with the best interaction ratios, the durations Facebook recommends have a greater presence than in the group as a whole. In addition, the increase in video duration is significantly, albeit slightly, related to the fall in interactions per view (RQ3). Therefore, we considered, although Facebook gives precedence to the reach formats have and durations which increase the amount of time users spend on the platform, fact-checkers create content in which they must combine the rules of the “visibility game” (Cotter, 2019) in different proportions according to what they consider will be better received by their audiences and more appropriate for their purposes.

Looking at the conclusions concerning RQ3 in more detail, the results in terms of the interaction ratios (average success in relation to Facebook) show that this indicator is especially linked to the number of likes and shares, which stand out from the other interactions. In turn, the act of sharing, an essential action in the organic distribution of contents on networks, is closely related to likes and comments, but also to love, wow and angry, as opposed to other kinds of possible reactions. These results are in keeping with previous research on the link between sharing and emotions, which grant contents which elicit intense emotions (both positive and negative ones) greater capacity to go viral (Dafonte, 2018). They are also related, to a certain extent, with the fact that the Facebook algorithm for some years now has rewarded controversial content with greater visibility (Horwitz; Seetharaman, 2020). Other elements which show the interest a content creates are the comments, which are linked to likes, shares and ha ha, love and angry.

“Software for automated contents filtering on the social networks is crucial for examining the changes in the information ecosystem model but constitutes a variable which is difficult to access and understand”

As for the content analysis applied to the sample of the most successful videos, the presence of native videos increased with respect to the total number of videos, while those hosted on external platforms such as *YouTube* fell remarkably. The presence of those between 1:00 and 4:59 were also prominent in the samples of the most successful videos which corresponded proportionately to the volume of production. Both circumstances coincided with formal aspects rewarded by the *Facebook* algorithm, and the results from Spearman's correlation coefficient, a longer duration did not have a negative effect on the views or interactions ratios (although it did on the interactions per view).

“ In the sample of videos with the highest interaction rate per fact-checker, short videos without a presenter but with at least one expert tend to dominate. However, both the relevance of the horizontal format and the high proportion of videos that are incomprehensible without sound denote a production not particularly adapted to mobile consumption ”

Regarding the variables related to video production we believe it is remarkable that in most there was no presenter and in the vast majority (82.13%) there was at least one expert, data which are a far cry from research from videos on other social networks (**García-Marín; Salvat-Martinrey**, 2022). Both the predominance of the horizontal format, and the large amount of videos that are incomprehensible without sound, show production is not especially adapted to consumption on mobile phones.

A clear dominance of posts was observed that were considered to be informative as opposed to those strictly concerning fact-checking: rebuttals, positive verifications and literacy. The results did not match those obtained by **Míguez-González** and **Dafonte-Gómez** (2022) in terms of the predominant category (rebuttals, in their study) but they did coincide with a greater proportion of rebuttals than positive verifications. However, it must be stated that while **Míguez-González** and **Dafonte-Gómez** (2022) base their results on constructed week sampling over a year, we focus on the most successful posts, and in this respect, the information category in our research obtained the best interaction ratios, comments, shares and views, which explains their greater presence in the sample. With respect to the topics politics and health, they are clearly predominant. They have percentages which are very similar to those obtained by **Míguez-González** and **Dafonte-Gómez** (2022) (41.84% and 29.59 respectively) and also match those from **Ceron, De-Lima-Santos** and **Quiles** (2021), and **García-Vivero** and **López** (2021). Likewise, politics stands out as the main topic in the research by **Bernal-Triviño** and **Clares-Gavilán** (2019); **Blanco-Alfonso, Chaparro-Domínguez and Repiso** (2021); and **Guallar et al.**, (2020). As for other topics which are less abundant in the sample, the high ratios of interaction and views of events/catastrophes stand out, which may be linked to relevance, unpredictability and the tragic nature of the topics dealt with and the high rate of views- and, to a lesser extent, interactions, the topics of racism and xenophobia receive.

Regarding RQ4, the results on the ratio of views and interaction rates match the data provided by **Newberry** (2012) on the fall in the organic reach which has accelerated since 2018. The trend in the ratio of views and the rate of interaction gained by videos posted by the fact-checkers suggest that the changes made by *Facebook* in the settings criteria for the *News Feed* which show each user over the years (and especially within the study period) have had important consequences. Firstly, the ratio of views obtained by the fact-checkers for their videos underwent a fall of 87.18% between 2017 and 2021. Secondly, while the variation in the general interaction ratio shows a positive albeit slight trend within the period the variation in the interactions ratio for the videos fell by 87.71% in the same period, with especially sharp falls in 2018, 2020 and 2021.

Although, generally speaking the interaction ratio is low in the contents posted by the fact-checkers, with video, the declining reach of the results between 2017 and 2021 is patent. We also see that the same videos which obtain a higher amount of interactions in terms of the amount of followers do not stand out in terms of the proportion of interactions concerning views, although looking at the videos as a whole, there is a significant positive, slight correlation between both ratios (0.387). Even so, the interactions per view ratio improves year by year since 2019, both in the sample and the videos as a whole. In short, more interactions per view are obtained, but fewer views and interactions per follower. This may be linked to a better distribution among audiences who do not follow the fact-checker. With this “algorithm reward” there is a loop in which views feed interactions, which feed views, in turn.

Lastly, it is especially remarkable that, from our point of view, which although **Mosseri** insisted in 2018 (**Vogelstein**, 2018) that video is a type of content which creates fewer interactions, our results show that this format obtained a greater rate of interaction than that for the posted contents as a whole from the fact-checkers until *Facebook* made changes to its *News Feed* in 2018. From then onwards it began to decline until falling below those for the posts as a whole from 2020.

Although the research provides a topical and formal characterisation of the most successful videos posted by Ibero-American fact-checkers on *Facebook* between 2017 and 2021, and explores in detail previous research results, we believe its main contribution comes from its comprehensive analysis of video production since 2016 in order to plot a trend in categories, durations and results obtained in terms of interactions and views. Although with the methodology used (and the opaqueness of *Facebook* regarding how their algorithms work) we cannot rule out other types of variables that have not been contemplated in this study may have had an influence on the trend in video production from the fact-checkers

and on their reach, we do believe it manages to give data-based indications on the influence of criteria outside the inherent features of the contents, which affect their visibility on *Facebook*, in accordance with successive changes in its algorithm. In this way, although we believe it is important to keep exploring the features of the most successful contents for fact-checking, we believe it is essential in future research to bear in mind the control any social network has on what reaches users in an environment the latter perceive as being neutral.

Although Mosseri insisted in 2018 that video is a type of content that, by its nature, generates fewer interactions, the results show that it was obtaining a higher interaction rate than the set of content published by the fact-checkers until *Facebook* made changes to its *News Feed* in 2018

7. References

- Amazeen, Michelle A.; Thorson, Emily; Muddiman, Ashley; Graves, Lucas** (2018). "Correcting political and consumer misperceptions: the effectiveness and effects of rating scale versus contextual correction formats". *Journalism & mass communication quarterly*, v. 95, n. 1, pp. 28-48.
<https://doi.org/10.1177/1077699016678186>
- Ardèvol-Abreu, Alberto; Delponti, Patricia; Rodríguez-Wangüemert, Carmen** (2020). "Intentional or inadvertent fake news sharing? Fact-checking warnings and users' interaction with social media content". *Profesional de la información*, v. 29, n. 5.
<https://doi.org/10.3145/epi.2020.sep.07>
- Bapna, Abhishek; Park, Seyoung** (2017). *Updating how we account for video completion rates*.
<https://about.fb.com/news/2017/01/news-feed-fyi-updating-how-we-account-for-video-completion-rates>
- Bechmann, Anja** (2020). "Tackling disinformation and infodemics demands media policy changes". *Digital journalism*, v. 8, n. 6, pp. 855-863.
<https://doi.org/10.1080/21670811.2020.1773887>
- Bernal-Triviño, Ana; Clares-Gavilán, Judith** (2019). "Uso del móvil y las redes sociales como canales de verificación de fake news. El caso de Maldita.es". *Profesional de la información*, v. 28, n. 3, e280312.
<https://doi.org/10.3145/epi.2019.may.12>
- Blanco-Alfonso, Ignacio; Chaparro-Domínguez, María-Ángeles; Repiso, Rafael** (2021). "El fact-checking como estrategia global para contener la desinformación". *Estudios sobre el mensaje periodístico*, v. 27, n. 3, pp. 779-791.
<https://doi.org/10.5209/esmp.76189>
- Bozdog, Engin** (2013). "Bias in algorithmic filtering and personalization". *Ethics and information technology*, v. 15, n. 3, pp. 209-227.
<http://doi.org/10.1007/s10676-013-9321-6>
- Ceron, Wilson; De-Lima-Santos, Mathias-Felipe; Quiles, Marcos** (2021). "Fake news agenda in the era of COVID-19: identifying trends through fact-checking content". *Online social networks and media*, v. 21.
<https://doi.org/10.1016/j.osnem.2020.100116>
- Cotter, Kelley** (2019). "Playing the visibility game: how digital influencers and algorithms negotiate influence on Instagram". *New media & society*, v. 21, n. 4, pp. 895-913.
<https://doi.org/10.1177/1461444818815684>
- CrowdTangle Team. CrowdTangle: Facebook.*
<https://apps.crowdtangle.com/universidaddevigofacebook/lists/1456283>
- Dafonte-Gómez, Alberto** (2018). "Audience as medium: motivations and emotions in news sharing". *International journal of communication*, v. 12, pp. 2134-2152.
<https://ijoc.org/index.php/ijoc/article/view/6790/2351>
- Dafonte-Gómez, Alberto; Míguez-González, María-Isabel; Corbacho-Valencia, Juan-Manuel** (2020). "Viral dissemination of content in advertising: emotional factors to reach consumers". *Communication & society*, v. 33, n. 1, pp. 107-120.
<https://doi.org/10.15581/003.33.33448>
- Dafonte-Gómez, Alberto; Míguez-González, María-Isabel; Ramahí-García, Diana** (2022). "Fact-checkers on social networks: analysis of their presence and content distribution channels". *Communication & society*, v. 35, n. 3, pp. 73-89.
<https://doi.org/10.15581/003.35.3.73-89>
- DeVito, Michael A.** (2017). "From editors to algorithms: a values-based approach to understanding story selection in the Facebook News Feed". *Digital journalism*, v. 5, n. 6, pp. 753-773.
<https://doi.org/10.1080/21670811.2016.1178592>

- Ecker, Ullrich K. H.; O'Reilly, Ziggy; Reid, Jesse S.; Chang, Ee-Pin** (2019). "The effectiveness of short-format refutational fact-checks". *British journal of psychology*, v. 111, n. 1, pp. 36–54.
<https://doi.org/10.1111/bjop.12383>
- Elizabeth, Jane** (2016). *The week in fact-checking: The fact is, fact-checking can get better*.
<https://www.americanpressinstitute.org/factchecking-project/week-fact-checking-fact-checking-can-better>
- Epstein, Robert; Robertson, Ronald E.** (2015). "The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections". In: *Proceedings of the National Academy of Sciences*, v. 112, n. 33, pp. E4512-E4521.
<https://doi.org/10.1073/pnas.1419828112>
- Ericsson** (2022). *Ericsson mobility report*.
<https://www.ericsson.com/4ae28d/assets/local/reports-papers/mobility-report/documents/2022/ericsson-mobility-report-november-2022.pdf>
- Facebook** (2018). *Video best practices*.
<https://www.facebook.com/facebookmedia/blog/video-best-practices>
- Facebook** (2019). *Understanding video distribution on Facebook*.
<https://www.facebook.com/facebookmedia/blog/understanding-video-distribution-on-facebook>
- Facebook** (2021). *CEO Mark Zuckerberg on Q3 2021 results. Earning call transcript*.
<https://seekingalpha.com/article/4461908-facebook-inc-fb-ceo-mark-zuckerberg-on-q3-2021-results-earning-call-transcript>
- Facebook** (2022). *Widely viewed content report: what people see on Facebook*.
<https://transparency.fb.com/es-es/data/widely-viewed-content-report>
- Fletcher, Richard; Nielsen, Rasmus-Kleis** (2017). "Are people incidentally exposed to news on social media? A comparative analysis". *New media & society*, v. 20, n.7, pp. 2450-2468.
<http://doi.org/10.1177/1461444817724170>
- García-Marín, David; Salvat-Martinrey, Guiomar** (2022). "Viralizing the truth: predictive factors of fact-checkers' engagement on TikTok". *Profesional de la información*, v. 31, n. 2, pp. 1699-2407.
<https://doi.org/10.3145/epi.2022.mar.10>
- García-Vivero, Gloria; López-García, Xosé** (2021). "La verificación de datos en Europa. Análisis de 5 iniciativas europeas: Maldita.es, Newtral, Pagella Política, Les Décodeurs y BBC Reality Check". *adComunica*, v. 21, pp. 235-264.
<https://doi.org/10.6035/2174-0992.2021.21.12>
- Gillespie, Tarleton** (2014). "The relevance of algorithms". In: Gillespie, Tarleton; Boczkowski, Pablo J.; Foot, Kirsten A. (eds.). *Media technologies: essays on communication, materiality, and society*. Cambridge, Massachusetts: MIT Press, pp. 167-193 ISBN: 978 0 2623194 6 1
<https://doi.org/10.7551/mitpress/9780262525374.001.0001>
- Grygiel, Jennifer; Lysak, Suzanne** (2021). "Police social media and broadcast news: an investigation into the impact of police use of Facebook on journalists' gatekeeping role". *Journalism practice*, v. 15, n. 7, pp. 994-1011.
<https://doi.org/10.1080/17512786.2020.1759123>
- Guallar, Javier; Codina, Luis; Freixa, Pere; Pérez-Montoro, Mario** (2020) "Desinformación, bulos, curación y verificación. Revisión de estudios en Iberoamérica 2017-2020". *Telos: revista de estudios interdisciplinarios en ciencias sociales*, v. 22, n. 3, pp. 595-613.
- Honan, Mat** (2016). "Why Facebook and Mark Zuckerberg went all in on live video". *BuzzFeed*, April 6.
<https://www.buzzfeednews.com/article/mathonan/why-facebook-and-mark-zuckerberg-went-all-in-on-live-video>
- Horwitz, Jeff; Seetharaman, Deepa** (2020). "Facebook executives shut down efforts to make the site less divisive". *Wall street journal*, May 26.
<https://www.wsj.com/articles/facebook-knows-it-encourages-division-top-executives-nixed-solutions-11590507499>
- Humprecht, Edda** (2020). "How do they debunk fake news? A cross-national comparison of transparency in fact checks". *Digital journalism*, v. 8, n. 3, pp. 310-327.
<https://doi.org/10.1080/21670811.2019.1691031>
- Kant, Vibhi; Xu, Jie** (2016). *Taking into account live video when ranking feed*.
<https://about.fb.com/news/2016/03/news-feed-fyi-taking-into-account-live-video-when-ranking-feed>
- Kemp, Simon** (2022). *Digital 2022 global overview report state of digital*.
<https://wearesocial.com/es/blog/2022/01/digital-2022>

- Kite, James; Foley, Bridget C.; Grunseit, Anne C.; Freeman, B.** (2016). "Please like me: Facebook and public health communication". *PloS one*, v. 11, n. 9.
<https://doi.org/10.1371/journal.pone.0162765>
- Madrigal, Alexis C.; Meyer, Robinson** (2018). "How Facebook's chaotic push into video cost hundreds of journalists their jobs". *The Atlantic*, October 18.
<https://www.theatlantic.com/technology/archive/2018/10/facebook-driven-video-push-may-have-cost-483-journaliststheir-jobs/573403>
- Mantzarlis, Alexios** (2016). *Can the worldwide boom in digital fact-checking make the leap to TV?*
<https://www.poynter.org/fact-checking/2016/can-the-worldwide-boom-in-digital-fact-checking-make-the-leap-to-tv>
- Margolin, Drew B.; Hannak, Aniko; Weber, Ingmar** (2017). "Political fact-checking on Twitter: when do corrections have an effect?". *Political communication*, v. 35, n. 2, pp. 196-219.
<https://doi.org/10.1080/10584609.2017.1334018>
- Meese, James; Hurcombe, Edward** (2020). "Facebook, news media and platform dependency: the institutional impacts of news distribution on social platforms". *New media & society*, v. 23, n. 8, pp. 2367-2384.
<https://doi.org/10.1177/1461444820926472>
- Míguez-González, María-Isabel; Dafonte-Gómez, Alberto** (2022). "La labor de verificación de los fact-checkers iberoamericanos en Facebook: análisis de las temáticas de la desinformación". En: *VIII Congreso internacional de la AE-IC: comunicación y ciudad conectada*, pp. 2053-2072.
<http://hdl.handle.net/11093/4527>
- Miller, David** (2019). *Updates to video ranking*.
<https://about.fb.com/news/2019/05/updates-to-video-ranking>
- Mitchelstein, Eugenia; Boczkowski Pablo J.** (2021). "What a special issue on Latin America teaches us about some key limitations in the field of digital journalism". *Digital journalism*, v. 9, n. 2, pp. 130-135.
<https://doi.org/10.1080/21670811.2021.1873813>
- Molina-Cañabate, Juan-Pedro; Magallón-Rosa, Raúl** (2021). "Desinformación y fact-checking en las elecciones uruguayas de 2019. El caso de Verificado Uruguay". *Perspectivas de la comunicación*, v. 14, n.1, pp. 89-112.
- Newberry, Christina** (2022). "Cómo funciona el algoritmo de Facebook en 2022 y cómo trabajar con él". *Hootsuite*, 25 marzo.
<https://blog.hootsuite.com/es/algoritmo-facebook-como-funciona>
- Newman, Nic; Fletcher, Richard; Robertson, Craig T.; Eddy, Kirsten; Nielsen, Rasmus-Kleis** (2022). *Reuters Institute digital news report 2022*. Oxford: Reuters Institute for the Study of Journalism. ISBN: 978 1 907384 98 1
<https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2022>
- Newman, Nic; Fletcher, Richard; Schultz, Anne; Andi, Simge; Robertson, Craig T.; Nielsen, Rasmus-Kleis** (2021). *Reuters Institute digital news report 2021*. Oxford: Reuters Institute for the Study of Journalism. ISBN: 978 1 907384 90 5
https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2020-06/DNR_2020_FINAL.pdf
- Nyhan, Brendan; Reifler, Jason** (2010). "When corrections fail: the persistence of political misperceptions". *Political behavior*, v. 32, n. 2, pp. 303-330.
<https://doi.org/10.1007/s11109-010-9112-2>
- Peterson-Salahuddin, Chelsea; Diakopoulos, Nicholas** (2020). "Negotiated autonomy: the role of social media algorithms in editorial decision making". *Media and communication*, v. 8, n. 3, pp. 27-38.
<https://doi.org/10.17645/mac.v8i3.3001>
- Powers, Elia** (2017). "My News Feed is filtered?: awareness of news personalization among college students". *Digital journalism*, v. 5, n. 10, pp. 1315-1335.
<https://doi.org/10.1080/21670811.2017.1286943>
- Rein, Katharina; Venturini, Tommaso** (2018). "Ploughing digital landscapes: how Facebook influences the evolution of live video streaming". *New media & society*, v. 20, n. 9, pp. 3359-3380.
<https://doi.org/10.1177/1461444817748954>
- Robertson, Craig T.; Mourão, Rachel R.; Thorson, Emily** (2020). "Who uses fact-checking sites? The impact of demographics, political antecedents, and media use on fact-checking site awareness, attitudes, and behavior". *The international journal of press/politics*, v. 25, n. 2, pp. 217-237.
<https://doi.org/10.1177/1940161219898055>
- Ryan, Eryca** (2022). *From Sofia to Springfield, fact-checking extends its reach*.
<https://reporterslab.org/tag/fact-checking-census>

Shawcross, Alistair (2016). *Facts we can believe in: how to make fact-checking better*. London: Legatum Institute.
<https://www.lse.ac.uk/iga/assets/documents/arena/archives/facts-we-can-believe-in-how-to-make-fact-checking-better-web-pdf.pdf>

Shin, Jieun; Thorson, Kjerstin (2017). "Partisan selective sharing: the biased diffusion of fact-checking messages on social media". *Journal of communication*, v. 67, n. 2, pp. 233-255.
<https://doi.org/10.1111/jcom.12284>

Smith, Maria (2014). *Updates to video distribution and monetization*.
<https://www.facebook.com/formedia/blog/updates-to-video-distribution-and-monetization>

Subramaniam, Tara (2021). "Fact check: 3 times internal docs contradicted Facebook's public stance". *CNN*, September 28.
<https://edition.cnn.com/2021/09/28/politics/facebook-public-internal-contradict-wsj/index.html>

Tandoc Jr., Edson C. (2021). "Tools of disinformation: how fake news gets to deceive". In: Jayakumar, Shashi; Ang, Benjamin; Anwar, Nur-Diyanah (eds.). *Disinformation and fake news*. Singapore: Springer, pp. 35-46. ISBN: 978 981 15 58757
https://doi.org/10.1007/978-981-15-5876-4_3

Tandoc Jr., Edson C.; Maitra, Julian (2018). "News organizations' use of native videos on Facebook: tweaking the journalistic field one algorithm change at a time". *New media & society*, v. 20, n. 5, pp. 1679-1696.
<https://doi.org/10.1177/1461444817702398>

Valenzuela, Sebastián; Correa, Teresa; Gil de Zúñiga, Homero (2018). "Ties, likes, and tweets: using strong and weak ties to explain differences in protest participation across Facebook and Twitter use". *Political communication*, v. 35, n. 1, pp. 117-134.
<https://doi.org/10.1080/10584609.2017.1334726>

Vogelstein, Fred (2018). "Your new newsfeed: why Facebook made its latest changes". *Wired*, January 13.
<https://www.wired.com/story/facebooks-adam-mosseri-on-why-youll-see-less-video-more-from-friends>

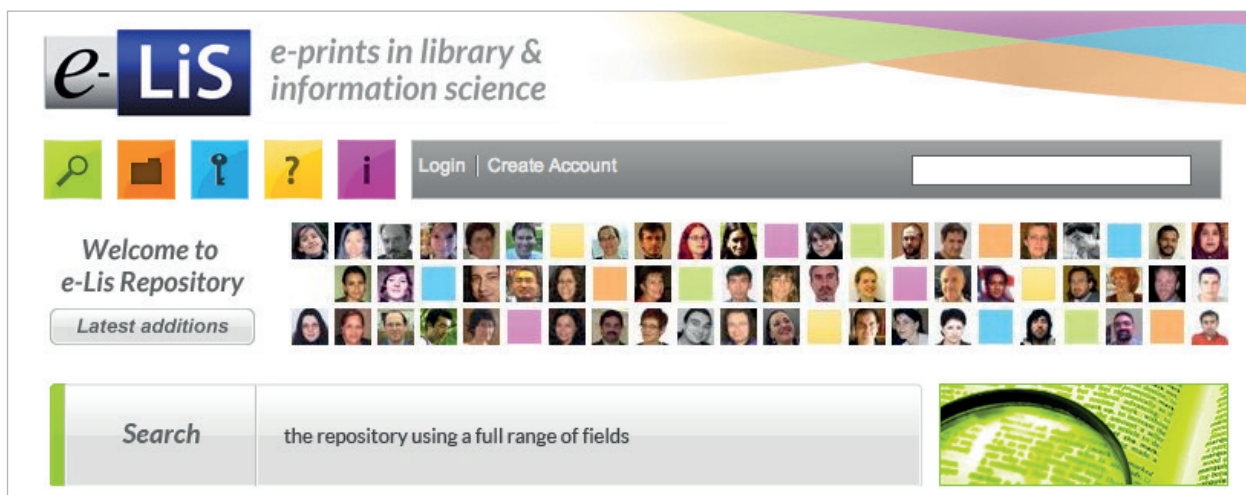
Wang, Meihong; Zhuo, Yue (2015). *Taking into account more actions on videos*.
<https://about.fb.com/news/2015/06/news-feed-fyi-taking-into-account-more-actions-on-videos>

Welch, Brett; Zhang, Xiaochen (2014). *News Feed FYI: showing better videos*.
<https://newsroom.fb.com/news/2014/06/news-feed-fyi-showing-better-videos>

West, Sarah-Myers (2017). "Raging against the machine: network gatekeeping and collective action on social media platforms". *Media and communication*, v. 5, n. 3, pp. 28-36.
<https://doi.org/10.17645/mac.v5i3.989>

Young, Dannagal G.; Jamieson, Kathleen-Hall; Poulsen, Shannon; Goldring, Abigail (2018). "Fact checking effectiveness as a function of format and tone: evaluating FactCheck.org and FlackCheck.org". *Journalism & mass communication quarterly*, v. 95, n. 1, pp. 49-75.
<https://doi.org/10.1177/1077699017710453>

Give visibility to your work by depositing it in e-LIS, the largest international repository on library & information science, and communication



<http://eprints.rclis.org>

Fact or fiction: An experiment on how information sources and message framing influence vaccine risk perception

Daniel Catalán-Matamoros; Enrique Prada; Andrea Langbecker

Nota: Este artículo se puede leer en español en:
<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87317>

Recommended citation:

Catalán-Matamoros, Daniel; Prada, Enrique; Langbecker, Andrea (2023). "Fact or fiction: An experiment on how information sources and message framing influence vaccine risk perception". *Profesional de la información*, v. 32, n. 5, e320510.

<https://doi.org/10.3145/epi.2023.sep.10>

Manuscript received on 10th March 2023
Accepted on 28th May 2023



Daniel Catalán-Matamoros ✉

<https://orcid.org/0000-0001-7783-4984>

Universidad Carlos III de Madrid
Madrid, 133
28903 Getafe (Madrid), Spain
dacatala@hum.uc3m.es



Enrique Prada

<https://orcid.org/0000-0002-1116-3097>

Universidad Carlos III de Madrid
Madrid, 133
28903 Getafe (Madrid), Spain
Universitat Autònoma de Barcelona
08193 Bellaterra (Barcelona), Spain
enrique.prada@uab.cat



Andrea Langbecker

<https://orcid.org/0000-0001-5292-8220>

Universidad Carlos III de Madrid
Madrid, 133
28903 Getafe (Madrid), Spain
alangbec@hum.uc3m.es

Abstract

In view of the growing disinformation about vaccines on social media since the beginning of the Covid-19 pandemic, effective communication strategies encouraging vaccine uptake are needed. We conducted an experiment through an online, preregistered survey to explore which types of information sources are more trusted by the population regarding the risks of the Covid-19 booster, and which types of message frames are more effective in influencing the perception of risks for children. We surveyed a representative sample composed of 1,800 Spaniards in June 2022. The two dependent variables were respondents' perceptions of (1) the Covid-19 booster vaccine effectiveness and (2) the safety of the Covid-19 vaccine for children. Participants were randomly exposed to different messaging regarding these vaccines, with different sources of information (scientific consensus, scientific dissensus, governmental, influencers and medical doctors), and different message framing (pro- and anti-vaccine storytelling and pro- and anti-vaccine scientific data). Additionally, some respondents who did not receive any messaging formed a control group. Our findings suggest that different information sources and frames can influence people's risk perception of vaccines. The source 'medical doctors' had a positive effect on risk perception of the Covid-19 booster vaccine ($p < 0.05$), and pro-vaccine messages, in the form of both storytelling and scientific expository frames, had a positive effect on respondents' risk perception of the vaccine for children ($p < 0.1$ and $p < 0.05$, respectively). On the one hand, male and older respondents rated booster vaccines as more effective than female and younger respondents. On the other hand, right-wing respondents believed vaccines are somewhat less safe for children than left-wing respondents. These findings might support the development of strategic communication in vaccination programmes by public health departments to improve immunization rates in the general population. The practical and theoretical implications are discussed.



Keywords

Health information; Vaccines; Vaccine hesitancy; Disinformation; Fake news; Social media; Social networks; Covid-19; Experiments; Surveys; Sources of information; Narratives; Expository message.

Funding

Study financed through the agreement signed between the *Community of Madrid (Ministry of Education, Universities and Science)* and the *Madrid University Carlos III* for the direct grant of EUR 4.859.000 to fund the development of research activities on SARS-CoV 2 and COVID-19 disease financed with React-EU resources from the *European Regional Development Fund “A way of making Europe”*.

This study has also received funds from the 2022 call for “Knowledge generation projects” of the *Ministry of Science and Innovation*. Project reference: PID2022-142755OB-I00. Comsalud Project: “Pseudoscience, conspiracy theories, fake news, and media literacy in health communication”.

1. Introduction

Full vaccination against Covid-19 has been received by 85.9% of the Spanish population, and 55.7% of adults have received the first booster dose. In the case of childhood vaccination, 46% of children aged 5-12 years have been fully vaccinated (*Ministerio de Sanidad, 2022*). However, this coverage is not enough, according to the *Spanish Association of Paediatrics (AEP, 2022)*. Vaccine hesitancy towards booster and child vaccination represents a challenge for Spanish public health departments, worsened by the fact that coverage rates, by age and by region, are not uniform (*Ministerio de Sanidad, 2022*).

Individuals with vaccine hesitancy are characterised by being unclear about whether they want to be vaccinated or may be completely unwilling to be vaccinated, even if they have access to the vaccine (**MacDonald, 2015**). Thus, the level of vaccine hesitancy is closely related to vaccine acceptance or refusal behaviour (**Larson et al., 2014**). It is a complex phenomenon in which several factors play a role, including confidence in the efficacy of vaccines (**Thorpe et al., 2022**) and risk perception. People’s perception of risk does not necessarily represent the real risk. This is because perception is subjective and can influence decision-making in relation to vaccines (**Malecki; Keating; Safdar, 2021**).

According to **Allington et al. (2023)**, in a study conducted in the UK, vaccine-hesitant individuals were found to have a strong reliance on social media for information and a weak dependence on traditional mass media (television, radio and press). Social media consumption has been steadily increasing, experiencing a significant surge, especially during the pandemic. In 2022, the number of users accounted for 58% of the global population, marking a 10% increase compared with the previous year (*We are Social, 2022*). Notably, social media played a prominent role during the pandemic, facilitating the spread of misinformation, disinformation, and conspiracy theories (**Sánchez-Duarte; Magallón-Rosa, 2020**). Misinformation refers to messages not intended to cause harm, while disinformation is deliberately produced to harm individuals, institutions, or countries. According to Das and Ahmed,

“misinformation spreads on a lighter note, while disinformation destabilizes society by transforming consumers into active agents of interpretation and propagation” (**Das; Ahmed, 2022, p. 10**).

García-Marín (2020) highlights how news on social media during the pandemic often underwent modifications, presenting a blend of accurate and false information, manipulated context and misleading facts. Such content prevailed over entirely fabricated news, creating an environment of uncertainty and distrust.

Regarding preventive measures during the pandemic, exposure to misinformation had an adverse effect on attitudes and intentions toward mask usage, as demonstrated by **Mourali and Drake’s (2022)** research. However, the authors also found that prolonged exposure to false claims has led to less positive reactions to this measure. According to **Greene and Murphy (2021)**, even if a person is exposed to fake news only once and the immediate effects may not appear significant, this singular exposure could still have consequences for their behaviour.

In addition, traditional mass media have produced a large amount of content to answer the needs of a society eager to understand what was happening around it (**Costa-Sánchez; López-García, 2020**). Although this return of the public to the media during the pandemic has reaffirmed its social value in times of crisis (**Casero-Ripollés, 2020**), vaccine-related news can be both a promoter of vaccination and a barrier to its acceptance when the message is negative (**Larson et al., 2014**). Beyond this approach, previous studies highlight that the format of the message (personal narratives and/or presentation of statistics) may also exert some influence on people’s perception of risk in relation to health issues (**Shelby; Ernst, 2013; Vandenberg et al., 2022; Nan et al., 2015**).

“Vaccine reluctance is a complex phenomenon involving several factors, including confidence in the efficacy of vaccines and risk perception, which can influence decision-making in relation to vaccines”

In this context, both information sources and message format may have implications for the acceptability of Covid-19 vaccines. This study conducted two experiments with 1,800 residents in Spain and aimed to evaluate the influence on risk perception of

- the information source on the effectiveness of the Covid-19 booster vaccine, and
- message framing in conveying messages about the risk associated with childhood vaccination against Covid-19.

This objective was based on the following research questions:

RQ1: How can information sources influence public perception of the Covid-19 booster vaccine effectiveness?

RQ2: How can message framing influence public perception regarding the safety of the Covid-19 vaccine for children?

2. Literature review

2.1. Information sources and vaccine hesitancy

The sources that people use play an important role when the information they seek is health related. The relationship with sources is complex and dynamic. Throughout the pandemic, there was a noticeable surge in the volume of information disseminated through various channels, such as the media, social media, official institutions, health experts, as well as from close friends and family. Choosing one source or several is related to the level of trust that people place in them. In the case of the coronavirus pandemic, the tendency has been to look for a variety of information sources (Sallam *et al.*, 2021). On the other hand, several studies point to social media as one of the most used sources for Covid-19 and vaccine issues (Al-Daghastani *et al.*, 2021; Baig *et al.*, 2020; Sulistyawati *et al.*, 2021).

Previous studies have established a relationship between the information source and vaccine acceptance. Park, Massey and Stimpson (2021) compared people who rely more on media and government sources with those who rely on other sources and found that the latter perceived Covid-19 as a less serious risk and had less intention to get vaccinated. Moreover, individuals who relied on social media exhibited the lowest levels of Covid-19 risk perception and willingness to be vaccinated when compared to those who relied on health professionals.

Karabela, Coşkun and Hoşgör (2021) investigated the most trusted sources of information about vaccines and found that participants who were thinking about getting vaccinated trusted *YouTube*, and those who would not get vaccinated trusted *WhatsApp* groups. However, people that trusted the government and health professionals had more pro-vaccine attitudes. Reno *et al.* (2021) found similar results. Social media have increased vaccine hesitancy, whereas the opposite has happened with institutional sites. Sallam *et al.* (2021) also established a relationship between vaccine hesitancy and social media as the main information source about Covid-19 vaccines.

In addition, we should also consider the Spanish context which presents relevant events that may have influenced communication and vaccination against Covid-19. Firstly, it is crucial to consider how certain specific aspects pertaining to adhering to European-level instructions may have influenced communication strategies in Spain. The collective response of the European Union and the decisions made during the authorisation process for vaccine boosters and childhood vaccinations have been significant factors that demand effective communication to inform the public and foster trust among the population. Moreover, it is important to analyse the role of medical doctors as a source of information in this context. It is important to mention cases such as ‘Doctors for the truth,’ who spread disinformation and denialist theories. Another important context to consider is the trust in supranational institutions at the global level, such as the *World Health Organization* and European authorities who played an important role in the Covid-19 vaccination programmes. These cases could have impacted vaccine acceptance, as well as spread misperceptions and doubts in the public, hindering efforts to communicate and promote vaccination.

2.2. Message framing and risk perception

Previous experiments have investigated the perception of health-related risk and how different types of message framing can influence risk perception. In a mediatised society, these frames used to communicate may have implications for the understanding of the world or may exert some kind of influence on the audience (Nan *et al.*, 2015). For example, narratives such as storytelling (short or long) or use of testimonials and first-person accounts are considered more persuasive than other formats (Dahlstrom, 2010). Indeed, the biographical perspective considers how personal experiences are narrated (Pfeiffer-Castellanos, 2014). When the audience identifies themselves with the story, values and points of view, the story can help them to make decisions about an issue, stimulate a behaviour change, or encourage them to increase their interest in and attention to a certain issue (Joubert; Davis; Metcalfe, 2019). In the context of audience loss, the use of narratives in journalism has been seen as a key strategy due to its positive influence in engaging the audience. It should be noted that the use of this format in news is not free of dilemmas between objectivity and subjectivity, between what

Previous studies highlight that, although this return of the public to the media during the pandemic has reaffirmed its social value in times of crisis, vaccine-related news can be both a promoter of vaccination and a barrier to its acceptance when the message is negative

is fact and what is fiction (Van-Krieken; Sanders, 2021). Boyson, Zimmerman and Shoemaker (2015), for example, found that exposure to news stories with a personal story had a greater influence on risk perception of antiretroviral therapy related to human immunodeficiency virus (HIV) than news stories that presented statistical data. In another study related to Human Papillomavirus (HPV), Nan *et al.* (2015), noted that participants indicated a higher risk of contracting HPV after being exposed to a hybrid type of message (statistical and narrative) relative to messages containing only statistical or storytelling framing. In addition, first-person storytelling led to a higher perceived risk of HPV than third-person text. Both types of framing had an indirect effect on the intention to be vaccinated against HPV.

Choosing one source or several is related to the level of trust that people place on them, and, in the case of the coronavirus pandemic, the tendency has been to look for a variety of information sources

In an experiment to determine whether message framing (storytelling versus expository) influences risk perception about childhood vaccination, Vandenberg *et al.* (2022) found that anti-vaccination storytelling could reduce risk perception towards vaccination, especially for those who had a certain level of vaccine hesitancy. The anti-vaccine movement has used this narrative framing to influence parents not to vaccinate their children while public health professionals build their messages with statistics and scientific facts, evidence-based data, and research. Perhaps these approaches are not enough and storytelling strategies beyond these employed resources should also be used (Shelby; Ernst, 2013). To fight against vaccine hesitancy during the Covid-19 pandemic, it is necessary to invest in effective communication to 'reduce doubt, promote confidence and increase acceptance' (Thorpe *et al.*, 2022). Similarly, Dahlstrom (2014) highlights the importance of incorporating storytelling into public science communication, as it is a persuasive tool that not only enhances public interest in the topic but also improves understanding of the content being conveyed.

3. Methodology

We conducted an original survey in Spain during June 2022. The survey was administered online (CAWI) to a sample of 1,800 respondents over 18 years old. The respondents were recruited through the company *AsuFieldwork* and completed the survey in exchange for a small monetary compensation. Gender, age, and regional quotas were established to ensure that the sample was representative of the Spanish population in these relevant sociodemographic characteristics. Our two dependent variables were respondents' perceptions of (1) the Covid-19 booster vaccine effectiveness and (2) the safety of the Covid-19 vaccine for children. We focused on citizens' evaluations of these two characteristics of vaccines for theoretical and practical reasons. Theoretically, vaccine effectiveness and vaccine safety are two relevant characteristics that can influence vaccine hesitancy. Second, in practical terms, booster shot effectiveness and vaccine safety for children were two important topics regarding Covid-19 after the initial phases of the vaccination campaigns. Every survey respondent had to provide an evaluation of these two issues on a 1–10 scale. However, before answering each question, most of the respondents were randomly exposed to different vaccine messages. Additionally, some respondents did not receive any messages, as they constituted the control group that was used as a benchmark to compare the effectiveness of each message framing type. All survey respondents participated in both experiments but, as each of the experiments focused on different elements of vaccines, we minimised the chances that the first set of messages (about booster shot effectiveness) could influence respondents' answers to the second question (vaccine safety for children).

Table 1 summarises the types of treatment options that were provided to respondents for each experiment (the exact wording of all the treatments and framing for the dependent variables are presented in the Appendix). Table 1 also features the average time that respondents took to answer the question.

Table 1. Types of treatments and average time of answering

Experiment	Treatment version	Respondents that received the treatment	Average time spent answering (in seconds)
1. Perceptions of booster shot effectiveness	Version 1 (Control group): No treatment (just the question about vaccine effectiveness)	300	14
	Version 2: Scientific community (with consensus)	299	19
	Version 3: Scientific community (with dissensus)	300	17
	Version 4: Government	301	21
	Version 5: Social media influencer	300	16
	Version 6: Medical doctors	300	13
2. Perceptions of vaccine safety for children	Version 1 (Control group): Neutral message about children vaccination campaign	359	11
	Version 2: Storytelling message opposed to vaccination	360	23
	Version 3: Data-based message opposed to vaccination	360	17
	Version 4: Storytelling message in favour of vaccination	360	25
	Version 5: Data-based message in favour of vaccination	361	19

The fact that the survey was conducted online is interesting for our experiment. On the one hand, in this context, less attention is paid than in a laboratory setting (Mutz, 2011, pp. 12-13), which would go against detecting any effect of the treatments. As argued by Mutz,

“the distractions of everyday life can reduce the likelihood that a treatment will have an impact” (Mutz, 2011, p. 13).

On the other hand, however, an online survey that is answered via computers or smartphones takes place in a similar environment to that in which an important part of the population consumes information. Therefore, even if the treatments of the survey experiment are in some ways different to the actual messages received by citizens in traditional and social media, the context in which the messages are received is similar, which arguably increases the external validity of the experiment.

All respondents gave their permission to participate in the research, and the data were completely anonymised. To comply with ethical requirements, the project *PredCov* (Multi-source and multi-method prediction to support Covid-19 policy decision-making) in which this study was developed, had been approved by the *Ethics Committee of Madrid University Carlos III* under the CEI_22 protocol. At the end of the survey, the participants had access to a text related to the benefits of vaccination and the booster dose, as well as links to obtain more information on the subject. The research protocol was previously published in *AsPredicted* under code 97920.

Some authors argue that the most appropriate analysis of an experiment is the comparison of mean answers across treatment groups (Mutz, 2011). This amounts to checking to what extent the answers provided by the respondents vary depending on the message that they received before answering the question. If we detect statistically significant differences in the response of a group with respect to the control group, this would mean that the treatment had an effect on respondents' perceptions. In our case, as we were dealing with multiple experimental groups and dependent variables on a 10-point scale, we performed Mann-Whitney U tests (also known as Wilcoxon rank-sum tests). This test is the non-parametric counterpart of the t-test (it does not require a normal distribution of replies) and we used it because, as shown, the distributions of replies were not normal and were skewed, as most respondents considered that Covid-19 vaccines were effective and safe.

With the Wilcoxon rank-sum test, we compared if the distributions of answers to different experimental versions were the same or different. Alternatively, this test could be interpreted as detecting differences in the median answers of each distribution. Secondly, we also fitted linear regression models that included relevant covariates. Although the randomisation of the survey messages already ensured that the effects found were related to the messages themselves and not due to other factors, the use of linear regressions as an additional analysis enabled us to incorporate and assess the effect of other factors on the evaluations of booster shot effectiveness and of vaccine safety for children.

4. Results

First, regarding the sociodemographic composition of our sample, in the tables 1 and 2 we can see that there was a similar proportion of males and females (Table 2). Additionally, all the Autonomous Communities (Spanish regional entities) were represented (Table 3). Finally, regarding age, we had a certain over-representation of younger respondents (Table 4), which was also related to the over-representation of more highly educated respondents in our sample.

Table 2. Participants' gender

Sex	Freq.	Per cent
Female	911	50.61
Male	889	49.39
Total	1800	100

Table 4. Participants' age

Age	Freq.	Per cent
18–24	200	11.11
25–34	293	16.28
35–44	390	21.67
45–54	368	20.44
55–64	279	15.50
65–74	193	10.72
More than 74	77	4.28
Total	1800	100

Table 3. Participants' residence in Spain

Autonomous community	Freq.	Per cent
Andalucía	326	18.11
Aragón	50	2.78
Asturias	40	2.22
Islas Baleares	42	2.33
Islas Canarias	81	4.50
Cantabria	23	1.28
Castilla La Mancha	79	4.39
Castilla y León	94	5.22
Cataluña	294	16.33
Comunidad Valenciana	191	10.61
Extremadura	41	2.28
Galicia	105	5.83
La Rioja	13	0.72
Madrid	253	14.06
Murcia	58	3.22
Navarra	25	1.39
País Vasco	85	4.72
Total	1800	100

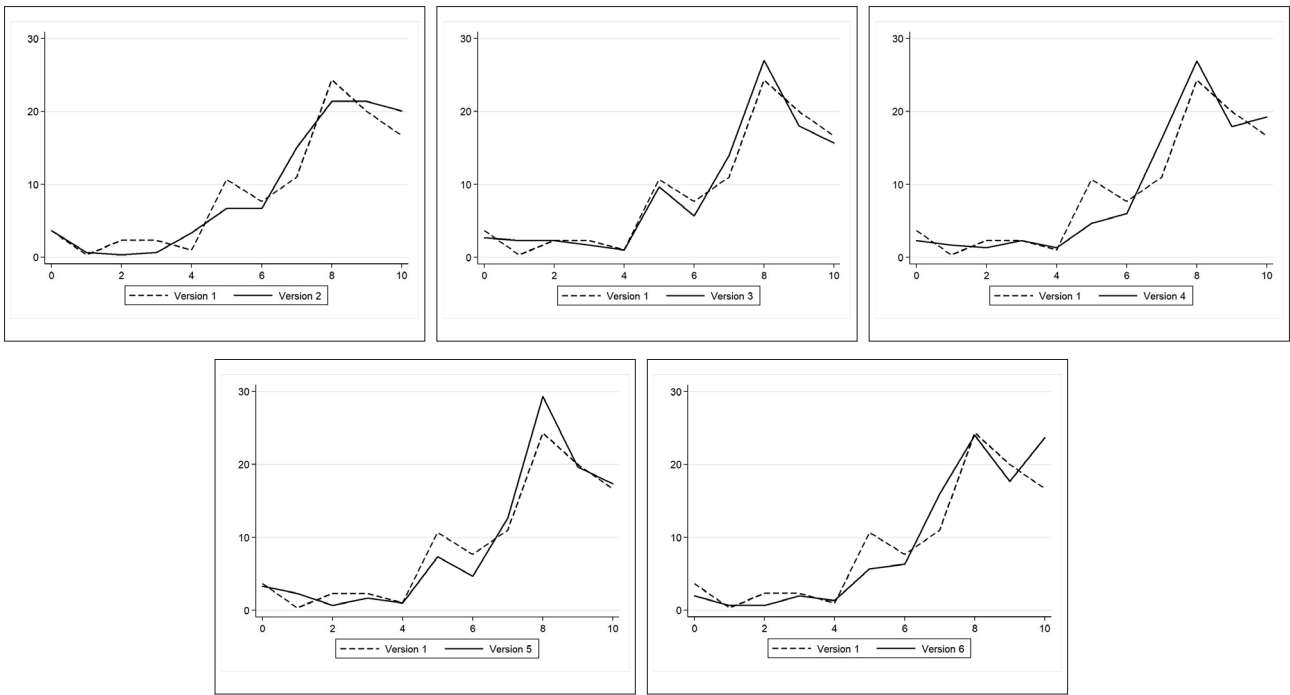


Figure 1. Information sources and perception regarding the Covid-19 booster vaccine effectiveness (by treatment group)

Concerning the answers to the experimental questions first, regarding the Covid-19 booster vaccine effectiveness (Figure 1), most respondents in all groups tended to consider the booster shots as effective or very effective (rating them with values higher than 5). In fact, the most common score across all groups was eight. Moreover, the answers of the groups that received messages were quite similar to the control group (Version 1, which only received the question without a previous message), which implies that the experimental messages were not significantly effective. The comparison between Version 1 and the other versions can be seen in the different panels of Figure 1. However, respondents that received Version 6 (the message with ‘medical doctors’ as sources) tended to value the booster shots more positively. In particular, they rated the booster shots with the maximum score to a greater extent than the control group (and the other groups).

Respondents’ perceptions about the Covid-19 vaccine safety for children were also predominantly positive, as shown in Figure 2. However, in this case, there were higher proportions of respondents who considered that Covid-19 vaccines were not safe for children. Additionally, the patterns of answers are similar across groups, but Versions 5 and 6 (the ones with pro-vaccination messages) showed a higher proportion of maximum scores than the control group (Version 1). This implies that the respondents who received these versions tended to value vaccine safety for children more than the respondents that received a neutral message (control group).

As shown in Table 5, the first experiment indicated that there was no statistically significant difference between the control group and the first four treatments, but the answers of those who received a message in which the source were ‘medical doctors’ were different to the answers of the control group at a statistically significant level ($p < 0.05$).

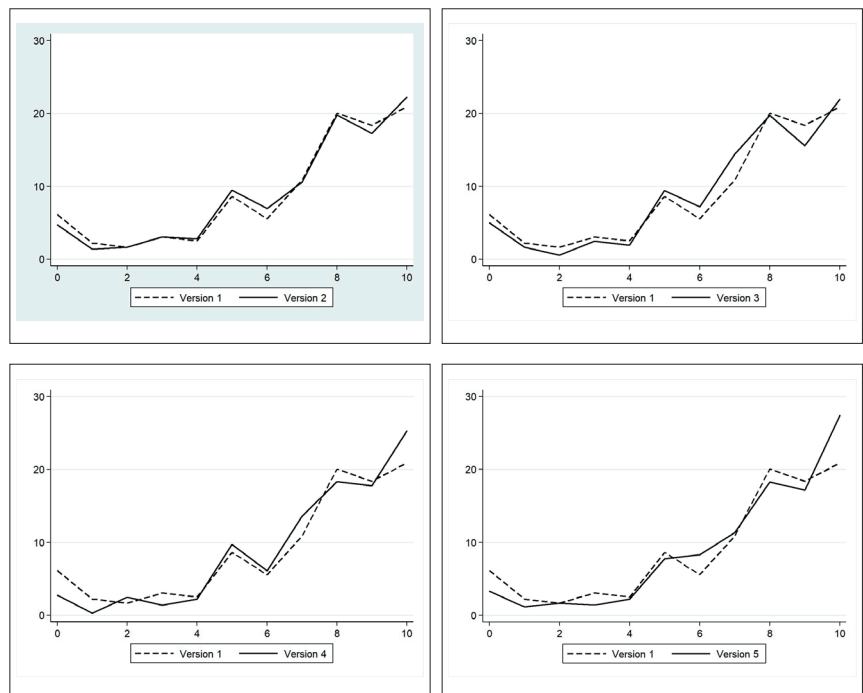


Figure 2. Message framing and perception of the Covid-19 vaccine safety for children (by treatment group)

Table 6 shows the results of the second experiment. We found that pro-vaccine messages (both storytelling and scientific messages) influenced respondents if we compared their answers with those of the control group that received a neutral message about child vaccination. In the case of pro-vaccine storytelling, there was a difference with respect to the control group with a 90% level of confidence ($p < 0.1$), while there was a difference between Group 5 (pro-vaccine scientific message) and the control group with a 95% level of confidence ($p < 0.05$).

The regression models (Table 7) align with the previous analysis. In the first experiment (Model 1), only the treatment that mentioned medical doctors (Version 6) had an influence on respondents' attitudes towards booster shot effectiveness. More specifically, this treatment resulted in 0.41 points increase in respondents' perceptions of vaccine effectiveness on a 1–10 scale, compared with the control group (which would be Version 1, the reference category).

Besides the effect of the experimental treatments, the regression coefficients also showed that males and older respondents regarded booster shots as more effective than females and younger respondents while holding the rest of the factors fixed (Wooldridge, 2016, p. 61). More specifically, males considered that booster shots were 0.23 points more effective, on average than females (on the 1–10 scale). Regarding the effect of age, the 0.006 coefficient means that, for instance, a difference of 45 years of age (between a 20-year-old and a 65-year-old individual) was associated with a 0.29 points more positive evaluation of booster shots. On the other hand, the educational level and ideology did not appear to be associated with attitudes towards booster shots in a statistically significant way.

Regarding the second experiment, we again saw an effect of pro-vaccine messages, both in the form of storytelling and scientific data message framing. This aligns with the results of the previous analysis (i.e. the comparison of each of the experimental groups' answers via Wilcoxon rank-sum tests). The effects of storytelling and scientific data framing were slightly higher than the effect of Version 6 of Experiment 1, and similar to each other. In addition, age was associated with more positive evaluations of vaccine safety. In this case, the difference between a 20-year-old and a 65-year-old individual was that the latter made evalua-

Table 5. Booster shot effectiveness

Experiment 1 (Pairwise comparisons)	p-value
Control and Group 2 (scientific consensus)	0.172
Control and Group 3 (scientific dissensus)	0.783
Control and Group 4 (government)	0.251
Control and Group 5 (influencers)	0.344
Control and Group 6 (medical doctors)	0.029**

Table 6. Vaccine safety for children

Experiment 2 (Pairwise comparisons)	p-value
Control and Group 2 (anti-vaccine storytelling)	0.739
Control and Group 3 (anti-vaccine scientific data)	0.872
Control and Group 4 (pro-vaccine storytelling)	0.093 *
Control and Group 5 (pro-vaccine scientific data)	0.036**

Table 7. Results of the regression models

Variables	Model 1 first experiment	Model 2 second experiment
Experiment (ref: Version 1)		
Version 2	0.246 (0.190)	0.133 (0.194)
Version 3	-0.0206 (0.190)	0.165 (0.194)
Version 4	0.254 (0.189)	0.505*** (0.194)
Version 5	0.219 (0.190)	0.483** (0.194)
Version 6	0.421** (0.190)	
Age	0.00638* (0.00358)	0.00742* (0.00400)
Sex	0.230** (0.112)	0.418*** (0.126)
Educational level (ref: Primary studies)		
Lower secondary (ESO)	-0.250 (0.368)	-0.550 (0.412)
Upper secondary (Bachillerato)	0.460 (0.344)	0.273 (0.385)
Vocational training (Middle)	-0.0573 (0.362)	-0.287 (0.406)
Vocational training (Upper)	0.0951 (0.349)	-0.321 (0.391)
Tertiary studies	0.460 (0.330)	0.103 (0.369)
Other studies	-0.0889 (0.768)	-0.877 (0.861)
Ideology	-0.0330 (0.0231)	-0.0548** (0.0259)
Constant	6.840*** (0.412)	6.897*** (0.465)
Observations	1,800	1,799
R-squared	0.024	0.030

Notes: Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

tions that were 0.33 points more positive, on average, than the former. According to the model results, males made evaluations 0.48 points more positive than females. Ideology appears to be related to the perception of vaccine safety for children; more right-wing respondents regarded vaccines as slightly less safe for children than

left-wing respondents. More specifically, a unit increase in the ideology scale (that is, having a more right-wing ideology) decreased the evaluation of vaccine safety for children by 0.05 points on the 1–10 scale. Note, however, that the effect was quite small. According to this regression coefficient, a nine-point difference in ideology (which is the same as going from extreme left to extreme right) resulted in a 0.45-point decrease in the evaluation of vaccine safety for children.

In sum, it is important to note that the experimental treatments had a slightly greater effect than the other predictive factors (age, gender, and ideology) on respondents' evaluations.

5. Discussion

Our experiment on disinformation and vaccines suggests that different information sources and narratives can influence people's trust in and acceptance of vaccines. The first experiment found that only the message from medical doctors influenced respondents' perceptions of the booster vaccine's effectiveness. The second experiment showed that pro-vaccine messages, both in the form of storytelling and scientific data message framing, had a positive effect on respondents' attitudes toward children's vaccine safety. These findings suggest that the source of the message is an important factor in influencing vaccine acceptance, with messages from medical doctors being particularly effective. Moreover, pro-vaccine messages can be effective in increasing vaccine acceptance. Furthermore, males and older respondents rated booster vaccines as more effective than females and younger respondents, while education level had no influence on the perception of the Covid-19 booster dosage. On the other hand, right-wing respondents believed vaccinations are somewhat less safe for children than left-wing respondents. These findings have further relevance as research suggests that exposure to disinformation and various narratives might influence people's trust and acceptance of vaccines. Studies have found that distrust in vaccines, mistrust in government and conspiracy beliefs can increase vaccine hesitancy (Jennings *et al.*, 2021). Additionally, exposure to Covid-19 vaccine misinformation on social media has been shown to impact vaccine acceptance (Cascini *et al.*, 2022; Loomba *et al.*, 2021).

The source of the message is an important factor in influencing vaccine acceptance, with messages from health professionals being particularly effective (Motta *et al.*, 2021; Unicef, 2020). Our results are in agreement with a recent publication in Nature (Bartoš *et al.*, 2022) that found that only the message from medical doctors influenced respondents' perceptions of booster vaccine effectiveness. That study found that communicating doctors' consensus persistently increases vaccine uptake, especially among those who underestimated doctors' trust and vaccination recommendations. Indeed, health professionals and public health experts hoping to encourage vaccination may be effective by appealing to personal and collective health risks, or the economic imperatives of vaccinating. This is because a key factor in gaining acceptance of a new vaccine is trust (AAFP, 2021). Media disinformation can cause public doubts about disease spread, prevention, lethality, and vaccine safety, and can promote mistrust of the government, policymakers, health authorities and pharmaceutical companies. Direct messaging from other health-care workers (HCWs), whether in the capacity of a personal physician or co-worker, may be especially effective in decreasing vaccine hesitancy. The population generally has a high level of trust in their physician (AAFP, 2021). Confidence among HCWs can be heightened through discussion, eliciting their concerns and involving them in vaccine recommendations. Indeed, the *Vaccine Messaging Guide* by Unicef (2020) suggests that credible communicators are essential for building trust with the audience. Health-care professionals are among the most trusted sources of health information and provider recommendations of vaccination significantly increase vaccine uptake. This makes them excellent messengers about vaccination, especially if they vaccinate themselves and their children. However, Motta *et al.* (2021) found that messages originating from expert sources are not necessarily less effective than messages from ordinary citizens, which are also effective at increasing intended vaccine uptake. Here we can see that both health professionals and ordinary people sharing positive stories about vaccines may be good allies in increasing vaccine uptake.

Another important factor influencing vaccine acceptance is message framing, i.e. storytelling or scientific data message framing. In this regard, our results are well aligned with previous studies which demonstrated that when presented alone, statistical information supplied by an expert (i.e. the science-supporting message) on vaccine safety and efficacy resulted in higher pro-vaccine attitudes than the control neutral message (Kuru *et al.*, 2021). However, while pro-vaccine messages can have a positive effect on respondents' attitudes toward children's vaccine safety, their effectiveness may vary depending on existing parental attitudes toward vaccines. For some parents, pro-vaccine messages may increase misperceptions or reduce vaccination intention (Nyhan *et al.*, 2014). Therefore, it is crucial that public health messages on vaccines be evaluated before being distributed to determine their effectiveness, particularly among resistant or sceptical parents.

“ Our experiment on disinformation and vaccines suggests that different narratives can influence people's trust and acceptance ”

“ The source of the message is an important factor in influencing vaccine acceptance, with messages from medical doctors being particularly effective ”

The regression coefficients also showed that males and older respondents were more likely to perceive vaccines as effective, while ideology was related to the perception of vaccine safety for children. Similar research also found that males and older respondents were more likely to perceive vaccines as effective (El-Mohandes *et al.*, 2021; Lindholt *et al.*, 2021). Specifically, females

have a five percentage point lower acceptance of an approved vaccine compared with males (Lindholt *et al.*, 2021). Unwillingness to vaccinate was significantly lower among respondents age 60 and older compared to younger respondents (El-Mohandes *et al.*, 2021). In addition, ideology was also noted in another study with similar findings in relation to the perception of vaccine safety for children, showing how conservative respondents are less likely to express pro-vaccine attitudes than liberal respondents (Baumgaertner; Carlisle; Justwan, 2018). These findings suggest that public health campaigns aimed at increasing vaccine acceptance may need to consider demographic and ideological factors in addition to the source and content of the message. In relation to these findings, a previous study (Jensen; Ayers; Koskan, 2022) was also aligned with our study showing that:

- males and older respondents were more likely to perceive vaccines as effective;
- ideology was related to the perception of vaccine safety for children (parents expressing conservative ideology were less likely to vaccinate their children against Covid-19);
- concerns about the vaccine's long-term side effects were uniquely associated with an increased likelihood of parents being unsure about vaccinating their children compared to positive vaccine intention.

However, in relation to political ideology, a recent study (Rasul; Ahmed, 2023) found that misinformation exposure increases Covid-19 vaccine hesitancy for conservatives and moderates but not for liberals. However, perceived misinformation exposure influences Covid-19 vaccine hesitancy among conservatives only if they are also 'flu vaccine-hesitant. This means, that ideology may produce a generalised vaccine hesitancy, not only against the Covid-19 vaccine.

We would like to declare some limitations that should be considered when interpreting the abovementioned results. First and foremost, the survey was self-administered, which does not guarantee that all respondents understood the questions or were totally honest in their responses. Second, our sample included a large proportion of university graduates, which focuses on a certain demographic and may have consequences for the results. Aside from that, the cross-sectional aspect of this study and the descriptive technique of the analysis should be highlighted. This limits us from proving the possible causal impacts of certain vaccine-related behaviours or attitudes in connection to sociodemographic characteristics, prior Covid-19 experience and usage of both conventional and social media. Future studies might use cross-tabulation analysis to investigate the association between these various factors. Moreover, regarding our first experiment, another factor that will likely moderate the effect of messages issued by different sources is the channel through which respondents receive those messages. Our design cannot assess the extent to which different channels will increase or decrease the effect of these messages, so this interaction will need to be explored in future analyses. Despite these limitations, these results highlight the importance of effective communication strategies in promoting vaccine acceptance, particularly in the face of disinformation and lack of scientific consensus, highlighting the responsibility of the media and institutions in communicating with the public during a health crisis (Costa-Sánchez; López-García, 2020). Our study might support the development of vaccination campaigns by public health departments as well as increase vaccination rates in the general population. Future research may need to investigate the long-term effectiveness of these strategies, as well as the effectiveness of interventions targeted at specific demographic and ideological groups. In addition, other lines of research may focus on: 1) comparing these results with those of other similar diseases, to develop more effective messages to encourage adult vaccination; and 2) an experiment to assess the effect of type of content (totally false, misleading message, manipulation of information) on people's behaviour towards vaccines.

“Males and older respondents rate booster vaccines as more effective than women and younger respondents, while education level had no influence on the perception of the Covid-19 booster dosage”

“It is crucial that public health messages on vaccines be evaluated before being distributed to determine their effectiveness, particularly among resistant or sceptical parents”

6. References

AAFP (2021). *Four reasons for Covid-19 vaccine hesitancy among health care workers, and ways to counter them*. American Academy of Family Physicians.

https://www.aafp.org/pubs/fpm/blogs/inpractice/entry/countering_vaccine_hesitancy.html

AEP (2022). “Evolución de la vacunación frente a la Covid de niños y adolescentes en España”. *Comité Asesor de Vacunas de la Asociación Española de Pediatría*, 2 mayo.

<https://vacunasaep.org/print/profesionales/noticias/covid-19-estado-vacunacion-ninos-y-adolescentes6-29abr2022>

- Al-Daghastani, Tamara; Tadros, Odate; Arabiyat, Shereen; Jaber, Deema; AlSalamat, Husam** (2021). "Pharmacists' perception of the coronavirus pandemic (Covid-19) in Jordan: a cross-sectional study". *International journal of environmental research and public health*, v. 18, n. 21, 11541.
<https://doi.org/10.3390/ijerph182111541>
- Allington, Daniel; McAndrew, Siobhan; Moxham-Hall, Vivienne; Duffy, Bobby** (2023). "Coronavirus conspiracy suspicions, general vaccine attitudes, trust and coronavirus information source as predictors of vaccine hesitancy among UK residents during the Covid-19 pandemic". *Psychological medicine*, v. 53, n. 1, pp. 236-247.
<https://doi.org/10.1017/S0033291721001434>
- Baig, Mukhtiar; Jameel, Tahir; Alzahrani, Sami; Mirza, Ahmad A.; Gazzaz, Zohair H.; Ahmad, Tauseef; Baig, Fizzah; Almurashi, Saleh H.** (2020). "Predictors of misconceptions, knowledge, attitudes, and practices of Covid-19 pandemic among a sample of Saudi population". *Plos one*, v. 15, n. 12, p. e0243526.
<https://doi.org/10.1371/journal.pone.0243526>
- Bartoš, Vojtěch; Bauer, Michal; Čahlíková, Jana; Chytilová, Julie** (2022). "Communicating doctors' consensus persistently increases Covid-19 vaccinations". *Nature*, v. 606, n. 7914, pp. 542-549.
<https://doi.org/10.1038/s41586-022-04805-y>
- Baumgaertner, Bert; Carlisle, Juliet; Justwan, Florian** (2018). "The influence of political ideology and trust on willingness to vaccinate". *Plos one*, v. 13, n. 1, e0191728.
<https://doi.org/10.1371/journal.pone.0191728>
- Boyson, Aaron R.; Zimmerman, Rick S.; Shoemaker, Sarah** (2015). "Exemplification of HAART and HIV/AIDS: a news experiment". *Health communication*, v. 30, n. 9, pp. 901-910.
<https://doi.org/10.1080/10410236.2014.903222>
- Cascini, Fidelia; Pantovic, Ana; Al-Ajlouni, Yazan A.; Failla, Giovanna; Puleo, Valeria; Melnyk, Andriy; Lontano, Alberto; Ricciardi, Walter** (2022). "Social media and attitudes towards a Covid-19 vaccination: a systematic review of the literature". *eClinicalMedicine*, v. 48, 101454.
<https://doi.org/10.1016/j.eclinm.2022.101454>
- Casero-Ripollés, Andreu** (2020). "Impact of Covid-19 on the media system: communicative and democratic consequences of news consumption during the outbreak". *Profesional de la información*, v. 29, n. 2, e290223.
<https://doi.org/10.3145/epi.2020.mar.23>
- Costa-Sánchez, Carmen; López-García, Xosé** (2020). "Comunicación y crisis del coronavirus en España: primeras lecciones". *Profesional de la información*, v. 29, n. 3, e290304.
<https://doi.org/10.3145/epi.2020.may.04>
- Dahlstrom, Michael F.** (2010). "The role of causality in information acceptance in narratives: an example from science communication". *Communication research*, v. 37, n. 6, pp. 857-875.
<https://doi.org/10.1177/0093650210362683>
- Dahlstrom, Michael F.** (2014). "Using narratives and storytelling to communicate science with nonexpert audiences". In: *Proceedings of the National Academy of Sciences*, pp. 13614-13620.
<https://doi.org/10.1073/pnas.1320645111>
- Das, Ronnie; Ahmed, Wasim** (2022). "Rethinking fake news: disinformation and ideology during the time of Covid-19 global pandemic". *IIM Kozhikode society & management review*, v. 11, n. 1, pp. 146-159.
<https://doi.org/10.1177/22779752211027382>
- El-Mohandes, Ayman; White, Trenton M.; Wyka, Katarzyna; Rauh, Lauren; Rabin, Kenneth; Kimball, Spencer H.; Ratzan, Scott C.; Lazarus, Jeffrey V.** (2021). "Covid-19 vaccine acceptance among adults in four major US metropolitan areas and nationwide". *Scientific reports*, v. 11, n. 1.
<https://doi.org/10.1038/s41598-021-00794-6>
- García-Marín, David** (2020). "Infodemia global: desórdenes informativos, narrativas fake y fact-checking en la crisis de la Covid-19". *Profesional de la información*, v. 29, n. 4, e290411.
<https://doi.org/10.3145/epi.2020.jul.11>
- Greene, Ciara M.; Murphy, Gillian** (2021). "Quantifying the effects of fake news on behavior: Evidence from a study of Covid-19 misinformation". *Journal of experimental psychology: Applied*, v. 27, n. 4, pp. 773-784.
<https://doi.org/10.1037/xap0000371>
- Jennings, Will; Stoker, Gerry; Bunting, Hannah; Valgarðsson, Viktor-Orri; Gaskell, Jennifer; Devine, Daniel; McKay, Lawrence; Mills, Melinda C.** (2021). "Lack of trust, conspiracy beliefs, and social media use predict Covid-19 vaccine hesitancy". *Vaccines*, v. 9, n. 6, 593.
<https://doi.org/10.3390/vaccines9060593>

- Jensen, Ulrich T.; Ayers, Stephanie; Koskan, Alexis M.** (2022). "Video-based messages to reduce Covid-19 vaccine hesitancy and nudge vaccination intentions". *Plos one*, v. 17, n. 4, e0265736.
<https://doi.org/10.1371/journal.pone.0265736>
- Joubert, Marina; Davis, Lloyd; Metcalfe, Jenni** (2019). "Storytelling: the soul of science communication". *Journal of science communication*, v. 18, n. 5.
<https://doi.org/10.22323/2.18050501>
- Karabela, Şemsi-Nur; Coşkun, Filiz; Hoşgör, Haydar** (2021). "Investigation of the relationships between perceived causes of Covid-19, attitudes towards vaccine and level of trust in information sources from the perspective of infodemic: the case of Turkey". *BMC public health*, v. 21, n. 1, p. 1195.
<https://doi.org/10.1186/s12889-021-11262-1>
- Kuru, Ozan; Stecula, Dominik; Lu, Hang; Ophir, Yotam; Chan, Man-Pui-Sally; Winneg, Ken; Jamieson, Kathleen-Hall; Albarraçín, Dolores** (2021). "The effects of scientific messages and narratives about vaccination". *Plos one*, v. 16, n. 3, e0248328.
<https://doi.org/10.1371/journal.pone.0248328>
- Larson, Heidi J.; Jarrett, Caitlin; Eckersberger, Elisabeth; Smith, David M. D.; Paterson, Pauline** (2014). "Understanding vaccine hesitancy around vaccines and vaccination from a global perspective: A systematic review of published literature, 2007-2012". *Vaccine*, v. 32, n. 19, pp. 2150-2159.
<https://doi.org/10.1016/j.vaccine.2014.01.081>
- Lindholt, Marie-Fly; Jørgensen, Frederik; Bor, Alexander; Petersen, Michael-Bang** (2021). "Public acceptance of Covid-19 vaccines: cross-national evidence on levels and individual-level predictors using observational data". *BMJ open*, v. 11, n. 6, e048172.
<https://doi.org/10.1136/bmjopen-2020-048172>
- Lomba, Sahil; De-Figueiredo, Alexandre; Piatek, Simon J.; De-Graaf, Kristen; Larson, Heidi J.** (2021). "Measuring the impact of Covid-19 vaccine misinformation on vaccination intent in the UK and USA". *Nature human behaviour*, v. 5, n. 3, pp. 337-348.
<https://doi.org/10.1038/s41562-021-01056-1>
- MacDonald, Noni E.** (2015). "Vaccine hesitancy: definition, scope and determinants". *Vaccine*, v. 33, n. 34, pp. 4161-4164.
<https://doi.org/10.1016/j.vaccine.2015.04.036>
- Malecki, Kristen M. C.; Keating, Julie A.; Safdar, Nasia** (2021). "Crisis communication and public perception of Covid-19 risk in the era of social media". *Clinical infectious diseases*, v. 72, n. 4, pp. 697-702.
<https://doi.org/10.1093/cid/ciaa758>
- Ministerio de Sanidad (2022). GIV Covid-19. Gestión integral de la vacunación Covid-19. Informe de actividad.
https://www.sanidad.gob.es/profesionales/saludPublica/ccayes/alertasActual/nCov/documentos/Informe_GIV_comunicacion_20221104.pdf
- Motta, Matt; Sylvester, Steven; Callaghan, Timothy; Lunz-Trujillo, Kristin** (2021). "Encouraging Covid-19 vaccine uptake through effective health communication". *Frontiers in political science*, v. 3, e630133.
<https://doi.org/10.3389/fpos.2021.630133>
- Mourali, Mehdi; Drake, Carly** (2022). "The challenge of debunking health misinformation in dynamic social media conversations: online randomized study of public masking during Covid-19". *Journal of medical internet research*, v. 24, n. 3, e34831.
<https://doi.org/10.2196/34831>
- Mutz, Diana C.** (2011). *Population based survey experiments*. Princeton: Princeton University Press. ISBN: 978 0 691144528
- Nan, Xiaoli; Dahlstrom, Michael F.; Richards, Adam; Rangarajan, Sarani** (2015). "Influence of evidence type and narrative type on HPV risk perception and intention to obtain the HPV vaccine". *Health communication*, v. 30, n. 3, pp. 301-308.
<https://doi.org/10.1080/10410236.2014.888629>
- Nyhan, Brendan; Reifler, Jason; Richey, Sean; Freed, Gary L.** (2014). "Effective messages in vaccine promotion: a randomized trial". *Pediatrics*, v. 133, n. 4, e83542.
<https://doi.org/10.1542/peds.2013-2365>
- Park, Sungchul; Massey, Philip M.; Stimpson, Jim** (2021). "Primary source of information about Covid-19 as a determinant of perception of Covid-19 severity and vaccine uptake: source of information and Covid-19". *Journal of general internal medicine*, v. 36, n. 10, pp. 3088-3095.
<https://doi.org/10.1007/s11606-021-07080-1>
- Pfeiffer-Castellanos, Marcelo-Eduardo-Pfeiffer** (2014). "A narrativa nas pesquisas qualitativas em saúde". *Ciência & saúde coletiva*, v. 19, n. 4, pp.1065-1076.
<https://doi.org/10.1590/1413-81232014194.12052013>

Rasul, Muhammad-Ehab; Ahmed, Saifuddin (2023). "Not all conservatives are vaccine hesitant: examining the influence of misinformation exposure, political ideology, and flu vaccine acceptance on Covid-19 vaccine hesitancy". *Vaccines*, v. 11, n. 3, 586.

<https://doi.org/10.3390/vaccines11030586>

Reno, Chiara; Maietti, Elisa; Di-Valerio, Zeno; Montalti, Marco; Pia-Fantini, Maria; Gori, Davide (2021). "Vaccine hesitancy towards Covid-19 vaccination: investigating the role of information sources through a mediation analysis". *Infectious disease reports*, v. 13, n. 3, pp. 712-723.

<https://doi.org/10.3390/idr13030066>

Sallam, Malik; Dababseh, Deema; Eid, Huda; Al-Mahzoum, Kholoud; Al-Haidar, Ayat; Taim, Duaa; Yaseen, Alaa; Ababneh, Nidaa A.; Bakri, Faris G.; Mahafzah, Azmi (2021). "High rates of Covid-19 vaccine hesitancy and its association with conspiracy beliefs: a study in Jordan and Kuwait among other Arab countries". *Vaccines*, v. 9, n. 1, pp. 42-58.

<https://doi.org/10.3390/vaccines9010042>

Sánchez-Duarte, José-Manuel; Magallón-Rosa, Raúl (2020). "Infodemia y Covid-19: evolución y viralización de informaciones falsas en España". *Revista española de comunicación en salud*, sup. 1, pp. S31-S41.

<https://doi.org/10.20318/recs.2020.5417>

Shelby, Ashley; Ernst, Karen (2013). "Story and science: how providers and parents can utilize storytelling to combat anti-vaccine misinformation". *Human vaccines & immunotherapeutics*, v. 9, n. 8, pp. 1795-1801.

<https://doi.org/10.4161/hv.24828>

Sulistyawati, Sulistyawati; Rokhmeyant, Rokhmeyant; Aji, Budi; Wijayanti, Siwi-Pramatama-Mars; Hastuti, Siti-Kurnia-Widi; Sukesi, Tri-Wahyuni; Mulasari, Surahma-Asti (2021). "Knowledge, attitudes, practices and information needs during the Covid-19 pandemic in Indonesia". *Risk management and healthcare policy*, v. 14, pp. 163-175.

<https://doi.org/10.2147/RMHP.S288579>

Thorpe, Alistair; Fagerlin, Angela; Drews, Frank A.; Butler, Jorie; Stevens, Vanessa; Riddoch, Marian S.; Scherer, Laura D. (2022). "Communications to promote interest and confidence in Covid-19 Vaccines". *American journal of health promotion*, v. 36, n. 6, pp. 976-986.

<https://doi.org/10.1177/08901171221082904>

Unicef (2020). *Vaccine messaging guide*.

<https://www.unicef.org/media/93661/file/Vaccinemessagingguide.pdf>

Van-Krieken, Kobie; Sanders, José (2021). "What is narrative journalism? A systematic review and an empirical agenda". *Journalism*, v. 22, n. 6, pp. 1393-1412.

<https://doi.org/10.1177/1464884919862056>

Vandenberg, Lisa; Meppelink, Corine S.; Sanders, José; Fransen, Marieke L. (2022). "Facts tell, stories sell? Assessing the availability heuristic and resistance as cognitive mechanisms underlying the persuasive effects of vaccination narratives". *Frontiers in psychology*, v. 13, e837346.

<https://doi.org/10.3389/fpsyg.2022.837346>

We are Social (2022). *Digital report 2022: el informe sobre las tendencias digitales, redes sociales y mobile*.

<https://wearesocial.com/es/blog/2022/01/digital-report-2022-el-informe-sobre-las-tendencias-digitales-redes-sociales-y-mobile>

Wooldridge, Jeffrey (2016). *Introductory econometrics: a modern approach*. 5th edition. Mason: Cengage Learning. ISBN: ISBN: 978 1 111 53104 1

https://economics.ut.ac.ir/documents/3030266/14100645/Jeffrey_M._Wooldridge_Introductory_Econometrics_A_Modern_Approach__2012.pdf

7. Appendix

Original questions and messages used in the two experiments. For each experiment, we show the translated text in English. Participants allocated in the control group (V1) did not receive any message prior the question (experiment 1) or received a neutral message (experiment 2). The remaining participants read a vaccine-related message prior the question.

Table 1. Sources of information and the Covid-19 booster vaccine effectiveness (Experiment 1)

<p>Control Group: Did not receive any prior message V1 Point out, according to you opinion, the effectiveness of Covid-19 booster shots in order to prevent hospitalizations:</p>
<p>Source: 'Scientific consensus' V2 There is scientific consensus in that the Covid-19 booster shot is very effective against the coronavirus variants. Point out, according to you opinion, the effectiveness of Covid-19 booster shots in order to prevent hospitalizations:</p>
<p>Source: 'Scientific dissensus' V3 Some scientific research shows that Covid-19 booster shots are very effective against coronavirus variants, while other research shows that booster shots are not very effective. Point out, according to you opinion, the effectiveness of Covid-19 booster shots in order to prevent hospitalizations:</p>
<p>Source: 'Governmental' V4 The Spanish Government has declared that Covid-19 booster shots are very effective against coronavirus variants. Point out, according to you opinion, the effectiveness of Covid-19 booster shots in order to prevent hospitalizations:</p>
<p>Source: 'Influencer' V5 A message by an Instagram influencer explaining that Covid-19 booster shots are very effective against coronavirus variants has gone viral. Point out, according to you opinion, the effectiveness of Covid-19 booster shots in order to prevent hospitalizations:</p>
<p>Source: 'Medical doctors' V6 Medical doctors have declared that the Covid-19 booster shot is very effective against coronavirus variants. Point out, according to you opinion, the effectiveness of Covid-19 booster shots in order to prevent hospitalizations:</p>

Table 2. Message framing and the Covid-19 vaccine safety for children (Experiment 2)

<p>Control Group: Neutral frame V1 Control group. Neutral message about children vaccination: The Covid-19 vaccination campaign continues this week, with the aim of vaccination children. In the afternoons, vaccination centres will receive children between 5-11 years of age. Point out to what extent do you agree with the following statement: "It is safe to vaccinate children":</p>
<p>Message frame: Anti-vaccine storytelling V2 Little Laura, of 7 years of age, can't stay still in the vaccination queue. Like the rest of the children, she doesn't want to be there. She would rather be at the park, playing with her friends. Her mother, María, also looks anxious with her arms crossed. She isn't sure if it is really necessary to vaccinate her daughter, she has read that the disease barely affects children. Her daughter was all right during the whole pandemic, why should she get the vaccine now? It doesn't make any sense. Her classmates haven't received the vaccine either. Perhaps she should leave without vaccinating her daughter. She grabs her by the hand and goes to the exit. She is sure that her daughter will continue to be all right, as always. Point out to what extent do you agree with the following statement: "It is safe to vaccinate children":</p>
<p>Message frame: Anti-vaccine scientific data V3 According to the latest epidemiological report by the <i>Carlos III Institute of Madrid</i>, children hospitalizations are rare and represent 0.78% of the overall hospitalizations since the beginning of the pandemic. Regarding deaths, seven children have passed away due to the Covid-19 during this period. The Spanish Society of Paediatric Infectiology considers that most children are safe against the effect of the pandemic, if compared with adults and older people. Point out to what extent do you agree with the following statement: "It is safe to vaccinate children":</p>
<p>Message frame: Pro-vaccine storytelling V4 Sara looks at the phone, she has received a message requiring her to go to the appointment to vaccinate Fernando this afternoon. Finally! She is relieved. She is always checking the news to see when the children vaccination campaign will start. She stops to think how this took so long... she was worried something bad could have happened to Fernando, but now she is more relaxed. She hugs her son, who is beside her in the couch. He has already suffered a lot due to his grandfather's death. We don't want this to happen also to Amelia, his grandmother. Point out to what extent do you agree with the following statement: "It is safe to vaccinate children":</p>
<p>Message frame: Pro-vaccine scientific data V5 The sixth wave has caused a lot of infections among children less than 12 years old, which now amounts to 13% of the cases. There are 393.394 cases since the beginning of the pandemic. Experts point out that even if children are less affected by hospitalizations and deaths, the vaccine will protect them against severe cases and deaths. Besides individual benefits, children vaccination will slow down cases of infections. Centres for Disease Control point that children vaccination will help stop the spread of the coronavirus towards "other people, even at home and at the school". Point out to what extent do you agree with the following statement: "It is safe to vaccinate children":</p>

Trust, disinformation, and digital media: Perceptions and expectations about news in a polarized environment

Javier Serrano-Puche; Natalia Rodríguez-Salcedo; María-Pilar Martínez-Costa

Nota: Este artículo se puede leer en español en:
<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87313>

Recommended citation:

Serrano-Puche, Javier; Rodríguez-Salcedo, Natalia; Martínez-Costa, María-Pilar (2023). "Trust, disinformation, and digital media: Perceptions and expectations about news in a polarized environment". *Profesional de la información*, v. 32, n. 5, e320518.

<https://doi.org/10.3145/epi.2023.sep.18>

Article received on March 9th 2023

Approved on August 31st 2023



Javier Serrano-Puche

<https://orcid.org/0000-0001-6633-5303>

Universidad de Navarra
Facultad de Comunicación
Departamento de Comunicación Pública
31009 Pamplona, Spain
jserrano@unav.es



Natalia Rodríguez-Salcedo

<https://orcid.org/0000-0003-0968-0447>

Universidad de Navarra
Facultad de Comunicación
Departamento de Marketing y Empresas
de Comunicación
31009 Pamplona, Spain
nrodriguez@unav.es



María-Pilar Martínez-Costa ✉

<https://orcid.org/0000-0002-6564-0207>

Universidad de Navarra
Facultad de Comunicación
Departamento de Periodismo
31009 Pamplona, Spain
marcosta@unav.es

Abstract

The transformation that the communicative environment has undergone in recent decades poses a challenge for the media in relation to their audiences, as trust is sensitive to social, economic, cultural, and technological changes. The aim of this research is to deepen the understanding of the reasons and factors that influence the loss of trust in the media by audiences who traditionally trusted news more (young people and adults aged 25-54), examining the relevance of disinformation and polarization in discrediting the media. Firstly, the state of the issue is reviewed from the theoretical point of view and the data provided, among others, such as the *Digital News Report* and several studies about the global loss of trust in other institutions, with special reference to the *Edelman Trust Barometer*. Secondly, the article adopts a qualitative methodology to investigate the motivations and expectations of citizens regarding the media. Specifically, three discussion groups were held in various Spanish cities. To ensure representativeness, sociodemographic diversity was considered, including gender, age, and educational level criteria, among others. Among the findings, it stands out that one of the main reasons for distrust towards the media is the perception that news is biased for political or economic reasons. The Covid-19 pandemic, which was rife with disinformation, has influenced attitudes towards the media and the way news is consumed. Once the pandemic was overcome, trust in the media decreased and participants sought alternative sources of information. However, some differences in perceptions and consumption habits are noted depending on age and educational level. Finally, the research indicates that distrust extends well beyond the media ecosystem, affecting all institutions.

Keywords

Trust; Credibility; Disinformation; Media; Polarization; Covid-19; Coronavirus; Pandemics; Reputation; Press; Television; Radio; Digital media; Focus groups; Audiences; Spain; *Digital News Report*; *Edelman Trust Barometer*.



Funding

This article is a result of the projects “DigiNativeMedia II – Digital Native Media in Spain: journalistic typologies, strategies, competencies and sustainability (2022-2025)” PID2021-122534OB-C22, funded by the *State Research Agency, Ministry of Science, Innovation and Universities* of Spain and the *European Regional Development Fund*; and “RRSsalud Dynamics of dissemination of fake health news on social networks,” funded by the *BBVA Foundation* (2020-2022).

1. Introduction

Trust is a basic fact of social life (Luhmann, 1968), which represents an integral support for the members of a society (Hawley, 2012; Simmel, 1987). However, it is a fragile asset, since it entails risk and lack of certainty (Narbona; Pujol; Gregory, 2020), since, according to Han,

“trust is only possible in a middle state between knowing and not knowing. Trust means: despite not knowing in relation to the other, building a positive relationship with it” (Han, 2016, p. 91).

It is therefore a relational concept, determined as much by the expectations of the one who conveys it as by the actions of its receiver (Garusi; Splendore, 2023) and which, when applied beyond the original interpersonal context, functions as an “institutional economizer” (Rosanvallon, 2008), as it reduces the need for proof and verification of the goods and services that it provides.

Along the same lines, trust is also considered a basic component of the organizations’ so-called social license, a concept developed in the business world at the end of the 20th century by Jim Cooney (Bice; Brueckner; Pforr, 2017), although Morrison (2014) shrewdly suggested that it could be incorporated into any institution representing civil society, such as the media. The social license is not issued and is difficult to quantify and, consequently, to measure. However, going beyond the permits or licenses that are legally and politically necessary for the development of organizations’ business activities, social license complements them, as it represents citizens’ tacit approval of an organization (Morrison, 2014), without which it is difficult to survive in an environment of severe public scrutiny. Said social license is based on three factors –legitimacy, consent, and trust– which stem from what the organization does (Morrison, 2014) and not so much from what it is or claims to be. Therefore, trust appears to be a key element based, in addition, on the maintenance of active and two-way relationships with audiences (Morrison, 2014).

The cultural, economic, social, and technological changes experienced during the last two decades have led to a crisis of trust that affects a variety of institutions, including multinational companies, political parties, nongovernmental organizations (NGOs), religious institutions, and media companies (Edelman, 2023; Pérez-Latre, 2022). In the public communication sphere, trust is one of the elements that shape the connection between citizens and the media (Coleman, 2012). To be exact, it is one of the factors that influence people’s relationship with news (Moran; Nechushtai, 2022) and thus may impact their news consumption (Fletcher; Park, 2017; Tsfaty; Ariely, 2014).

Since trust is an integral factor that relies on the activity developed by the media and the way they relate to society, the objective of this research is to delve into the reasons behind and factors that influence the current loss of trust. For this reason, we will delve into the perceptions of audiences that traditionally trusted the news more (young people and adults aged 25-54 years), examining the importance of disinformation when it comes to the discrediting of the media.

2. State of affairs**2.1. Trust in the media**

Media trust is a central topic in academic research, where it has generated extensive literature (Kohring; Matthes, 2007; Meyer, 1988; Otto; Köhler, 2018; Strömbäck *et al.*, 2020). As Jakobsson and Stiernstedt (2023) point out, three main arguments regarding the importance of trust in the media can be found in the scientific literature. First, this is significant as it regards the situation of media companies as well as media as an industry (Picard, 2018; Vanacker; Belmas, 2009). Moreover, trust in the media is essential for democracy, as it helps to have informed citizens with the capacity for political engagement (Delli-Carpini; Keeter, 1996; Uslaner, 2002). Finally, trust in the media is related to other forms of trust (or social, political, and institutional license) and to a broader existential discussion on ontological security (Giddens, 1991). However, Jakobsson and Stiernstedt (2023) argue that these normative beliefs about the importance of trust are not always consistent with and do not always correlate with empirical findings. Other authors such as Strömbäck *et al.* (2020) or Prochazka and Schweiger (2019) also point out that research on trust lacks conceptual (what trust is and the level at which it operates) and methodological (how to measure it) precision, where, in addition, the preponderance of unidimensional quantitative techniques leads to an incomplete understanding of this phenomenon (Engelke; Hase; Winterlin, 2019; Garusi; Splendore, 2023).

At the theoretical level, research on trust has been developed together with studies on credibility, since both are interrelated and multidimensional concepts that are sometimes used synonymously. Studies on credibility are deeply rooted in the Communication discipline: research on the credibility of sources (Hovland; Weiss, 1951), on the credibility of di-

fferent media (press, radio, television, and online media; **Kiousis**, 2001; **Shaw**, 1973; **Westley**; **Severin**, 1964), or on the measurable dimensions that make up credibility (**Gaziano**; **McGrath**, 1986; **Meyer**, 1998). As with trust, there is a lack of consensus on how to conceptualize it and a lack of clear measurement scales (**Hanimann et al.**, 2023). In any case, it should be noted that credibility is more objective, insofar as it is based on a series of traits or characteristics that, if fulfilled, confer credibility (and thus provide confidence). The veracity of the information disseminated, impartiality, the media's independence from external agents, and their commitment to defending the audience's interests are some of the predictor variables of media credibility, making media outlets that have these worthy of trust (**Lee**, 2010). In other words, trust is the effect it has on the audience as a result of the quality of the news product (**Medina**; **Etayo-Pérez**; **Serrano-Puche**, 2023), since it depends on news being created using the appropriate procedures to ensure that it corresponds with the reality of the events that have occurred (**Fawzi et al.**, 2021; **Hanitzsch**; **Van-Dalen**; **Steindl**, 2018).

In sum, trust deals with an attitudinal issue, resulting from a cognitive and relational process in which the person subjectively examines and evaluates the media system, the qualities of an information source (the type of media outlet, the brand, or a journalist), or the content of its messages (**Lucassen**; **Schraagen**, 2012; **Otto**; **Köhler**, 2018; **Strömbäck et al.**, 2020). Therefore, it is worth taking into account the political-cultural context in which this relationship develops and citizens' expectations when it comes to the media, as they affect their perceptions and attitudes toward media institutions (**Gil de Zúñiga et al.**, 2019; **Kohring**; **Matthes**, 2007; **Tsfati**; **Ariely**, 2014).

2.2. Trust in the face of the challenges of today's media and social landscape

The transformation of the communicative environment poses a challenge for traditional media, since the new ecosystem is characterized by, among other factors, a blurring of the classical boundaries of journalism (**Carlson**; **Lewis**, 2015), the hybridization between classical and digital media logics (**Chadwick**, 2013), and users' high level of choice (**Van-Aelst et al.**, 2017), leading to very diverse players battling to catch their attention (**Wu**, 2016).

Among these new media actors, partisan and alternative news sources run rampant, which specifically rely on attacking the traditional media, accusing them of being untrustworthy, as part of their editorial strategy (**Thorbjørnsrud**; **Figenschou**, 2022). Also populist political bodies carry out campaigns to discredit and delegitimize journalistic work (**Carlson**; **Robinson**; **Lewis**, 2021; **Van-Dalen**, 2021), which helps to increase skepticism toward the media (**Tsfati**, 2003), and even hostility (**Gunther et al.**, 2017). On a larger scale, all this fuels social polarization, both ideologically –increasing the gap between the ends of the political spectrum– and affectively, encouraging among citizens a feeling of animosity toward those who are not part of their reference social group (**Rodríguez-Virgili**; **Portilla-Manjón**; **Sierra-Iso**, 2022; **Serrano-Puche**, 2021).

In parallel, the rise of disinformation poses a challenge for the media, as it increases citizens' uncertainty about the trustworthiness of the content circulating in the public sphere, and this leads to both less trust in the media (**Vaccari**; **Chadwick**, 2020) and a high percentage of citizens –ranging from 70% to 80% depending on the country– who fear that fake news will be used as a weapon (**Edelman**, 2018, p. 16). Based on a survey in 10 European countries, **Hameleers**, **Brosius**, and **De-Vreese** (2022) conclude that those users who have a stronger perception of disinformation are more likely to consume news on social networks and alternative, nonconventional media. Other research shows that those who trust the media less are more susceptible to believing online disinformation (**Zimmermann**; **Kohring**, 2020), and conversely, citizens of countries with high levels of trust in the media and low levels of polarization and populist communication are more resilient to false information (**Humprecht**; **Esser**; **Van-Aelst**, 2020). Therefore, structural tensions in the media environment are related to the breakdown of trust in democratic institutions, which paves the way for disinformation to act as a disruptive element in the public sphere (**Bennet**; **Livingstone**, 2018).

This is clearly evidenced by phenomena such as the coronavirus crisis. In February 2020, The *World Health Organization* (*WHO*) already warned of the danger of an “infodemic,” insofar as the overabundance of information about the topic, whether accurate and truthful or false or confusing, makes it difficult for people to find trustworthy sources (*World Health Organization*, 2020). In fact, in the first months of the pandemic, a large amount of fake news circulated, and not just related to health and science; numerous false political and governmental pieces of content were also spread (**Salaverría et al.**, 2020). This may explain why the consumption of traditional media increased during the months of the lockdown (**Casero-Ripollés**, 2020; **Masip et al.**, 2020; **Rodero**, 2020), as did the activity of fact-checking organizations during the first wave of coronavirus disease 2019 (Covid-19) (**Dafonte-Gómez**; **Míguez-González**; **Martínez-Rolán**, 2022).

These challenges that face the media today are not unrelated to the growing distrust of other global social, economic, cultural, and technological actors, including governments. Since 2001, the consulting firm *Edelman* has put out its *Trust Barometer*, one of the most comprehensive and recognized endeavors to measure trust in these institutions, on an annual basis (**Edelman**, 2023, p. 2). From the beginning, the evolution of the data has shown growing gaps in the trust granted to the four most recognized social institutions: governments, companies, nongovernmental organizations (NGOs), and the media.

“After the first stage of the pandemic, trust in the media has decreased, and citizens are looking for alternative sources of information”

The 2008 financial crisis led to the breakdown of trust in the business world, followed by a decline in global trust in governments, while NGOs stood out as the most trusted institutions for 19 years. Other experts and international studies indicate the same thing and –to the long shadow of said financial crisis– add the perception of an unfair distribution of wealth, growing anxiety about an uncertain job future, and political corruption (Lagarde, 2018). The factors affecting the loss of trust in institutions are complex, though reports highlight two that seem to be interrelated in this progressive deterioration of trust: economic uncertainty and government corruption (Perry, 2021). Unsurprisingly, several international studies confirm the loss of trust in institutions and governments in recent years (Pew Research Center, 2022; Eurofound, 2022; Saad, 2023) and warn of the difficulty of enforcing accountability in environments with high political polarization (Martínez-Bravo; Sanz, 2023) while advocating restoring trust to achieve institutional legitimacy, which is key in building constructive and harmonious societies (Sapienza, 2021), particularly in developing countries (Khemani, 2020) and in post-Covid-19 environments.

Edelman's latest global report on trust focuses in particular on the phenomenon of polarization, which it considers to be both a cause and a consequence of global distrust. Specifically, Spain is listed as one of the four most polarized countries in the world, only surpassed by Argentina, Colombia, and the United States (Edelman, 2023, p. 57). According to the *Barometer*, extreme polarization is motivated by ideological division, the inability to generate consensus, and the lack of agreements aiming to address essential issues in these countries.

Edelman's historical index (2018; 2019a; 2019b; 2020a; 2020b; 2021a; 2021b; 2022; 2023) thus confirms the existence of a cycle of distrust centered around the media and governments. In a global environment characterized by polarization and, consequently, social division, citizens rely more on companies than on any other institution to overcome it. Thus, after the pandemic, only companies have improved in their index and seem to be the only institution deserving of trust, based on a positive rating of both their ethical behavior and their competence. NGOs, although ethical, were not considered competent. In contrast, governments and the media had negative ratings in both competence and ethics (Edelman, 2023, p. 26).

In fact, the report points to an alleged “battle” or fight for truth in the media as one of the reasons behind polarization (Edelman, 2023, p. 4). Citizens do not trust the media, especially social networks. The most recent findings indicate that journalists and government leaders are divisive rather than unifying social forces (Edelman, 2023, p. 21). In countries such as Spain, 53% of respondents –at the median of the study– think that the country is more divided than in the past (Edelman, 2023, p. 20). Although companies and NGOs are perceived as global forces that unify society in a group of countries, the report points out that, in Spain, only 49% of the population considers companies to be trustworthy, compared with the overall average of 62% in the rest of the countries in the study.

This year's data also indicate that, globally, 59% of citizens trust traditional media, whereas social networks receive a rating of 41% for trust (Edelman, 2023, p. 52). The data are in line with other European estimates that give a confidence index of 4 out of 10 to the media (Eurofound, 2022, p. 10). If we look at the specific data by country, Spain is the country with the fourth lowest level of trust in the media, with a percentage of 38%, two points lower than in the previous year's edition (Edelman, 2023, p. 44). At this point, it should be noted that Edelman considers a percentage below 49% to be distrust, whereas a score between 50-59% is considered neutral, and any institution that receives more than 60% of favorable ratings from respondents is considered trustworthy.

2.3. Trust in the media in Spain

With regard to Spain, other previous studies also show that generally Spanish citizens have a low level of trust in their media, not only because they are considered to be more a source of social division than of unity, as mentioned above (Eurofound, 2022; Edelman, 2023), but also because of their political and ideological bias (Baumgartner; Chaqués-Bonafont, 2015; Roses, 2012). At the same time, it seems that the media's credibility is not homogeneous but varies depending on a variety of sociodemographic variables. Thus, according to research by Roses and Fariás-Batlle (2012), for most Spaniards, television is the most credible form of media, but the most educated people believe that the type of media with the greatest credibility is the press. Also, based on telephone surveys of a sample of 1,200 individuals across Spain, Roses and Gómez-Calderón (2015) concluded that Spaniards trust the media across the board, although the most skeptical citizens tend to be middle-aged men, with a high level of education, living in large cities in the northern half of the country.

In contrast, other research shows that skepticism is higher among young people, those disinterested in current affairs, and those who prefer to use social networks for news, compared with those who receive news through traditional media, who have greater trust (Serrano-Puche, 2017; Vara-Miguel, 2018; 2020). Young Spaniards' distrust toward the media is also one of the findings of Pérez-Escoda and Pedrero-Esteban (2021). This distrust extends to politicians and social networks, although paradoxically, while admitting their misgivings about them, young people do not forgo massive consumption of them.

For its part, the *Digital News Report 2022* notes that trust in news has declined in almost half of the countries studied (46), reversing the growth experienced at the height of the pandemic. Out of the total country sample, 4 in 10 people (42%) say they trust most news most of the time (Newman et al., 2022).

In Spain, news users' loss of trust is worsening, and for the first time since 2015 (Table 1), the percentage of people who do not trust it (39%) is higher than that of people who generally trust it (32%; **Vara-Miguel et al.**, 2022).

Table 1. Trust in the news (2015-2022)

	2015	2016	2017	2018	2019	2020	2021	2022
Survey base, <i>n</i> =	2,026	2,104	2,006	2,023	2,005	2,006	2,019	2,028
Trust	34.2%	42.9%	50.6%	44.5%	43.5%	36.4%	36.5%	32.5%
Neither agree nor disagree	32.6%	27.4%	25.2%	25.5%	25.3%	28.8%	28.7%	28.9%
Distrust	33.2%	29.7%	24.3%	30.1%	31.3%	34.8%	34.8%	38.6%

Source: *Digital News Report*, 2022

[Q6_2016] Next, we're going to ask you about your trust in the news. First, we ask you to what extent you trust the news as a whole in your country. We then ask you to what extent you trust the news you choose to consume. Indicate the extent to which you agree with the following statements: "I think most news can be trusted most of the time."

In 2022, distrust was higher among younger users: More than half (51%) of those younger than 35 years did not trust the news, compared with 21% who did. The trend extended to older age groups, those who have traditionally been more trusting of the news (Table 2). Thus, among respondents between 45 and 54 years of age, trust was lower, and the majority did not trust the news (34%; **Vara-Miguel et al.**, 2022).

Table 2. Trust in news by age

	Total (<i>n</i> = 2,028)	18–24 years	25–34 years	35–44 years	45–54 years	55–64 years	65+ years
	32.5%	24.3%	19.4%	28.8%	30.3%	41.9%	41.6%
Agree (total/partial)	28.9%	27.9%	28.3%	32.2%	36.0%	23.5%	25.6%
Neither agree nor disagree	38.6%	47.9%	52.3%	39.0%	33.8%	34.6%	32.9%
Disagree (total/partial)	38,6%	47,9%	52,3%	39,0%	33,8%	34,6%	32,9%

Source: *Digital News Report*, 2022

[Q6_2016] Next, we're going to ask you about your trust in the news. First, we ask you to what extent you trust the news as a whole in your country. We then ask you to what extent you trust the news you choose to consume. Indicate the extent to which you agree with the following statements: "I think most news can be trusted most of the time."

In addition, Spain continues to be one of the countries most concerned about disinformation (62%), although less so than in previous years. In 2018, when this question was first asked in the *Digital News Report*, 69% of respondents reported being alarmed by the spread of fake news (**Vara-Miguel et al.**, 2022).

The 2023 report confirms this trend: The percentage of Spaniards who declared that they generally do not trust the news reached 40%, with young people between 25 and 34 years of age being the most distrustful of the news (**Amoedo et al.**, 2023).

Concern about disinformation is one of the factors that affects attitudes toward the media, but other aspects also play a role. For example, users who consume news mainly from digital media tend to distrust news in general more than those who receive news through traditional media (**Rodríguez-Virgili; Sierra; Serrano-Puche**, 2022). Ideological inclination is also important as a trustworthiness variable, as **Moreno-Moreno** and **Sanjurjo-San-Martín** (2020) have shown, in the sense that the user's political orientation curbs or increases the consumption of some journalistic brands by mediating trust. Also based on the *Digital News Report*, these authors conclude that users who are politically right-leaning trust the brand less and, therefore, consumption is reduced. In contrast, users who consider themselves to be on the left of the ideological spectrum more loyally consume the media they trust the most (**Moreno-Moreno; Sanjurjo-San-Martín**, 2020).

Along the same lines, based on a representative survey of the Spanish population (*n* = 1,000), **Masip, Suau, and Ruiz-Caballero** (2020) point out that ideology plays a significant role in trust in the media and note a clear polarization in consumption, giving rise to media ecosystems differentiated according to ideology, which is typical of a polarized pluralistic media system such as the Spanish one (**Hallin; Mancini**, 2004; **Salaverría; Martínez-Costa**, 2023).

3. Objective, research questions, and methodology

Based on the state of affairs described above, the general objective of this article is to delve into the reasons behind and factors that influence the loss of trust in the media among the public that traditionally trusted the news the most according to the different studies reviewed: the group of young people and adults between 25 and 54 years of age. Specifically, the following research questions are posed:

RQ1: What are the reasons for the crisis of trust in the media, and of them, what weight does disinformation have?

RQ2: Is the level of trust in the media different when it comes to traditional media and digital media?

RQ3: What were the reasons for the decline in trust in the media after the first wave of the Covid-19 pandemic?

To address these issues, the qualitative methodology of focus groups was used, which allowed us to answer the research questions, delving into the reasons behind perceptions and behaviors. In this way, it was possible to contextualize the quantitative data from previous studies.

One of the main reasons for distrust towards the media is the perception that they offer news that is biased due to their political or economic positioning

Socio-demographic diversity was sought for the focus groups, taking into account, among other criteria, sex, age, level of education, and ideological position. Thus, the sample included a balanced number of men and women, both young adults (25-40 years) and adults (41-54 years), and included both university and non-university participants.

A sample of 23 participants was recruited and divided into three focus groups. The sample was considered sufficient by the researchers, given that the focus groups' dynamics evaluate the participants' individual interactions and not their representativeness with respect to a population group. We collaborated with *CIES*, a company dedicated to market and opinion analysis. The first two focus groups were face-to-face and were held in Pamplona, Navarra, on October 6, 2021. The sample of subjects who participated in these two groups was heterogeneous in sex and age, but homogeneous in level of education: The first group included only participants with completed university studies, whereas the second group included people without completed university studies. This decision was made based on previous studies (*Sáda-ba; Salaverría, 2023; Corbu; Oprea; Frunzaru, 2021; Wang et al., 2020*) that had indicated that the level of education was a determining factor in attributing greater or lesser trust to the media and identifying disinformation actions. Finally, to broaden the geographic sample, the third group was held on October 7, 2021, via videoconference, with participants from Madrid and Barcelona. The sociodemographic profile of this third group was heterogeneous in all variables and acted as a control group.

To recruit participants, a questionnaire including the composition and filter variables presented in Table 3 was designed.

Table 3. Composition of the focus groups

	Total	Group 1 IDs 1, 3, 5-7, 10, and 23	Group 2 IDs 4, 13, 15-18, 20, and 22	Group 3 IDs 2, 8, 9, 11, 12, 14, 19, and 21
Number of participants	23	7	8	8
Age (years)				
25-40 (Y1)	7	1	2	4
41-54 (Y2)	16	6	6	4
Sex				
Male (M)	11	3	4	4
Female (F)	12	4	4	4
Education				
University (U)	13	1	8	4
No university (U)	10	6	0	4

The 90-minute discussion groups were moderated by a professional from *CIES*, who also prepared a qualitative preliminary report analyzing the responses obtained and grouping them thematically. The objectives of the study were explained, and participants' consent was requested for video and audio recording as a means of transcribing their speech and analyzing the results anonymously, assigning each participant an ID that corresponds with their profile. For the discussion, we used a script prepared by the research team that grouped the questions into four thematic blocks: (1) general news consumption before and after the declaration of the pandemic, (2) attitude toward disinformation, (3) type of sources, media, and experts to whom authority was attributed, and (4) actions and motivations for sharing content and/or reporting false content. A previous study with the same sample exclusively analyzed individuals' overconfidence or excessive confidence in their own abilities to detect the falseness of news content and, therefore, the perception of being somewhat immune to the threat of disinformation (*Martínez-Costa et al., 2022*). In this research, the results obtained in blocks 1–3 were taken into account, and the analysis focused on the reasons behind and factors that influence the loss of trust in the media.

4. Results

The qualitative study provided information relevant in identifying the reasons that lead citizens to distrust the media, with university-educated participants being the most critical. Among these motivations we can highlight the following:

4.1. Information saturation

The overabundance of news, of direct access to sources, and of media was seen as a factor that does not aid in trusting the media:

“We have become distrustful from having so much access to information” (ID 3).

Somehow it was felt that, when there was less media available, belief in their trustworthiness was greater:

“what you saw on a newscast was true” (ID 3).

Adult participants affirm that they distrust the media more after the pandemic, and the interests that move them

Distrust was also heightened because participants mentioned receiving contradictory news updates in the same day on digital media:

“You’re getting conflicting news all day long” (ID 1),

“We’ve gotten so much excessive disinformation that we’ve reached a point where I think we’ve all gotten fed up with all that excessive information and then realized that a lot of it was not real or was disinformation or the next day was totally contradictory, and we’ve said: enough” (ID 3).

In addition, they pointed out that the disinformation phenomenon affects the media and citizens alike:

“I have not felt deceived by the media, because they’ve received the same information, too” (ID 10),

“I think that, more than the media, politicians and international institutions knew much earlier and knew the consequences of the reach that this could have and haven’t warned us” (ID 7).

4.2. The pandemic experience

The Covid-19 pandemic has influenced the perception of the media and the way news is consumed. As noted in the quantitative studies reviewed above, media consumption and reliance on the media increased during the early stage of the pandemic. Consumption was higher among people who had more direct personal experiences and who turned to sources they deemed trustworthy:

“If it’s a scientist you’ve heard of before or who works in a hospital or on benchmark sites, it carries more credibility for you” (ID 7),

“If I’m looking up something, I try to find a more reliable or more medical article rather than just a headline” (ID 20).

Who told the story and how they told it also had a positive influence:

“In general, I think that how they have felt and how these people are when they tell you has a lot of influence (...) it depends on who tells you the same story, it depends on the each one’s experience” (ID 5),

“There are people who have been in the emergency room and have experienced the coronavirus firsthand and tell you in a way... ‘dramatic’ isn’t the word, but more affected and more... in a rawer way” (ID 7).

Additionally, the effect of oversaturation and disinterest in the news has increased:

“I think they have overwhelmed us too much already” (ID 10),

“At the beginning I used to seek out a lot of news and then I stopped seeking out news, that is, it overwhelmed me” (ID 7),

“I read the headlines, and generally I no longer get into the news story” (ID 16).

Meanwhile, once the first stage of the pandemic had passed, reliance on the media, especially television, had diminished, and participants were seeking alternative news sources:

“I trust the press headlines much less; I prefer to look for news that seems more truthful to me” (ID 20),

“before I did search out news, because I had the time, through television, and I’m sick of television because of the way it presents a topic; it’s overwhelmed me” (ID 14),

“Now it’s more difficult for me to believe anything they say” (ID 3),

“It has gotten worse, because if the level of truthfulness was already bad, now it is worse because the misconduct of many communication professionals combines with the foolishness of the masses (...)” (ID 22).

Other participants mentioned that they already had little trust in the media before the pandemic and that this had now been exacerbated:

“If there were few perspectives, I had little trust in it, well now less” (ID 18),

“It is more or less the same, it was not very high before, it is not very high now either, whenever I look at a news item I’ve searched out because I’m interested in fact-checking the accuracy a bit, I usually look at several sources” (ID 13).

In the aftermath of the pandemic, adult participants said they were more distrustful of the media and the interests that drive them:

“With this and with age, I have become more aware of things (...) I’ve seen that there are economic, political, social interests, movements... and I see that there are these interests and some say things one way, others another, and all of them are looking out for their own interests” (ID 10),

“I don’t know if the news about the pandemic has reached us correctly or what, but at home (...) the pandemic, the news, especially on television, has been called into question more, we knew they exaggerated, but now we believe that they lie more. They have used it to lose credibility” (ID 14).

However, others indicated that their level of trust was the same as before the pandemic because they understood that this was an exceptional situation:

“It’s something historic that has happened for the first time. I understand that one day they would say it’s white, it’s black (...) I understood it, I understood that there was so much contradiction, so my mindset has not changed, I still believe little” (ID 1).

One participant noted that after the pandemic he preferred to stay constantly informed and updated through digital media and, therefore, had stopped reading the printed press,

“Because it is true that, ever since, I haven’t trusted it, I prefer to stay constantly informed virtually” (ID12).

4.3. Lack of media neutrality

In all the focus groups, it was stated that the media have identifiable ideological biases and that they provide news that has been filtered through their political or economic lens, with this being one of the main reasons behind distrust:

“I don’t trust a single piece of news, no one and no source gives me absolute trustworthiness, none, because always... to begin with there are no neutral media outlets, I believe they do not exist, all media outlets and newspapers of all kinds have an ideology” (ID 22),

“I believe that news is manipulated based on the media outlet that publishes it, whether it is sympathetic or politically, or it is convenient for economic purposes” (ID 17),

“The media deceive us. The media are made up of people, people have interests, and the interests are dominated by money” (ID 6),

“For me, more than planning to deceive, it’s that they cover their mouths” (ID 10),

“I know that there are media outlets that lean one way and media that lean the other, so you have to look at those on one side, those on the other, and draw your conclusions” (ID 23).

Some participants went so far as to state outright that they assume that the media lies or manipulates:

“I start out with the assumption that the media can deceive, I mean, I don’t feel deceived, I know for a fact that they lie” (ID 6),

“It doesn’t matter which party, it’s the same thing, they always, always manipulate” (ID 22).

Also because they have been able to verify it by going to other sources:

“(I) have discovered many things that have been lies” (ID 10),

“They publish many things that are lies” (ID 7),

“You only have to look for the same news item and see different sources, and you see that the content is totally different” (ID 22).

In this regard, the type of traditional media that comes out worst is television: because

“They exaggerate things” (ID 14),

“They don’t tell the whole truth” (ID2).

This bias is also present in formal elements of the presentation of the news, such as headlines and graphics:

“I have noticed that, in a lot of news, both on television and in newspapers, they manipulate the graphics enormously, especially the tables, the percentages, they use it for their own benefit” (ID 13),

“The way in which they have given the news, which has already made you give in and read it, is a manipulation” (ID 1),

“Because we buy through our eyes too (...), because if the news is well written, with clearer print, you buy it sooner because we are visual and it comes to you that way” (ID 7).

However, for some participants, knowing the ideological position of a media outlet was not a problem:

“I think that almost all of us usually know where each newspaper or each radio station stands; then, if you set aside a little bit the political tinge they are going to give to a news item, in general they seem trustworthy to me, always knowing that, if they lean more toward one side or the other, they are going to focus it one way or another” (ID 20).

On the other hand, another participant stressed that we live in a polarized society:

“We read what we like to hear or what we like to see, it’s all so polarized that you feel comfortable in one echo chamber or the other, and we already know that it’s not all real” (ID 18).

In general, participants tended to trust the newspapers that they were used to reading regularly and that had the same ideological line as them more, and rejected those that did not align with their beliefs

“Where the news may be true or not, but many are biased and give opinion” (ID 2).

Even so, for groups without a university education, traditional media were more reliable despite these biases:

“If it’s from a relatively serious media outlet, it can be mainly trustworthy” (ID 6),

“Those that have been around for more years, who you have known for longer, give us more confidence” (ID 3),

“If it is a prestigious media outlet, it is more trustworthy” (ID 7),

“Regardless of the fact that in general they may have a tendency, I see that they argue... perhaps because they argue for what I want, maybe. But I have the impression that they argue in a much more logical way than many media outlets that systematically attack” (ID 6).

In contrast, they are more distrustful of digital media:

“It’s not the same if the news is being told by one person or one media outlet or another. When you open a digital newspaper, you already know which way it is going to lean, whether it is one, the other, or the media outlet, you already go in with the idea that there are certain opinions that you know are going to filter through the lens of one side or the other and you have to, with your ideology or whatever, bring them back to the center or I’ll keep it or I won’t keep it” (ID 7).

The belief that there are no non-ideologized media outlets also increases the need to fact-check the news consumed:

“When there is something that shocks me, first I’m going to fact-check it, I always fact-check, always, I have a complete lack of trust” (ID 22),

“When you are searching for news, you want them to speak with credibility, so you look for trustworthy, credible sources” (ID 19),

“I at least look, if it is a newspaper, I look at several newspapers, if it is on the radio, then on several radio shows” (ID 4),

“If I see a news item that interests me a lot, I look for it maybe eight times, I look at it eight times in different places to make sure that it is true or not true” (ID 1),

“The truth is that, when you hear a news item, if you are interested, you have to fact-check it using several media outlets, none will tell you the truth because each one will give you their opinion from their side and has their ideology” (ID 23).

There was a greater need to turn to supplementary sources when dealing with more specialized topics. For news about specific topics, the participants opted for specialized media:

“I prefer (to check) scientific articles directly taken from more specialized sites” (ID 3).

There was also an increase in the active search for news of particular interest:

“Now I look for, I’m referring to before, if something came up that interested me, I would read it, and now there are things that I look for, before I didn’t look for anything” (ID 15).

It should also be noted that distrust goes beyond the media ecosystem and affects all institutions. Specifically with regard to news about the Covid-19 pandemic, participants stated that they felt deceived by the government and international organizations:

“Politicians, the *WHO*, everyone, because they knew about this long before and they knew the consequences of the spread that this could have” (ID 7),

“(We are) quite aware that in the end we will never know the truth of everything” (ID 16).

4.4. Interest in engaging the audience

When faced with an overabundance of news and access channels, each media outlet must establish a strategy that allows it to grow its audience, in some cases with lower quality content or with news bait that entices people to click on it using eye-catching headlines. One participant stated that they felt deceived when they search for news and what appears are

“Trivial news items, news items that are not relevant” (ID 18).

There was a generalized belief that

“All the media and press are simply looking for news to draw people in” (ID 6),

“That they tend to be a bit more sensationalist perhaps, precisely so you fall into the trap, so to speak, and click on that news item” (ID 12).

One participant pointed out that the whole structure of media production and presentation of content is geared toward engaging the audience:

“In regard to how digital media is organized, there are newspapers that you start going through and it’s km after km after km to get more advertising in. The last half of all the newspapers you can avoid, the fake news is almost obligatory, it is filler, they have armies of gazetteers with computer interns filling in, filling in more news because that’s the way they work. You can believe a newspaper’s masthead and little else (...) I don’t trust anything you see in the second half of the paper” (ID 18).

5. Discussion and conclusions

The purpose of this article was to examine the motivations and expectations of citizens with regard to the media, focusing on the reasons that influence the loss of trust in the media. We have taken Spain as a case study and, specifically, the age groups that have traditionally trusted the news the most, i.e., people between 25 and 54 years of age.

Throughout the article, it has been shown that lack of trust is a global phenomenon and that it is not exclusive to the media (Edelman, 2018; 2019a; 2020a; 2021a; 2022; 2023; Perry, 2021; Sapienza, 2021; Eurofound, 2022; Pew Research Center, 2022; Martínez-Bravo; Sanz, 2023; Saad, 2023). There is a global context of distrust toward traditional institutions, which are apparently incapable of proposing solutions to social problems. Moreover, polarization appears to be a sign of democratic deterioration (Edelman, 2023) and incubator of serious democratic problems (Khemani, 2020; Rizzi, 2023; Martínez-Bravo; Sanz, 2023). Thus, it is understood that, in response to the first research question (RQ1), disinformation does not appear to be the main factor of distrust in the media, since the participants of the study considered it to affect both the media and the citizens. They were more concerned about information overload, political polarization of the media, and the proliferation of news bait to suck the audience in. For this reason, they claimed to have a more active approach when it comes to searching out news, turning to a variety of sources to balance out a news item, which they believed was always biased according to not only ideological but also economic interests.

These perceptions about the lack of independence and polarization of the media in Spain are in line with the trend reported in the *Digital News Report 2022* (Newman et al., 2022) and with other research (Martínez-Costa et al., 2019; Eurofound, 2022; Martínez-Bravo; Sanz, 2023). The problem arises when the journalistic media's alignment with opposing groups is reflected in society, in terms of not only opinions but also feelings and attitudes. This increases the difference between the attachment to the social and media collective with which one sympathizes and the rejection of the one perceived to be a rival, giving rise to the formation of in-groups and out-groups. Moreover, polarization and disinformation feed back on each other in a vicious circle, which can affect trust in the news. On the one hand, disinformation widens both ideological and affective social gaps (Au; Ho; Chiu, 2021; Serrano-Puche, 2021). On the other hand, it is worth asking whether a growing social polarization may not in turn have an impact on perceptions of disinformation, either when the news is not aligned with one's ideological orientation (Bessi et al., 2016; Rao; Morstatter; Lerman, 2022) or when the news is negative and there are also ideological differences between central and regional governments, such that they end up hindering the accountability of public authorities (Martínez-Bravo; Sanz, 2023). Along these lines, the latest *Edelman Barometer* data also confirm that the most polarized countries have low levels of trust in the media and in governments (34% and 27%, respectively; Edelman, 2023, p. 19). This is a question that remains to be explored in future work.

Regarding the second research question (RQ2), about whether citizens trust traditional media or digital media more, there was no consensus among the participants.

“ Distrust permeates the media ecosystem and affects all institutions ”

There were those who preferred the latter's constant updating and those who only relied on the well-known brands, especially among non-university participants. However, television was seen as the least trustworthy type of traditional media, in clear contrast to studies from a decade ago that found this type of media to be the most trustworthy (Roses; Farias-Batlle, 2012). It also emerged in the focus groups that audiences place more trust in expert sources and also in those who have had personal experience or strong emotional involvement in the story they tell. Therefore, this is a matter that depends on sociodemographic factors as well as on habits to satisfy diverse news needs and that give rise to “media repertoires” that articulate consumption practices embedded in daily routines (Hasebrink; Domeyer, 2012).

Finally, in relation to the last question directing this research (RQ3), it can be concluded that the pandemic changed most participants' perception of the media, increasing distrust, as also pointed out by quantitative studies (Newman et al., 2022; Vara-Miguel et al., 2022; Eurofound, 2022). In Spain, institutional trust indices for traditional media showed a negative evolution—even within margins considered as distrustful, since they were lower than 59%—from 2021, the year in which 42% of the population said they trusted them (Edelman, 2021a). In 2022, only 40% of citizens said they trusted the media (Edelman, 2022; Eurofound, 2022), and in 2023, only 38% of the population according to Edelman (2023) and 33% according to the *Digital News Report* considered them to be a trustworthy institution (Vara-Miguel et al., 2022; Amoedo et al., 2023).

This qualitative study sheds light on citizens' perceptions and expectations of news and media in a global context of a systemic decline in trust. However, this is a phenomenon that must continue to be studied using new methodological approaches and taking into account social changes.

Trust is an integrating factor based on the activity carried out by the media and the way they relate to society. However, citizens noticed that there was a “lack of information hygiene” (Edelman, 2021b) and that journalists and political leaders were “divisive social forces” (Edelman, 2023, p. 21). It seems, then, that there is a difference between what the media are or should be and what is perceived. If the media wish to regain social license and thus trust, it is worth remembering that trust

“is established and maintained through active relationships” (Morrison, 2014, p. 63)

with citizens, based on truthful and trustworthy news. Trust, in fact, is considered a fragile, intangible, and highly relational asset. First, you have to do and show what you really are, and not just say what you are. And although trust is not synonymous with reputation, since

“For participants without university studies, traditional media are more reliable than digital media”

“it does not imply recognition of excellence, nor does it carry with it the goal of recommendation” (Mora, 2020, p. 21), there is a certain relationship. It must not be forgotten that reputation includes a sum of intangibles among which we find “legitimacy, credibility, honesty, competence and quality” (Mora, 2020, p. 24).

Therefore, from the media companies' perspective, it is urgent to take measures to restore their credibility and reputation, such as recovering space and time for research and good compliance with journalistic standards, fostering a climate of innovation that promotes professional growth and avoids job insecurity (Pérez-Latre, 2022), and giving more prominence to audiences, promoting an “architecture of listening” (McNamara, 2016) in the company and a positive relationship to get to know them better.

6. References

- Amoedo, Avelino; Moreno, Elsa; Negredo, Samuel; Kaufmann-Argueta, Jürg; Vara-Miguel, Alfonso** (2023). *Digital News Report España 2023*. Pamplona: Servicio de Publicaciones de la Universidad de Navarra. <https://doi.org/10.15581/019.2023>
- Au, Cheuk-Hang; Ho, Kevin; Chiu, Dickson** (2021). “The role of online misinformation and fake news in ideological polarization: barriers, catalysts, and implications”. *Information systems frontiers*, v. 24, n. 2. <https://doi.org/10.1007/s10796-021-10133-9>
- Baumgartner, Frank R.; Chaqués-Bonafont, Laura** (2015). “All news is bad news: newspaper coverage of political parties in Spain”. *Political communication*, v. 32, n. 2, pp. 268-291. <https://doi.org/10.1080/10584609.2014.919974>
- Bennett, W. Lance; Livingston, Steven** (2018). “The disinformation order: disruptive communication and the decline of democratic institutions”. *European journal of communication*, v. 33, n. 2, pp. 122-139. <https://doi.org/10.1177/0267323118760317>
- Bessi, Alessandro; Petroni, Fabio; Del-Vicario, Michela; Zollo, Fabiana; Anagnostopoulos, Aris; Scala, Antonio; Caldarelli, Guido; Quattrociochi, Walter** (2016). “Homophily and polarization in the age of misinformation”. *The European physical journal special topics*, v. 225, pp. 2047-2059. <https://doi.org/10.1140/epjst/e2015-50319-0>
- Bice, Sara; Brueckner, Martin; Pffor, Christof** (2017). “Putting social license to operate on the map: a social, actuarial and political risk and licensing model (SAP Model)”. *Resources policy*, v. 53, pp. 46-55. <https://doi.org/10.1016/j.resourpol.2017.05.011>
- Carlson, Matt; Lewis, Seth C.** (eds.). (2015). *Boundaries of journalism: professionalism, practices and participation*. London: Routledge. ISBN: 978 1 138020 67 2
- Carlson, Matt; Robinson, Sue; Lewis, Seth C.** (2021). *News after Trump: journalism's crisis of relevance in a changed media culture*. Oxford: Oxford University Press. ISBN: 978 0 197550 34 2
- Casero-Ripollés, Andreu** (2020). “Impact of Covid-19 on the media system. Communicative and democratic consequences of news consumption during the outbreak”. *Profesional de la información*, v. 29, n. 2, e290223. <https://doi.org/10.3145/epi.2020.mar.23>
- Chadwick, Andrew** (2013). *The hybrid media system: politics and power*. Oxford: Oxford University Press. ISBN: 978 0 190696 73 3
- Coleman, Stephen** (2012). “Believing the news: From sinking trust to atrophied efficacy”. *European journal of communication*, v. 27, n. 1, pp. 35-45. <https://doi.org/10.1177/0267323112438806>
- Corbu, Nicoleta; Oprea, Denisa-Adriana; Frunzaru, Valeriu** (2021). “Romanian adolescents, fake news, and the third-person effect: a cross-sectional study.” *Journal of children media*, v. 16, n. 3, pp. 387-405. <https://doi.org/10.1080/17482798.2021.1992460>
- Dafonte-Gómez, Alberto; Míguez-González, María-Isabel; Martínez-Rolán, Xavier** (2022). “Los fact-checkers iberoamericanos frente a la Covid-19. Análisis de actividad en Facebook”. *Observatorio (OBS*)*, v. 16, n. 1, pp. 160-182. <https://doi.org/10.13140/RG.2.2.35345.56168>
- Delli-Carpini, Michael X.; Keeter, Scott** (1996). *What Americans know about politics and why it matters*. New Haven: Yale University Press. ISBN: 0 300 06256 7

- Edelman (2018). *2018 Edelman Trust Barometer. Global report*.
https://www.edelman.com/sites/g/files/aatuss191/files/2018-10/2018_Edelman_Trust_Barometer_Global_Report_FEB.pdf
- Edelman (2019a). *2019 Edelman Trust Barometer. The battle for truth*.
<https://www.edelman.com/trust/2019-trust-barometer>
- Edelman (2019b). *Research. 2019 Edelman Trust Barometer*.
<https://www.edelman.com/trust/2019-trust-barometer>
- Edelman (2020a). *Research. Edelman 2020 Trust Barometer*.
<https://www.edelman.com/trust/2020-trust-barometer>
- Edelman (2020b). *Global Report. Edelman Trust Barometer. Trust: Competence and ethics*.
<https://www.edelman.com/trust/2020-trust-barometer>
- Edelman (2021a). *Global Report. Edelman Trust Barometer. Declaring information bankruptcy*.
<https://www.edelman.com/trust/2021-trust-barometer>
- Edelman (2021b). *The Trust 10. Edelman Trust Barometer*.
<https://www.edelman.com/sites/g/files/aatuss191/files/2021-01/2021%20Trust%20Top%2010.pdf>
- Edelman (2022). *Edelman Trust Barometer. The cycle of distrust*.
<https://www.edelman.com/trust/2022-trust-barometer>
- Edelman (2023). *Edelman Trust Barometer. Navigating a polarized world*.
<https://www.edelman.com/trust/2023/trust-barometer>
- Engelke, Katherine M.; Hase, Valerie; Wintterlin, Florian** (2019). "On measuring trust and distrust in journalism: reflection of the status quo and suggestions for the road ahead". *Journal of trust research*, v. 9, n. 1, pp. 66-86.
<https://doi.org/10.1080/21515581.2019.1588741>
- Eurofound (2022). *Fifth round of the Living, working and COVID-19 e-survey data: living in a new era of uncertainty*. Luxembourg: Publications Office of the European Union.
<https://doi.org/10.2806/190361>
- Fawzi, Nayla; Steindl, Nina; Obermaier, Magdalena; Prochazka, Fabian; Arlt, Dorothee; Blöbaum, Bernd; Dohle, Marco; Engelke, Katherine M.; Hanitzsch, Thomas; Jakob, Nikolaus; Jakobs, Ilka; Klawier, Tilman; Post, Senja; Reinemann, Carsten; Schweiger, Wolfgang; Ziegele, Marc** (2021). "Concepts, causes and consequences of trust in news media - a literature review and framework". *Annals of the international communication association*, v. 45, n. 2, pp. 154-174.
<https://doi.org/10.1080/23808985.2021.1960181>
- Fletcher, Richard; Park, Sora** (2017). "The impact of trust in the news media on online news consumption and participation". *Digital journalism*, v. 5, n. 10, pp. 1281-1299.
<https://doi.org/10.1080/21670811.2017.1279979>
- Garusi, Diego; Splendore, Sergio** (2023). "Advancing a qualitative turn in news media trust research". *Sociology compass*, v. 17, n. 4, e13075.
<https://doi.org/10.1111/soc4.13075>
- Gaziano, Cecillie; McGrath, Kristin** (1986). "Measuring the concept of credibility". *Journalism and mass communication quarterly*, v. 63, n. 3, pp. 451-462.
<https://doi.org/10.1177/107769908606300301>
- Giddens, Anthony** (1991). *Modernity and self-identity: Self and society in the late modern age*. Cambridge: Polity Press. ISBN: 978 0 804719445
- Gil De Zúñiga, Homero; Ardèvol-Abreu, Alberto; Diehl, Trevor; Gómez-Patiño, María; Liu, James H.** (2019). "Trust in institutional actors across 22 countries. Examining political, science, and media trust around the world". *Revista latina de comunicación social*, v. 74, pp. 237-262.
<https://doi.org/10.4185/RLCS-2019-1329>
- Gunther, Albert C.; McLaughlin, Bryan; Gotlieb, Melissa R.; Wise, David** (2017). "Who says what to whom: content versus source in the hostile media effect". *International journal of public opinion research*, v. 49, n. 3, pp. 363-383.
<https://doi.org/10.1093/ijpor/edw009>
- Hallin, Daniel C.; Mancini, Paolo** (2004). *Comparing media systems: three models of media and politics*. Cambridge: Cambridge University Press.
<https://doi.org/10.1017/CBO9780511790867>

- Hameleers, Michael; Brosius, Anna; De-Vreese, Claes H.** (2022). "Whom to trust? Media exposure patterns of citizens with perceptions of misinformation and disinformation related to the news media". *European journal of communication*, v. 37, n. 3, pp. 237-268.
<https://doi.org/10.1177/02673231211072667>
- Han, Byung-Chul** (2016). *La sociedad de la transparencia*. Barcelona: Herder. ISBN: 978 84 254 3252 1
- Hanimann, Anina; Heimann, Andri; Hellmueller, Lea; Trilling, Damian** (2023). "Believing in credibility measures: reviewing credibility measures in media research from 1951 to 2018". *International journal of communication*, v. 17, pp. 214-235.
<https://ijoc.org/index.php/ijoc/article/viewFile/18815/3992>
- Hanitzsch, Thomas; Van-Dalen, Arjen; Steindl, Nina** (2018). "Caught in the nexus: A comparative and longitudinal analysis of public trust in the press". *The international journal of press/politics*, v. 23, n. 1, pp. 3-23.
<https://doi.org/10.1177/1940161217740695>
- Hasebrink, Uwe; Domeyer, Hanna** (2012). "Media repertoires as patterns of behaviour and as meaningful practices. A multimethod approach to media use in converging media environments". *Participations. Journal of audience & reception studies*, v. 9, n. 2, pp. 757-779.
<https://www.semanticscholar.org/paper/Media-repertoires-as-patterns-of-behaviour-and-as-A-Hasebrink-Domeyer/f72eb59cbf8871955d73d7aa693f105361cafb75>
- Hawley, Katherine** (2012). *Trust: a very short introduction*. Oxford University Press. ISBN: 978 0 199697342
- Hovland, Carl I.; Weiss, Walter** (1951). "The influence of source credibility on communication effectiveness". *The public opinion quarterly*, v. 15, n. 4, pp. 635-650.
<http://www.jstor.org/stable/2745952>
- Humprecht, Edda; Esser, Frank; Van-Aelst, Peter** (2020). "Resilience to online disinformation: a framework for cross-national comparative research". *The international journal of press/politics*, v. 25, n. 3, pp. 493-516.
<https://doi.org/10.1177/1940161219900126>
- Jakobsson, Peter; Stiernstedt, Friedrik** (2023). "Trust and the media: arguments for the (irr)levance of a concept". *Journalism studies*.
<https://doi.org/10.1080/1461670X.2023.2169191>
- Khemani, Stuti** (2020). *An opportunity to build legitimacy and trust in public institutions in the time of COVID-19*. World Bank research and policy brief, n. 32, May.
<https://documents1.worldbank.org/curated/en/355311588754029852/pdf/An-Opportunity-to-Build-Legitimacy-and-Trust-in-Public-Institutions-in-the-Time-of-COVID-19.pdf>
- Kiousis, Spiro** (2001). "Public trust or mistrust? Perceptions of media credibility in the information age". *Mass communication & society*, v. 4, n. 4, pp. 381-403.
https://doi.org/10.1207/S15327825MCS0404_4
- Kohring, Matthias; Matthes, Jörg** (2007). "Trust in news media: development and validation of a multidimensional scale". *Communication research*, v. 34, n. 2, pp. 231-252.
<https://doi.org/10.1177/0093650206298071>
- Lagarde, Christine** (2018). "There's a reason for the lack of trust in government and business: corruption". *The Guardian*, 4 May.
<https://www.theguardian.com/commentisfree/2018/may/04/lack-trust-government-business-corruption-christine-lagarde-imf>
- Lee, Tien-Tsung** (2010). "Why they don't trust the media: an examination of factors predicting trust". *American behavioral scientist*, v. 54, n. 1, pp. 8-21.
<https://doi.org/10.1177/0002764210376308>
- Lucassen, Teun; Schraagen, Jean M.** (2012). "Propensity to trust and the influence of source and medium cues in credibility evaluation". *Journal of information science*, v. 38, n. 6, pp. 566-577.
<https://doi.org/10.1177/0165551512459921>
- Luhmann, Niklas** (1968). *Vertrauen: ein Mechanismus der Reduktion Sozialer Komplexität*. Stuttgart: UTB. ISBN: 978 3 825221850
- Martínez-Bravo, Mónica; Sanz, Carlos** (2023). *Trust and accountability in times of pandemics*. Documentos de trabajo n. 2306, Madrid: Banco de España.
<https://www.bde.es/f/webbde/SES/Secciones/Publicaciones/PublicacionesSerias/DocumentosTrabajo/23/Files/dt2306e.pdf>
- Martínez-Costa, María-Pilar; López-Pan, Fernando; Buslón, Nataly; Salaverría, Ramón** (2022). "Nobody-fools-me perception: influence of age and education on overconfidence about spotting disinformation". *Journalism practice*.
<https://doi.org/10.1080/17512786.2022.2135128>

- Martínez-Costa, María-Pilar; Serrano-Puche, Javier; Portilla, Idoia; Sánchez-Blanco, Cristina** (2019). "Young adults' interaction with online news and advertising". *Comunicar*, v. 59, pp. 19-28.
<https://doi.org/10.3916/C59-2019-02>
- Masip, Pere; Aran-Ramspott, Sue; Ruiz-Caballero, Carlos; Suau, Jaume; Almenar, Ester; Puertas-Graell, David** (2020). "Consumo informativo y cobertura mediática durante el confinamiento por el Covid-19: sobreinformación, sesgo ideológico y sensacionalismo". *Profesional de la información*, v. 29, n. 3, e290312.
<https://doi.org/10.3145/epi.2020.may.12>
- Masip, Pere; Suau, Jaume; Ruiz-Caballero, Carlos** (2020). "Percepciones sobre medios de comunicación y desinformación: ideología y polarización en el sistema mediático español". *Profesional de la información*, v. 29, n. 5, e290527.
<https://doi.org/10.3145/epi.2020.sep.27>
- McNamara, Jim** (2016). *Organizational listening. The missing essential in public communication*. New York: Peter Lang Verlag.
<https://doi.org/10.3726/978-1-4539-1739-8>
- Medina, Mercedes; Etayo-Pérez, Cristina; Serrano-Puche, Javier** (2023). "Categorías de confianza para los informativos televisivos e indicadores para su medición: percepciones de grupos de interés en Alemania, España e Italia". *Revista mediterránea de comunicación*, v. 14, n. 1, pp. 307-324.
<https://www.doi.org/10.14198/MEDCOM.23416>
- Meyer, Philip** (1988). "Defining and measuring credibility of newspapers: developing an index". *Journalism & mass communication quarterly*, v. 65, n. 3, pp. 567-574.
<https://doi.org/10.1177/107769908806500301>
- Mora, Juan-Manuel** (2020). *El valor de la reputación*. Pamplona: Eunsa. ISBN: 978 84 313 3461 1
- Moran, Rachel E.; Nechushtai, Efrad** (2022). "Before reception: trust in the news as infrastructure". *Journalism*, v. 24, n. 3, pp. 457-474.
<https://doi.org/10.1177/14648849211048961>
- Moreno-Moreno, Elsa; Sanjurjo-San-Martín, Elena-Luisa** (2020). "Consumo selectivo de las marcas periodísticas en España: confianza, orientación política y grado de consumo digital". *Profesional de la información*, v. 29, n. 5, e290524.
<https://doi.org/10.3145/epi.2020.sep.24>
- Morrison, John** (2014). *The social license: how to keep your organization legitimate*. Basingstoke: Palgrave MacMillan. ISBN: 978 1 137370716
- Narbona, Juan; Pujol, Jordi; Gregory, Anne** (2020). "The fragile tenets of trust". *Church, communication and culture*, v. 5, n. 3, pp. 293-297.
<https://doi.org/10.1080/23753234.2020.1825975>
- Newman, Nic; Fletcher, Richard; Robertson, Craig T.; Eddy, Kirsten; Nielsen, Rasmus-Kleis** (2022). *Digital news report*. Oxford: Reuters Institute for the Study of Journalism.
https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2022-06/Digital_News-Report_2022.pdf
- Otto, Kim; Köhler, Andreas** (eds.) (2018). *Trust in media and journalism: Empirical perspectives on ethics, norms, impacts and populism in Europe*. Wiesbaden: Springer Fachmedien.
<https://doi.org/10.1007/978-3-658-20765-6>
- Pérez-Escoda, Ana; Pedrero-Esteban, Luis-Miguel** (2021). "Retos del periodismo frente a las redes sociales, las fake news y la desconfianza de la generación Z". *Revista latina de comunicación social*, v. 79, pp. 67-85.
<https://doi.org/10.4185/RLCS-2021-1519>
- Pérez-Latre, Francisco-Javier** (2022). *Crisis de confianza (2007-2022). El descrédito de los medios*. Pamplona: Eunsa. ISBN: 978 84 313 3792 6
- Perry, Jonathan** (2021). *Trust in public institutions: trends and implications for economic security*. Policy brief n. 108. New York: United Nations Department of Economic and Social Affairs.
https://www.un.org/development/desa/dspd/wp-content/uploads/sites/22/2021/08/PB_108.pdf
- Pew Research Center** (2022). *Trust in government: 1958-2022*. Washington: Pew Research Center.
<https://www.pewresearch.org/politics/2022/06/06/public-trust-in-government-1958-2022>
- Picard, Robert G.** (2018). "The economics of journalism and news provision". In: Vos, Tim P. (ed.). *Journalism*. Boston: De Gruyter Mouton, pp. 281-296.
<https://doi.org/10.1515/9781501500084-014>

- Prochazka, Fabian; Schweiger, Wolfgang** (2019). "How to measure generalized trust in news media? An adaptation and test of scales". *Communication methods and measures*, v. 13, n. 1, pp. 26-42.
<https://doi.org/10.1080/19312458.2018.1506021>
- Rao, Ashwin; Morstatter, Fred; Lerman, Kristina** (2022). "Partisan asymmetries in exposure to misinformation". *Scientific reports*, v. 12, n. 15671.
<https://doi.org/10.1038/s41598-022-19837-7>
- Rizzi, Andrea** (2023). "No es solo Brasil. La erosión global de la democracia avanza". *El País*, 15 de enero, pp. 2-3.
<https://elpais.com/internacional/2023-01-15/no-es-solo-brasil-la-erosion-global-de-la-democracia-avanza.html>
- Rodero, Emma** (2020). "Radio: the medium that best copes in crises. Listening habits, consumption, and perception of radio listeners during the lockdown by the Covid-19". *Profesional de la información*, v. 29, n. 3, e290306.
<https://doi.org/10.3145/epi.2020.may.06>
- Rodríguez-Virgili, Jordi; Portilla-Manjón, Idoia; Sierra-Iso, Aurken** (2022). "Cuarenta años de polarización ideológica en España". *Revista empresa y humanismo*, v. 25, n. 2, pp. 75-103.
<https://doi.org/10.15581/015.XXV.2.75-103>
- Rodríguez-Virgili, Jordi; Sierra, Aurken; Serrano-Puche, Javier** (2022). "Percepciones sobre las funciones del periodismo en España". En: *Libro de resúmenes VIII congreso Internacional de la AE-IC 'Comunicación y ciudad conectada'*, p. 173. Asociación Española de Investigación de la Comunicación.
- Rosanvallon, Pierre** (2008). *Counter-democracy: politics in an age of distrust*. Cambridge: Cambridge University Press.
<https://doi.org/10.1017/CBO9780511755835>
- Roses, Sergio** (2012). "Análisis de indicadores de la confianza de los ciudadanos españoles en los medios en 2010". En: *Comunicación y Riesgo: III Congreso Internacional de la Asociación Española de Investigación de la Comunicación (AE-IC)*. Tarragona: Universitat Rovira i Virgili.
- Roses, Sergio; Farias-Batlle, Pedro** (2012). "Credibilidad de los medios: un análisis bivariado de las opiniones de los españoles". *Revista mediterránea de comunicación*, v. 3, pp. 79-104.
<https://doi.org/10.14198/MEDCOM2012.3.06>
- Roses, Sergio; Gómez-Calderón, Bernardo** (2015). "Credibilidad de los medios en España: divergencias de percepción y caracterización de los escépticos". *Profesional de la información*, v. 24, n. 4, pp. 432-439.
<https://doi.org/10.3145/epi.2015.jul.10>
- Saad, Lydia** (2023). *Historically low faith in US institutions continues*. Washington: Gallup.
<https://news.gallup.com/poll/508169/historically-low-faith-institutions-continues.aspx>
- Sádaba, Charo; Salaverría, Ramón** (2023). "Combatir la desinformación con alfabetización mediática: análisis de las tendencias en la Unión Europea". *Revista latina de comunicación social*, v. 81.
<https://doi.org/10.4185/RLCS-2023-1552>
- Salaverría, Ramón; Buslón, Natalia; López-Pan, Fernando; León, Bienvenido; López-Goñi, Ignacio; Erviti, María-Carmen** (2020). "Desinformación en tiempos de pandemia: tipología de los bulos sobre la Covid-19". *Profesional de la información*, v. 29, n. 3.
<https://doi.org/10.3145/epi.2020.may.15>
- Salaverría, Ramón; Martínez-Costa, María-Pilar** (2023). "Digital journalism in Spain: Technological, sociopolitical and economic factors as drivers of media evolution". *Journalism*. Special issue: Digital news histories in Europe.
<https://doi.org/10.1177/14648849231170519>
- Sapienza, Emanuele** (2021). *Trust in public institutions. A conceptual framework and insights for improved governance programming*. Oslo: United Nations Development Programme.
https://www.undp.org/sites/g/files/zskgke326/files/migration/oslo_governance_centre/Trust-in-Public-Institutions-Policy-Brief_FINAL.pdf
- Serrano-Puche, Javier** (2017). "Credibilidad y confianza en los medios de comunicación: un panorama de la situación en España". En: González, Marta; Valderrama, Mónica (eds.). *Discursos comunicativos persuasivos hoy*, pp. 427-438. Madrid: Tecnos. ISBN: 978 9978 375 65 5
- Serrano-Puche, Javier** (2021). "Digital disinformation and emotions: exploring the social risks of affective polarization". *International review of sociology*, v. 31, n. 2, pp. 231-245.
<https://doi.org/10.1080/03906701.2021.1947953>
- Shaw, Eugene F.** (1973). "Media credibility: Taking the measure of a measure". *Journalism quarterly*, v. 50, pp. 306-311.
<https://doi.org/10.1177/107769907305000213>

- Simmel, George** (1987). *Sociología, 1 Estudios sobre las formas de socialización*. Madrid: Alianza Editorial. ISBN: 978 84 206 2479 2
- Strömbäck, Jesper; Tsfati, Yari; Boomgaarden, Hajo; Damstra, Alyt; Lindgren, Elina; Vliegenthart, Rens; Lindholm, Torun** (2020). "News media trust and its impact on media use: toward a framework for future research". *Annals of the International Communication Association*, v. 44, n. 2, pp. 139-156.
<https://doi.org/10.1080/23808985.2020.1755338>
- Thorbjørnsrud, Kjersti; Figenschou, Tine U.** (2022). "The alarmed citizen: fear, mistrust, and alternative media". *Journalism practice*, v. 16, n. 5, pp. 1018-1035.
<https://doi.org/10.1080/17512786.2020.1825113>
- Tsfati, Yariv** (2003). "Media skepticism and climate of opinion perception". *International journal of public opinion research*, v. 15, n. 1, pp. 65-82.
<https://doi.org/10.1093/ijpor/15.1.65>
- Tsfati, Yariv; Ariely, Gal** (2014). "Individual and contextual correlates of trust in media across 44 countries". *Communication research*, v. 41, n. 6, pp. 760-782.
<https://doi.org/10.1177/0093650213485972>
- Uslaner, Eric M.** (2002). *The moral foundations of trust*. Cambridge: Cambridge University Press. ISBN: 978 0 511614934
<https://doi.org/10.1017/CBO9780511614934>
- Vaccari, Cristian; Chadwick, Andrew** (2020). "Deepfakes and disinformation: exploring the impact of synthetic political video on deception, uncertainty, and trust in news". *Social media + society*, v. 6, n. 1.
<https://doi.org/10.1177/2056305120903408>
- Vanacker, Bastian; Belmas, Genele** (2009). "Trust and the economics of news". *Journal of mass media ethics*, v. 24, n. 2-3, pp. 110-126.
<https://doi.org/10.1080/08900520902885277>
- Van-Aelst, Peter; Strömbäck, Jesper; Aalberg, Toril; Esser, Frank** (2017). "Political communication in a high-choice media environment: a challenge for democracy?". *Annals of the international communication association*, v. 41, n. 1, pp. 3-27.
<https://doi.org/10.1080/23808985.2017.1288551>
- Van-Dalen, Arjen** (2021). "Rethinking journalist-politician relations in the age of populism: how outsider politicians delegitimize mainstream journalists". *Journalism*, v. 22, n. 11, pp. 2711-2728.
<https://doi.org/10.1177/1464884919887822>
- Vara-Miguel, Alfonso** (2018). "Confianza en noticias y fragmentación de mercado: el caso español". *Comunicació: revista de recerca i d'anàlisi*, v. 35, n. 1, pp. 95-113.
<https://doi.org/10.2436/20.3008.01.168>
- Vara-Miguel, Alfonso** (2020). "Cross-national similarities and differences between legacy and digital-born news media audiences". *Media and communication*, v. 8, n. 2, pp. 16-27.
<https://doi.org/10.17645/mac.v8i2.2733>
- Vara-Miguel, Alfonso; Amoedo, Avelino; Moreno, Elsa; Negrodo, Samuel; Kaufmann-Argueta, Jürg** (2022). *Digital News Report España*. Pamplona: Servicio de Publicaciones de la Universidad de Navarra.
<https://doi.org/10.15581/019.2022>
- Wang, Peng-Wei; Lu, Wei-Hsin; Ko, Nai-Ying; Chen, Yi-Lung; Li, Dian-Jeng; Chang, Yu-Ping; Yen, Cheng-Fang** (2020). "COVID-19-related information sources and the relationship with confidence in people coping with COVID-19: Facebook survey study in Taiwan". *Journal of medical internet research*, v. 22, n. 6, e20021.
<https://doi.org/10.2196/20021>
- Westley, Bruce H.; Severin, Werner J.** (1964). "Some correlates of media credibility". *Journalism quarterly*, v. 41, n. 1, pp. 325-335.
<https://doi.org/10.1177/107769906404100301>
- WHO, World Health Organization** (2020). *Rolling updates on coronavirus disease (Covid-19)*. 2 February.
<https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200202-sitrep-13-ncov-v3.pdf>
- Wu, Tim** (2016). *The attention merchants. The epic scramble to get inside our heads*. New York: Knopf Doubleday Publishing Group. ISBN: 978 0 385352 01 7
- Zimmermann, Fabian; Kohring, Matthias** (2020). "Mistrust, disinforming news, and vote choice: a panel survey on the origins and consequences of believing disinformation in the 2017 German parliamentary election". *Political communication*, v. 37, n. 2, pp. 215-237.
<https://doi.org/10.1080/10584609.2019.1686095>

Research on selective media exposure in Spain: a critical review of its findings, application phases, and blind spots

Lidia Valera-Ordaz

Nota: Este artículo se puede leer en español en:

<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87236>

Recommended citation:

Valera-Ordaz, Lidia (2023). "Research on selective media exposure in Spain: a critical review of its findings, application phases, and blind spots". *Profesional de la información*, v. 32, n. 5, e320507.

<https://doi.org/10.3145/epi.2023.sep.07>

Manuscript received on 22th March 2023

Accepted on 24th July 2023



Lidia Valera-Ordaz

<https://orcid.org/0000-0002-1085-980X>

Universitat de València
Facultat de Filologia, Traducció i
Comunicació
Av. Blasco Ibáñez, 32
46010 Valencia, Spain
lidia.valera@uv.es

Abstract

This paper is the first review of selective exposure studies in Spain, offering comprehensive insights into the diverse findings and analytical strategies employed over almost three decades by studies that explore the politically oriented media consumption of Spanish audiences. The article divides the evolution of this research line in Spain into two main phases –an initial phase (1995-2016) and a consolidation phase (2017-present)– according to the methodological approaches used during each period. Despite the challenges posed by working with secondary data and the lack of experimental designs, we note a robust initiation of this research line in Spain, accompanied by increasing methodological sophistication and diversification and a substantial accumulation of evidence on how Spanish audiences selectively consume like-minded news media. The paper also identifies the blind spots of selective exposure research in Spain, such as the scarcity of primary data sources, which overlooks phenomena such as online selective exposure and selective avoidance, the need to supplement multivariate analyses with other techniques that allow the direction of causal influence between variables to be established, and the significance of assessing the role of public media as facilitators or inhibitors of cross-cutting exposure.

Keywords

Literature review; Selective exposure; Cross-cutting exposure; Selective avoidance; Media consumption; Political information; Television; Radio; Press; Internet; Social media.

Funding

This work has been funded by:

- 1) *Cibest* 2022-23 grant from the *Generalitat Valenciana* for research stays; and
- 2) National R&D project 'Disinformation flows, polarisation and crisis of media intermediation' (PID2020-113574RB-I00), of the Spanish *Ministerio de Ciencia e Innovación, Agencia Estatal de Investigación*, led by Eva Campos and Guillermo López-García.

Acknowledgements

The author thanks José-Luis Dader and Guillermo López-García for their valuable contributions to an early version of this paper, and María-Luisa Humanes for the enriching dialogue over the years on selective exposure research.



1. Introduction: origin and relevance of selective exposure

Selective exposure is a classic theory in communication studies, positing that individuals exhibit a preference to consume media and content that aligns with their pre-existing viewpoints, while shunning content that contradicts their established beliefs. Its origin goes back to the theories of Mass Communication Research on limited media effects (**Lazarsfeld; Berelson; Gaudet, 1948; Klapper, 1960**).

“Actual exposure to partisan propaganda does not parallel availability. Availability plus predispositions determine exposure –and predispositions lead people to select communications which are congenial; which support their prior position” (**Lazarsfeld; Berelson; Gaudet, 1948, p. 89**).

Within the framework of the two-step flow of communication, where the impact of media is influenced by social groups, interpersonal communication, and opinion leaders, as noted by **Klapper (1960)**, selective exposure emerged as a viable rationale for the limited effects observed, subsequently evolving into a foundational element of the reinforcement thesis. Even in laboratory situations (**Cooper; Jahoda, 1947; Postman; Bruner; McGinnies, 1948; Mills; Aronson; Robinson, 1959**), the social psychology research conducted during the 1940s and 1950s revealed that individuals exhibited a strong inclination to avoid, misinterpret, or swiftly forget messages that contradicted their pre-existing beliefs, indicating selective tendencies in exposure, perception, and memory processes.

The Selective Exposure Theory refers to the presence of any systematic bias in the composition of the audience (**Sears; Freedman, 1967, p. 165**), which explains the discrepancy between availability and effective consumption. Despite the frequent association of selective exposure with confirmation bias and its predominant examination in the context of politically-oriented media consumption, it is essential to understand that the selection of media content can be influenced not only by political biases, but also by factors such as sociodemographic and racial attributes (**Knobloch-Westerwick, 2015a, 2015b**). Nevertheless, it remains evident that selective exposure guided by political beliefs is unequivocally the most extensively documented phenomenon (**Hart et al., 2009; Stroud, 2011**).

The phenomenon of selective exposure has been researched through two primary methodological approaches (**Feldman et al., 2013**).

- The first and most common approach involves proving a correlation phenomenon, i.e., determining the ideological point of view of the message and then comparing it with that of the audience, usually through the combination of content analysis and survey or panel (**Stroud, 2008**).
- The second, less common, approach is focused on demonstrating an influence process through experimental designs that prompt individuals to choose between different types of media messages –ideologically neutral, aligned with their political preferences, or openly contrary (**Iyengar; Hahn, 2009; Knobloch-Westerwick et al., 2003**).

In general, studies based on correlation analysis have yielded more empirical evidence than those based on experiments, whose results are not so unanimous. In their well-known literature review, **Sears and Freedman (1967)** argued that correlation alone is insufficient to assert that political leanings genuinely motivate selective exposure. The correspondence only proves the existence of *de facto selectivity*. Nonetheless, subsequent meta-analyses have demonstrated that experimental studies conducted between 1956 and 1996 also support the existence of selective exposure (**D’Alessio; Allen, 2002**).

Decades after the first conceptualisations by **Lazarsfeld, Berelson and Gaudet (1948)**, the profound transformations in media environments, such as the emergence of social networks and digital media, the multiplication of media supply, and fragmentation and hybridisation of consumption, have sparked renewed interest in this line of research in the context of new informational environments (**Stroud, 2018; Camaj, 2019; Humanes, 2019**), and as a result, it is now one of the eight most extensively utilised perspectives in international communication studies (**Bryant; Miron, 2004**). Its contemporary relevance is further evidenced in the publication of monographs in high-impact journals over the past decade, such as:

- *Communication methods and measures* (2013)
<https://www.tandfonline.com/toc/hcms20/7/3-40>
- *Media and communication* (2019)
<https://www.cogitatiopress.com/mediaandcommunication/issue/view/130>

In Spain, there has also been a robust initiation of research in this field, with numerous studies exploring the impact of political leanings on consumer preferences (**Martín-Llaguno; Berganza, 2001; Fraile; Meilán, 2012; Humanes, 2014; Humanes; Mellado, 2017; Ramírez-Dueñas; Vinuesa-Tejero, 2020**). Indeed, the expansion of media supply presents a widened opportunity structure for selective exposure to information (**Skovsgaard; Shehata; Strömback, 2016**), consequently increasing the likelihood of selective exposure (**Fischer; Schulz-Hardt; Frey, 2008; Iyengar; Hahn, 2009**).

“ In Spain, there has also been a robust initiation of research in this field, with numerous studies exploring the impact of political leanings on consumer preferences ”

What are the underlying reasons that lead individuals to prefer media and content that align with their existing beliefs? Selective exposure has found support from various theories in cognitive psychology, which have served as the foundation for understanding this phenomenon (Stroud, 2011). According to the theory of cognitive dissonance, individuals have an inherent tendency to seek psychic-cognitive coherence, striving to maintain internal consistency in their beliefs. As a result, they systematically prefer messages that are congruent with their existing beliefs while avoiding content that induces dissonance (Festinger, 1957).

Subsequent theories, such as motivated reasoning (Kunda, 1990) and epistemic motivations (Kruglanski, 1989), centre on the underlying motivations that drive information-seeking behaviours, which, depending on the circumstances, may be influenced by the desire for accuracy or directionality, meaning the inclination to confirm pre-existing beliefs (Hart *et al.*, 2009). Finally, the concept of “cognitive misers” (Kahneman; Slovic; Tversky, 1982) describes how individuals employ heuristics to optimise decision-making processes and how party identification and ideology play a particularly significant role as drivers of media consumption due to their personal relevance for individuals (Donsbach, 1991).

From a normative perspective, the significance of selective exposure can be understood in light of democracies’ historical aim to provide individuals with exposure to a certain diversity (Sampedro, 2000). This exposure is intended to foster social cohesion and political tolerance among citizens (Mutz, 2002). However, the evolving media landscape, characterised by digital technologies and the vast array of user choices in current media environments, has shifted the traditional function of media in producing public space (Sunstein, 2003; Precht, 2010).

In this sense, there are well-founded arguments regarding the effects of politically-oriented media consumption on democracy. On one hand, we know that consuming political information increases citizens’ political knowledge, increases their sense of political efficacy, their interest in public affairs, and fosters their political participation (Cantijoch; Jorba; San-Martín, 2008; Meilán, 2010). But, on the other hand, selective exposure to information has an impact on the perception of the public issues deemed relevant (Stroud, 2011). It also increases the intensity of political attitudes and their cognitive accessibility (Knobloch-Westerwick, 2012; Matthes; Schmuck, 2017), and contributes to processes of political polarisation (Stroud, 2010) and affective polarisation (Levendusky, 2013).

This phenomenon occurs because political opinions are underpinned by emotions (Stroud, 2011). Furthermore, political ideologies not only offer individuals a way of perceiving the world, but also grant them entry into an affective community (Arias-Maldonado, 2016, p. 100). Indeed, the growing prominence of affects in shaping subjectivity and political mobilisation (Arias-Maldonado, 2016), along with the emergence of identity politics, cannot be fully comprehended without considering the role of selective exposure in promoting and reinforcing identities (Knobloch-Westerwick; Meng, 2011). Because of the above, researching how various affective communities are exposed to information sources in hybrid media environments (Chadwick, 2013) becomes essential. This involves exploring how

“identification with politically relevant groups, including political parties and national, ethnic, linguistic, or gender groups” (Huddy, 2013)

influences the consumption of information, making it an imperative undertaking in communication studies.

2. Selection of the corpus

Carrying out a literature review requires defining the criteria used to gather the bibliographic corpus. In this particular review, the focus is on studies that:

- investigate politically-oriented media consumption patterns of audiences, and
- prioritise Spanish audiences as the primary focus of analysis and reference population.

This entails research on both online and offline consumption, as well as studies focused on one or multiple types of media. In essence, the review considers research whose primary objective is to explore the relationship between ideological/political predispositions and media consumption patterns among Spanish audiences.

This decision implies encompassing studies conducted by both foreign academics and researchers from the fields of political and social sciences. In disciplinary terms, the review incorporates studies not only from the field of communication but also from the political and social sciences (Fraile; Meilán, 2012; Ramírez-Dueñas; Vínuesa-Tejero, 2020; Cardenal *et al.*, 2019a, etc.). The authors’ academic background matters less; what’s crucial is how their work contributes to our understanding of how Spanish audiences make selective choices, considering almost three decades of empirical research. The interdisciplinary nature of the research subject, which involves exploring the relationship between political beliefs and media consumption, along with its multidisciplinary empirical development both within and outside of Spain, necessitates this approach.

However, this approach excludes international comparative studies, even those that involve Spain as part of their analysis. These studies are invaluable for testing hypotheses in diverse contexts and advancing theoretical knowledge by identifying universal patterns and explaining differences between countries. However, their primary focus is not on specifically describing or explaining the behaviour of Spanish audiences. News consumption is, nevertheless, highly contextual (Feldman *et al.*, 2013), meaning that the differences between media environments at the macro level significantly in-

fluence individual decisions about media exposure (**Clay; Barber; Shook**, 2013; **Skovsgaard; Shehata; Strömbäck**, 2016; **Steppat; Castro-Herrero; Esser**, 2022).

Because of the above, this review includes those research articles and book chapters that meet the specified criteria. Although these publications may not be subject to the same review processes, we find it necessary to include them because the review spans a significant period, from the 1990s to the present, when journal publication was not as prevalent in Spanish academia (**Fernández-Quijada; Masip**, 2013), and when many of the initial works were published in the form of book chapters. Their quality should not be underestimated as a result. To avoid duplication, we excluded conference communications that subsequently resulted in publications in the form of articles and book chapters. With one exception: the work of **Humanes** and **Mellado** (2017), which was presented at an *AIMCR* conference, but was never published as an article or book chapter.

Sampling employed diverse strategies. First, a search was carried out in the bibliographic portal *Dialnet* using the Spanish terms “exposure”, “selective” and “media”, with a “social sciences” filter, following the strategy of other literature reviews (**Ardèvol-Abreu**, 2015). The search returned 39 documents, of which only 7 were relevant (**Martín-Llaguno; Berganza-Conde**, 2001; **Humanes**, 2014; 2016; **Valera-Ordaz**, 2018; **Ramírez-Dueñas; Vinuesa-Tejero**, 2020; **Moreno-Moreno; Sanjurjo-San-Martín**, 2020; **Sintes-Olivella; Fondevila-Gascón; Berbel-Giménez**, 2023). Secondly, a comprehensive bibliography of all these articles was thoroughly examined to identify additional works that met the specified criteria, following a snowball approach. Thirdly, recent publications were searched for on *Google Scholar*, focusing on authors who had previously conducted selective exposure studies. Finally, the corpus of papers analysed totalled 26 documents: seven book chapters, one communication at an international conference and 19 research articles.

Table 1. Bibliographic corpus

Authors and year	Type
Díez-Nicolás and Semetko (1995)	Book chapter
Díez-Nicolás and Semetko (1999)	Book chapter
Gunther, Montero and Wert (2000)	Book chapter
Martín-Llaguno and Berganza-Conde (2001)	Article
Alcalde-Villacampa (2003)	Article
Fraile and Meilán (2012)	Book chapter
Humanes (2014)	Article
Lacasa-Mas, Jandura and Cano-Castells (2014)	Article
Humanes (2016)	Book chapter
Humanes and Mellado (2017)	Communication
Valera-Ordaz (2018)	Article
Masip, Suau and Ruiz-Caballero (2018)	Article
Cardenal et al. (2019a)	Article
Cardenal et al. (2019b)	Article
Moreno-Moreno and Sanjurjo-San-Martín (2020)	Article
Masip, Suau and Ruiz-Caballero (2020a)	Article
Masip, Suau and Ruiz-Caballero (2020b)	Article
Ramírez-Dueñas and Vinuesa-Tejero (2020)	Article
Ramírez-Dueñas and Vinuesa-Tejero (2021)	Article
Valera-Ordaz, Requena-Mora and Pop (2021)	Book chapter
Valera-Ordaz and Humanes (2022)	Book chapter
Rodríguez-Virgili, Sierra and Serrano-Puche (2022)	Article
Valera-Ordaz (2022)	Article
Sintes-Olivella, Fondevila-Gascón and Berbel-Giménez (2023)	Article
Humanes y Valera-Ordaz (2023)	Article
Victoria-Mas et al. (2023)	Article

3. The application of the selective exposure theory in Spain: emergence and consolidation

Compared to the extensive trajectory of the selective exposure theory in North America, starting from the initial theorisations by **Lazarsfeld, Berelson** and **Gaudet** (1948), **Cooper** and **Jahoda** (1947) or **Postman, Bruner** and **McGinnes** (1948), in Spain, its trajectory is significantly more recent. It is important to specify that the theory of selective exposure initially made its way to Spain in the 1980s and 1990s through informative works that discussed the theories of political

Researching how various affective communities are exposed to information sources in hybrid media environments becomes essential

and sociocultural effects of the media (**Dader**, 1983; **Roda**, 1990; **Muñoz-López**, 1993; **Canel**, 1998). However, it wasn't until the mid-1990s that the first empirical studies on the politically-oriented consumption of Spanish audiences were published.

Next, we will explore how and when the first empirical research on selective exposure emerged in Spain and how scientific production can be categorised into different phases around this line of research, paying attention to the dominant methodological strategies employed at each stage.

The empirical research on selective exposure in Spain can be categorised into two periods: a) an initial emerging phase (1995-2016), characterised by bivariate studies exploring how the political leanings on the left-right axis of audiences influence their consumption of various media types (primarily press and television, with some consideration for radio); and b) a second consolidation phase (2017-2023), characterised by a higher methodological sophistication, where bivariate studies gave way to multivariate analyses, the use of primary data sources advanced, and the predictors of media consumption studied became more diversified: positions on the left-right axis are complemented with variables associated with the centre-periphery cleavage, interest in politics, political polarisation processes, various sociodemographic attributes, and even the nature of the news (Table 2). The main characteristics and studies of each phase are described in greater detail below.

We note here that this categorisation is based on the methodological strategies dominant during each of the periods. This does not mean there are no other ways to categorise them, nor does it exclude the possibility of diagnosing development phases through methodological approaches that consider criteria such as the social and institutional history of the Spanish scientific community.

3.1. Emergency of research into selective exposure in Spain (1995-2016)

Pioneering studies on selective exposure in Spain were published in the late 1990s and early 2000s, and were the first bivariate explorations of the partisanship of Spanish press and television audiences during elections in the 1990s (**Díez-Nicolás**; **Semetko**, 1995; 1999; **Gunther**; **Montero**; **Wert**, 2000; **Martín-Llaguno**; **Berganza-Conde**, 2001). Most of them used contingency tables and the chi-square statistic to analyse whether there is a statistical association between the ideology and/or vote of individuals, and their media consumption preferences (**Fraile**; **Meilán**, 2012; **Humanes**, 2014). In addition to these studies, there is some isolated work on selective exposure by political elites, which also confirms the occurrence of politically-oriented media consumption among members of the Parliament of Catalonia (**Lacasa-Mas**; **Jandura**; **Cano-Castells**, 2014).

Gunther, **Montero** and **Wert** (2000) conducted a pioneering study analysing the political leanings of readers of major Spanish newspapers during the 1993 election. The research revealed significant differences in newspaper consumption based on electoral behaviour. They found that readers of *El País* in 1993 had a leftist ideological orientation and voted more often for *PSOE* and *Izquierda Unida (IU)*, those of *El Mundo* were located on the centre-left and usually voted for *Partido Popular (PP)* and *IU*, and readers of *ABC* were centre-right and more often voted for *PP*. However, it is important to note that their findings were derived from a very small sample of readers (N=175) from the 1993 *CNEP* study, mainly due to the relatively low press consumption in Spain. As a result, several more years of research with larger samples from the Spanish population would be necessary to conclusively establish the partisanship of Spanish audiences.

The initial analyses on television's come from **Díez-Nicolás** and **Semetko** (1995; 1999). They investigated the correlation between voting patterns and viewers' preferred television channel during the 1993 and 1996 elections. These studies were conducted when television had emerged as the primary source of information for the Spanish population, and private channels like *Antena 3* and *Telecinco* had gained significant market influence. Their approach does not solely fit within the scope of selective exposure, as they examine the relationship between voting behaviour and media consumption in both directions. They explore whether television consumption influences voting choices or if the vote influences the selection of a particular television channel. Nevertheless, their findings indicate

“a highly significant correlation between watching *TVE1* and voting for *PSOE*, and watching *Antena 3* and voting for *PP*” (**Díez-Nicolás**; **Semetko**, 1999, p. 167).

Their study's merit lies in conducting content analysis to examine the informational treatment provided by both channels towards the two parties, revealing a more favourable treatment of *PSOE* by *TVE1* and a similar pattern by *Antena 3* for *PP*.

The analysis carried out by **Alcalde-Villacampa** (2003) concurs with the results found by **Díez-Nicolás** and **Semetko**, and reflects how the public perceives a favourable bias of *TVE1* towards *PSOE* and a similar bias of *Antena 3* in favour of *PP*. Their multivariate analysis further confirmed the correlation between media consumption and voting in the reverse direction: watching *TVE1* increases the likelihood of voting for *PSOE*, and consuming *Antena 3* increases the likelihood of voting for *PP* (**Alcalde-Villacampa**, 2003, pp. 163-164).

“The theory of selective exposure initially made its way to Spain in the 1980s and 1990s through informative works that discussed the theories of political and sociocultural effects of the media”

Table 2. Application stages of the theory of selective exposure in Spain

	Emergence (1995-2016)	Consolidation (2017-2023)
Main studies	Díez-Nicolás and Semetko (1995; 1999) Gunther, Montero and Wert (2000) Martín-Llaguno and Berganza (2001) Alcalde-Villacampa (2003) Fraile and Meilán (2012) Humanes (2014; 2016) Lacasa-Mas, Jandura and Cano-Castells (2014)	Humanes and Mellado (2017) Masip, Suau and Ruiz-Caballero (2018; 2020a; 2020b) Valera-Ordaz (2018; 2022) Cardenal et al. (2019a; 2019b) Moreno-Moreno and Sanjurjo-San-Martín (2020) Ramírez-Dueñas and Vinuesa-Tejero (2020; 2021) Rodríguez-Virgili, Sierra and Serrano-Puche (2022) Valera-Ordaz and Humanes (2022) Sintes-Olivella, Fondevila-Gascón and Berbel-Giménez (2023) Humanes and Valera-Ordaz (2023) Victoria-Mas et al. (2023)
Research techniques	Quantitative: Bivariate analysis (contingency table, chi-square, corrected standardised residuals)	Quantitative: Bivariate analysis (contingency table, chi-square) Multivariate analysis (binomial and multinomial logistic regression, factor analysis) Qualitative: Discussion groups
Data sources	Primary (telephone and face-to-face surveys designed <i>ad hoc</i>) and secondary (<i>ASEP</i> , <i>CNEP</i> and <i>CIS</i>) Strong predominance of secondary sources	Primary (online surveys designed <i>ad hoc</i> , discussion groups and web crawling data) and secondary (<i>CIS</i> , <i>Digital News Report</i> , <i>CEO</i>) Further progress of the primary sources
Independent variables	Political identity on the left-right axis: ideology, vote recall, sympathy and closeness to parties	Political identity on the left-right axis: ideology, vote recall and closeness to parties (Ramírez-Dueñas; Vinuesa-Tejero , 2020; Cardenal et al. , 2019a; 2019b; Humanes; Valera-Ordaz , 2023) Political identity on the centre-periphery axis: national identity, nationalism and support for independence (Valera-Ordaz , 2018; 2022 Sintes-Olivella; Fondevila-Gascón; Berbel-Giménez , 2023) Interest in politics (Humanes; Mellado , 2017; Rodríguez-Virgili; Sierra; Serrano-Puche , 2022; Valera-Ordaz; Humanes , 2022) Perceived confidence and quality of the media (Moreno-Moreno; Sanjurjo-San-Martín , 2020; Masip; Suau; Ruiz-Caballero , 2020b; Victoria-Mas et al. , 2023) Political context, polarisation, longitudinal studies (Ramírez-Dueñas; Vinuesa-Tejero , 2020; Valera-Ordaz , 2022; Humanes; Valera-Ordaz , 2023, Sintes-Olivella; Fondevila-Gascón; Berbel-Giménez , 2023) Nature of the news (Cardenal et al. , 2019a)
Dependent variables	Traditional media consumption: press, radio and television	Traditional media consumption: press, radio and television Digital media consumption: digital press, social networks, internet browsing and news search engines

Martín-Llaguno and Berganza (2001) investigated the relationship between vote recall and the consumption of print media and television during the 1996 general election. The study took place in a media landscape marked by the emergence of private television channels, namely *Antena 3*, *Telecinco*, and *Canal+*. Through a telephone survey involving over 300 individuals in Pamplona, the researchers observed significant differences in media preferences among different groups of voters. This outcome lends support to the hypothesis of selective exposure in Spain concerning both print media and television consumption. Similarly, the researchers confirmed that the influence of voting behaviour on preferences for the written press primarily occurs with national newspapers but not with regional publications. However, it's worth noting that these studies approached their interpretations with caution. While they establish a significant relationship between voting and media consumption, they refrain from determining the direction of influence: causal relationships can only be adequately addressed through experimental studies (**Alcalde-Villacampa**, 2003; **Martín-Llaguno; Berganza-Conde**, 2001; **Díaz-Nicolás; Semetko**, 1995; 1999).

The works of **Fraile and Meilán** (2012) and **Humanes** (2014) would significantly enhance and deepen the study of politically-motivated media consumption in Spain for two crucial reasons.

- Firstly, because they use data sources that allow analysis replication. In fact, both studies used post-election surveys, either from the *Centre for Sociological Research (CIS)* or from the *Comparative National Elections Project¹ (CNEP)*, and Fraile and Meilán also incorporated a *TNS* survey conducted during the 2004 election. These surveys offered data from substantially larger and more representative samples of the Spanish population.
- Secondly, because they analysed the connection between political leanings and media exposure, considering a wider range of media: television, press, and radio, and incorporating multiple time points of analysis.

The study conducted by **Fraile and Meilán** (2012) researched the association between ideological self-placement on the left-right axis and media consumption of press, radio, and television during the 1993 and 2004 general elections, as well as the 2009 European elections. Their findings offer substantial empirical evidence in support of the thesis of

selective exposure in Spain. The study reveals that press, radio, and television audiences in Spain are significantly segmented by ideology. Furthermore, it indicates a certain political polarisation occurring over time within the main newspapers (*El País, El Mundo, El Periódico y La Vanguardia*), radio stations (*Cadena SER, COPE, Onda Cero* and *RNE*) and, to a lesser extent, television channels (*Antena 3, Telecinco, Canal+*), measured as an increased proportion of the audience who are ideologically aligned between 1993 and 2009. These results, however, rely on contingency tables and do not incorporate a specific statistical test to demonstrate these temporal differences.

Pioneering studies on selective exposure in Spain were published in the late 1990s and early 2000s, and were the first bivariate explorations of the partisanship of Spanish press and television audiences during elections in the 1990s

Humanes (2014; 2016) explored the selective exposure of Spaniards to television, radio, and the press, employing three indicators of political leanings –vote recall, closeness to political parties, and ideological self-positioning– during the 2008 and 2011 elections. Her results broadly confirm that politically-oriented media consumption occurs in Spain, in addition to two other major findings: 1) that the written press is the medium most conditioned by selective exposure when compared to radio and television, and 2) that the relationship between political beliefs and selective consumption is particularly robust for partisan identification variables:

“vote recall and proximity to a party exhibit stronger associations than ideological leaning across all the three types of media” (**Humanes**, 2014, pp. 788–789).

3.2. The consolidation of the study of politically-oriented media consumption (2017–2023)

These 20 years of slow but steady development in empirical research on selective exposure in Spain is followed by clear consolidation, driven by four circumstances:

- studies now display greater methodological sophistication primarily attributed to the incorporation of multivariate techniques;
- the analysis of variables predicting media consumption becomes more diverse;
- there is growing research into online selective exposure; and
- studies become less reliant on secondary data, with a stronger emphasis on the design of ad hoc questionnaires and other data collection strategies.

This phase commenced in 2017 with the research conducted by **Humanes** and **Mellado**, which introduced the first multivariate analysis of the selective exposure of Spanish audiences, and it continues today. Firstly, studies increasingly apply multivariate techniques, such as binomial and multinomial logistic regression, to larger and more representative samples of the Spanish population through the use of CIS or Digital News Report surveys (**Humanes; Mellado**, 2017; **Ramírez-Dueñas; Vinuesa-Tejero**, 2020; **Moreno-Moreno; Sanjurjo-San-Martín**, 2020; **Rodríguez-Virgili; Sierra; Serrano-Puche**, 2022; **Valera-Ordaz**, 2022; **Humanes; Valera-Ordaz**, 2023). Here, the application of multivariate analysis techniques contributes significantly to the progress of this research field in Spain.² On one hand, it enables researchers to ensure that the correlation between political attitudes and media consumption is not influenced by other latent variables, such as sociodemographic attributes or interest in politics. On the other hand, it quantifies the impact of political leanings on exposure while accounting for the influence of other relevant variables, aligning with international studies (**Castro-Herrero; Nir; Skovsgaard**, 2018; **Stroud**, 2008; **Skovsgaard; Shehata; Strömbäck**, 2016). These control variables encompass age, sex, educational level, interest in politics, trust in institutions, social class, and frequency of media consumption. In other words, during this second phase, the studies confirm a significant correlation between the political leanings of Spanish audiences and their media consumption preferences, regardless of age, sex, educational level, social class, trust in institutions, interest in politics, and frequency of media consumption (**Humanes; Mellado**, 2017; **Valera-Ordaz**, 2018; **Ramírez-Dueñas; Vinuesa-Tejero**, 2020; **Cardenal et al.**, 2019a; 2019b; **Valera-Ordaz; Humanes**, 2022).

It should be noted that this methodological advancement, which may include the use of qualitative techniques such as focus groups (**Masip; Suau; Ruiz-Caballero**, 2018), needs to be contextualised within the fact that Spanish studies on selective exposure have primarily relied on correlation analyses, and there are no experimental designs that provide individuals with various options that allow us to unambiguously determine their preference for ideologically related information (**Feldman et al.**, 2013; **Iyengar; Hahn**, 2009; **Knobloch-Westerwick; Meng**, 2011).

As highlighted by **Knobloch-Westerwick** (2015a), selective exposure research relying on correlations, even when employing panel data and introducing multivariate controls, cannot completely rule out the possibility of other unknown factors influencing the observed patterns of covariance between political beliefs and media preferences. Only experimental research enables the manipulation of stimuli in controlled environments and permits making causal inferences between variables (**Knobloch-Westerwick**, 2015a, pp. 111–112). In Spain, however, there is only one study (**Cardenal et al.**, 2019a) that uses real news events as quasi-experiments to study online information consumption. Nonetheless, there is a notable absence of laboratory designs or anything comparable to the well-established track record of experimental research in the US. On the other hand, Spanish correlation analyses are based on large and representative samples, enhancing their validity and enabling the generalisation of results to the general population (**Clay; Barber; Shook**, 2013, **Knobloch-Westerwick**, 2015a).

Secondly, in recent years there has been a significant diversification of the independent variables studied in Spain. Thus, along with the classic positions on the left-right axis and partisan identification (**Humanes; Mellado, 2017; Ramírez-Dueñas; Vinuesa-Tejero, 2020; Humanes; Valera-Ordaz, 2023**), studies have revealed that media consumption in Catalonia is also influenced by factors such as national identity and positions on the centre-periphery axis (**Valera-Ordaz, 2018; 2022; Sintés-Olivella; Fondevila-Gascón; Berbel-Giménez, 2023**). Indeed, national identity, nationalism, and support for independence serve as robust predictors of media consumption in Catalonia. Individuals who feel more Catalan than Spanish and support independence display a significantly higher preference for regional and local media, while those who identify more nationally with Spain and oppose independence tend to consume state media to a much greater extent (**Valera-Ordaz, 2018; Sintés-Olivella; Fondevila-Gascón; Berbel-Giménez, 2023**).

“ The application of multivariate analysis techniques contributes significantly to the progress of this research field in Spain.² On one hand, it enables researchers to ensure that the correlation between political attitudes and media consumption is not influenced by other latent variables, such as sociodemographic attributes or interest in politics. On the other hand, it quantifies the impact of political leanings on exposure while accounting for the influence of other relevant variables, aligning with international studies ”

Furthermore, research indicates that a climate of political polarisation contributes to the intensification of selective exposure. During the *procés* in Catalonia, for instance, the attitudes of Catalans towards independence progressively gained prominence as predictors of radio and television consumption. This occurred concurrently with the intensification of the conflict surrounding independence (**Valera-Ordaz, 2022**). The reasons behind this trend can be attributed to the continuous activation of these attitudes and the subsequent increase in their cognitive accessibility (**Knobloch-Westerwick; Meng, 2011; Knobloch-Westerwick, 2012**). This was further fuelled by the extensive media coverage of the conflict, leading to a notable polarisation of radio and television audiences (**Martínez-Amat, 2020**). In this sense, the Spanish context still offers unexplored cases of analysis, where multiple axes of political conflict coexist in regions such as the Basque Country, Galicia, and the Valencian Community. Investigating these cases further could deepen the understanding of how selective exposure, driven by identity-related factors, shapes media consumption patterns.

Other studies have raised questions about the significance of interest in politics as a predictor of selective exposure (**Humanes; Mellado, 2017; Rodríguez-Virgili; Sierra; Serrano-Puche, 2022**). These studies aim to ascertain whether selective exposure in the Spanish media system is predominantly driven by ideological and partisan factors or whether it is primarily guided by interest in politics, such that audiences are segmented based on their preference for information-oriented content versus entertainment-oriented content (**Dahlgren, 2019; Ksiazek; Malthouse; Webster, 2010; Skovsgaard; Shehata; Strömbäck, 2016; Strömbäck; Djerf-Pierre; Shehata, 2013**). In the Spanish context, research has confirmed that political selective exposure significantly prevails (**Humanes; Mellado, 2017; Valera-Ordaz; Humanes, 2022**), and that this is especially the case when citizens are interested in public affairs (**Cardenal et al., 2019b; Rodríguez-Virgili; Sierra; Serrano-Puche, 2022**).

However, the data sources used by Spanish researchers do not provide detailed information on the specific contents consumed, they only capture the preferred media of the audiences. This limitation hinders a comprehensive evaluation of the extent of selective exposure guided by interest in politics. Despite this, certain studies do suggest that an interest in public affairs leads to increased exposure to channels like *La Sexta* regardless of political leaning. This may be due to *La Sexta*'s higher presence of political programmes on its schedule compared to *Antena 3* and *Telecinco* (**Humanes; Valera-Ordaz, 2023**).

Other studies suggest that political beliefs exert a direct effect—as postulated by selective exposure—but also an indirect effect—mediated by factors such as trust in the media and the perceived quality of journalistic brands—on media behaviour (**Moreno-Moreno; Sanjurjo-San-Martín, 2020; Victoria-Mas et al., 2023**). In this sense, findings have shown that ideology plays a role in shaping trust in the media, and it can even influence how individuals perceive the media as distributors of disinformation (**Masip; Suau; Ruiz-Caballero, 2020b**). Thus, those who align with left-wing positions have much more trust in media outlets such as *La Sexta*, *El País*, and *Cuatro*, while those who align with right-wing positions have more trust in *El Mundo*, *Antena 3*, and *ABC*. In fact, not many media sources are perceived as reliable by the entire Spanish population. And citizens of both the left and the right believe that non like-minded media sources are creators of disinformation (**Masip; Suau; Ruiz-Caballero, 2020b**). This polarised and agonistic perception of the Spanish media makes the Spanish context an ideal case study to investigate the role of ideological extremism and political antagonism as drivers of both exposure and selective avoidance.

Other studies highlight that during situations where there is greater need for cognitive guidance—such as the COVID-19 pandemic—information utility may outweigh confirmation bias as a driver of media consumption (**Knobloch-Westerwick; Kleinman, 2012; Zoizner et al., 2022**). The study carried out by **Cardenal et al. (2019a)** underscores that the nature of news impacts individuals' motivations when seeking information, concluding that selective exposure tends to increase when there is a surge in information consumption during significant news events (**Cardenal et al., 2019a**).

In addition to the diversification of independent variables, initial studies have been published analysing the extent of online selective exposure in Spain (**Moreno-Moreno; Sanjurjo-San-Martín**, 2020; **Rodríguez-Virgili; Sierra; Serrano-Puche**, 2022), albeit in a very limited way compared to the vast international literature. In general, the studies challenge the hypotheses of self-reinforcing information loops, echo chambers (**Sunstein**, 2003), and filter bubbles. They emphasise that there is significantly more cross-cutting exposure than expected (**Masip; Suau; Ruiz-Caballero**, 2018; **Cardenal et al.**, 2019a; 2019b). For example, **Cardenal et al.** (2019b) demonstrate, using web crawl data from 408 Spanish internet users, that while direct browsing on the internet does lead to an increase in selective exposure, the influence of Google in reducing selective exposure and promoting cross-cutting exposure is more significant. Along the same lines, other studies conclude that Spanish internet users take advantage of the ideological diversity of the media offered by the internet to display transversal consumption (**Cardenal et al.**, 2019a).

“ In Spain, however, there is only one study (**Cardenal et al.**, 2019a) that uses real news events as quasi-experiments to study online information consumption. Nonetheless, there is a notable absence of laboratory designs or anything comparable to the well-established track record of experimental research in the US ”

Similar conclusions were found by **Masip, Suau** and **Ruiz-Caballero** (2018; 2020a). Their study, based on a survey of more than 6,000 digital media users and data from discussion groups, showed that, despite 50% of internet users subscribing to like-minded media, 33% also read news from opposing media to cross-reference and gain a better understanding of reality. Their findings further validate that incidental exposure to content shared by friends and acquaintances on social networks plays a crucial role in promoting cross-cutting consumption of political information (**Masip; Suau; Ruiz-Caballero**, 2018).

In other words, Spanish studies align with international findings that refute the existence of ideological insularity in online news consumption (**Garrett**, 2013), despite initial concerns (**Doval-Avenidaño; Badía-Lago**, 2009). Instead, social networks serve as diversification engines, enabling individuals to discover new information (**Barnidge; Peacock**, 2019), and access a variety of news sources (**Fletcher; Nielsen**, 2018).

Finally, this stage of consolidation is marked by significant advancements in the use of primary sources, reflecting the growing capability of the research community to employ data collection strategies to explore selective exposure. These developments are often facilitated through research project funding. These methodologies involve designing and conducting online surveys targeting representative samples of the general population (**Masip; Suau; Ruiz-Caballero**, 2020b), internet users (**Masip; Suau; Ruiz-Caballero**, 2018; **Cardenal et al.**, 2019a) and individuals registered with digital newspapers (**Masip; Suau; Ruiz-Caballero**, 2020a; **Victoria-Mas et al.**, 2023), carrying out discussion groups, and web crawling data to directly collect data from internet users on their media consumption patterns (**Cardenal et al.**, 2019a; 2019b).

Thus, the study of selective exposure in Spain has become less reliant on general surveys and barometers conducted periodically by various research institutes and companies. These offer the advantage of being based on well-constructed and broadly representative samples of the population. However, they do not provide researchers with the ability to test hypotheses concerning the factors that underlie media consumption beyond the variables collected in the pre-designed questionnaires. It is worth clarifying, however, that secondary sources remain the most commonly used resource in selective exposure research in Spain (see, for example, **Moreno-Moreno; Sanjurjo-San-Martín**, 2020; **Valera-Ordaz**, 2018; 2022; **Ramírez-Deñás; Vinuesa-Tejero**, 2020; **Rodríguez-Virgili; Sierra; Serrano-Puche**, 2022; **Humanes; Valera-Ordaz**, 2023; **Sintes-Olivella; Fondevila-Gascón; Berbel-Giménez**, 2023).

Thus, the study of selective exposure in Spain has become less reliant on general surveys and barometers conducted periodically by various research institutes and companies. These offer the advantage of being based on well-constructed and broadly representative samples of the population. However, they do not provide researchers with the ability to test hypotheses concerning the factors that underlie media consumption beyond the variables collected in the pre-designed questionnaires. It is worth clarifying, however, that secondary sources remain the most commonly used resource in selective exposure research in Spain (see, for example, **Moreno-Moreno; Sanjurjo-San-Martín**, 2020; **Valera-Ordaz**, 2018; 2022; **Ramírez-Deñás; Vinuesa-Tejero**, 2020; **Rodríguez-Virgili; Sierra; Serrano-Puche**, 2022; **Humanes; Valera-Ordaz**, 2023; **Sintes-Olivella; Fondevila-Gascón; Berbel-Giménez**, 2023).

4. Limitations in the application of selective exposure in Spain

Despite this prolific trajectory, selective exposure research in Spain also faces certain limitations which will be outlined below. We will set out some of the blind spots and argue critically why it is essential to expand research in specific directions to develop a research agenda that empowers the academic community and gives it a deeper insight into selective exposure in Spain.

We will focus on five interrelated issues:

- the challenges related to available data sources;
- the limited research on online selective exposure;
- the ongoing issue of endogeneity and the lack of experimental designs;
- the necessity to explore selective avoidance through suitable techniques; and
- the relatively understudied role of public media in fostering or inhibiting cross-cutting exposure.

4.1. Available data: reliance on secondary sources and scarcity of primary sources

Research on selective exposure in Spain has demonstrated a significant reliance on secondary data sources, primarily surveys and barometers conducted periodically by different research institutes and consortia. Table 3 systematically lists the data sources used in selective exposure studies in Spain, together with the type of source and the specific nature of the data. As is evident from the information presented, there is a notable scarcity of primary sources; only a limited

number of studies are based on data specifically designed and collected by researchers, such as a telephone survey (Martín-Llaguno; Berganza-Conde, 2001), one face-to-face survey with members of the *Parliament of Catalonia* (Lacasa-Mas; Jandura; Cano-Castells, 2014) and various online questionnaires and web crawl data (Masip; Suau; Ruiz-Caballero, 2018; 2020a; 2020b; Cardenal *et al.*, 2019a; 2019b; Victoria-Mas *et al.*, 2023). The significant advancement in using primary sources in recent studies on selective exposure can be seen.

Table 3. Data sources used in selective exposure studies in Spain

Authors and year	Title	Data sources
Gunther, Montero and Wert (2000)	The media and politics in Spain: From dictatorship to democracy	Secondary: CNEP (1993 Election)
Diez-Nicolás and Semetko (1995; 1999)	Television news programmes and the 1993 and 1996 electoral campaigns: ownership, content, and influence	Secondary: ASEP (<i>Sociological, Economic and Political Analysis</i>) 1993 and 1996 Elections
Martín-Llaguno and Berganza-Conde (2001)	Voters and the media in the 1996 Spanish general election: Selective exposure or media influence?	Primary: Ad hoc telephone survey (N=363)
Alcalde-Villacampa (2003)	The media as vote intermediaries: the influence of TVE1 and Antena 3 in the 1993 legislative elections	Secondary: CNEP (1993 General election)
Fraile and Meilán (2012)	The media and political information in the 2009 European elections	Secondary: CNEP (1993 General election) TNS (2004 General election) CIS (2009 European elections)
Humanes (2014; 2016)	Selective exposure and partisanship of audiences in Spain. Consumption of political information during the 2008 and 2011 electoral campaigns	Secondary: CIS (2008 and 2011 General elections)
Lacasa-Mas, Jandura and Cano-Castells (2014)	Fragmented exposure to journalistic information, political polarisation and representative democracy: The case of Catalan Parliament members and their use of the media	Primary: Ad-hoc face-to-face survey (N=99)
Humanes and Mellado (2017)	Explanatory models of selective exposure to political information and partisanship of audiences in Spain	Secondary: CIS (2015 and 2016 General elections)
Masip, Suau and Ruiz-Caballero (2018)	Questioning the selective exposure to news: Understanding the impact of social networks on political news consumption	Primary: Ad hoc online survey (N=6679) 12 discussion groups
Valera-Ordaz (2018)	The media, national identity and selective exposure: predictors of the media preferences of Catalan people	Secondary: CIS (2015 Catalonia regional elections)
Cardenal <i>et al.</i> (2019a)	Digital technologies and selective exposure: How choice and filter bubbles shape news media exposure	Primary: Ad hoc, online panel survey and web crawl data (N=408)
Cardenal <i>et al.</i> (2019b)	Echo-chambers in online news consumption: Evidence from survey and navigation data in Spain	Primary: Ad hoc, online panel survey and web crawl data (N=408)
Masip, Suau and Ruiz-Caballero (2020a)	Incidental exposure to non-like-minded news through social media: Opposing voices in echo-chambers' news feeds	Primary: Ad hoc online survey (N=6625)
Masip, Suau and Ruiz-Caballero (2020b)	Perceptions about the media and disinformation: ideology and polarisation in the Spanish media system	Primary: Ad hoc online survey (N=1001)
Ramírez-Dueñas and Vinuesa-Tejero (2020)	Selective exposure and its effects on the electoral behaviour of citizens: the influence of media consumption on voting in the 2015 and 2016 general elections in Spain	Secondary: CIS (2015 and 2016 General elections)
Moreno-Moreno and Sanjurjo-San-Martín (2020)	Selective consumption of journalistic brands in Spain: trust, political leaning and degree of digital consumption	Secondary: 2019 <i>Digital News Report</i>
Valera-Ordaz, Requena-Mora and Pop (2021)	Ideology and selective exposure: Spain's media consumption of during the April 2019 electoral campaign	Secondary: CIS (2019 April General election)
Ramírez-Dueñas and Vinuesa-Tejero (2021)	How does selective exposure affect partisan polarization? Media consumption on electoral campaigns	Secondary: CIS (2019 April General election)
Rodríguez-Virgili, Sierra and Serrano-Puche (2022)	Motivations for news exposure in different media systems: A comparative study of Germany, Spain and the United Kingdom	Secondary: 2020 <i>Digital News Report</i>
Valera-Ordaz and Humanes (2022)	What drives selective exposure to political information in Spain? Comparing political interest and ideology	Secondary: CIS (2019 November General election)
Valera-Ordaz (2022)	Political identity and news media choice: The polarizing logic of selective exposure during the Catalan independence conflict	Secondary: CIS (2010, 2012, 2015 and 2017 regional elections in Catalonia)
Sintes-Olivella, Fondevila-Gascón and Berbel-Giménez (2023)	Television and independence in Catalonia. Selective exposure of Catalan viewers during the independence process (2015–2020)	Secondary: CEO (2015, 2016, 2017, 2018, 2019 and 2020 <i>Barometers</i>)

Authors and year	Title	Data sources
Humanes and Valera-Ordaz (2023)	Partisanship, ideology, and selective exposure: A longitudinal analysis of media consumption in Spain (2008–2019)	Secondary: CIS (2008, 2011, 2015, 2016, April 2019 and November 2019 general elections)
Victoria-Mas et al. (2023)	News media brands' value within polarized media markets: Perceived quality vs. political congruence	Primary: Ad hoc online survey (N=15000)

Specifically, the most frequently used data sources come from post-election studies by CIS (**Humanes**, 2014; 2016; **Humanes; Mellado**, 2017; **Valera-Ordaz**, 2018; 2022; **Ramírez-Dueñas; Tejero-Vinuesa**, 2020; 2021; **Humanes; Valera-Ordaz**, 2023), barometers of its Catalan counterpart *Centre de Estudis d'Opinió* (**Sintes-Olivella; Fondevila-Gascón; Berbel-Giménez**, 2023), and studies by the *Comparative National Elections Project* (**Gunther; Montero; Wert**, 2000; **Alcalde-Villacampa**, 2003; **Fraile; Meilán**, 2012), an international academic consortium based at *Ohio State University*, that has conducted election polls since the 1980s in countries around the world, and which has been often used by political science researchers.

More recently, researchers from the *University of Navarra* have published studies based on data pertaining to Spain collected by the *Digital News Report*. This report is yearly compiled by the *Reuters Institute for the Study of Journalism* at the *University of Oxford*, who maintains a collaborative partnership with the *University of Navarra*. The group of researchers associated with the *Center for Internet Studies and Digital Life*, which includes Alfonso Vara-Miguel, Samuel Negrodo, Avelino Amoedo-Casais, Elsa Moreno-Moreno and Jürg Kaufmann-Arqueta, compiled the aforementioned annual report spanning 2014 to today through an online survey carried out by *YouGov* on a sample of approximately 2,000 internet users (**Vara-Miguel**, 2022). Whilst confined to the Spanish online user population, this internet survey encompasses a range of questions on media consumption, sociodemographic attributes, and political factors, and has facilitated comparative analyses concerning individual motivations behind news exposure (**Rodríguez-Virgili; Sierra; Serrano-Puche**, 2022), and research exploring the influence of political leanings and journalistic branding on news consumption (**Moreno-Moreno; Sanjurjo-San-Martín**, 2020).

Likewise, surveys of demoscopic companies, such as *Sociological, Economic and Political Analysis* (**Díez-Nicolás; Semetko**, 1995, 1999) and *TNS Demoscopia* (**Fraile; Meilán**, 2012), have also been used sporadically. Such instances are less common and can be attributed to the direct affiliations of certain researchers with these companies, as the data they use remain inaccessible to the public. Contrarily, using studies from these companies is expensive and only within the reach of well-funded research groups (**Masip; Suau; Ruiz-Caballero**, 2018; 2020; **Cardenal et al.**, 2019a; 2019b). This is why a significant portion of the Spanish academic community has chosen to use freely accessible data generated by research institutes for statistical analysis.

Secondary sources, nevertheless, have certain benefits that have contributed to the robustness of the study of selective exposure in Spain.

- Firstly, its use safeguards against the numerous issues introduced by confirmation bias that exists in all scientific research. In essence, they provide greater assurance against researchers' inclination to confirm preconceived notions and unintentionally introduce biases in questionnaire design and data collection processes.
- Secondly, surveys like *CIS*, *CEO*, or *CNEP* are built upon extensive and well-constructed population samples, thereby enhancing the potential for generalisations.
- Finally, the media consumption questions found in these surveys, which are employed as dependent variables in the analyses, can be readily regarded as measures of habitual exposure rather than singular exposure. Because they typically inquire about interviewees' overall media preferences, this enhances the capacity to generalise the findings (**Clay; Barber; Shook**, 2013; **Stroud**, 2008).

Despite these advantages, working with secondary data also has obvious limitations. For instance, this type of survey, where interviewees are questioned about past behaviours, is more vulnerable to post hoc rationalisations compared to experiments where individual actions can be directly observed. Moreover, employing general surveys introduces limitations in the empirical investigation of selective exposure, as it confines researchers to test hypotheses within the pre-defined boundaries set by questionnaires. For example, the *CIS* had only included questions about preferred media (radio, television, and newspaper) since the 2008 election; it was only in the November 2019 post-election survey when a specific question about interviewees' favourite digital newspaper was introduced, which was not retained in subsequent surveys. This has led to a focus on traditional media consumption in most research, with limited exploration of online selective exposure.

The majority of electoral surveys of this nature do not collect specific media content and programmes consumed by individuals, nor the time they dedicate to them,

This stage of consolidation is marked by significant advancements in the use of primary sources, reflecting the growing capability of the research community to employ data collection strategies to explore selective exposure. These developments are often facilitated through research project funding

only preferred media. A lack of disaggregated content data makes evaluating the extent of selective exposure driven by political interest in Spain challenging, unlike research conducted in other countries (Skovsgaard; Shehata; Strömbäck., 2016). Nor do they account for the consumption of digital media (such as digital natives, social networks, etc.) or the deliberate avoidance of certain media due to ideological reasons.

Employing general surveys introduces limitations in the empirical investigation of selective exposure, as it confines researchers to test hypotheses within the predefined boundaries set by questionnaires

The fact that Spanish researchers depend on general surveys carried out by other organisations has not facilitated an analysis of the phenomenon of selective exposure beyond the level of individuals' preferred media. As a result, there is a lack of understanding regarding the variables that explain why certain individuals prefer television news, others opt for morning magazines, and some choose radio talk shows as their sources of information. Similarly, the need to further diversify the independent variables remains unaddressed. To enable researchers to conduct in-depth analyses of how factors like the nature of news or trust in the media influence selective exposure, having access to data sources that include these types of questions is essential.

What is the situation of studies on audiences? Unlike surveys, audience studies do collect detailed data on content consumption, encompassing various media formats, types, duration, and frequency of consumption.

- Firstly, the issue lies in the expensive nature of access to this data, making it unavailable to researchers, as it is reserved for members only. For instance, gaining full access to the *General Media Study* conducted annually by the *AIMC* (*Association for Communication Media Research*), which involves a consortium of media, agencies, and advertisers, requires a minimum annual fee of 7,500 euros.³
- Secondly, the methodology of these studies is frequently opaque as they are designed primarily for commercial purposes rather than scientific research. Indeed, this is the case with measurements of television audiences conducted by *Kantar Media* using audimeters, and measurements of digital media audiences by *Comscore*,⁴ whose methodologies are not made public (Zain, 2014).
- Lastly, audience studies typically include sociodemographic variables such as sex and age, but they lack political variables like ideology and voting behaviour. This limitation makes these data sources unsuitable for selective exposure studies. As an example, the *General Media Study* conducted annually by the *AIMC* involves interviews with approximately 30,000 individuals, and while it is likely the most transparent audience study in Spain, providing both a summary of the questionnaire and a detailed technical file,⁵ it records sociodemographic classification data but lacks variables related to the political positions of individuals. Establishing collaboration between the communication research community and the *AIMC*, possibly through some of its reference associations (such as *AE-IC*, *SEP*), could be highly beneficial, as such a collaboration could aim to include questions of sociological and political significance. This would make the *General Media Study* a very valuable data source, not only for the industry, but also for the scientific community.

4.2. Online selective exposure

The second important limitation pertains to the scarcity of studies exploring selective consumption of online information among Spanish audiences. Most of the research has primarily focused on analysing the consumption of traditional media. The lack of studies addressing this phenomenon contrasts with the considerable consumption of digital media and the significance of social networks as information sources for Spaniards. As an example, according to the 2022 *Digital News Report*, the internet is a source of information for 79% of Spaniards, social networks are regular sources of information for 56%, while television is for 59%, and print media is for only 26%.

Despite this data, there is a notable scarcity of studies analysing the phenomenon of online selective exposure in Spain (Cardenal *et al.*, 2019a; 2019b; Masip; Suau; Ruiz-Caballero, 2018; 2020a). These studies cover aspects such as the consumption of digital media, exposure to content on social networks, or browsing facilitated by search engines like *Google*. The reasons have been outlined above, directing us to the available data. Nonetheless, two research teams from Catalan universities stand out as notable exceptions, as their scientific output significantly contributes to our current understanding of the extent of the phenomenon in Spain. The team led by the late Pere Masip from *Ramon Llull University*, stemming from the communication field, and the team of Ana S. Cardenal from *Universitat Oberta de Catalunya*, specialising in political science, published groundbreaking studies on the phenomenon of echo chambers and filter bubbles. Their research utilised *ad hoc* online surveys (Cardenal *et al.*, 2019a; 2019b; Masip; Suau; Ruiz-Caballero, 2020a), discussion groups (Masip; Suau; Ruiz-Caballero, 2018), and even web tracking data from the Spanish internet user population (Cardenal *et al.*, 2019a; 2019b).

Cardenal *et al.* (2019a) used web browsing data tracking from a sample of 408 Spanish internet users and concluded that there are no echo chambers in the consumption of online political information. Their findings show that individuals use the diversity of online media to expose themselves to both like-minded and non-like-minded content. This differs from the US scenario where there is a consumption of centrist media (Gentzkow; Shapiro, 2011; Nelson; Webster, 2017), in Spain, the exposure is truly transversal, encompassing all kinds of media. Furthermore, they highlight that while direct consumption of digital media via the internet slightly increases selective exposure, the influence of *Google* in reducing

selective exposure is more significant (**Cardenal et al.**, 2019b). In essence, Google not only prevents filter bubbles but also fosters cross-cutting exposure. Their findings also reveal that online selective exposure becomes more pronounced as the volume of information consumed increases during specific news events (**Cardenal et al.**, 2019a). This emphasises the need to expand the study of selective exposure by considering the nature of the news, a task that remains pending in the Spanish academic realm.

Establishing collaboration between the communication research community and the *AIMC*, possibly through some of its reference associations (such as *AE-IC*, *SEP*), could be highly beneficial, as such a collaboration could aim to include questions of sociological and political significance

Studies by **Masip, Suau** and **Ruiz-Caballero** (2018; 2020a) yield similar findings. Using a triangulated approach, incorporating a survey of over 6,000 digital media users and discussion groups, they found that while half of the users exclusively subscribe to like-minded media, 33% also engage with news from non-like-media outlets to gain a better understanding of the reality. Furthermore, their work reveals incidental exposure to diverse content shared by friends and acquaintances on social networks, depicting a cross-cutting consumption scenario that deviates from self-reinforcing information (**Masip; Suau; Ruiz-Caballero**, 2018). In fact, in a subsequent study, the authors observed a notable association between the intensive use of social networks and consumption of information not aligned with ideological beliefs, particularly on platforms such as Twitter (**Masip; Suau; Ruiz-Caballero**, 2020a).

Similarly, the comparative study by **Rodríguez-Virgili, Sierra** and **Serrano-Puche** (2022) highlights that a majority of German, British and Spanish consumers of digital media prefer unbiased news. However, their results also demonstrate that Spanish citizens are more inclined than their British or German counterparts to prefer like-minded media. Meanwhile, the study conducted by **Victoria-Mas et al.** (2023) identified the key dimensions of journalistic brands for readers of the Catalan digital press. It concluded that political congruency is one of the most significant aspects, indicating that brands are perceived as valuable when they align with users' political beliefs.

In summary, the limited Spanish studies on online selective exposure align with international research findings. They reveal that ideological insularity in online news consumption is less pronounced than anticipated (**Garrett**, 2013), influenced by factors like search engine mediation (**Cardenal et al.**, 2019b), the role of social networks and weak social ties that foster diversification by facilitating exposure to diverse and plural information (**Barnidge; Peacock**, 2019; **Fletcher; Nielsen**, 2018).

4.3. The ongoing issue of endogeneity and the lack of experiments

Thirdly, researching selective exposure in Spain poses the problem of endogeneity, which refers to the challenge of establishing causal relationships between two phenomena—political beliefs and media consumption—where a single direction of influence cannot be definitively determined. Indeed, from a theoretical standpoint, it appears more plausible to view them as mutually influential phenomena. Political beliefs are not shaped in isolation from media exposure, and media consumption cannot be considered entirely impervious to the ideological orientations of individuals.

At the core of the endogeneity issue lies a crucial issue: Spanish studies rely solely on analysing correlations between political beliefs and media consumption. The vast majority use retrospective reports of media consumption through surveys, with direct observation measures, such as web tracking data, being used to a much lesser extent (**Cardenal et al.**, 2019). There are currently no experimental designs that expose individuals to different types of media content (related, neutral, contradictory) to observe their behaviour, as seen in the American academic tradition, which has a long history of experimental studies (**Iyengar; Hahn**, 2009; **Knobloch-Westerwick**, 2003).

Experiments are the only way to rule out that the correlation between two phenomena is due to a latent variable. For this reason, when dealing with cases that involve endogeneity, estimating the parameters of the regressor variables may lack consistency, leading to potentially significant discrepancies between the estimated effects and the real ones. It is important to remember that in the history of empirical research on selective exposure, studies based on correlations have tended to yield stronger effects than those based on experiments (**D'Alessio; Allen**, 2002; **Stroud**, 2011). It is not a question of empirical research being hindered by methodological dogmatism; however, due to the lack of experimental designs, it is highly probable that the effects of selective exposure in Spain have been overestimated.

Most Spanish authors acknowledge that endogeneity is a limitation in their studies (**Martín-Llaguno; Berganza-Conde**, 2001; **Díez-Nicolás; Semetko**, 1999; **Humanes**, 2014), and some use panel data to measure political leanings before media consumption, thereby mitigating potential problems of reverse causality (**Cardenal et al.**, 2019a; 2019b). In fact, certain studies explicitly acknowledge the endogeneity issues and suggest analysing the relationship between media consumption and voting in both directions: from a selective exposure perspective and exploring how media consumption influences electoral behaviour (**Díez-Nicolás; Semetko**, 1999; **Alcalde-Villacampa**; 2003; **Ramírez-Dueñas; Vinuesa-Tejero**, 2020) or partisan polarisation (**Ramírez-Dueñas; Vinuesa-Tejero**, 2021).

It is worth noting, however, that experimental designs are not exempt from challenges and limitations. Firstly, experimental designs carry the risk of overestimating variables that researchers can control within limited laboratory settings, whe-

re subjects participate for a restricted period. Secondly, experimental designs require individuals select media in highly unrealistic choice contexts, often with a limited number of created media messages by researchers. These messages are categorised as consistent, neutral, or inconsistent, but there are always fewer options than what exists in the real world (Clay; Barber; Shook, 2013). The fact that these laboratory choice contexts do not resemble the real situation in which individuals make their decisions about media consumption constitutes a major limitation of the experimental approach (Knobloch-Westerwick, 2015a), since media selection is highly context-sensitive (Feldman *et al.*, 2013).

Lastly, other techniques, while not as robust as randomised experiments, can help establish the direction of influence between political attitudes and media consumption. Panel studies and time series analyses, which involve monitoring a sample of individuals over time, enable the identification of behaviour patterns and changes in individuals' media consumption over time. These approaches present a valuable opportunity to advance selective exposure research in Spain.

4.4. Selective avoidance, a pending task

Research conducted by Ramírez-Dueñas and Vinuesa-Tejero (2020) provided confirmation of the influence of media consumption on voting behaviour during the 2015 and 2016 elections. In addition to discovering evidence of individuals selecting media outlets that align with their opinions, they emphasise

“the limited openness of voters to engage with media that are not aligned, and even a growing inclination to reject media that are contrary to their political position (ideological or territorial)” (Ramírez-Dueñas; Vinuesa-Tejero, 2020, p. 23).

Humanes and Valera-Ordaz (2023) arrived at similar conclusions in their examination of the variables influencing media consumption among Spaniards over an 11-year period (from 2008 to 2019). They discovered that certain voter groups are more likely to consume related media, and voting is also linked to a notably reduced willingness to be exposed to non-related content.

“The findings of this study also suggest that certain media produce more polarisation because they simultaneously inspire exposure and selective avoidance by certain ideological groups (...) Although both selective exposure and avoidance occur simultaneously only at specific times, the existence of these media is likely to compromise exposure to diversity” (Humanes; Valera-Ordaz, 2023, p. 11).

This leads us to another limitation in the study of selective exposure in Spain: the scarcity of research on selective avoidance. In Spain, there are no studies exploring whether individuals intentionally avoid certain media due to political or ideological reasons. The absence of specific questions on explicit media rejection in the data sources used contributes to this limitation. Indeed, the CIS questionnaires used in the aforementioned analyses (Ramírez-Dueñas; Vinuesa-Tejero, 2020, 2021; Valera-Ordaz; Humanes, 2022; Humanes; Valera-Ordaz, 2023) only gather information on individuals' preferred media (e.g., “Which newspaper/radio station/television station do you prefer to follow political and electoral information during the campaign?”). However, they do not capture data on comprehensive media consumption patterns, nor do they include specific inquiries related to media avoidance.

While selective exposure studies in Spain often highlight that certain ideological groups or voters tend to avoid consuming non-like-minded media, this interpretation is based on the negative sign coefficients observed in logistic regression models (Ramírez-Dueñas; Vinuesa-Tejero, 2020; Valera-Ordaz; Humanes, 2022; Humanes; Valera-Ordaz, 2023). However, it is important to note that just because *PSOE* voters preferentially consume *El País*, this does not necessarily imply that they avoid *El Mundo*, *La Vanguardia* or *ABC*. To determine which media individuals avoid consuming, one would need to know their media diets exhaustively or inquire about them in detail.

So, while the interpretation of the coefficients is technically valid (indicating that certain ideological groups are significantly less likely to be exposed to certain media), the fundamental issue lies in the design of the CIS post-election questionnaires and the nature of the questions they contain, which do not allow such inferences to be made. In fact, from a theoretical perspective, it is more likely that individuals simultaneously consume multiple media sources. For instance, they may read both a national and a regional newspaper or a newspaper and a television channel, as well as being exposed to messages and content when browsing the internet (Masip; Suau; Ruiz-Caballero, 2018; Cardenal *et al.*, 2019a). Empirically validating the occurrence of selective avoidance necessitates experimental designs or, at the very least, correlational analysis that includes explicit questions about which media individuals intentionally avoid.

Therefore, the fact that some groups of voters have a lower probability of preferring media that are situated at their ideological antipodes does not constitute a general verification of the phenomenon of selective avoidance. As highlighted by certain researchers (Garrett, 2009; 2013; Camaj, 2019), selective exposure does not necessarily entail systematic

Despite this data, there is a notable scarcity of studies analysing the phenomenon of online selective exposure in Spain (Cardenal *et al.*, 2019a; 2019b; Masip; Suau; Ruiz-Caballero, 2018; 2020a). These studies cover aspects such as the consumption of digital media, exposure to content on social networks, or browsing facilitated by search engines like Google

avoidance of non-like-minded information; inferring the absence of cross-cutting and incidental exposure from variables that solely include a preferred media outlet is not possible. Ultimately, it is more plausible that the effects of selective exposure are offset by media diets that include the consumption of all types of media.

These studies do not diminish the significance of the aforementioned research; the findings indicating the lower likelihood of certain ideological groups consuming specific media (**Ramírez-Dueñas; Vinuesa-Tejero, 2020, Humanes; Valera-Ordaz, 2023**) prompt the research community to further explore how variables such as ideological extremism, political interest, (dis)trust in institutions, or partisan loyalty influence selective avoidance. To achieve this, questionnaires should be designed to include questions on avoidance and comprehensively capture interviewees' media diets. This is a particularly pertinent task for Spain, since the media are considerably politicised (**Hallin; Mancini, 2004**), its journalistic style tends to emphasise opinion over information, and professional autonomy is limited by the strong interdependence of the political and media systems (**Humanes; Martínez-Nicolás; Saperas, 2013; Humanes; Roses, 2018**).

It is not a question of empirical research being hindered by methodological dogmatism; however, due to the lack of experimental designs, it is highly probable that the effects of selective exposure in Spain have been overestimated

4.5. Public media: facilitators of selective or cross-cutting exposure?

The quality of democracy is significantly influenced by cross-cutting exposure, which refers to citizens' encounters with diversity. Such exposure plays a crucial role in enhancing political tolerance (**Mutz, 2002**) and promoting consensus-building in complex societies (**Mutz; Young, 2011**). Although the recent proliferation of media options has made individual interest a significant driver of media consumption, research has shown that the structure of opportunities provided by each media system also impacts the likelihood of individuals being exposed to diversity (**Skovsgaard; Shehata; Strömbäck, 2016**).

For example, **Castro-Herrero, Nir and Skovsgaard (2018)** conducted a study to explore the influence of public service broadcasters' dominance in the media market of 27 European countries on individuals' cross-sectional exposure. They found that public media's tendency to offer a greater amount of political information contributes to increasing the likelihood of individuals encountering non-like-minded information. Their results indicate that there are fewer disparities in cross-cutting exposure between individuals with varying levels of political interest in countries with a robust public service. In contrast, in countries with a weaker public service, cross-cutting exposure tends to be more unequal. In the context of this study, Spain recorded mean values of cross-cutting exposure well below Scandinavian countries, but higher than those in Baltic countries. In a similar vein, **Bos, Kruike-meier and De-Vreese (2016)** demonstrated that Dutch public television mitigated citizens' selective exposure. As the only medium consumed by all ideological groups, it acted as a unifying force for social cohesion, despite the considerable levels of selective exposure observed in other public media.

Numerous studies have established that European public service broadcasters exhibit varying levels of popularity and politicisation (**Brüggemann et al., 2014**). In systems of polarised pluralism like Spain, public media tend to be politicised (**Hallin; Mancini, 2004**), and it is common for governments to use the state and regional broadcasting systems they administer to their advantage (**Arriaza-Ibarra, 2013; Jones, 2007**), leading to perceptions of partisanship by the public (**Azurmendi; Muñoz-Saldaña; Ortega-Mohedano, 2018**). Even the granting of licenses to private broadcasters has been used as a strategy by national and regional governments to favour specific media groups (**Arriaza-Ibarra, 2013; Jones, 2007**).

In Spain, there have been relatively few studies analysing the role of public broadcasting systems as facilitators or inhibitors of cross-cutting exposure, despite them often being criticised for their instrumentalisation. In the case of *RTVE*, several studies find patterns of selective exposure based on the party in government, but there has been no systematic review, only cross-sectional studies that analyse specific moments.

For example, **Fraile and Meilán (2012)** found that the ideological composition of Televisión Española audiences changed over time according to whichever party was in government.

"The most intriguing finding pertains to the composition of the public television audience, which shifted from predominantly left-wing in 1993 to right-wing in 2004, and then returned to being primarily left-wing in 2009, aligning with the governments in power during those periods" (**Fraile; Meilán, 2012, p. 125**).

The same trend was found for *Radio Nacional de España (RNE)*, so that the majority ideology of its audiences changed with the changes of government between 1993 and 2009.

Subsequent studies have endorsed this trend. Thus, according to the *2010 Report on the journalistic profession in Spain* (cited in **Humanes, 2014, p. 788**), the Spanish placed *RNE* and *TVE* on the centre-left in 2010 (4.7 and 4.5, respectively, on a scale from 1 to 10) after six years of Zapatero's socialist government. Furthermore, the bivariate analyses conducted by **Humanes (2014)** revealed statistically significant associations between voting for *PSOE* and feeling an affinity to this party, and the consumption of *TVE* in both 2008 and 2011 (**Humanes, 2014, p. 791**), although this pattern was not observed with the public channel *La 2*. In the case of *RNE*, there was also an association with voting for *PSOE*, but effects were smaller.

In the 2015 election, following four years of Rajoy's government, voting for the *PP* was associated with an increased likelihood of consuming *TVE1* (Humanes; Valera-Ordaz, 2023), a trend that persisted in the 2016 election (Ramírez-Dueñas; Vinuesa-Tejero, 2020). Similarly, in April 2019, individuals located on the centre and on the right exhibited a higher probability of preferring *TVE1* compared to those situated on the left (Valera-Ordaz; Requena; Pop, 2021).

Empirically validating the occurrence of selective avoidance necessitates experimental designs or, at the very least, correlational analysis that includes explicit questions about which media individuals intentionally avoid

Nevertheless, the pattern differed for *RNE*. In 2015 and 2016, the consumption of *RNE* was evenly distributed among groups of voters (Ramírez-Dueñas; Vinuesa-Tejero, 2020), and in April 2019, it also “has a transversal presence in the different segments of the ideological spectrum” (Valera-Ordaz; Requena; Pop, 2021, p. 167), and “it is not a reference for any specific political faction” (Valera-Ordaz; Requena; Pop, 2021, p. 171). In summary, exposure to the public television channel varies depending on the party in government, while public radio has been consistently consumed transversally by all Spaniards since 2015, regardless of their ideology and vote.

However, at a regional level, only the case of the *Corporació Catalana de Mitjans Audiovisuals (CCMA)* has been investigated, even though some studies clearly indicate that, in general, regional television systems –such as *Telemadrid* and *TV3*– are the least trusted media sources for the Spanish population in terms of reliability (Azurmendi; Muñoz-Saldaña; Ortega-Mohedano, 2018; Masip; Suau; Ruiz-Caballero, 2020). It is worth noting that most of the regional public media in Spain were established with the aim of reinforcing regional identities and promoting linguistic and cultural integration (Holtz-Bacha, 2015). In some cases, they were even conceived as tools for national construction (Guimerà, 2014; 2018), with the *CCMA* serving as an exemplary case.

In this context, studies have shown that watching *TV3* is associated with an increased likelihood of supporting Catalan independence (Serrano, 2013) and a feeling of national identification with Catalonia (Hierro-Hernández, 2012). Additionally, research has found that individuals who support independence are significantly more inclined to prefer Catalan public television over non-independence supporters (Valera-Ordaz, 2018; Sintés-Olivella; Fondevila-Gascón; Berbel-Giménez, 2023), and that selective exposure to *TV3* and *Catalunya Radio* has grown substantially over the last few years (Valera-Ordaz, 2022). In summary, studies concur that in recent years, the Catalan public broadcasting system has been selectively consumed by individuals who support independence, identify exclusively as Catalan, and endorse nationalism. Conversely, it has been avoided by those who oppose independence and exclusively identify as Spanish.

Indeed, beyond the Catalan case, there is a lack of longitudinal or multivariate studies exploring the role of most regional broadcasting services (e.g., *Telemadrid*, *Euskal Telebista*, *Corporación Radio e Televisión de Galicia*, the old *Canal Nou* and the new *À Punt*, *Radio y Televisión de Andalucía*, etc.) as promoters or inhibitors of cross-cutting exposure. Despite being administered by the same party for extended periods, and being perceived as politicised by the public, research on their impact remains unexplored (Azurmendi; Muñoz-Saldaña; Ortega-Mohedano, 2018). Humanes (2014) highlights that in 2008 and 2011, there was a significant statistical association between voting for a specific party and consuming the regional public television channel in that area, such as *Canal Sur* (associated with *PSOE*), *TV3 (CiU)*, *Canal Nou (PP)*, *Telemadrid (PP)*, *TVG (BNG)*, and *ETB1 (PNV and Amaiur)*.

These data provide a compelling invitation to conduct in-depth analyses of the role of public media. In this sense, the research community in Spain must make an effort to determine the extent to which these systems encourage cross-cutting exposure or are selectively consumed only by some ideological groups. After all, the idea of public service is to provide quality journalistic information to all citizens beyond their social, economic and political differences. This should mean that public television and radio stations serve the whole of society, not only the sectors related to the government that administers them. Conducting a comprehensive analysis of the factors that influence the consumption of public media in Spain, both at national and regional levels, and examining their trends over time, is an essential task in selective exposure studies.

5. Conclusions: an agenda for the future

Selective exposure theory remains highly relevant in international communication studies (Camaj, 2019; Humanes, 2019) due to the recent evolving information environments (Chadwick, 2013) and the increased availability of media options, enabling individuals to engage in politically-oriented and selective media consumption (Skovsgaard; Shehata; Strömbäck, 2016).

Over the past three decades, selective exposure research in Spain has witnessed significant growth and contributions from the fields of communication, political, and social sciences; with an initial emerging phase of bivariate studies (Gunter; Montero; Wert, 2000; Díez-Nicolás Semetko, 1999; Fraile; Meilán, 2012; Humanes, 2014), followed by a consolidated stage marked by methodological advancements, diversified independent variables, exploration of online selective exposure, and an increased commitment to primary data sources (Humanes; Mellado, 2017; Cardenal *et al.*, 2019a; Masip; Suau; Ruiz-Caballero, 2020b; Moreno-Moreno; Sanjurjo-San-Martín, 2020; Ramírez-Dueñas; Vinuesa-Tejero,

2020; **Rodríguez-Virgili; Sierra; Serrano-Puche**, 2022; **Humanes; Valera-Ordaz**, 2023). Through numerous studies, it has been established with certainty that Spanish press, radio, and television audiences choose their media according to their political leaning. Additionally, there is provisional evidence suggesting that online information consumption exhibits less ideological insularity and more cross-cutting exposure than initially presumed (**Masip; Suau; Ruiz-Caballero**, 2020b; **Cardenal et al.**, 2019b).

In this article we have argued that, despite the significant progress in empirical research of selective exposure in Spain, there are still some gaps in our understanding of the phenomenon. These limitations stem from the reliance on secondary data and the absence of experiments, which have hindered the further development of this research area in the country.

To ensure effective advancement in the study of selective exposure in the future, two critical steps must be taken. Firstly, researchers should prioritise primary data sources and secure adequate funding to collect original data with diversified variables of interest, using methods such as questionnaires, discussion groups, or experimental designs. Secondly, public research institutes should incorporate a wider range of questions related to media consumption, considering the increasingly complex contemporary informational environments. These measures will enhance the quality and scope of research in this field in Spain. Ultimately, we will only be able to adequately assess the influence of various variables (e.g., interest in politics, ideological extremism, nature of the news) as drivers of selective exposure, particularly in the context of media proliferation, by analysing disaggregated consumption data.

We emphasised the need for the research community in Spain to focus more on selective avoidance and online selective exposure phenomena. Presently, there is a lack of studies specifically addressing selective avoidance using appropriate techniques, partially due to limitations imposed by post-election surveys as data sources. However, some existing studies offer evidence of significantly lower consumption of certain media by specific ideological groups (**Ramírez-Dueñas; Vinuesa-Tejero**, 2020; **Valera-Ordaz; Humanes**, 2022; **Humanes; Valera-Ordaz**, 2023).

Online selective exposure has been explored in a preliminary and constrained manner in Spain through certain studies (**Masip; Suau; Ruiz-Caballero**, 2020b; **Cardenal et al.**, 2019a; 2019b). However, considering the significant influence of the internet and social networks as primary information sources for Spaniards, further comprehensive research is necessary. Continued research in this direction will strengthen initial findings and deepen our understanding of the factors influencing internet browsing patterns and digital media consumption.

We believe that the theory of selective exposure can be used comprehensively to evaluate the role of both national and regional public broadcasting services in the Spanish public sphere, aligning with international research (**Bos; Kruijemeier; de Vreese**, 2016; **Castro-Herrero; Nir; Skovsgaard**, 2018). Ultimately, democracies bear the responsibility of exposing citizens to quality journalistic information and fostering diversity, a task specifically entrusted to public media systems.

Advancing the understanding of politically-oriented media consumption will shed light on its relation to phenomena such as misinformation, populism, and political and affective polarisation. Recent studies indicate that restrictive media diets, characterised by exposure to a limited number of media with highly partisan content, can impact voting decisions and diminish tolerance towards other political parties (**Ramírez-Dueñas; Vinuesa-Tejero**, 2021).

In conclusion, we would like to note some of the limitations of this study and propose potential avenues for further research. Firstly, it is worth noting that certain studies on the selective exposure of Spanish audiences may not be captured in the results, as the *Dialnet* database might not encompass all scientific publications and this study has a broad timespan (1995-2023). Secondly, the growing number of international comparative studies that have not been included must be noted (there is an increasing integration of Spanish researchers into international teams financed by solid transnational research projects, such as Ana S. Cardenal, Laia Castro, and the researchers at the *University of Navarra* who collaborate with the *Reuters Institute for the Study of Journalism*). Future reviews should consider the social, institutional, and intellectual history of the Spanish scientific community; employing additional research techniques such as interviews with experts, the Delphi method, or discussion groups⁶ that can provide insights into the direction of research on selective exposure in Spain.

Other noteworthy areas for advancing the analysis of selective exposure in Spain include the significant role played by the social and political sciences, evident in various publications (**Díez-Nicolás; Semetko**, 1995; 1999; **Fraille; Meilán**, 2012; **Ramírez-Dueñas; Vinuesa-Tejero**, 2020; **Cardenal et al.**, 2019a; 2019b). This influence might extend to the adoption of analytical approaches and techniques in communication studies. Indeed, from the perspective of the sociology of science, exploring this line of work to understand the evolution of selective exposure studies in Spain would be of interest.

Additionally, analysing the impact of early empirical work on subsequent research as the consolidation phase progresses and more empirical evidence accumulates

“ In Spain, there have been relatively few studies analysing the role of public broadcasting systems as facilitators or inhibitors of cross-cutting exposure, despite them often being criticised for their instrumentalisation ”

will help identify the most influential papers. One last issue worthy of attention is the scientific production of complementary lines of research, such as selective perception and memory, which draws from psychology. Integrating findings from these areas will broaden our understanding and provide a more comprehensive view of the phenomenon of selective exposure. Indeed, according to its original formulation, selectivity encompasses not only exposure to the media but also information processing, thus contributing to ideological reinforcement mechanisms (Camaj, 2019).

6. Notes

1. The *Comparative National Elections Project (CNEP)* is an association of academics founded in the late 1980s, based at *Ohio State University*, which has conducted electoral surveys on five continents for decades. The *CNEP* includes 66 surveys carried out between 1990 and 2022 in 30 different countries. In the case of Spain, the surveys include the 1993, 2004, 2011, and 2015 elections.

2. The consolidation stage aligns with the transformation of the Spanish political landscape from a consolidated bipartisanship to a multiparty system, marked by the emergence of parties like *Podemos* and *Ciudadanos* initially, and later *Vox*. This transformation also facilitates the application of techniques such as logistic regression and multinomial regression, enabling the operationalisation of the vote for parties beyond *PSOE* and *PP*.

3. Information on the *AIMC* fee:

<https://www.aimc.es/aimc/como-asociarse>

4. More about *Comscore's* method of measuring audiences:

<https://mynews.es/comscore-el-medidor-de-audiencias-en-medios-digitales/#:~:text=En%20primer%20lugar%2C%20ComScore%20utiliza,cProxy%2C%20instalado%20en%20sus%20ordenadores>

5. Here is an example of a technical data sheet:

<https://www.aimc.es/egm/que-es-el-egm/ficha-tecnica/ficha-tecnica-egm-3a-ola-2019>

and a summary of the questionnaire content:

<https://www.aimc.es/egm/que-es-el-egm/temas-tratados-cuestionarios>

6. The excellent study by **Requena-Mora et al.** (2016) on the origin and evolution of discussion groups as a research technique in critical qualitative sociology can serve as a valuable example to follow.

7. References

Alcalde-Villacampa, Javier (2003). "Medios de comunicación como intermediarios del voto: la influencia de *TVE1* y *Antena 3* en las elecciones legislativas de 1993". *Revista española de investigaciones sociológicas*, v. 104, n. 3, pp. 145-178. <https://doi.org/10.2307/40184572>

Ardèvol-Abreu, Alberto (2015). "Framing o teoría del encuadre en comunicación: Orígenes, desarrollo y panorama actual en España". *Revista latina de comunicación social*, v. 70, pp. 423-450. <https://doi.org/10.4185/RLCS-2015-1053>

Arias-Maldonado, Manuel (2016). *La democracia sentimental: Política y emociones en el siglo XXI*. Barcelona: Página Indómita. ISBN: 978 84 94481659

Arriaza-Ibarra, Karen (2013). "The situation of national and regional public television in Spain". *Nordicom review*, v. 34, n. 1, pp. 145-156. <https://doi.org/10.2478/nor-2013-0048>

Azurmendi, Ana; Muñoz-Saldaña, Mercedes; Ortega-Mohedano, Félix (2018). "Civic participation and independence from political power: two requirements of public service broadcasters in Spain". *Communication & society*, v. 31, n. 3, pp. 111-140. <https://doi.org/10.15581/003.31.3.111-138>

Barnidge, Matthew; Peacock, Cynthia (2019). "A third wave of selective exposure research? The challenges posed by hyperpartisan news on social media". *Media and communication*, v. 7, n. 3, pp. 4-7. <https://doi.org/10.17645/mac.v7i3.2257>

Bos, Linda; Kruikemeier, Sanne; De-Vreese, Claes (2016). "Nation binding: how public service broadcasting mitigates political selective exposure". *PloS one*, v. 11, n. 5, e0155112. <https://doi.org/10.1371/journal.pone.0155112>

Brüggemann, Michael; Engesser, Sven; Büchel, Florin; Humprecht, Edda; Castro, Laia (2014). "Hallin and Mancini revisited: Four empirical types of western media systems". *Journal of communication*, v. 64, n. 6, pp. 1037-1065. <https://doi.org/10.1111/jcom.2014.64.issue-6>

Bryant, Jennings; Miron, Dorina (2004). "Theory and research in mass communication". *Journal of communication*, v. 54, n. 4, pp. 662-704. <https://doi.org/10.1111/j.1460-2466.2004.tb02650.x>

- Camaj, Lindita** (2019). "From selective exposure to selective information processing: A motivated reasoning approach. *Media and communication*, v. 7, n. 3, pp. 8-11.
<https://doi.org/10.17645/mac.v7i3.2289>
- Canel, María-José** (1998). "Los efectos de las campañas electorales". *Communication & society*, v. 11, n. 1, pp. 47-67.
<https://doi.org/10.15581/003.11.36422>
- Cantijoch, Marta; Jorba, Laia; San-Martin, Josep** (2008). "Exposure to political information in new and old media: Which impact on political participation?". In: *Annual meeting of the American political science association*, pp. 28-31.
- Cardenal, Ana S.; Aguilar-Paredes, Carlos; Cristancho, Camilo; Majó-Vázquez, Silvia** (2019a). "Echo-chambers in online news consumption: Evidence from survey and navigation data in Spain". *European journal of communication*, v. 34, n. 4, pp. 360-376.
<https://doi.org/10.1177/0267323119844409>
- Cardenal, Ana S.; Aguilar-Paredes, Carlos; Galais, Carol; Pérez-Montoro, Mario** (2019b). "Digital technologies and selective exposure: How choice and filter bubbles shape news media exposure". *The international journal of press/politics*, v. 24, n. 4, pp. 465-486.
<https://doi.org/10.1177/1940161219862988>
- Castro-Herrero, Laia; Nir, Lilach; Skovsgaard, Morten** (2018). "Bridging gaps in cross-cutting media exposure: The role of public service broadcasting". *Political communication*, v. 35, n. 4, pp. 542-565.
<https://doi.org/10.1080/10584609.2018.1476424>
- Chadwick, Andrew** (2013). *The hybrid media system: Politics and power*. Oxford: Oxford University Press. ISBN: 978 0190696733
<https://doi.org/10.1093/acprof:oso/9780199759477.001.0001>
- Clay, Russ; Barber, Jessica M.; Shook, Natalie J.** (2013). "Techniques for measuring selective exposure: A critical review". *Communication methods and measures*, v. 7, n. 3-4, pp. 147-171.
<https://doi.org/10.1080/19312458.2013.813925>
- Cooper, Eunice; Jahoda, Marie** (1947). "The evasion of propaganda: How prejudiced people respond to anti-prejudice propaganda". *The Journal of psychology: Interdisciplinary and applied*, v. 23, n. 1, pp.15-25.
<https://doi.org/10.1080/00223980.1947.9917316>
- D'Alessio, Dave; Allen, Mike** (2002). "Selective exposure and dissonance after decisions". *Psychological reports* v. 91, n. 2, pp. 527-532.
<https://doi.org/10.2466/pr0.2002.91.2.527>
- Dader, José-Luis** (1983). *Periodismo y pseudocomunicación política*. Pamplona. Eunsa. ISBN: 978 84 31307974
- Dahlgren, Peter M.** (2019). "Selective exposure to public service news over thirty years: The role of ideological leaning, party support, and political interest". *The international journal of press/politics*, v. 24, n. 3, pp. 293-314.
<https://doi.org/10.1177/1940161219836223>
- Díez-Nicolás, Juan; Semetko, Holli A.** (1995). "La televisión y las elecciones de 1993". En: Muñoz-Alonso, Alejandro; Rospir, Juan-Ignacio. *Comunicación política*. Madrid: Universitas, pp. 243-304. ISBN: 978 84 79910396
https://juandieznicolas.es/phocadownload/1_ARTICULOS/art._1995-03_La_Television_y_las_Elecciones_de_1993.pdf
- Díez-Nicolás, Juan; Semetko, Holli A.** (1999). "Los programas de noticias de televisión y las campañas electorales de 1993 y 1996: propiedad, contenido e influencia". En: Muñoz-Alonso, Alejandro; Rospir, Juan-Ignacio. *Democracia mediática y campañas electorales*. Barcelona: Ariel, pp. 151-199. ISBN: 978 84 34412774
- Donsbach, Wolfgang** (1991). "Exposure to political content in newspapers: The impact of cognitive dissonance on reader's selectivity". *European journal of communication*, v. 6, n. 2, pp. 155-186.
<https://doi.org/10.1177/0267323191006002003>
- Doval-Avendaño, Montserrat; Badía-Lago, Blanca** (2009). "Exposición selectiva de noticias en Internet: servicios personalizados y sitios de promoción de noticias". En: Pérez-Herrero, Pilar; Rivas-Nieto, Pedro; Gelado-Marcos, Roberto. *Estudios de periodística XIV. Periodismo ciudadano: posibilidades y riesgos para el discurso informativo*. Salamanca: Universidad Pontificia de Salamanca, pp. 281-286. ISBN: 978 84 72998421
- Feldman, Lauren; Stroud, Natalie-Jomini; Bimber, Bruce; Wojcieszak, Magdalena** (2013). "Assessing selective exposure in experiments: The implications of different methodological choices". *Communication methods and measures*, v. 7, n. 3-4, pp. 172-194.
<https://doi.org/10.1080/19312458.2013.813923>
- Fernández-Quijada, David; Masip, Pere** (2013). "Three decades of spanish communication research: Towards legal age. [Tres décadas de investigación española en comunicación: hacia la mayoría de edad]". *Comunicar*, v. 41, n. 2, pp. 15-34.
<https://doi.org/10.3916/C41-2013-01>

- Festinger, Leon** (1957). *A theory of cognitive dissonance*. Stanford University Press. ISBN: 978 1503620766
<https://doi.org/10.1515/9781503620766>
- Fischer, Peter; Schulz-Hardt, Stefan; Frey, Dieter** (2008). "Selective exposure and information quantity: How different information quantities moderate decision makers' preference for consistent and inconsistent information". *Journal of personality and social psychology*, v. 94, n. 2, pp. 231-244.
<https://doi.org/10.1037/0022-3514.94.2.94.2.231>
- Fletcher, Richard; Nielsen, Rasmus-Kleis** (2018). "Are people incidentally exposed to news on social media? A comparative analysis". *New media & society*, v. 20, n. 7, pp. 2450-2468.
<https://doi.org/10.1177/1461444817724170>
- Fraile, Marta; Meilán, Xabier** (2012). "Los medios de comunicación y la información política en las elecciones europeas de 2009". En: Font, Joan; Torcal, Mariano. *Las elecciones europeas de 2009*. Madrid: CIS, pp. 109-138. ISBN: 978 84 74765946
- Garrett, R. Kelly** (2009). "Politically motivated reinforcement seeking: Reframing the selective exposure debate". *Journal of communication*, v. 59, n. 4, pp. 676-699.
<https://doi.org/10.1111/j.1460-2466.2009.01452.x>
- Garrett, R. Kelly** (2013). "Selective exposure: New methods and new directions". *Communication methods and measures*, v. 7, n. 3-4, pp. 247-256.
<https://doi.org/10.1080/19312458.2013.835796>
- Gentzkow, Matthew; Shapiro, Jesse M.** (2011). "Ideological segregation online and offline". *Quarterly journal of economics*, v. 126, n. 4, pp. 1799-1839.
<https://doi.org/10.3386/w15916>
- Guimerà, Josep-Àngel** (2014). *Les polítiques de mitjans de comunicació durant els governs de Jordi Pujol. Premsa, ràdio i televisió en el procés de construcció nacional de Catalunya*. Barcelona: Proa. ISBN: 978 84 75884974
- Guimerà, Josep-Àngel** (2018). "El papel de la comunicación en la construcción nacional de Cataluña: Jordi Pujol y la instrumentalización política de los medios (1968-1989)". *Historia y política*, n. 40, pp. 363-387.
<https://doi.org/10.18042/hp.40.13>
- Gunther, Richard; Montero, José-Ramón; Wert, José-Ignacio** (2000): "The media and politics in Spain: From dictatorship to democracy". En: Gunther, Richard; Mughan, Anthony. *Democracy and the media: A comparative perspective*. Cambridge: Cambridge University Press, pp. 28-84. ISBN: 978 0521777438
<https://doi.org/10.1017/CBO9781139175289.002>
- Hallin, Daniel C.; Mancini, Paolo** (2004). *Comparing media systems: Three models of media and politics*. Cambridge: Cambridge University Press. ISBN: 978 0 521543088
<https://doi.org/10.1017/CBO9780511790867>
- Hart, William; Albarracín, Dolores; Eagly, Alice H.; Brechan, Inge; Lindberg, Matthew J.; Merrill, Lisa** (2009). "Feeling validated versus being correct: A meta-analysis of selective exposure to information". *Psychological bulletin*, v. 135, n. 4, pp. 555-588.
<https://doi.org/10.1037/a0015701>
- Hierro-Hernández, María-José** (2012). *Changes in national identification: the case of Catalonia*. Tesis doctoral. Universidad Autónoma de Madrid.
https://repositorio.uam.es/bitstream/handle/10486/10285/52319_Hierro_Maria_Jose.pdf?sequence=1
- Holtz-Bacha, Cristina** (2015). "The role of public service media in nation-building". En: Arriaza-Ibarra, Karen; Nowak, Eva; Kuhn, Raymond. *Public service media in Europe: A comparative approach*. Abingdon, UK: Routledge, pp. 27-40. ISBN: 978 1 138020689
- Huddy, Leoni** (2013). "From group identity to political cohesion and commitment". En: Huddy, Leoni; Sears, David O.; Levy, Jack S. (eds.). *The Oxford handbook of political psychology*. Oxford: Oxford University Press, pp. 737-773.
<https://doi.org/10.1093/oxfordhb/9780199760107.013.0023>
- Humanes, María-Luisa** (2014). "Exposición selectiva y partidismo de las audiencias en España. El consumo de información política durante las campañas electorales de 2008 y 2011". *Palabra clave*, v. 17, n. 3, pp. 773-802.
<https://doi.org/10.5294/pacla.2014.17.3.9>
- Humanes, María-Luisa** (2016). "Exposición selectiva, partidismo y polarización de las audiencias de los medios en España". En: Casero-Ripollés, Andreu. *Periodismo y democracia en el entorno digital*. Madrid: SEP, pp. 37-52. ISBN: 978 84 60669456
<https://dialnet.unirioja.es/servlet/articulo?codigo=6411744>

- Humanes, María-Luisa** (2019). "Selective exposure in a changing political and media environment". *Media and communication*, v. 7, n. 3.
<https://doi.org/10.17645/mac.v7i3.2351>
- Humanes, María-Luisa; Martínez-Nicolás, Manuel; Saperas-Lapiedra, Enric** (2013). "Political journalism in Spain. Practices, roles and attitudes". *Estudios sobre el mensaje periodístico*, v. 19, n. 2, pp. 715-731.
https://doi.org/10.5209/rev_ESMP.2013.v19.n2.43467
- Humanes, María-Luisa; Mellado, Claudia** (2017). "Modelos explicativos de la exposición selectiva a la información política y partidismo de las audiencias en España". En: *IAMCR Conference Transforming culture, politics & communication: New media, new territories, new discourses*. Cartagena de Indias.
- Humanes, María-Luisa; Roses, Sergio** (2018). "Journalistic role performance in the Spanish national press". *International journal of communication*, v. 12, pp. 1032-1053.
<https://ijoc.org/index.php/ijoc/article/view/8109>
- Humanes, María-Luisa; Valera-Ordaz, Lidia** (2023). "Partisanship, ideology, and selective exposure: A longitudinal analysis of media consumption in Spain (2008-2019)". *Media and communication*, v. 11, n. 2, pp. 113-126.
<https://doi.org/10.17645/mac.v11i2.6280>
- Iyengar, Shanto; Hahn, Kyu S.** (2009). "Red media, blue media: Evidence of ideological selectivity in media use". *Journal of communication*, v. 59, n. 1, pp. 19-39.
<https://doi.org/10.1111/j.1460-2466.2008.01402.x>
- Jones, Daniel E.** (2007). "Pujol y la construcción de un espacio catalán de comunicación: interacciones entre instituciones políticas y empresas mediáticas (1980-2003)". *Revista ámbitos*, v. 16, pp. 499-524.
<https://www.redalyc.org/pdf/168/16801627.pdf>
- Kahneman, Daniel; Slovic, Paul; Tversky, Amos** (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press. ISBN: 978 0 521284141
<https://doi.org/10.1017/CBO9780511809477>
- Klapper, Joseph T.** (1960). *The effects of mass communications*. Glencoe: The Free Press. ISBN: 978 0 029173800
- Knobloch-Westerwick, Silvia** (2012). "Selective exposure and reinforcement of attitudes and partisanship before a presidential election". *Journal of communication*, v. 62, n. 4, pp. 628-642.
<https://doi.org/10.1111/j.1460-2466.2012.01651.x>
- Knobloch-Westerwick, Silvia** (2015a). *Choice and preference in media use. Advances in selective exposure theory and research*. New York, NY: Routledge. ISBN: 978 0 805855159
<https://doi.org/10.4324/9781315771359>
- Knobloch-Westerwick, Silvia** (2015b). "The selective exposure self- and affect-management (SESAM) model: Applications in the realms of race, politics, and health". *Communication research*, v. 42, n. 7, pp. 959-985.
<https://doi.org/10.1177/0093650214539173>
- Knobloch-Westerwick, Silvia; Hastall, Matthias; Zillmann, Dolf; Callison, Coy** (2003). "Imagery effects on the selective reading of internet newsmagazines". *Communication research*, v. 30, n. 1, pp. 3-29.
<https://doi.org/10.1177/0093650202239023>
- Knobloch-Westerwick, Silvia; Kleinman, Steven B.** (2012). "Pre-election selective exposure: Confirmation bias versus informational utility". *Communication research*, v. 39, n. 2, pp. 170-193.
<https://doi.org/10.1177/0093650211400597>
- Knobloch-Westerwick, Silvia; Meng, Jingbo** (2011). "Reinforcement of the political self through selective exposure to political messages". *Journal of communication*, v. 61, n. 2, pp. 349-368.
<https://doi.org/10.1111/j.1460-2466.2011.01543.x>
- Kruglanski, Arie W.** (1989). *Lay epistemics and human knowledge: Cognitive and motivational Bases*. New York: Springer. ISBN: 978 0306430787
<https://doi.org/10.1007/978-1-4899-0924-4>
- Ksiazek, Thomas B.; Malthouse, Eduard C.; Webster, James G.** (2010). "News-seekers and avoiders: Exploring patterns of total news consumption across media and the relationship to civic participation". *Journal of broadcasting & electronic media*, v. 54, n. 4, pp. 551-568.
<https://doi.org/10.1080/08838151.2010.519808>
- Kunda, Ziva** (1990). "The case for motivated reasoning". *Psychological bulletin*, v. 108, n. 3, pp. 480-498.
<https://doi.org/10.1037/0033-2909.108.3.480>

- Lacasa-Mas, Iván; Jandura, Olaf; Cano-Castells, Francesc** (2014). "Exposición fragmentada a la información periodística, polarización política y democracia representativa: el caso de los diputados del Parlamento catalán y su uso de los medios informativos". *Estudios sobre el mensaje periodístico*, v. 20, n. 1, pp. 431-450.
https://doi.org/10.5209/rev_ESMP.2014.v20.n1.45241
- Lazarsfeld, Paul F.; Berelson, Bernard; Gaudet, Hazel** (1948). *The people's choice: How the voter makes up his mind in a presidential campaign* (2nd edition). New York: Duell, Sloan and Pearce. ISBN: 978 0 231197953
- Levendusky, Matthew** (2013). "Partisan media exposure and attitudes towards the opposition". *Political communication*, v. 30, n. 4, pp. 565-581.
<https://doi.org/10.1080/10584609.2012.737435>
- Martín-Llaguno, Marta; Berganza-Conde, M. Rosa** (2001). "Votantes y medios de comunicación en las elecciones nacionales españolas de 1996: ¿Exposición selectiva o influencia mediática?". *Communication & society*, v. 14, n. 1, pp. 51-70.
<https://doi.org/10.15581/003.14.36382>
- Martínez-Amat, Marc** (2020). "One country, two media systems: The evolution of the media audience in Catalonia in the context of the independence debate". *Comunicació: revista de recerca i d'anàlisi*, v. 37, n. 1, pp. 53-73.
<https://doi.org/10.2436/20.3008.01.190>
- Masip, Pere; Suau, Jaume; Ruiz-Caballero, Carlos** (2018). "Questioning the selective exposure to news: Understanding the impact of social networks on political news consumption". *American behavioral scientist*, v. 62, n. 3, pp. 300-319.
<https://doi.org/10.1177/0002764217708586>
- Masip, Pere; Suau, Jaume; Ruiz-Caballero, Carlos** (2020a). "Incidental exposure to non-like-minded news through social media: Opposing voices in echo-chambers' news feeds". *Media and communication*, v. 8, n. 4, pp. 53-62.
<https://doi.org/10.17645/mac.v8i4.3146>
- Masip, Pere; Suau, Jaume; Ruiz-Caballero, Carlos** (2020b). "Percepciones sobre medios de comunicación y desinformación: ideología y polarización en el sistema mediático español". *Profesional de la información*, v. 29, n. 5, e290527.
<https://doi.org/10.3145/epi.2020.sep.27>
- Matthes, Jörg; Schmuck, Desirée** (2017). "The effects of anti-immigrant right-wing populist ads on implicit and explicit attitudes: A moderated mediation model". *Communication research*, v. 44, n. 4, pp. 556-581.
<https://doi.org/10.1177/0093650215577859>
- Meilán, Xabier** (2010). *Causas y consecuencias del consumo de información política en España (2000-2009)*. Madrid: CIS. ISBN: 978 84 74765212
- Mills, John; Aronson, Elliot; Robinson, Harold** (1959). "Selectivity in exposure to information". *Journal of abnormal and social psychology*, v. 59, n. 2, pp. 250-253.
<https://doi.org/10.1037/h0042162>
- Moreno-Moreno, Elsa; Sanjurjo-San-Martín, Elena-Luisa** (2020). "Consumo selectivo de las marcas periodísticas en España: confianza, orientación política y grado de consumo digital". *Profesional de la información*, v. 29, n. 5, e270524.
<https://doi.org/10.3145/epi.2020.sep.24>
- Muñoz-López, Blanca** (1993). "El Estado actual de la sociología de la cultura y de la comunicación de masas en los Estados Unidos". *Ensayos: revista de la Facultad de Educación de Albacete*, n. 8, pp. 95-109.
<https://dialnet.unirioja.es/servlet/articulo?codigo=2282505>
- Mutz, Diana C.** (2002). "The consequences of cross-cutting networks for political participation". *American journal of political science*, v. 46, n. 4, pp. 838-855.
<https://doi.org/10.2307/3088437>
- Mutz, Diana C.; Young, Lori** (2011). "Communication and public opinion: Plus ça change?". *Public opinion quarterly*, v. 75, n. 5, pp. 1018-1044.
<https://doi.org/10.1093/poq/nfr052>
- Nelson, Jacob L.; Webster, James G.** (2017). "The myth of partisan selective exposure: A portrait of the online political news audience". *Social media + society*, v. 3, n. 3.
<https://doi.org/10.1177/2056305117729314>
- Postman, Leo; Bruner, Jerome S.; McGinnies, Elliott** (1948). "Personal values as selective factors in perception". *Journal of abnormal and social psychology*, v. 43, n. 2, pp. 142-154.
<https://doi.org/10.1037/h0059765>
- Precht, Richard-David** (2010). *Die kunst kein egoist zu sein: Warum wir gerne gut sein wollen und was uns davon abhält*. München: Goldmann Verlag. ISBN: 978 3 442156313

- Ramírez-Dueñas, José-María; Vinuesa-Tejero, María-Lourdes** (2020). "Exposición selectiva y sus efectos en el comportamiento electoral de los ciudadanos: la influencia del consumo mediático en el voto en las elecciones generales españolas de 2015 y 2016". *Palabra clave*, v. 23, n. 4, e2346.
<https://doi.org/10.5294/pacla.2020.23.4.6>
- Ramírez-Dueñas, José-María; Vinuesa-Tejero, María-Lourdes** (2021). "How does selective exposure affect partisan polarization? Media consumption on electoral campaigns". *Journal of international communication*, v. 27, n. 2, pp. 258-282.
<https://doi.org/10.1080/13216597.2021.1899957>
- Requena-Mora, Marina** (coord.); **Conde-Gutiérrez-del-Álamo, Fernando; Alonso-Benito, Luis-Enrique; Rodríguez-Victoriano, José-Manuel; Callejo-Gallego, Manuel-Javier; Martín-Criado, Enrique; Martínez-Sanz, Paula; Serrano-Pascual, Araceli; Betancor-Núñez, Gomer; Barbeta-Viñas, Marc; Prieto-Serrano, David; Pereda-Olarte, Carlos; De-Prada-Junquera, Miguel-Ángel** (2016). "Un grupo sobre el grupo de discusión: Entre la lógica instrumental y el eterno retorno a la sociología crítica". *Encrucijadas: revista crítica de ciencias sociales*, n. 12.
<https://dialnet.unirioja.es/servlet/articulo?codigo=5889889#>
- Roda, Rafael** (1990). "Televisión y cultura política". *Aprendizaje. Revista de psicología social*, v. 5, n. 1, pp. 63-70.
<https://doi.org/10.1080/02134748.1990.10821618>
- Rodríguez-Virgili, Jordi; Sierra, Aurken; Serrano-Puche, Javier** (2022). "Motivations for news exposure in different media systems: A comparative study of Germany, Spain and the United Kingdom". *Communication today*, v. 13, n. 1, pp. 60-72.
https://communicationtoday.sk/wp-content/uploads/05_RODRIGUEZ-VIRGILI-et-al_CT-1-2022.pdf
- Sampedro, Víctor** (2000). *Opinión pública y democracia deliberativa: medios, sondeos y urnas*. Madrid: Istmo. ISBN: 84 70903829
- Sears, David O.; Freedman, Jonathan L.** (1967). "Selective exposure to information: A critical review". *Public opinion quarterly*, v. 3, n. 12, pp. 194-213.
<https://doi.org/10.1086/267513>
- Serrano, Ivan** (2013). "Just a matter of identity? Support for independence in Catalonia". *Regional & federal studies*, v. 23, n. 5, pp. 523-545.
<https://doi.org/10.1080/13597566.2013.775945>
- Sintes-Olivella, Marçal; Fondevila-Gascón, Joan-Francesc; Berbel-Giménez, Gaspar** (2023). "Televisión e independentismo en Cataluña. Exposición selectiva de los televidentes catalanes durante el proceso independentista (2015-2020)". *Revista española de investigaciones sociológicas*, v. 181, pp. 101-116.
<https://doi.org/10.5477/cis/reis.181.101>
- Skovsgaard, Morten; Shehata, Adam; Strömbäck, Jesper** (2016). "Opportunity structures for selective exposure: Investigating selective exposure and learning in Swedish election campaigns using panel survey data". *The international journal of press/politics*, v. 21, n. 4, pp. 527-546.
<https://doi.org/10.1177/1940161216658157>
- Steppat, Desiree; Castro-Herrero, Laia; Esser, Frank** (2022). "Selective exposure in different political information environments: How media fragmentation and polarization shape congruent news use". *European journal of communication*, v. 37, n. 1, pp. 82-102.
<https://doi.org/10.1177/02673231211012141>
- Strömbäck, Jesper; Djerf-Pierre, Monika; Shehata, Adam** (2013). "The dynamics of political interest and news media consumption: A longitudinal perspective". *International journal of public opinion research*, v. 25, n. 4, pp. 414-435.
<https://doi.org/10.1093/ijpor/eds018>
- Stroud, Natalie-Jomini** (2008). "Media use and political predispositions: Revisiting the concept of selective exposure". *Political behaviour*, v. 30, n. 3, pp. 341-366.
<https://doi.org/10.1007/s11109-007-9050-9>
- Stroud, Natalie-Jomini** (2010). "Polarization and partisan selective exposure". *Journal of communication*, v. 60, n. 3, pp. 556-576.
<https://doi.org/10.1111/j.1460-2466.2010.01497.x>
- Stroud, Natalie-Jomini** (2011). *Niche news: The politics of news choice*. Oxford: Oxford University Press. ISBN: 978 0 199755516
<https://doi.org/10.1093/acprof:oso/9780199755509.001.0001>
- Stroud, Natalie-Jomini** (2018). "Selective exposure theories". In: Kenski, Kate; Hall-Jamieson, Kathleen (eds.). *The Oxford handbook of political communication*. Oxford: Oxford University Press, pp. 531-548. ISBN: 978 0 190090456
- Sunstein, Cass R.** (2003). *Republica.com. Internet, democracia y libertad*. Barcelona: Paidós. ISBN: 84 49313848
- Valera-Ordaz, Lidia** (2018). "Medios, identidad nacional y exposición selectiva: predictores de preferencias mediáticas de los catalanes". *Revista española de investigaciones sociológicas*, v. 164, pp. 135-165.
<https://doi.org/10.5477/cis/reis.164.135>

Valera-Ordaz, Lidia (2022). "Political identity and news media choice: The polarizing logic of selective exposure during the Catalan independence conflict". *Mass communication and society*, v. 26, n. 2, pp. 326-352.

<https://doi.org/10.1080/15205436.2022.2127366>

Valera-Ordaz, Lidia; Humanes, María-Luisa (2022). "What drives selective exposure to political information in Spain? Comparing political interest and ideology". In: Palau-Sampió, Dolors; López-García, Guillermo; Ianelli, Laura. *Contemporary politics, communication, and the impact on democracy*. IGI Global, pp. 93-112. ISBN: 978 1 799880578

<https://doi.org/10.4018/978-1-7998-8057-8.ch006>

Valera-Ordaz, Lidia; Requena-Mora, Marina; Pop, Ioana-Anastasia (2021). "Ideología y exposición selectiva: El consumo mediático de los españoles durante la campaña de abril de 2019". En: López-García, Guillermo; Campos-Domínguez, Eva. *Redes en campaña: liderazgos y mensajes en las elecciones de 2019*. Barcelona: Comunicación Social, pp. 157-174. ISBN: 978 84 17600617

Vara-Miguel, Alfonso; Negro, Samuel; Moreno-Moreno, Elsa; Kaufmann, Jürg (2022). *Digital news report 2022*. Reuters Institute for the Study of Journalism.

<https://reutersinstitute.politics.ox.ac.uk/es/digital-news-report/2022/espana>

Victoria-Mas, María; Lacasa-Mas, Ivan; Fernández-Planells, Ariadna; Justel-Vázquez, Santiago (2023). "News media brand's value within polarized markets: Perceived quality vs. Political congruence". *Catalan journal of communication and cultural studies*, v. 15, n. 1, pp. 3-25.

https://doi.org/10.1386/cjcs_00076_1

Zain, Ana-Laura (2014). "La medición de la audiencia digital: comScore". *Indice: revista de estadística y sociedad*, v. 61, pp. 32-34.

<http://www.revistaindice.com/numero61/p32.pdf>

Zoizner, Alon; Sheaffer, Tamir; Castro, Laia; Aalbert, Toril; Cardenal, Ana S.; Corbu, Nicoleta; De-Vreese, Claes; Esser, Frank; Hopman, David N.; Koc-Michalska, Karolina; Matthes, Jörg; Schemer, Christian; Spendore, Sergio; Stanyer, James; Stępińska, Agnieszka; Štětka, Václav; Strömbäck, Jesper; Theocharis, Yannis; Van-Aelst, Peter (2022). "The effects of the Covid-19 outbreak on selective exposure: Evidence from 17 countries". *Political communication*, v. 39, n. 5, pp. 674-696.

<https://doi.org/10.1080/10584609.2022.2107745>

Rediseñando lo que somos
para renovar lo que hacemos

<https://www.sedic.es>

<https://intranetsedic.es>

<https://formacionsedic.online>

c/Gargantilla 13, local 24

Madrid 28005

+34 639 186 570 | +34 91 593 40 59

sedic@sedic.es



<https://twitter.com/SEDIC20>



<https://www.facebook.com/AsociacionSEDIC>



<https://www.linkedin.com/company/sedic/>



<https://www.instagram.com/sedicasociacion/>

Informative reels on *Instagram* in the face of new ways of reporting to attract new audiences

María-Isabel Rodríguez-Fidalgo; Adriana Paíno-Ambrosio; Francisco-Javier Herrero-Gutiérrez

Nota: Este artículo se puede leer en español en:
<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87218>

Recommended citation:

Rodríguez-Fidalgo, María-Isabel; Paíno-Ambrosio, Adriana; Herrero-Gutiérrez, Francisco-Javier (2023). "Informative reels on *Instagram* in the face of new ways of reporting to attract new audiences". *Profesional de la información*, v. 32, n. 5, e320519.
<https://doi.org/10.3145/epi.2023.sep.19>

Article received on December 17th 2022
Approved on July 24th 2023



María-Isabel Rodríguez-Fidalgo

<https://orcid.org/0000-0001-5167-0272>

Universidad de Salamanca
Facultad de Ciencias Sociales
Depto. de Sociología y Comunicación
Pº Francisco Tomás y Valiente, s/n
37007 Salamanca, Spain
mrfidalgo@usal.es



Adriana Paíno-Ambrosio

<https://orcid.org/0000-0003-3817-6071>

Universidad de Salamanca
Facultad de Ciencias Sociales
Depto. de Sociología y Comunicación
Pº Francisco Tomás y Valiente, s/n
37007 Salamanca, Spain
adriana.paino@usal.es



Francisco-Javier Herrero-Gutiérrez

<https://orcid.org/0000-0002-1362-7406>

Universidad de Salamanca
Facultad de Ciencias Sociales
Depto. de Sociología y Comunicación
Pº Francisco Tomás y Valiente, s/n
37007 Salamanca, Spain
javiherrero82@usal.es ✉

Abstract

In recent years, the current media ecosystem has been transformed owing to the adaptation to digital media, affecting all genres and audiovisual formats. As a result, from the point of view of participation, the viewer-user has acquired an unprecedented protagonism owing to the possibility now offered by social networks. The informative genre, far from turning its back on this reality, has allied itself with it and is exploring new ways of reaching audiences, and one of the latest trends has to do with informative television channels and the use of *Instagram* reels. This research focuses on the case of *RTVE Noticias* (@rtvenoticias), with the purpose of analyzing the use and communication strategy that they make of *Instagram* reels on this platform, precisely because it is the feature that has achieved more popularity of late. The findings show that *RTVE Noticias* does not currently have a clear strategy when it comes to informing through this medium, which translates into content being reused from other media without really taking advantage of the narrative potential of the network, resulting in low interaction with followers. The main results obtained allow us to continue a rich academic debate related to the media's use of social networks from the perspective of information and consumption, mainly among the youngest users.

Keywords

Reels; *Instagram*; Social media; Social networks; *RTVE*; News; Digital media; Information; Engagement; Participation; Interaction.

Funding

This article is funded within the *University of Salamanca's* own program for Recognized Research Groups. Group *Narrativas Audiovisuales y Estudios Socioculturales (NAES)*.



1. Introduction

For more than two decades, the Internet has become a cohesive force of the media, and is probably one of the best channels for promotional spaces (Larrañaga, 2008, pp. 2-3; Piñeiro, 2015), which has in turn transformed it into a meta-medium, understood as

“a computer-generated online digital medium with properties of traditional media and other new potentialities based on computation, which has autonomy for the creation, editing, reproduction, dissemination, access, and downloading (browsing, searching, reading, listening, viewing, and interaction) of different multimedia content” (Campos-Freire, *et al.*, 2016, p. 451),

a definition arrived at by those authors on the basis of concepts previously used by Jensen (2013) and Campos-Freire (2015a; 2015b).

Social networks, as part of the Internet phenomenon, have been a real alternative to other promotional, advertising, or marketing spaces since their conception. These virtual social networks, which in many ways replace analog ones, are virtual communication spaces that, among other things, can put brands/companies in contact with their users without the need for physical presence.

This purpose or variable of “promotion” has not always been so: for example, when the *Six Degrees* social network was created in the 1990s, its aim was for users to be able to contact each other (person to person), focusing, to a large extent, on the circle of friends or people who they had previously met in person. Later, other social networks, such as *Facebook*, *MySpace*, *Twitter*, *Hi5*, or *Tuenti*, began to emerge with similar objectives, also underpinning the focus of success on the “person.”

Once social networks began to reach the thresholds of maximum success (from 2008 onward), more research began to be carried out on this new reality owing to its diversification and according to variables such as target group, objectives, horizontality/verticality, immediacy, and the short or long term. Thus, people ceased to be the sole “center of attention” of social networks and other possibilities entered the scene: horizontal/vertical, human/content-related, and nomadic/sedentary social networks.

Although most, if not all, social networks have not lost sight of their social character, from a business point of view the emphasis has been on taking advantage of them for promotion, marketing, advertising, direct and indirect sales, and brand image. In general, social networks have been used to increase the popularity of the users themselves, as well as to improve their positioning in virtual markets, since the Internet is still a medium

“governed by commercial purposes and market logic” (Díaz-Soloaga, 2002, p. 2).

In this sense, journalistic companies are governed by the same criteria as any other:

“Companies are aware that social media are becoming prescriptive spaces where users go to stay informed about brands or companies and express their adherence to them or raise their criticisms” (López-de-Ayala-López; Catalina-García; Fernández-Fernández, 2018, p. 72).

We are immersed in a virtual world that is framed

“in an increasingly complex communicative ecosystem” (Campos-Freire *et al.*, 2016, p. 452),

and

“the communication industry has no doubt: the present and the future is in the networked society. As new digital media enter their third decade of evolution, now clad in meta-media, no one wants to live apart from social networks” (Campos-Freire *et al.*, 2016, p. 452).

This is why social networks are a communicative space that journalists have been taking advantage of (Porto, 2011) to get the most out of their information and expand it within a more globalized framework, in addition to generating traffic. There are authors who even claim that traditional media have ceased to be the main focus of news production (Rodero, 2019) in favor of these social networks, something, however, that some journalistic sectors question. In any case, these more conventional media have also taken advantage of social networks to achieve greater impact and more reach (Marcos-García; Alonso-Muñoz; López-Meri, 2021), and have used them for several years already as platforms for distributing information (Pérez-Soler, 2017). The digital context, therefore, currently configures a media landscape that imposes new challenges to information professionals (Canavilhas, 2015) in terms of the use of social networks, no longer only under exclusively business parameters, but as previously indicated, as a journalistic support and source, with the latter being the main context where this research is framed.

1.1. Instagram, social networks, and information

Instagram is a social network that was first available for *iOS* on October 6, 2010 and for *Android* on April 3, 2012 (Caerols; Tapia; Carretero, 2013). It was subsequently purchased by *Facebook* on April 9, 2012 for approximately 1 billion USD (Caerols; Tapia; Carretero, 2013). It can be considered

“In Spain, according to data from *IAB Spain* in its 2023 social network study, *Instagram* is the second most known social network, only behind *Facebook*; and with *YouTube*, it is the third social network in terms of number of users, only behind *WhatsApp* and *Facebook*”

“an application whose main value has been attributed to the capacity of interaction between brand and user through photography. In addition, it allows for the sharing of these snapshots on different social networks [...]” (Caerols; Tapia; Carrero, 2013, p. 70).

“*Instagram*’s principal influence is through images, in either photographic or video format, and not in text, as on other social networks. Among the different functionalities, there is the *reel*, a short video that can be edited to add audio and effects. Reels are shared on the profile of the user who creates them, reaching their followers as well as the public, if the account on which it is shared is public”

In just two months it reached 1 million users (Bilton, 2010), then 5 million by June 2011 (Caerols; Tapia; Carrero, 2013) and 10 million by September 2011 (Protalinski, 2012). In its early years it grew at a rate of 5 million new users per week (Griggs, 2012). Currently, according to 2021 data collected by Mohsin (2022), *Instagram* has 100 million monthly active users and is the second social network after *Facebook*, with more than 70% of users under 35 years old and who spend an average of 53 minutes per day on the app. In addition, it is the social network that has grown the most globally (Carrasco; Sánchez; Trelles, 2020).

In Spain, according to data from *IAB Spain* in its 2023 social network study, *Instagram* is the second most known social network, only behind *Facebook*; and with *YouTube*, it is the third social network in terms of number of users, only behind *WhatsApp* and *Facebook*. The average rating given by users is 7.4 points on a scale of 1-10. *Instagram* is the second network in terms of frequency of use, while in terms of interaction, so far this year, it comes in first. It was also the social network in which the most was invested in terms of advertising in the last 3 years, according to the aforementioned study.

Today, it is difficult to refute that a large part of the information that is consumed is done so through the Internet, and within this considerable framework, through social networks, to the point of reconfiguring the information ecosystem (Pedrero-Esteban; Pérez-Escoda; Establés, 2021), given that social networks monopolize a communicative flow parallel to the formal discourse but reach many more users (Newman *et al.*, 2020). Social networks play a significant role on many levels (Bastos; Mercea, 2018) every day of the year (Brems *et al.*, 2017), in both the personal as well as the business spheres, as previously noted.

If we focus on the journalistic world, social networks are important links between media companies and digital users. Taking into account the main market studies carried out in Spain, there are three social networks with the highest media consumption by number of users: *Facebook*, *Twitter*, and *Instagram*. The media themselves have put the most focus on social networks to reach this type of audience. In fact, if you look at the websites of the six most watched television channels in Spain, they are the ones that most reference and most use social media. To date, no results have been found as to the traffic generated by these media or rankings per se, so we simply start from a probability, a hypothesis on the basis of the social networks that the media themselves advertise on their websites.

It is on these social networks that a hypertextual, multimedia, and interactive language (Ruiz-Paz; Rodríguez-Fidalgo; Paíno-Ambrosio, 2022; Salaverría, 2005), as well as multidirectional (Arjona-Martín, 2012), is combined. Therefore,

“journalistic practices have consequently shifted to the use of such channels and applications” (Longhi-Heredia, 2022, p. 19),

paying special attention to social networks (Herrero-Gutiérrez, 2019; Salaverría, 2019).

Within these social networks, this study puts particular emphasis on *Instagram*, as it is the one that “gives preference to audiovisual image culture” (Longhi-Heredia, 2022, p. 19) over textual or argumentative languages (Longhi-Heredia, 2022), thus constructing a visual journalistic narrative (Mañas; Veloso; Cuesta, 2019).

“This is how *Instagram* has managed to position itself as a reference space in the consumption of fast-content-type news” (Longhi-Heredia, 2022, p. 19).

In terms of news consumption, *Instagram* is the second most used social network by young people for getting online news, and is first if only the 18-24-year-old age group is considered, according to the *Digital News Report 2021*. In terms of motivations, according to the same study,

“*Instagram* is associated with a ‘fun and entertaining way to pass the time’ (26%), incidental news consumption, (18%) and offering ‘perspectives not available in mainstream media’ (16%)” (*Digital News Report*, 2021).

1.2. RTVE and its commitment to social networks

Within this framework, European public service media have undertaken in recent years

“an accelerated process of incorporation into social networks and the development of mobile applications” (Cru-safon; González-Saavedra; Murciano, 2020, p. 50).

RTVE has also been able to use social networks to obtain greater visibility, as mentioned previously. The first major social network that the RTVE joined was *YouTube*, on March 30, 2006, and as of May 2023, it has 2.31 million subscribers and currently around 1.5 billion views.

Three years later, on February 4, 2009, RTVE joined *Facebook*, according to its official and verified page, where it currently has more than 936,000 followers. In March 2009, RTVE joined *Twitter*, where it currently has 1.2 million followers. RTVE also has accounts on social networks such as *TikTok*, *Twitch*, and *LinkedIn*, among others, as well as others that were important in the past but have since disappeared, such as *Tuenti*. In July 2013, RTVE joined *Instagram*, the social network addressed in this article, and since its inception, has published more than 10,000 posts, has more than 436,000 followers, and follows 444 accounts.

Within RTVE, there are many programs and/or sections, apart from journalists, that also use their own social network accounts, and are aimed at more specific audiences, for example, *RTVE Noticias*. *RTVE Noticias* has accounts on networks such as *Facebook* since March 24, 2017, with more than 139,000 followers; *Twitter* since May 2009, with more than 1.5 million followers; *YouTube* since March 18, 2020, with more than 1.63 million subscribers (Figure 1); and *Instagram* since January 2018, with more than 210,000 followers.

Given RTVE's commitment to social networks, as well as many other journalistic companies, it was deemed necessary to carry out a more precise and limited study of the use of *Instagram*, which had not been done until now, although other studies have also used other cases as their object of analysis (Arreguez-Manozzo and Merlo, 2020; Castelló-Martínez, et al., 2023; Rodríguez-Hernández et al., 2022; Román-San-Miguel, Olivares-García, and Jiménez-Zafra, 2022; Sancho-Belichón, 2023).

2. Methodological Approach

2.1. Sample object of study

The sample analyzed consisted of 358 reels, which were collected in two phases. The first phase included all reels published between February 11, 2021, when *RTVE Noticias* (@rtvenoticias) published its first reel, and June 30, 2022. This data collection yielded a sample of 179 reels for further analysis. In the second phase, a sample of another 179 reels published between August 1 and October 31, 2022 was randomly selected. The reason for this second selection lies in *Instagram* announcing a relevant change in July of this year that affected video-type publications: all videos with a duration of less than 90 seconds would be integrated into the "reels" category, with no differentiation between "videos" and "reels" as had been done until then; in fact, the "videos" category no longer exists (Figure 2).

Moreover, this measure has been retroactive, so that all videos published on *Instagram* since *RTVE Noticias*

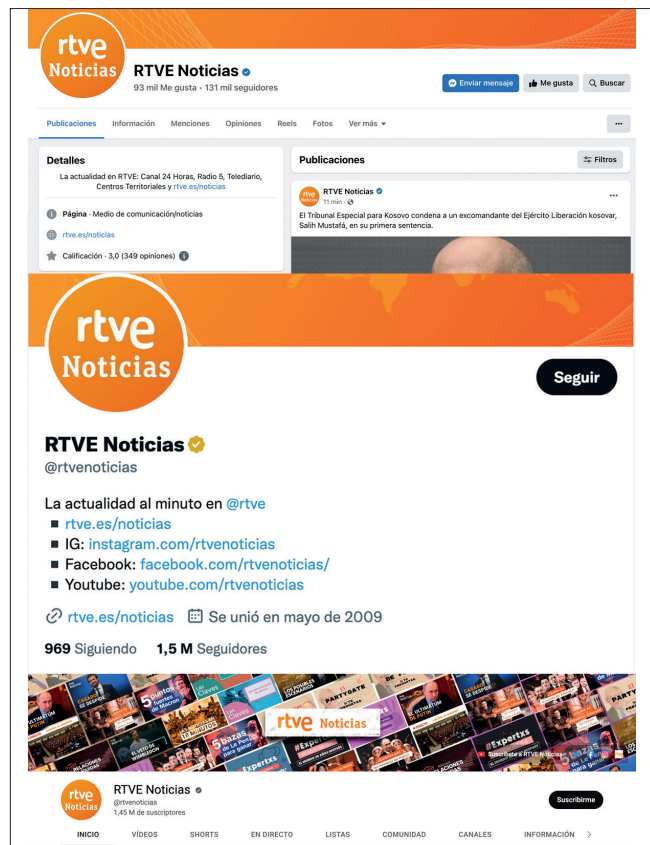


Figure 1. RTVE Noticias Facebook, Twitter, and YouTube profiles. Source: RTVE Noticias.

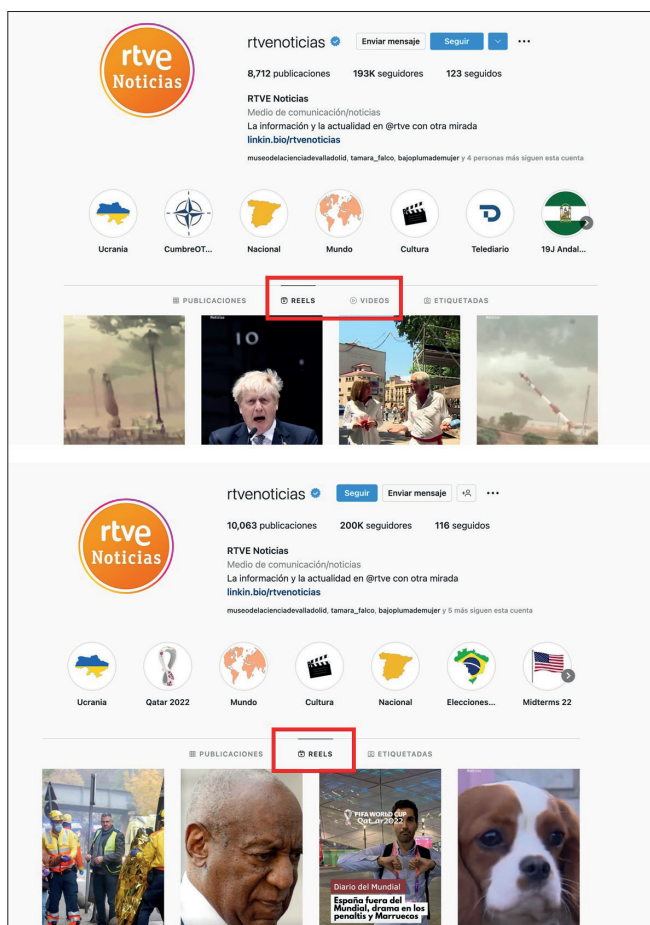


Figure 2. Comparison of the tabs available on *Instagram* before and after the change of cataloging videos as reels. Source: RTVE Noticias' profile on *Instagram* (@rtvenoticias).

opened its account and that met the aforementioned duration requirement now appear in the “reels” category. As a result, the number of reels has increased significantly, so much so that in the 3 months consulted in the second phase, a total of 733 reels were published. For this reason, and to have a volume equivalent to that of the first phase, it was decided to select a sample of 179, which represents 24.4% of the total number of reels published between August and October 2022. It is necessary to specify that the reels corresponding to the month of July 2022 have not been part of the sample, since this was the time when the changes were taking place and the intent was to avoid sampling errors.

“RTVE has also been able to use social networks to obtain greater visibility”

2.2. Research objectives

The general objectives of this research are as follows:

1. Analyze RTVE’s use of reels on its *Instagram* profile (@rtvenoticias).
2. Determine the communication strategy used by RTVE Noticias when publishing reels before and after the policy changed what *Instagram* considers to be reels.

From these general objectives arise the following specific objectives that will be approached from a comparative perspective:

1. Identify the characteristics of the reels, in terms of date and type of publication, topics addressed, and use of resources such as mentions, hashtags, emoticons, geolocation, and sound.
2. Analyze the interaction (views, likes, and comments) generated in the reels published in the periods under study.

2.3. Methods

To carry out this research, we have chosen the methodological technique of the case study (Savin-Baden; Howell, 2022; Yacuzzi, 2005; Yin, 1994). More specifically, this analysis focuses on the *Instagram* profile, @rtvenoticias, of the public broadcaster RTVE.

For this purpose, from the quantitative point of view, a series of ad hoc variables have been designed for this study, scientifically proven in previous works (Bonilla-del-Río; Figuereo-Benítez; García-Prieto, 2022; Paíno-Ambrosio; Rodríguez-Fidalgo; Rodríguez-Monge, 2021; Rodríguez-Monge; Rodríguez-Fidalgo; Paíno-Ambrosio, 2022). A content analysis will thus be carried out because it is a

“method of studying and analyzing communication in a systematic, objective, and quantitative way, with the purpose of measuring certain variables” (Wimmer; Dominick, 1996, p. 70);

the same ideas, in this sense, are defended by Berelson (1952). In addition, content analysis is a technique that can be replicated with some ease because it is considered by many authors to be the most objective, as it is quantifiable (Riffe; Lacy; Watson; Fico, 1998). Thus, Krippendorff defines it as

“a research technique aimed at formulating, from certain data, reproducible and valid inferences that can be applied to its context” (1998, p. 20).

To extract the activity data on *Instagram*, the social network analysis and management tool *Fanpage Karma* was used, and to carry out the aforementioned quantitative content analysis, a file was prepared that included the variables shown in Table 1.

Table 1. Analysis sheet.

Reel identification	Specify exact date
	Day of the week (publication strategy)
	Type of publication (original, TikTok, newscast, or other RTVE program)
Reel content and resources	Topic
	Includes text
	Sound (original audio, music, no sound)
	Includes geolocation
Content and resources of the text accompanying the reel	Includes hashtags
	Includes emoticons
	Includes mentions
Activity generated	Number of views
	Number of likes
	Number of comments
	Engagement rate

Sample coding was performed between July 1 and July 15, 2022 for the first phase of analysis, and between November 1 and November 15, 2022 for the second phase. This was carried out by two researchers. To calculate process reliability, a subsample of 50 reels, corresponding to 14% of the total number of reels under study, was randomly selected and analyzed simultaneously by the two coders to calculate Krippendorff's alpha (Krippendorff, 2011). The results yielded an average of $K\alpha = 0.95$ for the 15 variables analyzed.

The sample analyzed consisted of 358 reels, which were collected in two phases. The first phase included all reels published between February 11, 2021, when RTVE Noticias (@rtvenoticias) published its first reel, and June 30, 2022. This data collection yielded a sample of 179 reels for further analysis

3. Analysis and Results

3.1. Characteristics of the reels that were the object of study

3.1.1. Identification of reels

The analysis of the distribution of reel publication by months shows that, although RTVE Noticias started publishing a high number of reels between February and March 2021, its interest in this type of publications quickly decreased until January 2022, when a significant increase was observed, which remained constant in the following 5 months (Graph 1).

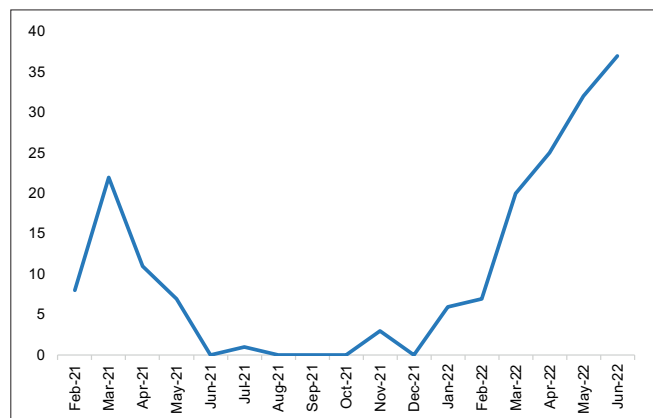
This frequency varies depending on the current news. As will be seen in the analysis of the topics addressed, certain events lead to a greater publication of reels covering that event. This happens most notably with the *Benidorm Fest* (the music contest held in 2022 for the selection of the Spanish song that participates in *Eurovision* that year), with 6 reels being published on consecutive days (from Monday to Saturday); the *Goya Awards* gala, with 5 reels (on two days, Saturday and Sunday); and the war in Ukraine, with reels beginning to appear at the end of February 2022 and appearing continuously during the period under study. Furthermore, in the study of the distribution of publications by days of the week, the results do not identify a clear trend in any "preferred" day for publishing: Monday (13%), Tuesday (12%), Wednesday (11%), Thursday (19%), Friday (12%), Saturday (19%), and Sunday (14%). These very similar percentages indicate that the day of the week is not a criterion used by RTVE Noticias when uploading reels to Instagram.

In the second phase of this study, a significant increase in the number of reels was noted, coinciding with the policy change made by Instagram. The data analyzed show that there is an average publication rate in the months of August–October 2022 of 8.5 reels per day, while in the previous period the average publication rate is 0.3 reels per day. However, RTVE Noticias still does not have a preferred day of publication. This means that in general, this type of publication is more linked to current news, and therefore does not follow the logic of social networks whereby there are days more conducive to obtaining greater engagement.

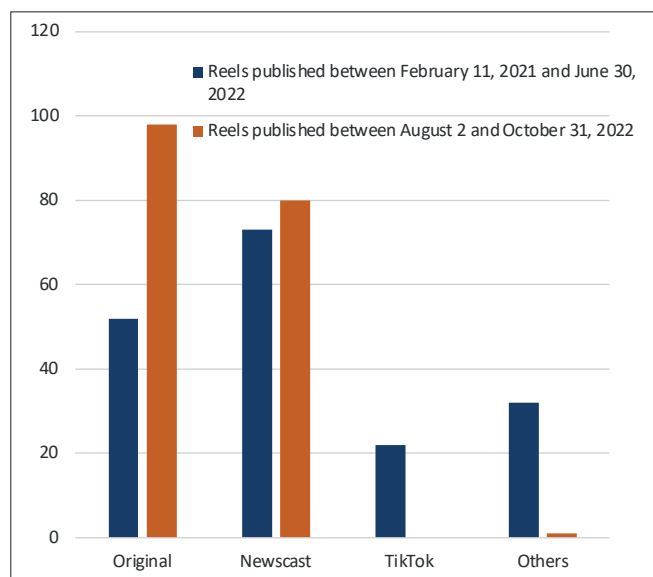
The origin of the 358 reels under study is unevenly distributed, as shown in Graph 2.

When taking into account the two phases of analysis, there is a difference in relation to origin. While in the first period the reels published come from the news, that is, they are small fragments that had previously been broadcast on the news, in the second period analyzed, the largest volume of reels corresponds to an original format that is created specifically for publication on Instagram or another social media platform.

The reels that come from TikTok show that there is a change when publishing this type of content before and after Instagram's implementation of the policy that converted all videos with a duration of less than 90 seconds into reels, such that, as the previous graph shows, they have completely stopped being used. Something similar happens with the reels of another type of origin, catalogued as "other," where a marked



Graph 1. Distribution of reel publication by month in the first period under study.



Graph 2. Reel origin.

decrease is also evident. In this regard, it should be noted that these publications in the first phase of analysis corresponded to content from *Radio 5 Actualidad*, specifically the chronicles sent by Fran Sevilla, published on the web as well as broadcast on the 24-hour channel.

The following two sections address the reels themselves, taking into account on the one hand, the analysis of audiovisual content, and on the other hand, the descriptive text that accompanies it from a double perspective: content and resources used.

3.1.2. Content and resources of published reels

When it comes to the audiovisual content of the reels, the topics identified differ very little between the two periods of analysis (Table 2).

In the first period, there are differences in the content catalogued within the thematic areas of *RTVE Igualdad* and *RTVE Verifica*, and there are no reels on the environment or health; while on the contrary, the opposite is the case in the second period. Furthermore, and taking into account the percentages for each period of analysis, the differences found are more pronounced. The preferred topics from February 11, 2021 to June 30, 2022 were the war in Ukraine, as well as aspects of culture and science and technology; and from August 1 to October 31, 2022 the topics of politics, events, and society were preferred, respectively.

Another of the characteristics of the content related to the topics is their connection with immediate news, an aspect that can be observed in the reels on the war in Ukraine; in fact, the first reel of this type was published on February 25, 2022, one day after the invasion by Russia took place. Ukrainian war reels are short video clips of chronicles that, in most cases, have been previously broadcast. They feature well-known news correspondents (Figure 3), including Fran Sevilla (*Radio 5 Actualidad*) and Almudena Ariza, Óscar Mijallo, Carlos Frangani-lo, and Ebbaba Hameida (*TVE*).

This aspect has also been observed in cultural content, in which the *Goya* gala and *Eurovision* have been identified as protagonists in several reels; and in the case of political content, with the death of Queen Elizabeth II. The same applies to sports, coronavirus, and economics. On the contrary, in the case of science and technology and other culture content of *RTVE Verifica* and *RTVE Igualdad*, mostly reels that do not have to do with current events are found. Among the particularities of science and technology and other culture, there is the specific case of three collections of reels: #AlHilo, #ConCiencia, and #CuandoNoEraViral (figure 4). These are groups of content created by three women who gather news in an informative tone for a young audience accustomed to the consumption of social networks. The reels are characterized by using a medium-short shot speaking directly to the camera (mobile) and recorded in vertical format,

The analysis of the distribution of reel publication by months shows that, although *RTVE Noticias* started publishing a high number of reels between February and March 2021, its interest in this type of publications quickly decreased until January 2022, when a significant increase was observed, which remained constant in the following 5 months

Table 2. Characteristics of the reels that were the object of study.

	Reels published between February 11, 2021 and June 30, 2022		Reels published between August 1, 2022 and October 31, 2022	
	n	%	n	%
Science and technology	21	11.73	6	3.35
Coronavirus	10	5.59	5	2.79
Culture	30	16.76	9	5.03
Sports	10	5.59	7	3.91
Economy	6	3.35	10	5.59
War in Ukraine	65	36.31	21	11.73
Environment	0	0	3	1.68
Politics	9	5.03	46	25.70
Self-promotion	2	1.12	2	1.12
<i>RTVE Igualdad</i>	3	1.68	0	0.00
<i>RTVE Verifica</i>	1	0.59	0	0.00
Health	0	0	10	5.59
Society	8	4.47	24	13.41
Current events	14	7.82	36	20.11
Total	179	100	179	100



Figure 3. Reels of correspondents in the Ukrainian War. Source: Instagram, *RTVE Noticias* (@rtvenoticias).

with the use of emoticons and resources that attract the user’s attention. However, after the sharp decline in the publication of reels in May 2021, this type of content does not appear again.

This section is completed by the analysis of the resources used in the reels, which specifically addresses the use of text, sound, music, geolocation, and hashtags, as presented in Table 3.

Table 3. Frequency and percentage of resources used.

	Reels published between February 11, 2021 and June 30, 2022		Reels published between August 1, 2022 and October 31, 2022		Total	
	n	%	n	%	n	%
Text	66	36.9	94	52.5	160	44.7
Sound (original audio)	173	96.6	76	42.5	249	69.6
Music (<i>Instagram</i> library)	5	2.8	7	3.9	12	3.4
Geolocation	31	17.3	4	2.2	35	9.8
Hashtags	8	4.5	0	0	8	2.2

The incorporation of text in this type of audio-visual publication has been recurrent in both periods under study, and is more common in the second one. These are short texts that act either as small informative headlines or as subtitles. In the first period of analysis, they appear mainly in those related to science and technology, followed by those related to culture and the coronavirus; and in the second period they appear mainly in politics, current events, and the war in Ukraine. Regarding the use of headlines or subtitles, the particularity of a still photo being combined with a headline that appears and disappears when the reel is published has been identified in the second period of analysis. This is due to a new *Instagram* policy that allows for the use of a new “informative reel” format, which previously did not appear catalogued as a reel and relies precisely on the use of this text resource (Figure 5).

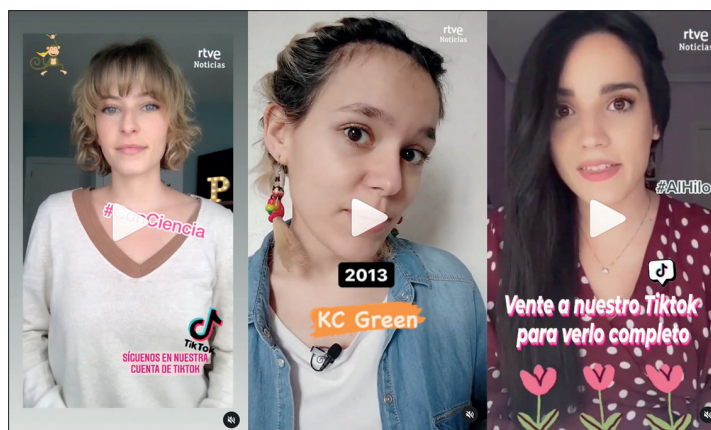


Figure 4. Reels #ConCiencia, #CuandoNoEraViral, and #AlHilo (from left to right). Source: *Instagram*, RTVE Noticias (@rtvenoticias).

Sound is the most widely used element in all the reels published, and is mainly an original type of audio, that is, unique audio created by RTVE Noticias in a video uploaded to its profile or in a reel made through *Instagram*. Among the reels analyzed, there is a high percentage –almost 30%– of content that does not use sound, an aspect that is identified with the new “informative reel” format. Also worth mentioning is the low use of the functionality that allows for the incorporation of music available through the platform’s library in both periods. There was no relationship found between the subject matter and the use of this resource.

The use of the geolocation resource is scarcely identified, although in the first period of analysis it appears to a greater extent, largely linked to the war in Ukraine, where special envoys indicate the location from which the video was uploaded to *Instagram*. A deeper analysis reveals that the geolocalized reels do not have a clear criterion in their usage, so that among the content of the same correspondent, some are geolocalized while others are not.

The use of hashtags is also not common within the reels, since, as will be seen later, their use is more common in the descriptive text that accompanies the video; in fact, they have only been identified in the first period of analysis. What’s worth highlighting here is the connection of this resource with the #AlHilo, #ConCiencia, #CuandoNoEraViral, #DatosRTVE, #VerificaRTVE or #LoDeHoy reel collections, and that they respond to original content from the channel.



Figure 5. Reels with animated text. Source: *Instagram*, RTVE Noticias (@rtvenoticias).

3.1.3. Content and resources of the descriptive text accompanying the reels

If there is one thing that characterizes the content published on *Instagram* besides the use of video, it is undoubtedly the text that accompanies it. This is seen to an even greater extent on the media accounts, since they use this resource to expand information about the news event they are describing. A more detailed analysis has identified the use of three types of resources: mentions, emoticons, and hashtags. Specifically, in the first period of analysis, hashtags were identified as the most used (171 publications), followed by mentions (85) and emoticons (64). These data differ when compared with those identified in the second period where, although the hashtag remains the most used resource (177), emoticons are now used in second place (82), and in last, mentions (27).

Regarding the use of headlines or subtitles, the particularity of a still photo being combined with a headline that appears and disappears when the reel is published has been identified in the second period of analysis. This is due to a new *Instagram* policy that allows for the use of a new “informative reel” format, which previously did not appear catalogued as a reel and relies precisely on the use of this text resource

A deeper analysis of these data shows that these resources appear abundantly within the same reel, which increases significantly in the second part of this analysis, as can be seen in Table 4.

Table 4. Frequency and average of the resources used in the descriptive text.

	Reels published between February 11, 2021 and June 30, 2022		Reels published between August 1, 2022 and October 31, 2022		Total	
	n	Average	n	Average	n	Average
Hashtags	926	5.2	1310	7.3	2236	6.2
Emoticons	114	0.6	144	0.8	258	0.7
Mentions	108	0.6	47	0.3	155	0.4

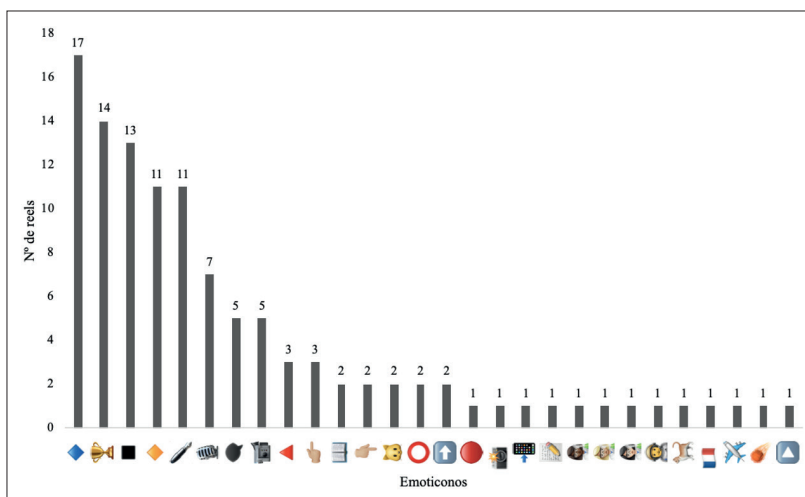
Furthermore, the analysis of the average indicates that, while the use of hashtags and emoticons is comparatively increasing, the use of mentions is decreasing.

A total of 1,094 of the hashtags used are different, while the rest are repeated. The most frequently used are those related to the war in Ukraine, which is explained by the fact that it is the most frequently addressed topic. Table 5 lists the 20 most frequently used hashtags.

Table 5. Most used hashtags.

Hashtag	N. of reels	Hashtag	N. of reels	Hashtag	N. of reels	Hashtag	N. of reels	Hashtag	N. of reels
#ucrania	81	#news	38	#dombás	9	#politica	15	#uk	11
#rusia	78	#ciencia	20	#tecnologia	9	#madrid	13	#zelenski	11
#guerraucrania	72	#europe	18	#guerraucrania	8	#BenidormFest	12	#canal24horas	10
#españa	57	#linkinbio	16	#madrid	8	#reinounido	12	#dombás	10
#conflictosarmados	44	#putin	16	#sabiasque	8	#economia	11	#zaporiyia	10

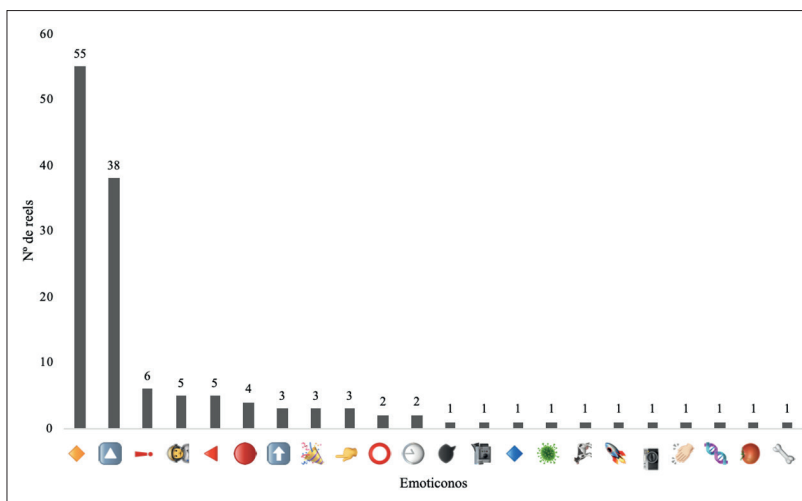
As for the use of emoticons, among the total number identified, 40 are different. Several issues should be mentioned here: firstly, emoticons are frequently used to point out or highlight some relevant information or data, or to establish categories or classifications (squares, diamonds, dots, arrows, pointing fingers). Secondly, emoticons linked to the journalistic profession are also common, indicating who is recording (the video camera), who is speaking, or who is responsible for the information (microphone and pen). Lastly, it should be noted that, although in the first period there is a high number of “trophy” emoticons, these come from the same reel



Graph 3. Most used emoticons in reels published between February 11, 2021 and June 30, 2022.

(relating to Rafael Nadal’s victories), while the rest are distributed throughout the sample.

As can be seen in graphs 3 and 4, there have been some changes between these two periods of analysis. While in the first period a greater variety of emoticons is used, in the second period, some of them are repeated. Furthermore, perhaps the most significant change can be seen in the play button emoticon, very significant within the media context, which is little used in the first period of analysis but used recurrently in the second.



Graph 4. Most used emoticons in reels published between August 1 and October 31, 2022.

Finally, mentions allude both to the channel itself (@rtve) or any of its programs (@aquilatierra, @rtvenoticias, @eurovision, @eltiempovtve, @rtveplay, etc.), as well as to the journalists covering the information and to the protagonists of said mentions (athletes, politicians, singers, actors, etc.). Specifically, 61 different mentions are identified, some of which are repeated more frequently. For example, @rne has 31, which mainly appear in the reels on the war in Ukraine reported by journalist Fran Sevilla. Also noteworthy are the 15 mentions of correspondent Almudena Ariza (@almudena.ariza) and the 10 of journalist Rosa Frasquet (@rosafRASQUETgregori), in charge of covering Benidorm Fest and Mercedes Fashion Week, with the reels in which they appear including mention of their personal profiles.

3.2. Interactivity generated by reels of @rtvenoticias

Among Instagram users, especially those who watch reels, the main interaction of a video publication or reel are views, with likes and comments having much lower figures, and those published by RTVE Noticias are no exception. Some reels do not even have any comments at all during the observation period, as can be seen in Table 6.

Table 6. Interaction received.

	Reels published between February 11, 2021 and June 30, 2022		Reels published between August 1, 2022 and October 31, 2022	
	Total	Average	Total	Average
Views	3,800,000	21,229.1	5,132,056	28,670.7
Likes	110,307	616.2	141,130	788.4
Comments	3,886	21.7	6,835	38.4

The interaction aspect is completed by the calculation of the interaction rate, which indicates the number of social interactions that content receives in relation to the reach or audience figures. This is materialized in the following formula: number of likes + number of comments / total followers in the period under study. Thus, in the first period of analysis, this rate ranges between 0.005 and 14.02, and in the second period between 0.01 and 9.62.

These figures reveal that, in general, all the reels under study have obtained a very low interaction rate; i.e., Instagram users do not interact with these types of posts. Though it is true that, on an individual basis, there are some topics that arouse greater interest among an account’s followers, such as those related to sports, the war in Ukraine, or certain specific events such as the death of Queen Elizabeth II, those that have addressed other topics, such as the economy, or to a large extent, politics, barely generate any interaction.

4. Discussion and Conclusions

The results of this study demonstrate how RTVE has been adapting to new technologies, mainly through the utilization of various platforms and social networks to connect with all types of audiences, from the oldest to those encompassed by Generation Z to the present day. Therefore, RTVE has always been present in this type of digital platform, from the longest lasting, Facebook, to the one most popular in Spain currently, TikTok (Quiroz, 2020).

The analysis of Instagram shows that RTVE Noticias has been modifying the use it makes of this social network in terms of what and how it publishes, aspects that allow for the analysis of their communication strategy through this profile in the period analyzed.

RTVE Noticias’s use of reels, as well as other content it publishes, highlights the use of reels with topics relating

“ The reels under study have obtained a very low interaction rate, i.e., Instagram users do not interact with these types of posts ”

to current events, a fact that is in line with the same strategy traditional media utilizes, where certain issues such as cultural topics are not necessarily linked to immediacy. This is related to the first idea set out in this section, that is, the effort made to ensure that the same reality reaches different audiences, even if this requires the use of all types of audiovisual platforms, and especially taking into account the capacity that young people have to switch between social networks (Marcelino-Mercedes, 2015; Pérez-Escoda; Barón-Dulce; Rubio-Romero; 2021). The use of the content is a different matter, as will be explained later.

RTVE has been adapting to new technologies, mainly through the utilization of various platforms and social networks to connect with all types of audiences, from the oldest to those encompassed by Generation Z to the present day

The results also show that the way in which *RTVE Noticias* publishes, i.e., aspects having to do with the content's format, has changed: while at first it seemed that *RTVE Noticias* opted for a type of reels oriented to young audiences, in a more casual tone, with a more attractive esthetic, and adjusted to the medium (both in the audiovisual and in the text that accompanies it), a drastic change has been observed in which it now prefers the use of videos that come from content previously broadcast in news programs such as newscasts. One might ask why this is the case, as it does not seem entirely logical to be present on different platforms and then use the same content instead of adapting that content to the platform being used (in this case, *Instagram*).

This trend has been accentuated by the outbreak of the war in Ukraine, which for a long period has had virtually all the focus when it comes to reel content, and has become more noticeable in the second period of analysis. These data are especially relevant when compared with the second period of analysis, which coincides with the changes introduced by *Instagram* at the time of publication. As a result, *RTVE Noticias* not only continues to publish in the way it had been previously, but also publishes mostly in a new format that has been identified as "informative reels," which consists of a photograph with text as an informative headline reproduced in a loop. When taking into account the potential that *Instagram* has to tell stories in reel format, these "informative reels" are actually scarce in novelties, due to being "animated still photos" and not audiovisual content that corresponds to a classic narrative. As a result of this measure and given its ease of production, the @rtvenoticias reel section has now significantly increased in amount of content from a quantitative perspective. This increase can be characterized as impressive, because the modifications introduced by *Instagram* are still very recent and represent a fairly abrupt change in how publications are made on social media, given that before the focus was on photo format and now the preference is on video in order to compete with other social networks such as *TikTok*. This aspect is directly related to other research that shows that young people prefer to consume short videos, leading them to jump to platforms such as *TikTok* (Cuesta-Valiño; Gutiérrez-Rodríguez; Durán-Álamo, 2022).

The data analyzed make it clear that *RTVE Noticias* still does not have an explicit strategy when it comes to what it publishes, demonstrated both in the resources used in the videos and in the accompanying text. This lack of strategy can also be seen in the use of emoticons, hashtags, and mentions, clearly seen in examples such as reels using several resources to address the same topic, sometimes written in different ways (i.e., #covid, #covid_19, #COVID19); in the mention or non-mention of the journalists covering the information; the option to geolocate some reels made from the same locations but not all of them; the use of one or other emoticons; or the use of both sound and music.

In addition, there is a deficit in the interaction generated by this type of content, given that it is mainly focused on views and the scarce interaction achieved through likes and comments, with some reels not even having any comments at all. *RTVE Noticias* is concerned about what it posts, i.e., about being present on social media, but this effort does not translate into real interest on the part of its followers, since there is hardly any interaction with the content.

The issues raised thus far are discursively linked to the contributions of authors such as Sobrados-León and Ruiz-Acosta (2021), who allude to the fact that the use of social networks from a journalistic perspective is still in the process of configuration, especially if we take into account their rapid evolution, both technologically and in terms of use.

Another topic of discussion arises linked to the "target audience" variable and the interest that this audience may have in the information provided by *RTVE Noticias*, as well as the question of adapting public television information to the younger generations and the need not only to adapt to the new platforms—as *RTVE* seems to be doing—but also to create new and more attractive content. All of this, of course, should also emphasize the possibilities that social networks have to offer in terms of the narrative point of view and the contributions that information created within these journalistic formats provide.

One limitation of this study is that it is based on a case study, and thus not intended at any time to extrapolate the main findings to other media contexts. The hope is that it will serve to raise other future avenues of research that will contribute to the academic debate within this object of study: the social network *Instagram* from the journalistic perspective.

There is a deficit in the interaction generated by this type of content, given that it is mainly focused on views and the scarce interaction achieved through likes and comments, with some reels not even having any comments at all

5. References

- Arjona-Martín, José-Borja** (2012). "Los nuevos canales audiovisuales basados en web: rtve.es". *Icono 14*, v. 8, n. 1, pp. 98-113. <https://doi.org/10.7195/ri14.v8i1.283>
- Arreguez-Manozzo, Soledad; Merlo, María-Emilia** (2020). "Medios y prácticas periodísticas. Instagram como plataforma periodística. La experiencia de agencia AUNO". *Austral comunicación*, v. 9, n. 1, pp. 146-148. <https://doi.org/10.26422/aucom.2020.0901.arr>
- Bastos, Marco; Mercea, Dan** (2018). "The public accountability of social platforms: lessons from a study on bots and trolls in the Brexit campaign". *Philosophical transactions of the Royal Society A. Mathematical, physical, and engineering sciences*, v. 376, n. 2128. <https://doi.org/10.1098/rsta.2018.0003>
- Berelson, Bernard** (1952). *Content analysis in communication research*. Glencoe (Illinois): The free press. ISBN: 978 0 02 841 210 8
- Bilton, Nick** (2010). "Instagram quickly passes 1 million users". *The New York Times*, 21 December. <https://nyti.ms/3hvkfYa>
- Bonilla-del-Río, Mónica; Figueroa-Benítez, Juan C.; García-Prieto, Victoria** (2022). "Influencers with physical disabilities on instagram: features, visibility and business collaboration". *Profesional de la información*, v. 31, n. 6, e310612. <https://doi.org/10.3145/epi.2022.nov.12>
- Brems, Cara; Temmerman, Martina; Graham, Todd; Broersma, Marcel** (2017). "Personal branding on Twitter: how employed and freelance journalists stage themselves on social media". *Digital journalism*, v. 5, n. 4, pp. 443-459. <https://doi.org/10.1080/21670811.2016.1176534>
- Caerols-Mateo, Raquel; Tapia-Frade, Alejandro; Carretero-Soto, Andrés** (2013). "Instagram, la imagen como soporte de discurso comunicativo participado". *Vivat academia*, n. 124, pp. 68-78. https://www.vivatacademia.net/index.php/vivat/article/view/190/pdf_1
- Campos-Freire, Francisco** (2015a). "Adaptación de los medios tradicionales a la innovación de los metamedios". *Profesional de la información*, v. 24, n. 4, pp. 441-450. <https://doi.org/10.3145/epi.2015.jul.11>
- Campos-Freire, Francisco** (2015b). "Los sitios de redes sociales como paradigma del ecosistema digital". In: Campos-Freire, Francisco; Rúas-Araújo, José (eds.). *Las redes sociales digitales en el ecosistema mediático*. Cuadernos artesanos de comunicación. Sociedad latina de comunicación social, pp. 13-36. <https://doi.org/10.4185/cac92>
- Campos-Freire, Francisco; Rúas-Araújo, José; López-García, Xosé; Martínez-Fernández, Valentín-Alejandro** (2016). "Impacto de las redes sociales en el periodismo". *Profesional de la información*, v. 25, n. 3, pp. 449-457. <https://doi.org/10.3145/epi.2016.may.15>
- Canavilhas, João** (2015). "Nuevos medios, nuevo ecosistema". *Profesional de la información*, v. 24, n. 4, pp. 357-362. <https://doi.org/10.3145/epi.2015.jul.01>
- Carrasco-Polanco, Rafael; Sánchez-de-la-Nieta-Hernández, Miguel-Ángel; Trelles-Villanueva, Alicia** (2020). "Las elecciones al parlamento andaluz de 2018 en Instagram: partidos políticos, periodismo profesional y memes". *Revista mediterránea de comunicación*, v. 11, n. 1, pp. 75-85. <https://www.doi.org/10.14198/MEDCOM2020.11.1.19>
- Castelló-Martínez, Araceli; Segarra-Saavedra, Jesús; Hidalgo-Marí, Tatiana** (2023). "Análisis de la prescripción de marca y de la identificación de su naturaleza publicitaria por parte de rostros televisivos en Instagram". *Fonseca, journal of communication*, n. 26, pp. 165-186. <https://doi.org/10.14201/fjc.29776>
- Crusafon, Carmina; González-Saavedra, Carlos; Murciano, Marcial** (2020). "Las redes sociales y las aplicaciones móviles en las estrategias de transformación digital de los medios de servicio público europeos". *Comunicación: revista de recerca i d'anàlisi*, v. 37, n. 2, pp. 33-54. <https://doi.org/10.2436/20.3008.01.195>
- Cuesta-Valiño, Pedro; Gutiérrez-Rodríguez, Pablo; Durán-Álamo, Patricia** (2022). "¿Por qué la gente vuelve a las plataformas de video? millennials y centennials en TikTok". *Media y comunicación*, v. 10, n. 1, pp. 198-207. <https://doi.org/10.17645/mac.v10i1.4737>
- Díaz-Soloaga, Paloma** (2002). "Construcción de marca en Internet. Aplicación de un modelo interactivo". *Área abierta*, v. 4, n. 3, pp. 1-10. <https://revistas.ucm.es/index.php/ARAB/article/view/ARAB0202330003A/4260>

- Digital News Report España* (2021). *Los servicios de mensajería instantánea se consolidan para leer, ver, encontrar, compartir o comentar noticias*.
<https://bit.ly/3FzyOSC>
- Griggs, Brandon** (2012). "Instagram hits 27 million, says android app coming soon". *CNN*, 11 March.
<https://cnn.it/3WbwloE>
- Herrero-Gutiérrez, Francisco-Javier** (2019). "Las redes sociales como medidoras de audiencias de la prensa deportiva española". *Apunts. Educación, física y deportes*, n. 135, pp. 26-35.
[https://doi.org/10.5672/apunts.2014-0983.es.\(2019/1\).135.02](https://doi.org/10.5672/apunts.2014-0983.es.(2019/1).135.02)
- IAB Spain* (2022). *Estudio de redes sociales 2022*.
<https://bit.ly/3WqcpOe>
- Jensen, Klaus-Bruhn** (2013). "How to do things with data: meta-data, meta-media, and meta-communication". *First Monday*, v. 18, n. 10.
<https://doi.org/10.5210/fm.v18i10.4870>
- Krippendorff, Klaus** (1998). *Content analysis: an introduction to its methodology*. London: Sage. ISBN: 0 761 91044 3
- Krippendorff, Klaus** (2011). "Agreement and Information in the reliability of coding". *Communication methods and measures*, v. 5, n. 2, pp. 93-112.
<https://doi.org/10.1080/19312458.2011.568376>
- Larrañaga-Rubio, Julio** (2008). "Análisis económico de la evolución de las audiencias de los medios tradicionales e impacto de Internet en su consumo". *Área abierta*, n. 21, pp. 1-18.
<https://revistas.ucm.es/index.php/ARAB/article/view/ARAB0808330001A/4127>
- Longhi-Heredia, Sebastián-Alberto** (2022). "Periodismo en Instagram: tratamiento mediático y análisis del humor en las viñetas del diario 20 Minutos". *Textual & visual media*, n. 15, pp. 17-39.
<https://doi.org/10.56418/txt.15.2022.002>
- López, María-Cruz; Catalina, Beatriz; Fernández, José-Gabriel** (2018). "Estrategias de gestión de los clubes de golf de la Comunidad de Madrid en Twitter". *Cuadernos.info*, n. 42, pp. 71-84.
<https://doi.org/10.7764/cdi.42.1304>
- Mañas-Viniegra, Luis; Veloso, Ana-Isabel; Cuesta, Ubaldo** (2019). "Fashion promotion on Instagram with eye tracking: curvy girl influencers versus fashion brands in Spain and Portugal". *Sustainability*, v. 11, n. 14, 3977.
<https://doi.org/10.3390/su11143977>
- Marcelino-Mercedes, Georgina-Victoria** (2015). "Migración de los jóvenes españoles en redes sociales, de Tuenti a Facebook y de Facebook a Instagram. La segunda migración". *Icono 14*, v. 13, n. 2, pp. 48-72.
<https://doi.org/10.7195/ri14.v13i2.821>
- Marcos-García, Silvia; Alonso-Muñoz, Laura; López-Meri, Amparo** (2021). "Periodismo y nuevas narrativas. Storytelling como formato de difusión informativa en redes sociales". *Estudios sobre el mensaje periodístico*, v. 27, n. 2, pp. 553-567.
<https://doi.org/10.5209/esmp.71193>
- Mohsin, Maryam** (2022). "Estadísticas Instagram 2021: 10 datos curiosos de Instagram que debes conocer". *Oberlo*, 12 febrero.
<https://bit.ly/3j4z0Sd>
- Newman, Nic; Fletcher, Richard; Schulz, Anne; Andi, Simge; Nielsen, Rasmus-Kleis** (2020). *Digital news report 2020*. Reuters Institute. University of Oxford.
<https://bit.ly/3FDcQ1b>
- Paíno, Adriana; Rodríguez, María-Isabel; Rodríguez, Montserrat** (2021). "Contenido generado por los usuarios (CGU) en Instagram en torno a la serie Bridgerton (Netflix)". In: Vizcaíno, Arantxa; Bonilla, Mónica; Ibarra, Noelia (coords.). *Cultura participativa, fandom y narrativas emergentes en redes sociales*, pp. 353-378. ISBN: 978 84 1377 588 3
- Pedrero-Esteban, Luis-Miguel; Pérez-Escoda, Ana; Establés, María-José** (2021). "Brexit's impact on Spanish media discourse: news coverage and audience reaction on Twitter". *Profesional de la información*, v. 30, n. 6, e300610.
<https://doi.org/10.3145/epi.2021.nov.10>
- Pérez-Escoda, Ana; Barón-Dulce, Gema; Rubio-Romero, Juana** (2021). "Mapeo del consumo de medios en los jóvenes: redes sociales, 'fakes news' y confianza en tiempos de pandemia". *Index.comunicación*, v. 11, n. 2, pp. 187-208.
<https://doi.org/10.33732/ixc/11/02Mapeod>
- Pérez-Soler, Susana** (2017). *Periodismo y redes sociales: Claves para la gestión de contenidos digitales*. Barcelona, España: Editorial UOC. ISBN: 978 84 9116 986 4

- Piñeiro-Otero, Teresa** (2015). "De las ondas a los 140 caracteres. El uso de Twitter por los principales programas de la radio española". *Palabra clave*, v. 18, n. 3, pp. 815-841.
<https://doi.org/10.5294/pacla.2015.18.3.8>
- Porto-Renó, Denis** (2011). "Periodismo, redes sociales y transmediación". *Razón y palabra*, n. 78.
http://www.razonypalabra.org.mx/varia/N78/1a%20parte/17_Porto_V78.pdf
- Protalinski, Emil** (2012). "Instagram passes 50 million users". *ZDNET*, 1 May.
<https://zd.net/3Wf56ZQ>
- Quiroz, Natalia-Tamara** (2020). "TikTok: La aplicación favorita durante el aislamiento". *Revista argentina de estudios de juventud*, v. 14, e044.
<https://doi.org/10.24215/18524907e044>
- Riffe, Daniel; Lacy, Stephen; Watson, Brendan; Fico, Frederic** (1998). *Analyzing media messages: using quantitative content analysis in research*. Mahwah (New Jersey): Lawrence Erlbaum. ISBN: 978 0 42 946 428 7
- Rodero, Emma** (2009). "Y siempre, por último, con ustedes: la radio. Revisión de la producción bibliográfica y hemerográfica radiofónica en España". *Admira*, n. 1, pp. 98-126.
<https://bit.ly/3PwiO8y>
- Rodríguez-Monge, Montserrat; Rodríguez-Fidalgo, María-Isabel; Paíno-Ambrosio, Adriana** (2022). "Influencers de moda y feministas españolas y sus seguidores en Instagram. ¿Ante nuevas plataformas para fomentar el diálogo social?". *Observatorio (OBS*)*, v. 16, n. 2, pp. 64-91.
<https://doi.org/10.15847/obsOBS16220221908>
- Rodríguez-Fidalgo, María-Isabel; Ruiz-Paz, Yanira; Paíno-Ambrosio, Adriana** (2019). "Análisis de las potencialidades de Facebook como herramienta electoral: estudio de caso de las elecciones presidenciales mexicanas en 2018". *Communication & society*, v. 32, n. 4, pp. 193-205.
<https://doi.org/10.15581/003.32.35657>
- Rodríguez-Hernández, Jonattan; Ortega-Fernández, Eglée-Andreina; Padilla-Castillo, Graciela** (2022). "Medios de comunicación españoles en Twitch. Cambio de paradigma del periodismo hacia el streaming". *Espejo de monografías de comunicación social*, n. 13, pp. 137-155.
<https://doi.org/10.52495/c7.emcs.13.p99>
- Román-San-Miguel, Aránzazu; Olivares-García, Francisco J.; Jiménez-Zafra, Salud-María** (2022). "El discurso de odio en Twitter durante la crisis migratoria de Ceuta en mayo de 2021". *ICONO 14*, v. 20, n. 2.
<https://doi.org/10.7195/ri14.v20i2.1858>
- Salaverría, Ramón** (2005) (coord.). *Cibermedios. El impacto de internet en los medios de comunicación en España*. Sevilla: Comunicación social ediciones y publicaciones. ISBN: 978 849 608 233 5
- Salaverría, Ramón** (2019). "Digital journalism: 25 years of research". Review article. *Profesional de la información*, v. 28, n. 1.
<https://doi.org/10.3145/epi.2019.ene.01>
- Sancho-Belinchón, Celia** (2023). "La adaptación de los medios de comunicación tradicionales españoles al formato de Instagram". *Hipertext.net*, n. 26, pp. 135-148.
<https://doi.org/10.31009/hipertext.net.2023.i26.17>
- Savin-Baden, Maggi; Howell, Claire** (2022). *Qualitative research. The essential guide to theory and practice*. Routledge. ISBN: 978 1 03 245 637 9
- Sobrados-León, Maritza; Ruiz-Acosta, María-José** (2021) (coords.). *Tecnologías que condicionan el periodismo*. Salamanca: Comunicación social. ISBN: 978 841 760 052 5
- Wimmer, Roger D.; Dominick, Joseph R.** (1996). *La investigación científica de los medios de comunicación: una introducción a sus métodos*. Barcelona: Bosch. ISBN: 847 676 359 X
- Yacuzzi, Enrique** (2005). "El estudio de casos como metodología de investigación". *Revista nacional de administración*, v. 1, n. 2, pp. 31-54.
<https://ucema.edu.ar/publicaciones/download/documentos/296.pdf>
- Yin, Robert K.** (2018). *Case study: Research and applications*. SAGE publications. ISBN: 978 1 50 633 616 9

Interactive mapping of Covid-19 disinformation in Ibero-America

María-Jesús Colmenero-Ruiz; Francisco-Carlos Paletta; Audilio Gonzales-Aguilar

Nota: Este artículo se puede leer en español en:

<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87479>

Recommended citation:

Colmenero-Ruiz, María-Jesús; Paletta, Francisco-Carlos; Gonzales-Aguilar, Audilio (2023). "Interactive mapping of Covid-19 disinformation in Ibero-America". *Profesional de la información*, v. 32, n. 5, e320513.

<https://doi.org/10.3145/epi.2023.sep.13>

Manuscript received on 18th August 2023

Accepted on 12th September 2023



María-Jesús Colmenero-Ruiz

<https://orcid.org/0000-0002-1650-1091>

Universidad Complutense de Madrid
Santísima Trinidad, 37
28010 Madrid, Spain
mcolmene@ucm.es



Francisco-Carlos Paletta

<https://orcid.org/0000-0002-4112-5198>

Universidade de São Paulo
Departamento de Informação e Cultura
Cidade Universitária
05508-020 São Paulo, Brazil
fcpaletta@usp.br



Audilio Gonzales-Aguilar ✉

<https://orcid.org/0000-0001-8693-2076>

Université Paul Valéry Montpellier III.
Département de Documentation
Route de Mende
34090 Montpellier, France
audilio.gonzales@gmail.com

Abstract

Fake news, created with different intentions and disseminated through various social networks and messaging platforms, has become a significant means of disinformation in the current landscape. In the face of certain events or controversial situations, the number of fake news items increases, a situation that can become worrying. One such event was the arrival of the Covid-19 pandemic. To combat this phenomenon, various fact-checking bodies appeared with the objective of demonstrating and disseminating whether a certain piece of information was correct. At the international level, these organizations have joined together in the *#CoronaVirusFacts/#DatosCoronaVirus Alliance*. The organization *Chequeado* from Argentina has collected information coming from Latin American (Latam) countries plus Spain and Portugal in a specific database, which allows for their analysis on an independent basis. The objective of this work is to analyze and make graphically visible the thematic distribution, media typology, and digital formats of the fake news that circulated regarding the coronavirus pandemic in Ibero-America by means of the social network analysis technique (SNA) and others, using the information from the aforementioned database. Its final visualization is shown by means of interactive mapping, which can be manipulated by the user to analyze the results through the application of different integrated filtering strategies.

Keywords

Fake news; Disinformation; Misinformation; Hoaxes; Fact-checking; Post-truth; Data visualization; Verification; Covid-19; Pandemics; Health information; Ibero-America.

1. Introduction

The global COVID-19 coronavirus pandemic since early 2020 has been accompanied by an unprecedented wave of planetary disinformation.

"Misleading content –cures, 'medical' advice, conspiracy theories, etc.– disseminated on various media (socio-digital networks, SMS, private messaging, traditional media, etc.). What do we know today about this mass of disinformation? What does it tell us about the functioning of information processes in contemporary societies?" (Monnier, 2020).



Research publications show the importance of a second, unprecedented pandemic of fake news. Disinformation about COVID-19 has attracted a lot of attention.

“Much of the initial research has focused on the types, origins, and spread (Brennen *et al.*, 2020), as well as possible effects of false information and conspiracies around the coronavirus and the pandemic (Freeman *et al.*, 2020). Academic, industry, and journalistic accounts have examined the propagation of COVID-19 disinformation by state actors (Swan, 2020), the prevalence of COVID-19 conspiracy theories (Freeman *et al.*, 2020; Uscinski *et al.*, 2020), the spread of misinformation on social media on platforms such as Facebook and Twitter (Hollowood; Mostrous, 2020), as well as interventions to counteract misinformation on social media (Pennycook *et al.*, 2020).” (Brennen *et al.*, 2021).

1.1. Disinformation

The concept of disinformation, understood as the dissemination of false news as a technique of political/war attack, has been known since ancient times but became popular in the media from the 1980s onwards (Rivas-Troitiño, 1995; Martínez-Musiño, 2011). The *Royal Spanish Academy (RAE)* (2014) defines disinformation as

“to give information intentionally manipulated to serve certain purposes, or to give insufficient information or omit it” (RAE, 2014).

Recently the EU has defined it as

“false, inaccurate or misleading information intentionally designed, presented and promoted to cause public harm or for profit” (*High Level Group on Fake News and Online Disinformation*, 2018).

With the development of the web and the subsequent explosion of social networks, which have been joined by instant messaging platforms, the phenomenon has overflowed both in magnitude and impact on the knowledge and decisions of citizens worldwide (Bond, 2012). This impact has already been detected in several areas of social behavior, but with special incidence in politics and health. Its appearance in the context of politics brought to fame the term post-truth¹, collected by the *Royal Spanish Academy (RAE)* in 2017 as

“Deliberate distortion of a reality, which manipulates beliefs and emotions in order to influence public opinion and in social attitudes” (RAE, 2014).

Misinformation, whether intentional or not, is an ally of post-truth and makes objective facts lose influence in defining public opinion, giving precedence to the emotional over the rational (Subires-Mancera, 2017). In addition, fake news spreads faster and has greater reach in social networks than real news (Vosoughi; Roy; Aral, 2018).

In the area of health, misinformation has been growing to such an extent that Larson (2018) published in the journal *Nature* an article entitled “*The biggest pandemic risk? viral misinformation*”. In it he stresses that

“the avalanche of contradictory information, misinformation, and manipulated information on social networks should be recognized as a global threat to public health.”

It denounces the fact that the anti-vaccine movement, which has been fueled by the misinformation circulating on the Internet, is causing an alarming drop in the number of people vaccinated, especially children, and an increase in infant mortality due to this cause. The use of the Internet as a source of health information on the web has been of concern since the network became popular, driving the use of reliable sources and the development of certificates of quality of information, but the increase of hoaxes or fake news on social networks has turned disinformation into a problem to which solutions must be found to deal with it (Chou; Oh; Klein, 2018).

The *II Study on Health Hoaxes (Instituto #SaludsinBulos; Doctoralia, 2019)* reveals that almost two out of three healthcare professionals surveyed (66%) claim that health hoaxes increased (compared to 57% who believed so in 2018) and that misinformation is causing patients to mistrust the healthcare professional as a source of information. Healthcare practitioners have dubbed this effect “Dr. Google” (Rosenbaum, 2018; Segrelles-Calvo, 2016; Stein, 2011). The following *Health Hoax Studies (III and IV)* focused on COVID hoaxes specifically.

The increase in hoaxes or fake news and their effect has led various international organizations to begin to study the phenomenon and try to adopt measures to control it. One of the first was the EU, whose *European Council* promoted the creation of the *East StratCom Task Force* in 2015 (*European Council, 2015*) to “counter current Russian disinformation campaigns”. They disseminate their work on the web at <https://euvdisinfo.eu>

where their flagship publication is *Disinformation Review*. Since then, it has been keeping it in mind to fight against it (*European Commission, 2018a; 2018b; 2018c; 2022; High Level Group on Fake News and Online Disinformation, 2018*).

For its part *Unesco*, through its *International Programme for the Development of Communication (IPC)*, published a manual, *Journalism, “fake news” and disinformation - Handbook for journalism education and training* (Ireton; Posetti, 2018) to guide journalism professionals and other actors to fight disinformation. The number of organizations and media outlets that have been joining to tackle disinformation has been growing since then in all countries, especially those dedicated to verifying information.

Due to criticism:

<https://www.nytimes.com/es/2019/11/01/espanol/opinion/red-social-facebook-sorkin.htm>

also the companies behind these social networks have been forced to intervene, especially the social network *Facebook* that recognizes that part of the problem is that fake news is a business, because it attracts clicks and therefore advertising (**Jiménez-Cano**, 2017). The European Union got, in October 2018, the industry to agree for the first time in the world, a voluntary self-regulatory code of good practice to combat the spread of disinformation online (*European Commission*, 2018c). This code was revised and in June 2022 published as a strengthened *Code of practice on disinformation*, to which other stakeholders have adhered (*European Commission*, 2022).

However, it was only after the pressure caused by the COVID-19 pandemic and the subsequent US elections, which have made things worse in terms of disinformation, that companies began to take more far-reaching measures than those taken until then (*El Español*, 2020, **Raya**, 2020; *ABC*, 2020). In fact, a report published by *Avaaz* (2020), a global citizens' platform, pointed to *Facebook's* algorithm as the problem, given that at despite the efforts made by the company to warn users, the results have been minimal, highlighting that the contents of the most popular websites that disseminate false information are seen four times more than those that show truthful information on the platform.

Another strategy to curb hoaxes has been to promote the information literacy of the targeted citizens. Information literacy, metaliteracy as it is currently conceptualized, or Media and Information Literacy, as it is called by *Unesco*, involves the acquisition of skills that, among other things, enable one to evaluate information critically, weighing its veracity. Promoted by librarians and teachers since the end of the 20th century, it is a long-term strategy, complementary to journalistic rigor and technological performance. The EU also promotes the competences of digital citizens since its emergence, having in the *DigComp Framework for digital competence* (**Ferrari; Punie; Brečko**, 2013) its current strategy to boost them. It is also appealing to those people with these competences to collaborate by reporting the hoaxes they find

<https://theconversation.com/5-ways-to-help-stop-the-infodemic-the-increasing-misinformation-about-coronavirus-137561>

Faced with the advance of disinformation, the journalistic media also began to worry, as their work was beginning to suffer in terms of public credibility. To avoid this, they have launched various platforms, called verification platforms, which try to dismantle this type of "news," and which have been emerging all over the world. Journalists have joined the push for media literacy and in 2015, the verification organizations formed an international network, the *IFCN* (*International Fact-Checking Network*), under the auspices of the *Pointer Institute*, an American organization founded in 1975 that is now one of the most influential journalism schools in the world.

<https://www.poynter.org/mission-vision>

1.2. COVID-19 and disinformation

Although the problem of misinformation has reached its zenith in the months of pandemic produced by the SARS-CoV-2 virus (COVID-19), it has not been absent in other pandemics² such as Zika virus (**Almeida**, 2016) or Ebola (**Oyeyemi; Gabarron; Wynn**, 2014; **Wang et al.**, 2019). And equally in other plagues recorded throughout history occurred when social networks did not exist³ (Figure 1).

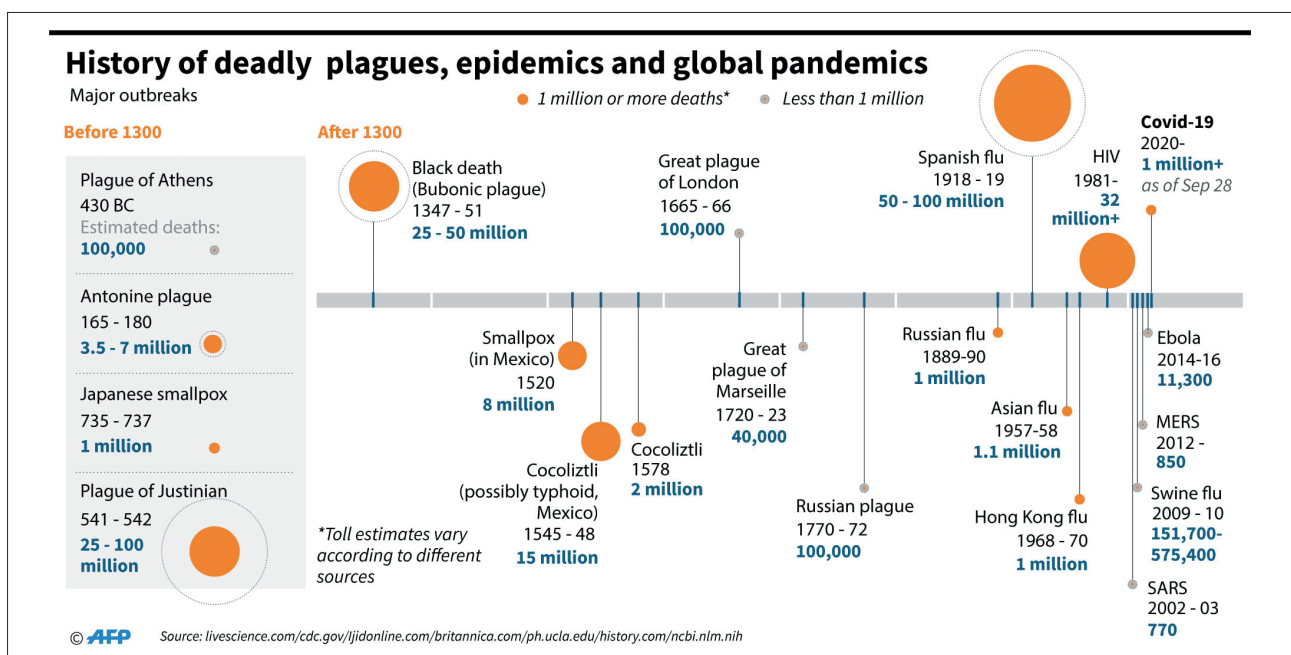


Figure 1. The great epidemics of history

<https://shorturl.at/cuMQX>

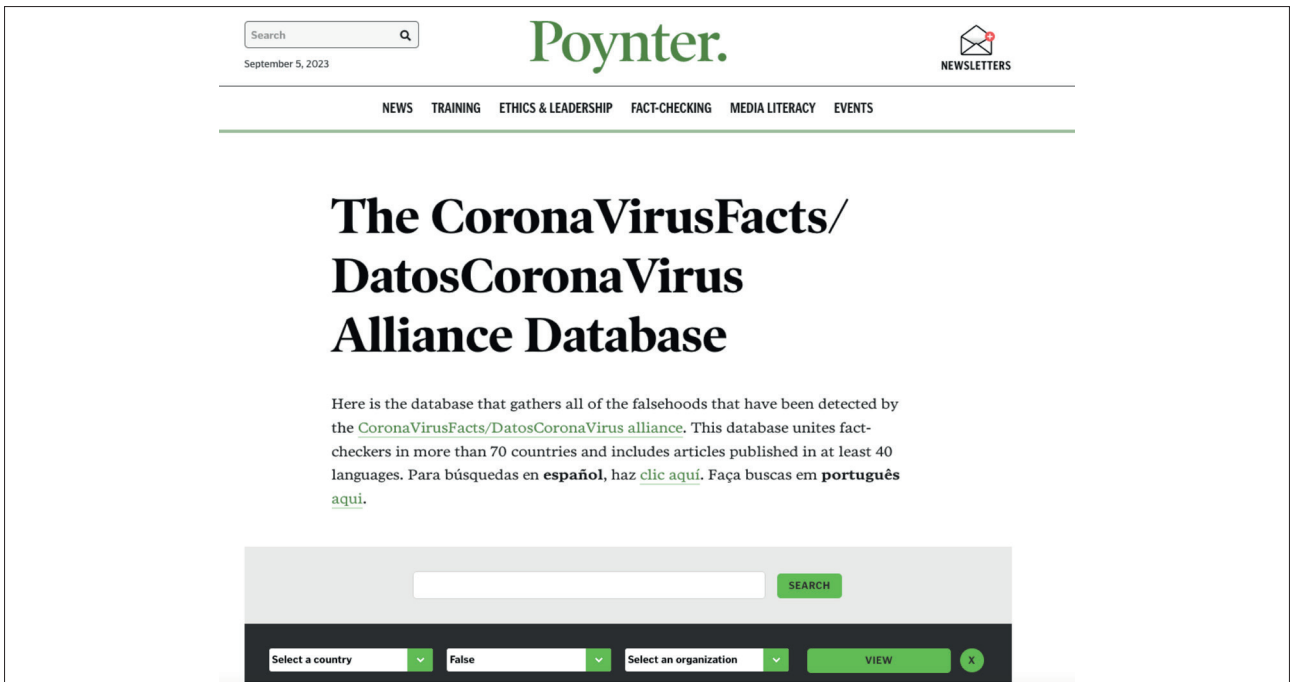


Figure 2. Poynter database web page on COVID-19. https://www.poynter.org/ifcn-covid-19-misinformation/?covid_countries=0&covid_rating=51174&covid_fact_checkers=0

The COVID-19 misinformation, however, involves many new elements, which makes it complex:

“the multiplication of media, channels and actors involved in the production and distribution of content (actors from the world of health, politics, the media, religious, ordinary citizens, etc.); the uncertainties and even contradictions that marked public communication almost everywhere in the world, particularly at the beginning of the epidemic, linked to the lack of data, the desire to reassure the public, the fear of triggering panic movements or of suffering the consequences of strong economic or political actions” (Monnier, 2020).

The Poynter Institute’s International Fact Checker Network (IFCN) launched the #CoronaVirusFacts Alliance in January 2020, at a time when the spread of the coronavirus was limited to China but already causing rampant misinformation around the world. This alliance united more than 100 fact-checkers from around the world to publish, share and translate facts related to the COVID-19 pandemic. The World Health Organization classifies this as an infodemic.

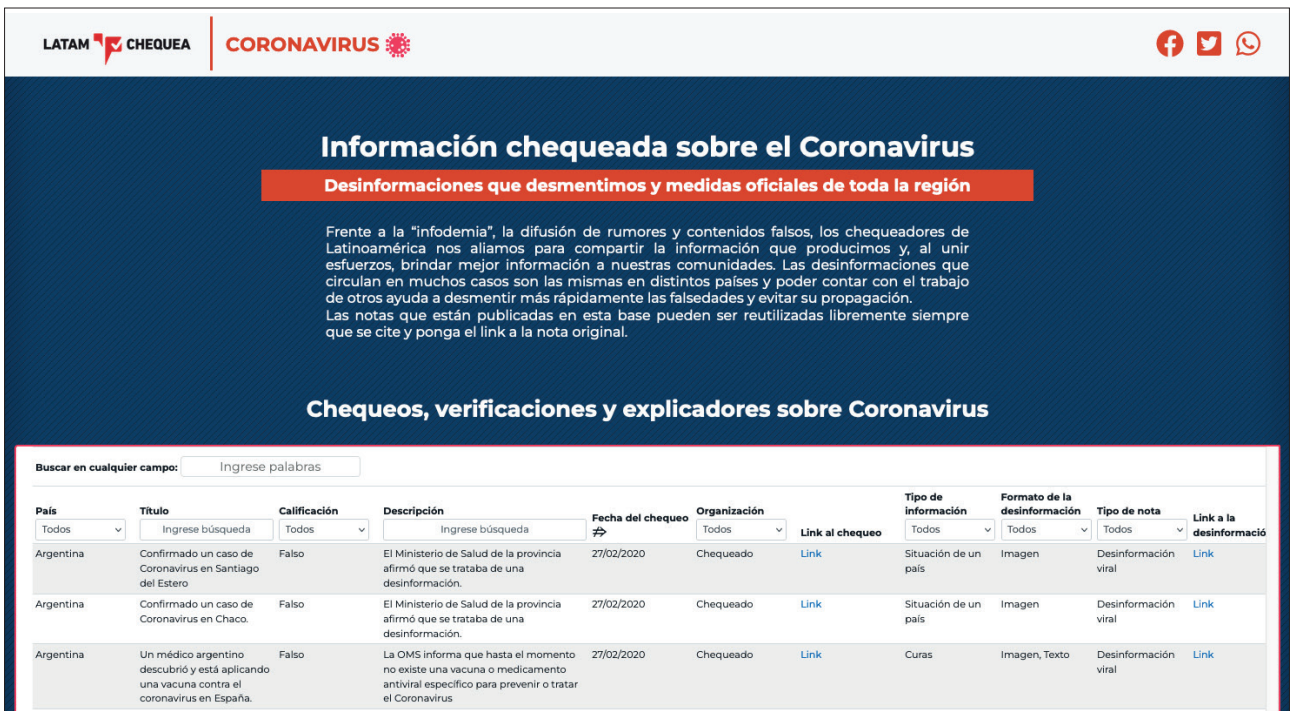


Figure 3. Latam Chequea web page on COVID-19. <https://chequeado.com/latamcoronavirus>

The *#CoronaVirusFacts* database

<https://www.poynter.org/coronavirusfactsalliance>

collects all hoaxes discovered by the *Alliance* and will be updated daily with new publications until early 2023. It can be sorted, filtered and searched by content. The database brings together fact-checkers from more than 70 countries and includes articles published in at least 40 languages (Figure 2).

In April of that same year, also under the auspices of the *#CoronaVirusFacts Alliance*, a specific database began to be published with verified information on coronavirus from the different verifying agencies of Latin American (Latam) countries, plus those of Spain and Portugal, a subset of the previous one. *Latam Chequea Coronavirus* integrated 35 organizations from 15 countries⁴, to disseminate in Spanish and Portuguese the misinformation circulating about COVID-19, Figure 3. The project was coordinated by the organization *Chequeado*⁵ (Argentina) and supported by *Google News Initiative*. Unlike the previous one, this database is free to download, available at:

<https://chequeado.com/latamcoronavirus>

The verification objectives of the organizations adhered to the *Latam Check Coronavirus* project can be summarized as follows:

1. To know the number of posts related to COVID-19 fake news, their volume with and frequency.
2. To analyze the most frequent interactions with false content published on COVID-19 and its relationship with the type of publication.
3. Determine the level of interaction of COVID-19 related content on a per verifier basis.

1.3. Visualization of fake news about COVID-19

Using visual elements such as graphs and maps, data visualisation provides an accessible way to identify and understand trends, outliers and patterns in data. It is used to provide simple summaries of large data sets. In visualisation, statistics is joined by other disciplines that help to highlight information to make it easier to understand.

“The raw material to develop an effective visualization is data, figures and the search for information. And, thanks to information design, it is possible to take these variables and transform them into graphic pieces that will help the general public to perceive the relationships they have more easily” (Bayas-Ramírez, 2020).

Thus, data visualisation on COVID-19 allows us to use another way of communicating information and methods that serve other disciplines (design, communication, data journalism, databases, etc.) to analyse and make data accessible.

Examples of visualisations of fake news data related to the SARS-CoV-2 pandemic include those using data collected by the *#CoronaVirusFacts Alliance* mentioned above. This alliance has created a database of more than 17,000 fake news verifications, which it presents in several types of interactive visualisations on its website. One of them (Figure 4) shows on a map the number of verifications per country since January 2020, a map updated at the end of each month. The distribution of the data logically reflects the contributors to the *Alliance*, as there is no direct data where there are no verifiers.

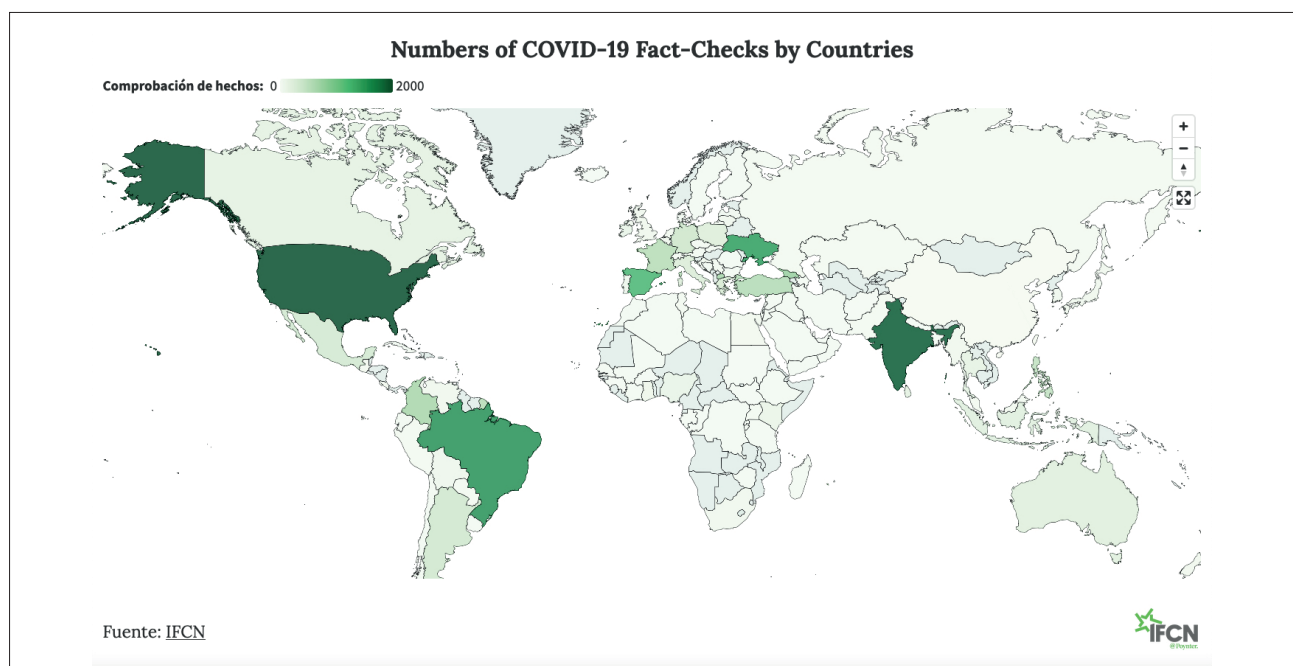


Figure 4. Fake news checks by country since January 2020.

<https://flo.uri.sh/visualisation/8541954/embed?auto=1>

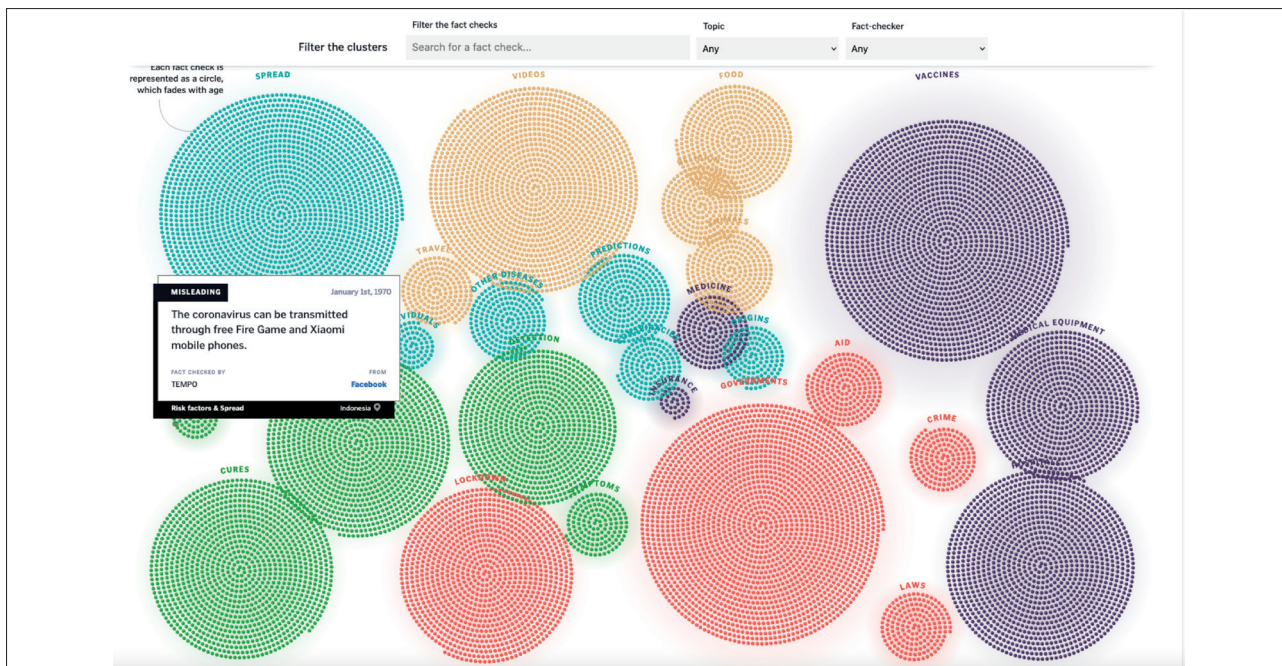


Figure 5. Categories of hoaxes.
<https://www.poynter.org/coronavirusfactsalliance>

Two other visualizations show the categories with which the alliance classifies the verifications made in its database. One of them shows their evolution over time, showing how they are changing and how some topics are appearing. Figure 5 shows them by size according to the number detected, allowing to read a small synthesis of each of the verifications. It also incorporates the possibility of filtering.

Another example to note is that five European verifiers⁶ made from a report they published on the spread of the infodemic in parallel with the spread of the COVID-19 pandemic across Europe and the imposition of large-scale public health measures, using the 645 COVID-19 related verifications they made during the months of March and April 2020. The results are displayed in visual form on the website:
<https://covidinfodemicurope.com/?lang=en>

As in the previous case, the different types stand out (in this case with an iconographic chart with the most common ones, Figure 6) and how they vary over time:

They also use the geographical distribution of some hoax subtypes (Figure 7) and timelines to show how specific hoaxes spread across different countries (Figure 8). To enlarge information also uses interactive visualizations (interactive version of the graphic in Figure 7 at:
<https://view.genial.ly/5ef071b6478cb20d63238325>

It is also worth mentioning the work published by Bayas-Ramírez (2020), which performs different visualizations with data from the *Ecuador Chequea* verifier.

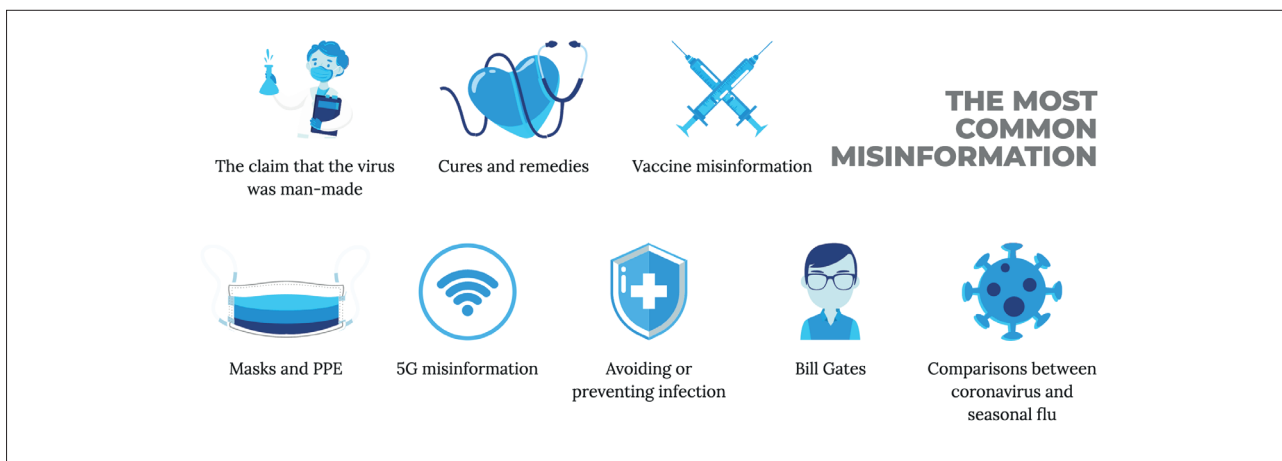


Figure 6. The most common hoaxes.
<https://covidinfodemicurope.com/?lang=en>

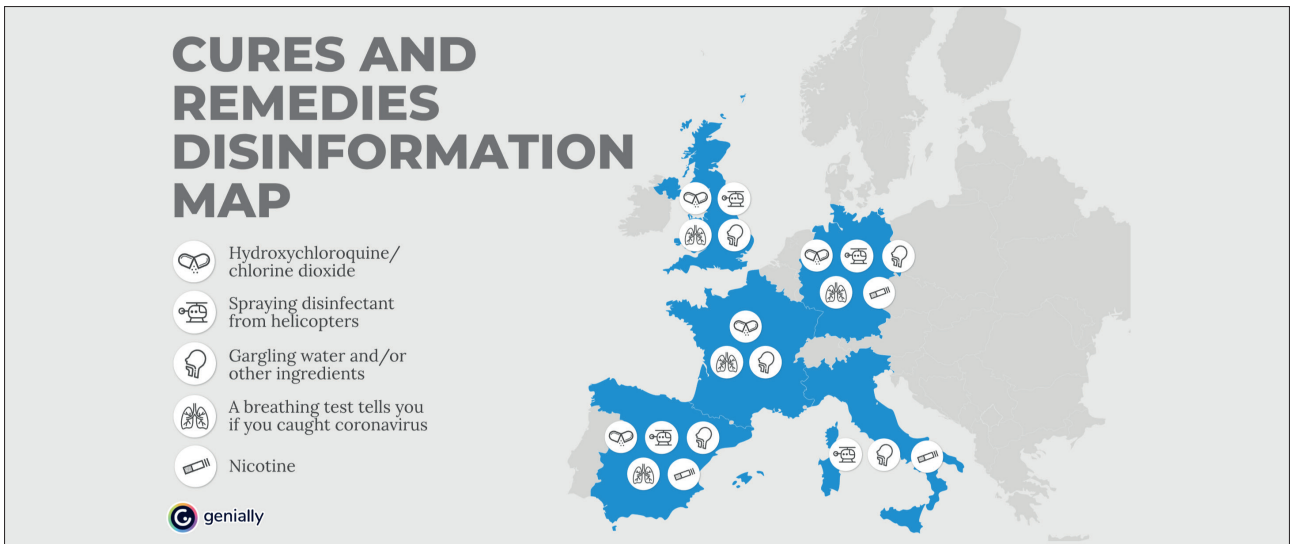


Figure 7. Map of misinformation about cures and remedies. <https://covidinfodemicseurope.com/?lang=en>

2. Target

The objective of this work is to analyze and make graphically visible the thematic and geographic distribution, the typology of media and digital formats of the fake news that circulated about the coronavirus pandemic in Iberoamerica, using statistical techniques and social network analysis (SNA).

3. Methodology and data source

Social network analysis (SNA) (Avila-Toscano, 2018) has been used as a methodology for part of the data visualization. SNA uses networks and graph theory (Andrienko *et al.*, 2020; Otte; Rousseau, 2002).

The software used to create these visualizations has been *Gephi* (Bastian; Heymann; Jacomy, 2009): <https://gephi.org>

Gephi is a program for visualizing, exploring, and understanding all types of graphs and networks (Cherven, 2015). It is free, open source and has ARS as the basis of its operation. The spatialization algorithms used have been *Atlas Force 2* and *Atlas 2-3D*. It has been combined with a viewer that allows exporting to the web the graphs made with *Gephi*, called *gexf.js* (Velt, 2011) and which is available on *Github*: <https://github.com/raphv/gexf.js>

This viewer enables interactivity for multiple users and supports the inclusion of several plots.

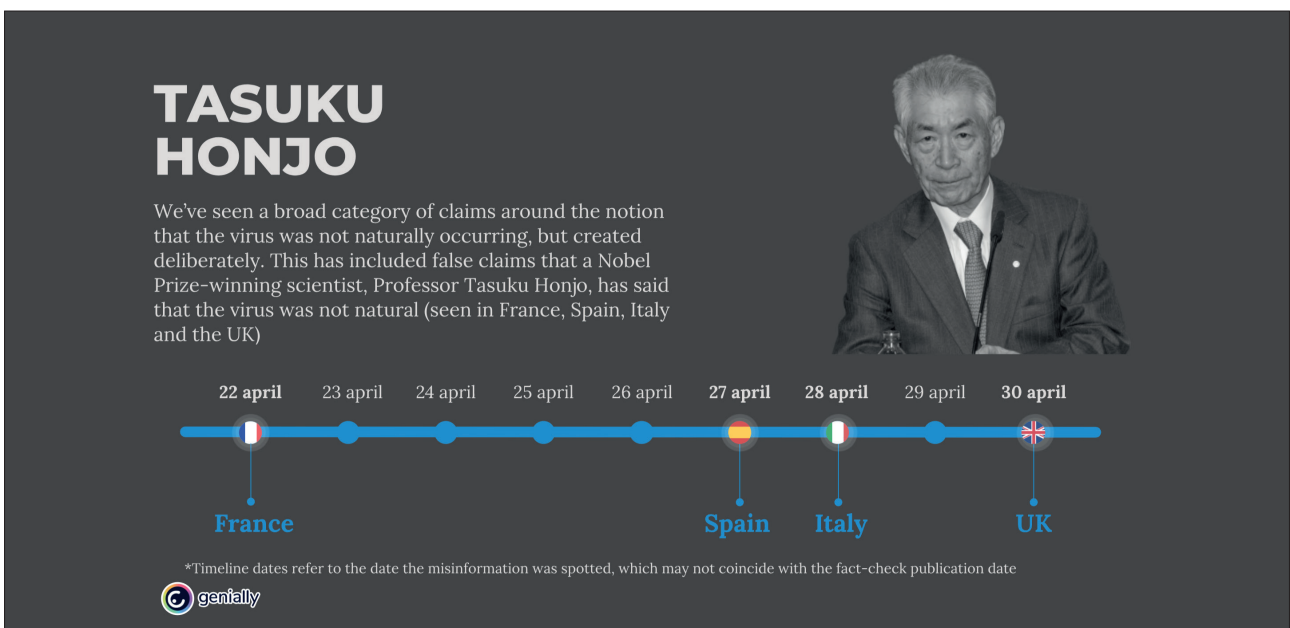


Figure 8. False claims referring to alleged statements by Nobel Prize winner Tasuku Honjo. <https://covidinfodemicseurope.com/?lang=en>

SCImago Graphica (Hassan-Montero; De-Moya-Anegón; Guerrero-Bote, 2022) , and *Excel* (dynamic graphics) have also been used for other statistical analyses.

<https://www.graphica.app>

The data used comes from the collection conducted by the Spanish, Portuguese, and Latin American verification agencies (*Latam Chequea Coronavirus* network) belonging to the *International #CoronaVirusFacts/ Network Alliance (IFCN)* of the *Poynter Institute* mentioned above.

Unlike the general data, this table can be downloaded in .csv format (comma separated values) for free use, with acknowledgement of authorship. It is the one that has been used in this work as the basis for data analysis, graphs, and visualizations (Figure 3).

Although it is called a database, it is a single table with structured data: the rows contain the different records, and the columns contain the different data (a database interrelates two or more tables). As it is in .csv format, it can be processed by different spreadsheet programs (*Excel* from the *Office* suite, *LibreOffice Calc* from the *LibreOffice* suite...) as well as visualization programs, as all of them take the data in this format for further processing.

The data collected from each contrasted information (registry) are described in Table 1.

Table 1. Data collected in the table of verified news data

Country	Corresponds to the country where the information is detected and where it is checked by a verifying organization.
Title	A short descriptive sentence of the information.
Rating	A word (or words) indicating the degree of truthfulness of the information.
Description	Text summarizing the information content.
Date of check	Date on which the information is analyzed by the verifying entity.
Organization	Verifying entity.
Link to the checkup	Link to the result of the information analysis.
Type of information	Describes the subject matter of the information using one or more descriptors from a controlled vocabulary.
Misinformation format	Whether it has been disseminated as text, image, video, audio or other, or any combination thereof.
Type of note	Describes the type of information generated by the verifying entity.
Link to disinformation	Link to the information being analyzed.
Date of detection misinformation	Date on which the information to be analyzed is detected.
Origin	The medium in which the information has been detected to be disseminated, whether traditional media or social networks, and which ones are being used, is collected.
Person	Collect the person or media source of the information, if known.
Updates	Date(s) on which the verifier’s analysis of the information has been updated.
Circulated in other countries	The countries where the analyzed information has also been detected are listed below.

Since 2022, the table has stabilised at a number of records slightly above 5000, obtained since January 2020. In addition to the standardisation of words (with and without accents, typographical errors, etc.), the table required some modifications to the data, which are detailed below.

Scoring: Qualification does not follow a controlled vocabulary shared by the verifiers. This leads to overlapping in some terms with similar or overlapping meanings. Thus, they have been assimilated to a single concept, the one closest to the term chosen to represent it (Table 2).

Table 2. Qualifying descriptors of the result of the study of the veracity of the information

False	Bulo, Fake, False, Unsustainable, Manipulated, Liar, Lie
True	True, True but..., Not fake...
Questionable	hasty, questionable, misrepresented, debatable, doubtful, misleading, exaggerated, out of context, inaccurate, half true, partially untrue, taken out of context
Satire	Satire
Not verifiable	Undecidable, No certainty, No evidence, Not verifiable, No data

There is also the option ‘Multiple verification’, which indicates that different statements in the information had to be verified. Based on the information provided by the verifier in the dataset, a consensus has been reached to assign it to

one of the grouped classification categories in the table above. Similarly, the blank cells have been filled in with the most approximate assessment based on the description provided by the verifier.

Type of information: This field collects the subject matter that the (alleged) news item is about. They also use a small controlled vocabulary that, in this case, does not present ambiguities: Authorities, Virus characteristics, Contagions, Coronavirus, Cures, Measures, Origin of the virus, Others, Prediction, Prevention, Spread, Tests, Symptoms, Situation of a country, Conspiracy theories, Antigen testing, Vaccines, COVID-19 variants.

Disinformation format: Whether text, image, video, audio, or any combination of these is used. No problems, except for some typographical errors.

Origin of the information: This refers to the digital medium through which it was detected that it was beginning to be disseminated. Although they use a vocabulary that is not excessively broad, it is not controlled either. It suffers from several problems: it incorporates specific concepts of some general ones that it also includes, for example: social networks (generic) and *Facebook, Instagram*, etc. (specific). This is due to the impossibility in some cases of pinpointing the exact origin. They also use different words to refer to a single concept (Website, Web page, Government website).

After cleaning up these problems the resulting vocabulary includes the following terms: Statement, Email, *Facebook, Instagram, Media, Other, Social Networks, SMS, Telegram, Television, TikTok, Twitter, Web, WhatsApp, YouTube, Zoom*.

The term Statement disappeared with the elimination of a note type category explained below.

Type of note: This field also follows a controlled vocabulary with only five categories: Media Check, Public Figure Check, Viral Disinformation, Explainer and Other, with more than one term being possible to assign (not frequent). The “Explainer” category refers to

“more general articles that are not likely to be verified or disproven (e.g., explanatory articles about the science behind the virus or about the different confinement rules in each country)”:

<https://covidinfodemicurope.com>

Since our focus is on misinformation, whether intentional or misinterpreted, these explanatory articles have been dismissed as an additional informational product of the verifying entity and not properly a verification of information. This leaves the total number of records analyzed at 3693.

4. Results

The web address where the interactive visualizations made with *Gephi* and *Gexf.js* can be consulted is: <http://metroteach.com/covid/index.html>

We will now analyze the results obtained for each of the aspects studied.

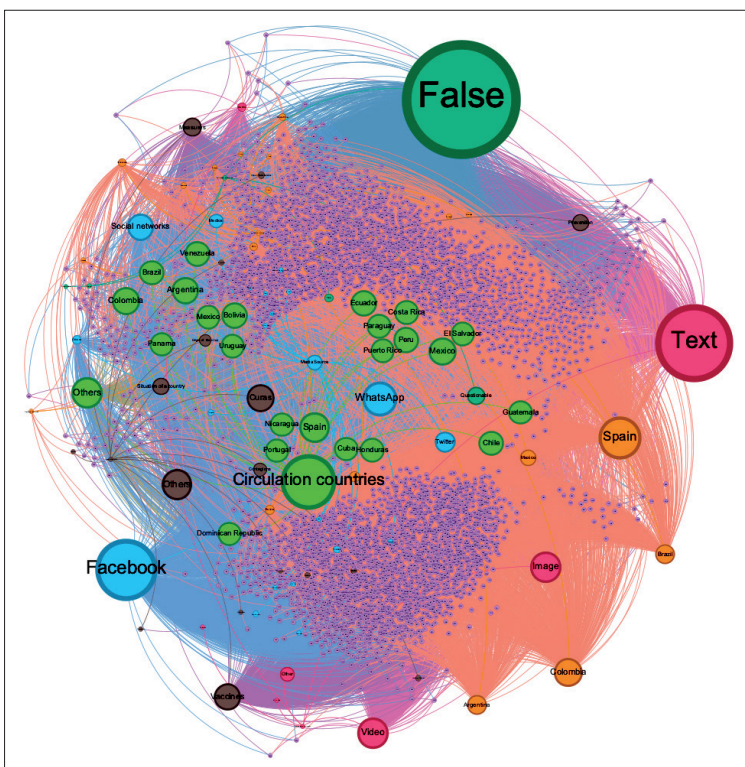


Figure 9. Total data set of verified news items.
Source: <http://metroteach.com/covid/index>

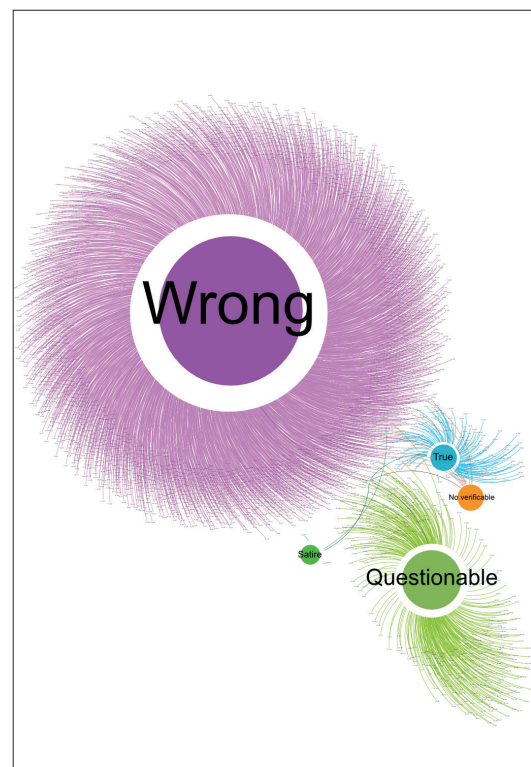


Figure 10. Final rating of the news after verification by the agencies (number of nodes: 3699, number of links: 3698).

Figure 9 shows the total set of verified news items by category with all the analyzed data and their respective nodes. It includes a total of 4690 nodes and 10040 links since the data for some criteria analyzed include more than one value. It represents the full set of unfiltered relationships.

4.1. Classification of verified news items

Figure 10 shows the result of the rating data category, which records the result of the analysis of the verification performed. As can be seen, the number of information determined as false is extremely high. The result is not surprising given that it has an initial bias: the agencies investigate and try to dismantle precisely those pieces of information that are not true. They also investigate those that citizens send them for verification, which they encourage. In both cases they are driven by suspicion of falsity. Hardly anyone is interested in verifying information if they believe it to be true.

Since the expected result of the verification is that it is false information, the focus of the analysis is on the rest, especially on those qualified as Questionable and True. It is noteworthy that in both sets, especially in true, one of the most prominent groups focuses about a country, in addition to cures and vaccines which are also important in the false ones.

The distribution of the rating of verified news by country also highlights that in some countries the percentage of questionable and true is high compared to others. Those countries with a small number of verified news items, such as the Dominican Republic, Cuba, Nicaragua, Paraguay, should be discarded.

4.2. Origin of verified news

Figure 13 shows the results of the source data category, which shows the channels through which the information was disseminated. *Facebook* is clearly the most used medium, followed by *WhatsApp*. It was already commented in

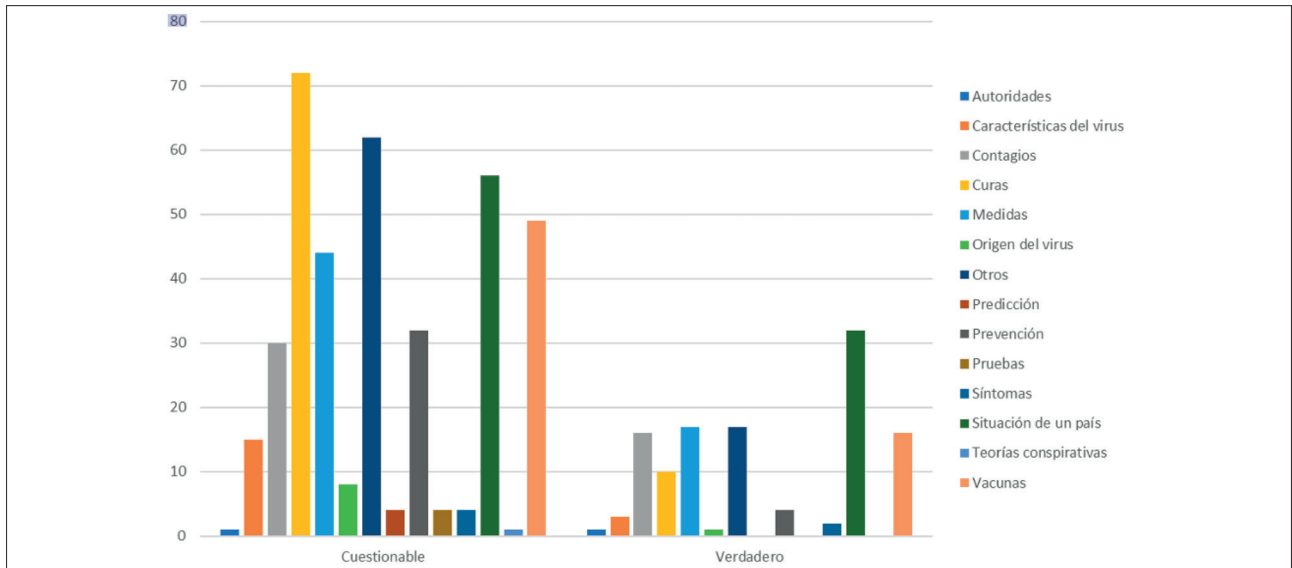


Figure 11. Thematic distribution of verified news items rated as questionable and true.

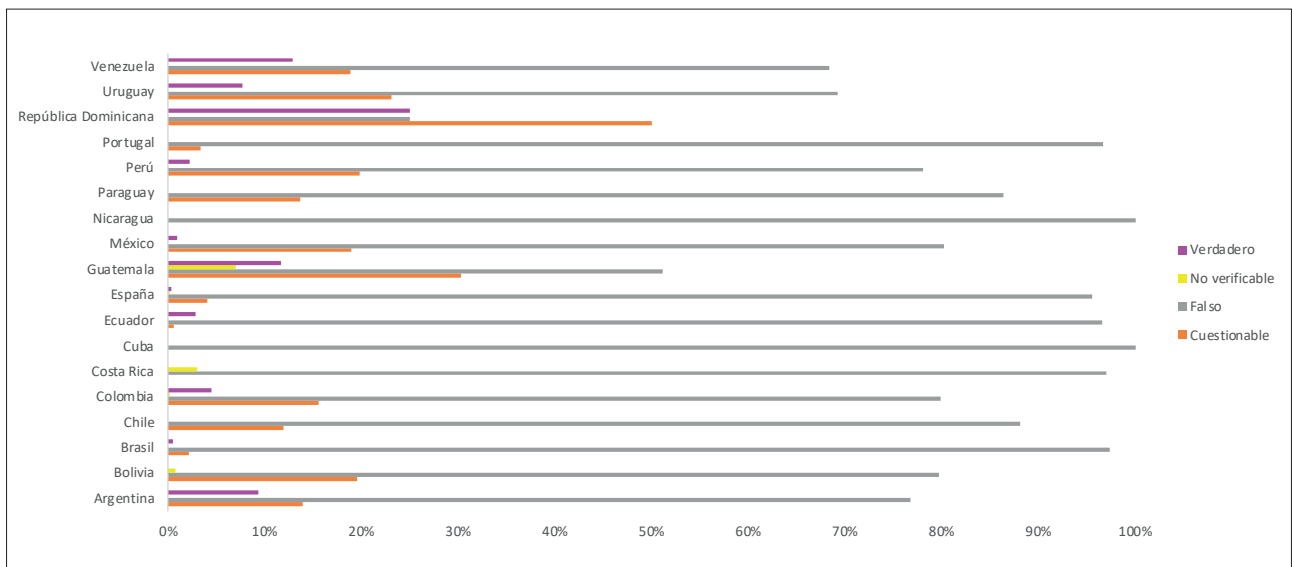


Figure 12. Percentage of the different ratings after verification of the information in relation to those conducted in each of the participating countries.

the introduction the problem that this social network has regarding the dissemination of fake news, a problem related to the profitability of the same. It also invites a reflection: *Facebook* is a social network used mostly by older people than the rest and *WhatsApp*. Regardless of other factors that may influence virality, the possibility of fitting with previous beliefs or prejudices in older people favors their dispersion without being questioned.

4.3. Geographic distribution by country of verified news items

Figure 14 shows the result of the geographic distribution of the verified news. It includes data from the Country and Circulated in other countries fields. It should be borne in mind that not all countries in the area have a verifying entity, but it gives us an idea of the impact of this problem. Although up to 62 countries are listed where some information was circulated, the great interconnection between the countries on both sides of the Atlantic, due to their common language

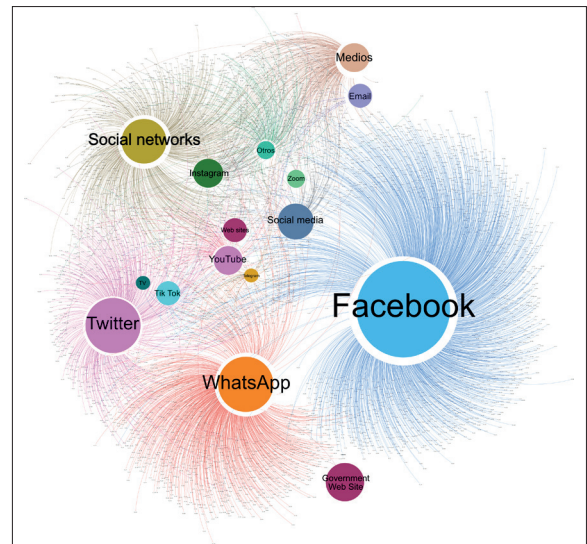


Figure 13. Initial distribution channels of the verified information (number of nodes: 4243, number of links: 4224)



Figure 14. Geographical distribution of verified news items (number of nodes: 62, number of links: 1643).

and culture, is evident. And, of course, between Latin American countries. The rest of the countries involved reflect the main destinations of emigration for work or studies.

Since we do not have data from all countries, nor other data that could provide more information, it is not possible to know whether, for example, the degree of digitalization of citizens has a positive or negative impact on their dispersion.

4.4. Type of information in verified news items

Figure 14 shows the subject matter of the news items checked. As it is a disease, health aspects were the subject of most of them, but the measures taken by governments and the situation of the country, which have political consequences, also stand out. The number of them devoted to vaccines is relevant, in line with what was denounced by the experts. A remarkably high amount to counteract, precisely, one of the relevant scientific milestones derived from the joint struggle of science against this global pandemic obtaining several safe vaccines in a brief period.

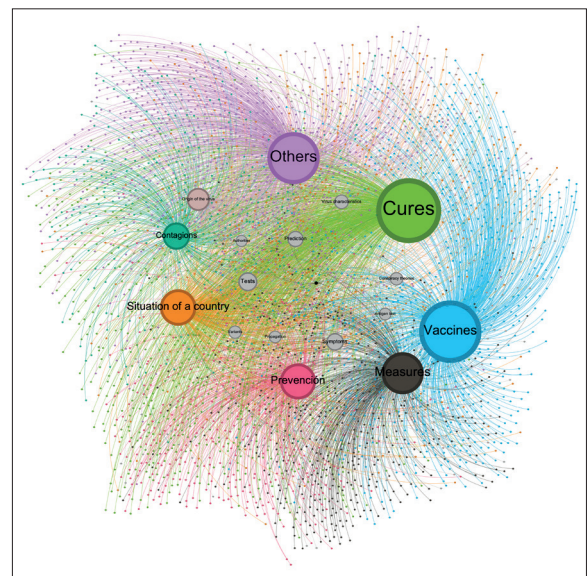


Figure 15. Distribution of the subject matter of the verified news (number of nodes: 3890, number of links: 3889).

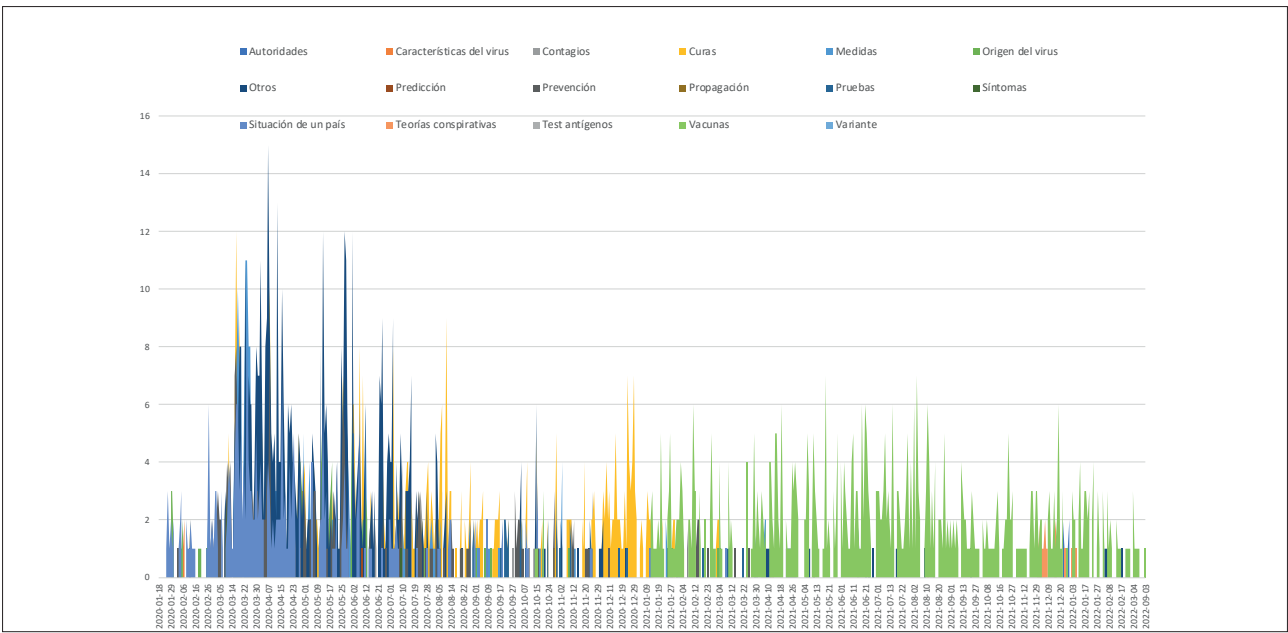


Figure 16. Evolution of verified news subject matter over time.

An interesting aspect is the evolution over time of the thematic content of the verified news. It adapts to the interest of the public, which reinforces the idea that part of the interest in the generation of fake news has an economic basis and manipulation of public opinion. Figure 16 shows the change in the colors (different subjects) in the three years in which the data is extended.

4.5. Format in which verified news is distributed

Figure 17 shows the distribution of formats in which verified information is transmitted. Text is the star, followed by video and image. This is because in many cases images and videos are accompanied by explanatory text to reinforce the idea.

5. Conclusions

The amount of false information that circulated on social networks in relation to the COVID-19 pandemic has been studied in numerous publications during and after the pandemic. This type of information has occurred at other similar moments in history, but today technology has facilitated it and allowed us to study the phenomenon. Fear of the unknown is a good breeding ground.

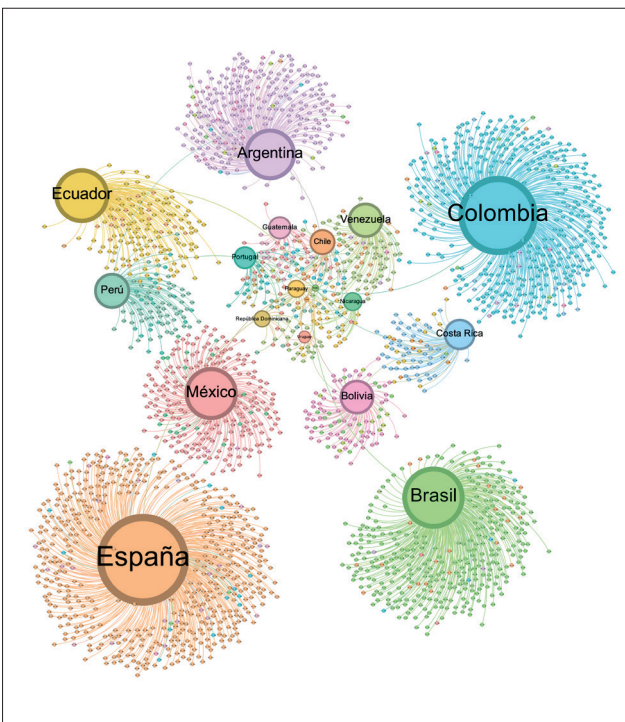


Figure 17. Distribution of news by countries.

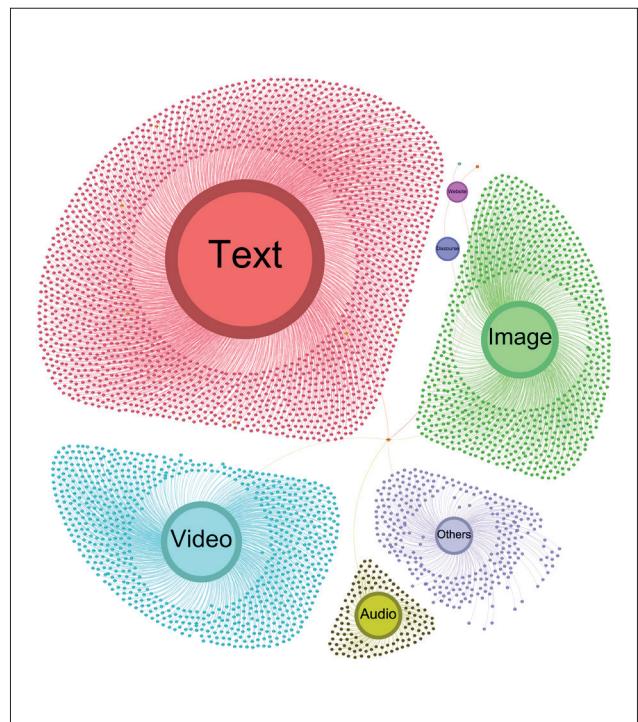


Figure 18. Evolution of verified news subject matter over time.

Although the data is initially skewed and most of the information verified is classified as false, a certain distrust is observed in news related to local situations. The results obtained help to understand the dynamics of fake news in a culturally related area through language, history and current relations, reflecting this connection in terms of geographical dispersion. The origin of this information suggests that the social network is the most common and that text is the format of choice, alone or in combination with images and videos.

The evolution over time of the theme they cover is in line with the change in interest that is also taking place, a strategy that is logically aimed at achieving their dissemination.

In order to achieve better results, it is advisable for organisations to work with uniform criteria from the moment they start collecting information, especially if they do so from different data entry points. The creation of controlled vocabularies for mandatory use improves overall consistency. The use of the 'other' category should also be minimised. In our case, this category represents too large a number in the information type data.

6. Notes

1. It was word of the year in 2016 for the *Oxford Dictionary*.

<https://languages.oup.com/word-of-the-year/2016>

2. See:

<https://elordenmundial.com/mapas-y-graficos/grandes-epidemias-de-la-historia>

<https://www.library-archives.cumc.columbia.edu/blog/2020-03/pandemics-history-short-bibliography>

<https://bcmj.org/blog/history-pandemics>

3. See, for example,

https://www.elespanol.com/cultura/historia/20200327/todas-fake-historia-pandemias-pestes-negra-coronavirus/477452796_0.html

<https://theconversation.com/bleach-bonfires-and-bad-breath-the-long-history-of-dodgy-plague-remedies-137658>

4. *AFP Factual*, *Salud con Lupa*; *Aos Fatos*, *Estadao Verifica* and *Lupa*, from Brazil; *Bolivia Verifica*; *Mala Espina Check*, from Chile; *La Silla Vacía* and *ColombiaCheck*, from Colombia; *#NoComaCuento* (*La Nación*) and *La Voz de Guanacaste*, from Costa Rica; *Periodismo de Barrio* and *El Toque*, from Cuba; *Ecuador Chequea* and *GK*, from Ecuador; *EFE Verifica*, *Maldita.es* and *Newtral*, from Spain; *Agencia Ocote*, from Guatemala; *Animal Político*, *Spondeo Media* and *Verificado*, from Mexico; *Despacho 505* and *La Lupa*, from Nicaragua; *El Surtidor*, from Paraguay; *Convoca*, *OjoPúblico* and *Verificador* (*La República*), from Peru; *Observador* and *Polígrafo*, from Portugal; *PoletikaRD*, from Dominican Republic; *UyCheck*, from Uruguay; and *Cotejo.info*, *Efecto Cocuyo* and *EsPaja*, from Venezuela.

5. *Chequeado* is a verification project of *Fundación La Voz Pública* de Argentina:

“We are a non-partisan, non-profit digital media dedicated to the verification of public discourse, the fight against disinformation, the promotion of access to information and the opening of data” (2022).

<https://chequeado.com/conocenos>

6. *Agence France-Presse* (AFP) in France, *Correctiv* in Germany, *Pagella Politica/Facta* in Italy, *Maldita.es* in Spain, and *Full Fact* in the United Kingdom.

7. References

ABC/EP (2020). “WhatsApp lanza una función para combatir los bulos dentro de la aplicación”. *ABC*, 4 agosto.

https://www.abc.es/tecnologia/moviles/aplicaciones/abci-whatsapp-lanza-funcion-para-combatir-bulos-dentro-aplicacion-202008041238_noticia.html

Almeida, Carla (2016). “Radar latinoamericano: La desinformación en tiempos de zika”. *SciDevNet*, 15 febrero.

<https://www.scidev.net/america-latina/enfermedades/blog-de-analistas/radar-latinoamericano-la-desinformacion-en-tiempos-de-zika.html>

Andrienko, Natalia; Andrienko, Gennady; Fuchs, Georg; Slingsby, Aidan; Turkay, Cagatay; Wrobel, Stefan (2020). “Visual analytics for understanding relationships between entities”. In: Andrienko, Natalia; Andrienko, Gennady; Fuchs, Georg; Slingsby, Aidan; Turkay, Cagatay; Wrobel, Stefan. *Visual analytics for data scientists*. Cham: Springer, pp. 201-218. ISBN: 978 3 030 56146 8

https://doi.org/10.1007/978-3-030-56146-8_7

Avaaz (2020). “Facebook’s algorithm: A major threat to public health”. *Avaaz*, 19 August.

https://secure.avaaz.org/campaign/en/facebook_threat_health

Ávila-Toscano, José-Hernando; Romero-Pérez, Ivón-Catherine; Marengo-Escuderos, Ailed; Saavedra-Guajardo, Eugenio (2018). “Identification of research thematic approaches based on keywords network analysis in Colombian social sciences”. In: Thomas, Ciza (ed.). *Data mining*. London: InTechOpen. ISBN: 978 1 789235975

<https://doi.org/10.5772/intechopen.76834>

- Bastian, Mathieu; Heymann, Sebastien; Jacomy, Mathieu** (2009). "Gephi: an open source software for exploring and manipulating networks". In: *Proceedings of the International AAAI Conference on web and social media*, v. 3, n. 1. *Third international AAAI conference on weblogs and social media*, pp. 361-362.
<https://ojs.aaai.org/index.php/ICWSM/article/view/13937>
- Bayas-Ramírez, Krushenka** (2020). "Visualización del contexto de las *fake news* para entender la infodemia". *#PerDebate*, v. 4, n. 1, pp. 88-108.
<https://doi.org/10.18272/pd.v4i1.1995>
- Bond, Robert M.; Fariss, Christopher J.; Jones, Jason J.; Kramer, Adam D. I.; Marlow, Cameron; Settle, Jaime E.; Fowler, James H.** (2012). "A 61-million-person experiment in social influence and political mobilization". *Nature* v. 489, pp. 295-298.
<https://doi.org/10.1038/nature11421>
- Brennen, J. Scott; Simon, Felix M.; Howard, Philip N.; Nielsen, Rasmus-Kleis** (2020). "Types, sources, and claims of COVID-19 misinformation". *RISJ Factsheet*. Reuters Institute for the Study of Journalism, 7 April.
<https://reutersinstitute.politics.ox.ac.uk/types-sources-and-claims-covid-19-misinformation>
- Brennen, J. Scott; Simon, Felix M.; Nielsen, Rasmus-Kleis** (2021). "Beyond (mis) representation: Visuals in COVID-19 misinformation". *The international journal of press/politics*, v. 26, n. 1, pp. 277-299.
<https://doi.org/10.1177/1940161220964780>
- Cherven, Ken** (2015). "Mastering Gephi network visualization: Produce advanced network graphs in Gephi and gain valuable insights into your network datasets". Birmingham, UK: Packt Publishing. ISBN: 978 1 783987344
<https://www.packtpub.com/product/mastering-gephi-network-visualization/9781783987344>
- Comisión Europea (2018a). "La lucha contra la desinformación en línea: un enfoque europeo". Comunicación de la Comisión al Parlamento europeo, al Comité económico y social europeo y al Comité de las regiones. COM(2018)236.
[https://ec.europa.eu/transparency/documents-register/detail?ref=COM\(2018\)236](https://ec.europa.eu/transparency/documents-register/detail?ref=COM(2018)236)
- Comisión Europea (2018b). "Action plan against disinformation". Join (2018) 36 final.
https://eeas.europa.eu/sites/eeas/files/action_plan_against_disinformation.pdf
- Comisión Europea (2018c). "Code of practice on disinformation".
<https://digital-strategy.ec.europa.eu/en/library/2018-code-practice-disinformation>
- Comisión Europea (2022). *Strengthened code of practice disinformation*.
<https://digital-strategy.ec.europa.eu/en/library/2022-strengthened-code-practice-disinformation>
- Consejo Europeo (2015). "Reunión del Consejo Europeo (19 y 20 de marzo de 2015) – conclusiones".
<https://www.consilium.europa.eu/media/21872/st00011es15.pdf>
- El Español* (2020). "Twitter censurará los mensajes que 'nieguen los consejos de expertos' sobre el coronavirus". *El Español*, 22 abril.
https://www.elespanol.com/omicron/20200422/twitter-eliminar-mensajes-nieguen-consejos-expertos-covid/484453080_0.html
- Ferrari, Anusca; Punie, Yves; Brečko, Barbara N.** (2013). *DigComp: A framework for developing and understanding digital competence in Europe*.
<https://publications.jrc.ec.europa.eu/repository/bitstream/JRC83167/lb-na-26035-enn.pdf>
- Hassan-Montero, Yusef; De-Moya-Anegón, Félix; Guerrero-Bote, Vicente P.** (2022). "SCImago Graphica: A new tool for exploring and visually communicating data". *Profesional de la información*, v. 31, n. 5.
<https://doi.org/10.3145/epi.2022.sep.02>
- High level Group on fake news and online disinformation (2018) "A multi-dimensional approach to disinformation: Report of the independent". Luxemburg: European Commission.
<https://digital-strategy.ec.europa.eu/en/library/final-report-high-level-expert-group-fake-news-and-online-disinformation>
- Instituto #SaludsinBulos; Doctoralia (2019). *II Estudio sobre bulos en salud*.
<https://saludsinbulos.com/wp-content/uploads/2019/11/es-II-estudio-bulos-salud.pdf>
- Ireton, Cherilyn; Posetti, Julie** (2018). *Journalism, 'fake news' & disinformation: handbook for journalism education and training*. Paris: Unesco Publishing. ISBN: 978 92 3 100281 6
- Jiménez-Cano, Rosa** (2017). "Facebook, contra las noticias falsas en España". *El País*, 21 septiembre.
https://elpais.com/tecnologia/2017/09/21/actualidad/1506019098_005465.html?rel=mas
- Larson, Heidi J.** (2018). "The biggest pandemic risk? Viral misinformation". *Nature*, v. 562, n. 309.
<https://doi.org/10.1038/d41586-018-07034-4>
- Martínez-Musiño, Celso** (2011). "Desinformar en la sociedad de la información". En: *Primeras jornadas virtuales iberoamericanas de ciencias de la información y la documentación*, Buenos Aires, octubre 10-30.
<http://eprints.rclis.org/16276/1/Desinformarenlasociedaddelainformaci%C3%B3n.pdf>

- Monnier, Angeliki** (2020). "Covid-19: de la pandémie à l'infodémie et la chasse aux fake news". *Recherches & educations*. HS, Juillet. <https://doi.org/10.4000/rechercheseducations.9898>
- Otte, Evelien; Rousseau, Ronald** (2002). "Social network analysis: a powerful strategy, also for the information sciences". *Journal of information science*, v. 28, n. 6, pp. 441-453. <https://doi.org/10.1177/016555150202800601>
- Oyeyemi, Sunday-Oluwafemi; Gabarron, Elia; Wynn, Rolf** (2014). "Ebola, Twitter, and misinformation: a dangerous combination?". *BMJ*, v. 349. <https://doi.org/10.1136/bmj.g6178>
- Raya, Adrián** (2020). "Facebook quiere ser más transparente: revela de dónde son las páginas". *El Español*, 22 abril. https://www.elespanol.com/omicrono/software/20200422/facebook-quiere-transparente-revela-paginas/484452514_0.html
- Real Academia Española* (2014). *Diccionario de la lengua española*, 23ª ed. <https://dle.rae.es>
- Rivas-Troitiño, José-Manuel** (1995). "Desinformación. Revisión de su significado. Del engaño a la falta de rigor". *Estudios sobre el mensaje periodístico*, n. 2. <https://revistas.ucm.es/index.php/ESMP/article/view/ESMP9595110075A>
- Rosenbaum, Peter** (2018). "Dr Google versus the health practitioner: can we still deliver?". *Developmental medicine and child neurology*, v. 60, n. 6, p. 530. <https://doi.org/10.1111/dmcn.13710>
- Segrelles-Calvo, Gonzalo** (2016). "«Dr. Google»: calidad de la información en la web, limitaciones e impacto en la relación médico-paciente". *Archivos de Bronconeumología*, v. 52, n. 11, p. 573. <https://doi.org/10.1016/j.arbres.2016.04.007>
- Stein, Andrew** (2011). "Dr Google o Dr Lazy ?", *QJM: An international journal of medicine*, v. 104, n. 4, pp. 373-377. <https://doi.org/10.1093/qjmed/hcr004>
- Subires-Mancera, María-Purificación** (2017). "La lucha contra las noticias falsas en Internet". En: Larrondo-Ureta, Aina; Meso-Ayerdi, Koldobika; Peña-Fernández, Simón (coord.), *IX Congreso Internacional de Cyberperiodismo: Innovación y emprendimiento al servicio de las audiencias*. Bilbao: Universidad del País Vasco, pp. 512-526. ISBN : 978 84 9082 783 3 <https://web-argitalpena.adm.ehu.es/listaproductos.asp?ldProducts=USPDF177833>
- Velt, Raphaël** (2011). "gexf-js". <https://github.com/raphv/gexf-js>
- Vosoughi, Soroush; Roy, Deb; Aral, Sinan** (2018). "The spread of true and false news online". *Science*, v. 359, n. 6380, pp. 1146-1151. <https://doi.org/10.1126/science.aap9559>
- Wang, Yuxi; McKee, Martin; Torbica, Aleksandra; Stuckler, David** (2019). "Systematic literature review on the spread of health-related misinformation on social media". *Social science & medicine*, v. 240, n. 112552. <https://doi.org/10.1016/j.socscimed.2019.112552>
- Wen-Ying; Sylvia-Chou; Oh, April; Kleinm, William-M.P.** (2018). "Addressing health-related misinformation on social media". *JAMA*. v. 320, n. 23, pp. 2417-2418. <https://doi.org/10.1001/jama.2018.16865>



SEDIC

Rediseñando lo que somos
para renovar lo que hacemos

<https://www.sedic.es>
<https://intranetsedic.es>
<https://formacionsedic.online>
 c/Gargantilla 13, local 24 Madrid 28005
 +34 639 186 570 | +34 91 593 40 59
sedic@sedic.es



<https://twitter.com/SEDIC20>



<https://www.facebook.com/AsociacionSEDIC>



<https://www.linkedin.com/company/sedic/>



<https://www.instagram.com/sedicasociacion/>

SCImago Media Rankings (SMR): situation and evolution of the digital reputation of the media worldwide

Magdalena Trillo-Domínguez; Ramón Salaverría; Lluís Codina; Félix De-Moya-Anegón

Nota: Este artículo se puede leer en español en:

<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87485>

Recommended citation:

Trillo-Domínguez, Magdalena; Salaverría, Ramón; Codina, Lluís; De-Moya-Anegón, Félix (2023). "SCImago Media Rankings (SMR): situation and evolution of the digital reputation of the media worldwide". *Profesional de la información*, v. 32, n. 5, e320521.

<https://doi.org/10.3145/epi.2023.sep.21>

Article received on 06-08-2023

Final acceptance: 12-09-2023



Magdalena Trillo-Domínguez ✉

<https://orcid.org/0000-0003-0647-2781>

Universidad de Granada
Campus Cartuja
18071 Granada, Spain
mtrillo@ugr.es



Ramón Salaverría

<https://orcid.org/0000-0002-4188-7811>

Universidad de Navarra
Edificio Ciencias Sociales
Campus Universitario
31080 Pamplona, Spain
rsalaver@unav.es



Lluís Codina

<https://orcid.org/0000-0001-7020-1631>

Universitat Pompeu Fabra
Roc Boronat, 138
08018 Barcelona, Spain
lluis.codina@upf.edu



Félix De-Moya-Anegón

<https://orcid.org/0000-0002-0255-8628>

SCImago Research Group
18220 Granada, España
felix.moya@scimago.es

Abstract

In this paper, we present the results of the first global ranking that assesses the status and evolution of the media according to their digital reputation. We analyzed 4,541 general information digital newspapers, both legacy and native, published in 200 countries and in more than 80 languages. Methodologically, we applied a new Digital Reputation Indicator (DRI) as a composite evaluation and measurement tool, weighting web metrics from globally recognized, balanced, and validated sources such as *Semrush*, *Ahrefs*, and *Majestic*. This establishes the foundations of the *SCImago Media Rankings (SMR)* as a balanced watchdog resource for analysis of the digital transformation of the media and as an international catalog of journalistic websites, with a consensus-based and homologous classification and typification that can be applied to different countries. The main conclusion of the first three quarterly waves, carried out between September 2022 and March 2023, is that there is a direct relationship between the degree of socioeconomic and cultural development of the countries and the digital reputation achieved by their media outlets in the ranking. Additionally, the influence of language is notable as a determining factor in the positioning of the media in the ranking. The top positions are occupied by reference digital newspapers located in Western countries of the Northern Hemisphere, especially those in the English language.

Keywords

SCImago Media Rankings (SMR); Digital Reputation Indicator (DRI); Global press; Media system; Digital observatory; Global media map; Digital newspapers; New media; Digitization; Literacy.



1. Introduction

Since their emergence in the 1990s, digital media have passed through various stages of a transformation process that continues to mark the roadmap of news media companies (Deuze; Witschge, 2018). They have moved from the traditional paradigm of the media industry to a horizon of transmedia convergence (Jenkins, 2008; Levinson, 2012; Renó; Flores, 2012; Jarvis, 2015), where new narratives and professional profiles coexist with the volatile consumption habits of user-prosumers (Toffler, 1980; Masip, 2016; Vázquez-Herrero, 2022), moving from print and audiovisual to multimedia.

All these changes take place in a complex context of challenges ranging from the search for business models that truly guarantee the viability of the media to the reinforcement of their democratic role as prescribers of information, watchdogs of power, and battle horses against disinformation. This is not to mention the impact of technological developments such as 3D and artificial intelligence (AI) that are blurring the very boundaries of the media ecosystem and shaking up the very meaning of the profession.

To evaluate this international digital media ecosystem (Scolari, 2021; Orihuela, 2021; Tubella; Alberich, 2012), *SCImago Media Rankings (SMR)* was created:

<https://www.scimagomedia.com>

SMR is a pioneering resource that measures the digital reputation of the media on a global scale with quarterly updates. In testing since September 2022 and publicly accessible since January 2023, *SMR* offers an analysis of the visibility and reputation of the general information digital press on a global scale, and is based on a sample of 4,987 digital newspapers published in more than 80 languages. The resource includes both formerly print (legacy) and native digital media from a total of 200 countries and includes publications with territorial coverage (*SMR* evaluates and compares media with regional, national, and international coverage).

The project has also emerged as a watchdog of the media ecosystem and the processes of transformation and digital evolution, with the purpose of growing in two complementary directions: on the one hand, to expand the size and diversity of the global sample of media, and on the other, to carry out periodic reviews with constant updating. *SMR* publishes all its data openly to facilitate collaborative and sectoral research.

Likewise, it will also be open to contributions from both academia and news media companies, in three directions: (1) to evaluate the consistency of the *Digital Reputation Indicator (DRI)* that classifies the media in the *SMR*; (2) to enrich the sample, deepening the specific knowledge of each media through metadata and reinforcing the presence of digital native media; and (3) to help journalistic media improve their positioning and web reputation, with a constructive and proactive approach. To this end, *SMR* will create a space for participation on the web and will activate a line of advice and consultation from the multidisciplinary team developing the project.

SMR has been prepared by *SCImago Research Group*, a Spanish company specialized in scientometrics, scientific publishing, and web visibility with an important international impact and extensive experience in the field of scientific publications. The project has a board of directors made up of researchers from the *Universidades de Granada, Navarra, and Pompeu Fabra*.

Having a method to rank these characteristics is considered appropriate given that numerous authors have been warning about the deficiencies of the current digital measurement methodologies (Maestro-Espínola, 2015; Solanas; Carreras, 2011) and the need to establish new, more transparent, comparable, and multiplatform systems (Nielsen, 2020; Hernández-Pérez; Rodríguez-Mateos, 2016; Papí-Gálvez, 2014) that connect audience measurement with big data, advancing context and interpretation (Rodríguez-Vázquez; Direito-Rebollal; Silva-Rodríguez, 2018; Livingstone, 2019).

From a perspective that is not only academic (Carlson, 2020), but also professional and commercial (Quintas-Froufe; González-Neira, 2021), the complexity of having adequate metrics, agreed upon and adapted to the digital reality, along with unaddressed challenges such as monitoring based on the individual and not on devices, or introducing massive analysis and AI techniques to increase the accuracy of the data, is noted. In short, we are faced with a scenario of gaps and challenges that justify the opportunity to tackle a project with the ambition and developments presented here.

2. Background and context of the project

SCImago Media Rankings (SMR), the global media ranking resource that is presented herein, provides initial evidence as to the usefulness and application possibilities of the *Digital Reputation Indicator (DRI)*. It consists of a composite evaluation and measurement tool that weights web metrics from global, balanced, and validated sources such as *Semrush, Ahrefs, and Majestic*. This composite metric tool has been designed with the precise purpose of incorporating objective and quantifiable criteria to analyze the media in the digital dashboard.

The objective of *SMR* is to provide a resource for the qualitative comparison of digital media according to a

“ *SCImago Media Ranking (SMR)* offers an analysis of the visibility and reputation of the general information digital press on a global scale, and is based on a sample of 4,541 digital newspapers published in more than 80 languages ”

webmetric model on the basis of their level of citation by other websites (citation flow), the quality of the sites linked to the media (trust flow), and the level of authority associated with their domain (domain rating and a-score). This composite of metrics aims to combine the rigor, objectivity, and precision offered by webmetrics to

present an indicator of the brand value of the media, and in parallel, as a quality factor in the competition for audiences. This can therefore be understood as an avenue of exploration and research that may be of interest to both the academic community and the journalism industry, as more than a quarter of a century of research in digital journalism has already been completed (Salaverría, 2019b).

On the one hand, the *DRI* allows for a uniform comparison to be made for digital media anywhere in the world, regardless of other characteristics, size, coverage, or language. The only specification for evaluating their digital reputation is that they are journalistic media that operate preferentially on the digital platform, that is, that they have at least a significant part of their audience on the Internet.

From this perspective, *SMR* presents an unprecedented international mapping and media classification. To date, there was no global catalog of digital news media based on web metric data, nor was there an analysis indicator that would allow for such an assessment to be made on a global scale. *SMR* structures its ranking according to multidimensional metrics capable of providing, for the first time, a comprehensive approach to the relative position of some media compared with others, as well as to gauge their level of adaptation to the digital environment. On the other hand, the *DRI* allows for analyses and comparisons to be made both on a global scale and in more restricted territorial frameworks, such as continents, countries, or even linguistic spaces.

The webmetric comparison proposed by the *DRI* solves the shortcomings pointed out by several authors (Olteanu *et al.*, 2019; Graves *et al.*, 2010) in previous digital media classification systems. Faced with sporadic sectoral reports on national and international media markets (Peters *et al.*, 2013) and with a general lack of tools to homogeneously compare digital media on a global scale (Carlson, 2020), *SMR* and its composite *DRI* indicator offer a novel, versatile, and reliable methodological resource.

The concept of “digital reputation” on which *SMR* is based measures the media’s efforts to adapt to the digital ecosystem. As explained in detail in the “Material and methods” section, it is presented as an objective and quantitative approximation as to the value of the medium in the digital context, which we construct from a composite indicator (Overall) on the basis of linking and access values (metrics of interest that the domain of the medium captures in the whole of the domains of the Network) through the weighting of four indicators from three validated sources. From this point of view, it is a concept without qualitative connotations that is limited to the behavior and positioning of the media on the basis of the assessment of the site by the digital audience as a whole.

Building on the development of the *DRI* outlined above, we propose to go a step further, with a first practical application focused on the global news media industry, and with the following research questions:

- Q1: Is it possible to generate a world map of the media system on the basis of the big data provided by the Internet?
- Q2: Using media mapping through internet big data as a basis, is it possible to develop longitudinal studies on the media as they adapt to the digital medium?
- Q3: Is it feasible to draw conclusions from the territorial distribution of the media included in the map and connect it to other global indicators such as the representation of Internet users and other socioeconomic factors related to the use of technology?
- Q4: Assuming that languages play a fundamental role in the world of communication, is the global digital reputation map affected by the language of each medium, and does language have any influence on the medium’s own digital reputation?

3. Material and methods

The *Global Media Ranking* brings together two complementary research and development actions: on the one hand, the design of the *DRI* as an original tool for analyzing journalistic sites in the digital ecosystem; and on the other hand, the application and generation of the media map that constitute the object of study. It is important to note that other available rankings do not have this webmetric approach nor global coverage, but rather respond to specific concerns in concrete geographical, historical, and socioeconomic contexts, or to strictly commercial conditioning factors more closely related to industry interests. With these precedents, we outline the process carried out to create the first media sample representative of the sector for the application of the *Digital Reputation Indicator (DRI)*.

3.1. Process and criteria for sample selection

In the absence of a global list of open access media, the following sources have been used for the *SMR* media search.

“The project has also emerged as a watchdog of the media ecosystem and the processes of transformation and digital evolution”

- Directory of printed newspapers of *Kiosko.net*
<https://es.kiosko.net>
The front pages of hundreds of newspapers worldwide are published daily. Each of them is accessed manually, and all domains are retrieved. This is, however, a very limited first approximation.
- Directory of *PrensaEscrita.com*
<http://www.prensaescrita.com>
A comprehensive list of daily newspapers (differentiating between legacy and digital natives) published worldwide. All domains are retrieved except for strictly local media. It has a marked Hispanic media catalog bias that is corrected with the incorporation of more sources.
- World list of *Online Newspapers*
<http://www.newspaperindex.com>
Major digital access newspapers in all countries and journalistic news sources, with a special focus on current affairs, politics, and economics. It is contrasted with the domains already retrieved. This is a list focused on the media with the greatest circulation and impact.
- Catalog of *W3 Newspapers*
<http://www.w3newspapers.com>
Extensive list of media including newspapers as well as news sites and magazines. This source is used to complete and expand the list in key countries such as the United States, United Kingdom, France, Germany, Italy, Russia, and China. Daily newspapers are selected (although it also includes an extensive list of non-daily newspapers).
- World directory of *Online Newspapers*
<http://onlinenewspapers.com>
It includes thousands of newspapers, as well as websites and audiovisual information sources (radio and television) and specialized media. It is used on an ad hoc basis for supplementation purposes.
- *ComScore* listing
<https://www.comscore.com>
In the case of Spain and Latin America, the media auditing company collaborates with the *SCImago Research Group* project by providing its internal lists.
- Database of digital media in Spain and Portugal from the *Iberifier* project (**Salaverría et al.**, 2022).
<http://map.iberifier.eu>
- In addition, the digital media listings of:
 - Media Information Association (AMI)*
<https://www.ami.info>
 - Office for the Justification of Diffusion (OJD)*
<https://www.ojd.es>
 - General Media Study (EGM) of the Association for Media Research (AIMC)*.
<https://www.aimc.es/egm>

From these digital media directories and databases, a sample was obtained according to the following selection criteria:

- General news media with an online presence, originating from both print media and digital native publications.
- Digital media with their own domain. Publications with only subdomains (i.e., *www.dominio.com/subdominio* or *subdomain.domain.com*) are excluded, as is often the case with vertical gateways or publications promoted by certain media groups.
- Media with national-global, regional, and “glocal” scope, as explained by **Gómez-Mompart** (2013).

SMR includes several descriptive variables of the media aimed at serving as a thermometer of the industry’s reality and its evolution:

- Typology: general, sports, economic, and other media (this last category is intended to serve as a resource to provisionally classify digital media specialized in other topics that will eventually be subject to future classification).
- Location: location of the media by country and by region (geographic and sociopolitical).
- Scope: international, national, regional/state.
- Language: includes both the main language used by the medium and additional languages.
- Categorization: in its initial phase, *SMR* classifies digital editions of print newspapers, digital native media,

“*SMR* has been prepared by *SCImago Research Group*, a Spanish company specialized in scientometrics, scientific publishing, and web visibility with an important international impact and extensive experience in the field of scientific publications. The project has a board of directors made up of researchers from the *Universidades de Granada, Navarra, and Pompeu Fabra*”

and large news agencies. In later development, specific categories for digital media originating from radio, television, and magazines will be included.

3.2. Sources for generating the *Digital Reputation Indicator (DRI)*

The design of the *Digital Reputation Indicator (DRI)* that was used for the preparation of the media ranking presented in this paper results from the triangulation of data obtained through *Semrush*, *Ahrefs*, and *Majestic*, three tools for analysis and data tracking for search engine optimization (SEO) that are consolidated and recognized by the web metrics sector at an international level and with recent studies corroborating their reliability in the analysis of domain authority (Reyes-Lillo; Morales-Vargas; Rovira, 2023).

- *Semrush* (2023), founded in 2008 by Oleg Shchegolev and Dmitry Melnikov as a comprehensive SEO analysis software, is used for keyword tracking, competitor research, search result position monitoring, and general website auditing. It also offers online advertising and social networking applications.
- *Ahrefs* (2023) has become one of the leading SEO analysis tools in the industry. Founded in 2010 by Dmitry Gerasimenko, it is known for its massive link database and its ability to trace links to websites and provide information about a website's link profile. Like *Semrush*, *Ahrefs* also offers applications for keyword research, competitor research, and search result position monitoring.
- *Majestic* (2023) is a SEO analysis tool that focuses on link research and backlink building. Founded in 2004 by Dixon Jones, *Majestic* also has a massive link database, and key to our study is its ability to trace links to websites and analyze them. In addition, *Majestic* also offers applications for keyword research and search result position monitoring.

"Positioning" refers to the position of a web site in the results page of a search engine for a certain keyword. The higher the position of a website, the more likely it is to be visited by users, and therefore the greater its digital influence. The algorithms of search engines such as *Google* use several factors to determine the position of a website, including relevance and site authority.

Connected to this concept is that of "visibility," i.e., the ability of a website to be found and seen by online users: the more visible a website is, the more likely it is that users will visit it. Visibility can be measured by the number of impressions (appearances) of the website on the first pages of search engine results.

In such searches, the appropriate use of keywords in news content is fundamental, since these are the terms or phrases with which users retrieve content from search engines to find information online. The use of applicable keywords is important in improving the positioning and visibility of a website. From this perspective, SEO analysis tools identify the suitability of keywords in content on the basis of their relevance and search volume.

Furthermore, backlinks (inbound links) are links that come from an external website. Backlinks are used by search engines to determine the authority and relevance of a website, which can translate into a better position on the search engine results page. In this case, SEO analysis tools identify and analyze the backlinks pointing to a website, as well as evaluate their quality and relevance.

Semrush, *Ahrefs*, and *Majestic* operate independently, use their own databases, and are not part of any holding or group, which is considered relevant in our study because of the data triangulation achieved by combining three approaches with independent development.

Although the most widespread use of these three tools is focused on the SEO and digital marketing industry, aimed at helping companies improve their online visibility and performance in search engines, they prove to be resources with great potential in academic research on the media system, including the more industrial dimension of media audits.

3.3. Composite indicator calculation and robustness testing

The *Digital Reputation Indicator (DRI)* combines the values of a total of four metrics obtained from *Semrush*, *Ahrefs*, and *Majestic*, with which it calculates the composite indicator (Overall¹). The metrics are as follows:

- **Authority Score (*Semrush*):** The A-Score measures a domain's authority on the basis of two main dimensions: the quantity and quality of a domain's inbound links and its estimated average monthly traffic.
- **Domain Rating (*Ahrefs*):** Measures the strength and authority of a website on the basis of its link profile. It is mainly calculated by evaluating a website's backlinks and social media posts. *Domain Rating* uses the number of unique domains linking to a website and the quality of those links. The higher the domain rating of a website, the higher its authority and the more likely it is to rank well in search results.
- **Citation Flow (*Majestic*):** Evaluates the authority of a website on the basis of the number of links it receives, without taking into account other factors.
- **Trust Flow (*Majestic*):** Measures the quality of links pointing to a website. *Trust Flow* is based on the idea that links from trustworthy websites carry more weight than links from low-quality websites. Thus, the higher the trust flow of a website, the higher the quality of the links pointing to it.

The weight assigned to these four metrics to calculate the composite indicator (Overall) is equal, at 25% each. This weighting provides triangulated data of high robustness and consistency, as shown in the “Results” section. In any case, *SMR* also offers disaggregated access to each of the four metrics by which the composite indicator (Overall) is calculated.

In short, the combination of the number of inbound links (citations), their quality (relative weight of each link), and the estimation of traffic are the basis of the integral *DRI* indicator that we propose to measure the web reputation of a media outlet in its digital aspect. With this precision, it is important to emphasize that “digital reputation” is what is being referred to at all times. The indicator does not make a statement as to the intrinsic journalistic quality of each media outlet, which requires different indicators.

In the following section, the results of three robustness tests of the *SMR* that have been carried out on the selected sample will be presented. The first corresponds to the pilot version of the ranking in September 2022 (Q3 edition 2022); the second was carried out in December of the same year (Q4 edition 2022); and the third is already in the first quarter of this year (Q1 edition 2023)². In all three cases, the results are published openly on *SMR*’s website, where the progress of the research is explained. This accessibility of the data allows for the verification of these tests by third parties, as well as for the eventual development of specific studies. As developed in the “Results” section, a global positioning scale is set from 0 to 100, and between the different tests to assess the consistency of the indicators, a period of 3 months is given to analyze the evolution of the media in the digital environment as well.

4. Analysis of results

In response to the dual usefulness of the *SCImago Media Rankings (SMR)* for both the academic and professional worlds, and given its international focus, the development of the project in which the *Global Digital Reputation Media Ranking* is framed is being carried out through a website in English that prioritizes transparency, open data, and participation: <https://scimagomedia.com>

The objective is to allow for personalized and corporate access to all content, to prioritize the usability of the interface, and to facilitate the downloading and visualization of data, as well as interaction and feedback with the research team as a constructive space for debate and discussion.

With these criteria, two key areas have been enabled on the project site: access to the ranking itself, which allows for individual media searches as well as searches by country, language, and region (political and geographic). This search proposal allows for fully customizable comparisons, useful for both news media companies and media researchers and consultants, as well as those that are evolving.

As mentioned in the “Material and methods” section, each media outlet includes a summary of the indicators applied to determine its digital reputation (*authority score, domain rating, citation flow, and trust flow*) and a series of identification and typification metadata³ (location, coverage, support, languages, and basic context information).

Complementary to this access, a visualization space (*Viz Tool*) has been developed with the idea of generating graphics that allow for analysis of the results, showing the evolution of the media system and contributing to the debate on the dynamics and opportunities of its evolution and digital development. In this sense, both the blog space opened on the web with observations made by the research team developing the project and the critical dissemination that is being carried out on social networks (especially on *Twitter*) are complementary.

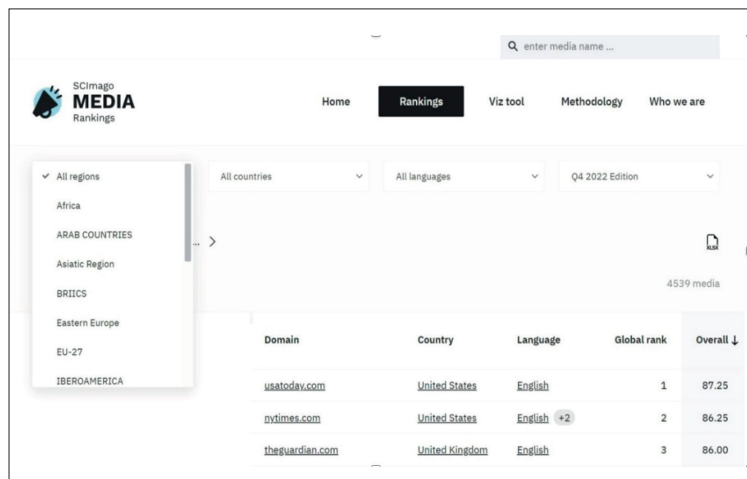


Figure 1. Project navigation menu interface. Source: *SCImago Media Rankings*.

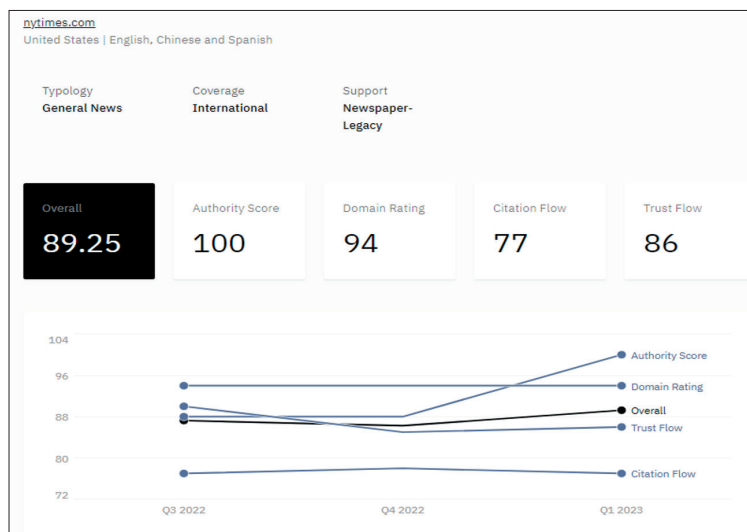


Figure 2. Detailed information for each media outlet (in this case *The New York Times*) with the *DRI*, evolution, and characterization and typification metadata. Source: *SCImago Media Rankings*.



Figure 3. Overall media distribution of the first *Global Digital Reputation Ranking* focused on the global press. Source: *SCImago Media Rankings*.

As a first analysis of the results of the *Global Digital Reputation Ranking* generated around the journalism industry, the general distribution of newspapers is focused on observing how a connection between the greater presence and weight of the sites analyzed and the factors and conditioning factors that have historically come into play in the development and evolution of media systems are generated (Reig; Labio, 2020). The socioeconomic position of the respective countries, the evolution of technology and Internet adoption in different regions, and more cultural aspects such as the level of digital literacy are all referred to. In this sense, it can be observed how a very powerful media system is emerging in Europe, the United States, and Latin America.

In the selection of the sample, due to a conditioning factor of the sources used for this purpose (which will be corrected for in future waves of analysis), there is still a greater weight of printed newspapers (compared with digital natives) and of Spanish- and English-language media, a nuance related to the typology of media analyzed.

In future waves, the sample should grow in native media and include a greater number of newspapers from regions such as Asia-Pacific and Africa (despite the difficulties posed by the characteristics and lack of transparency in their own media systems, including important legal and political development limitations) to improve global representativeness, and above all, to act as a watchdog in the intense and accelerated process of digitization in which the entire sector is immersed, especially following the impact of the coronavirus disease 2019 (Covid-19) pandemic.

The potential that the Asia-Pacific region is acquiring at a global level is not only indicative of the weight and projection of countries such as China at an economic level, with the international geopolitical debates questioning even the possibility that it might dispute the United States' world leadership (Bustelo, 2022), but also of that of India from the population point of view, with the recent news that it is already the most inhabited area on the planet (Sánchez, 2023). The technological aspect itself, focused on cell phones, is also an illustrative factor of this, as can be seen in Figure 4 on the use and prospects for the implementation of 5G technology worldwide.

In parallel to the factors that come into play in the digital transformation of the media, if the focus is on analyzing their digital reputation –on how they are facing the responsibility of making the leap into cyberspace and carrying out the best possible journalism for audiences in the web ecosystem– both the final ranking (Table 2) and the image on the territorial distribution of the media according to their digital reputation (Figures 9 and 10) present a very strong and direct connection with the global map on the implementation of the Internet (Figure 5).

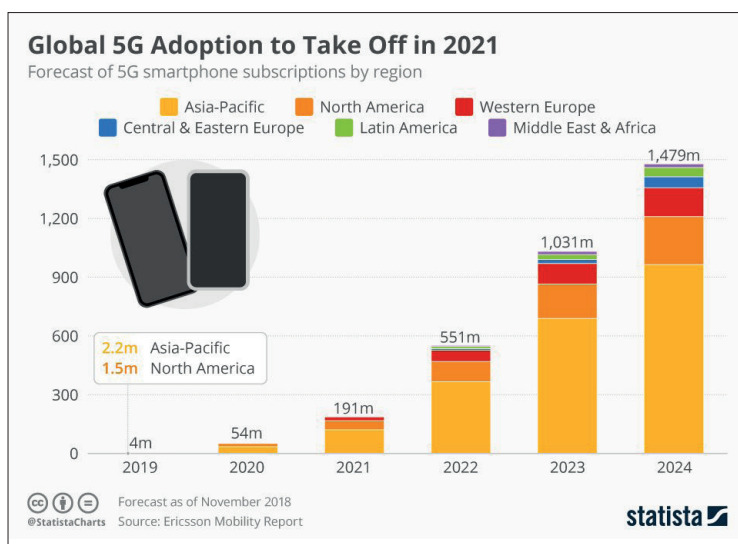


Figure 4. Global expansion of 5G mobile network technology by region. Source: *Ericsson Mobility Report*.

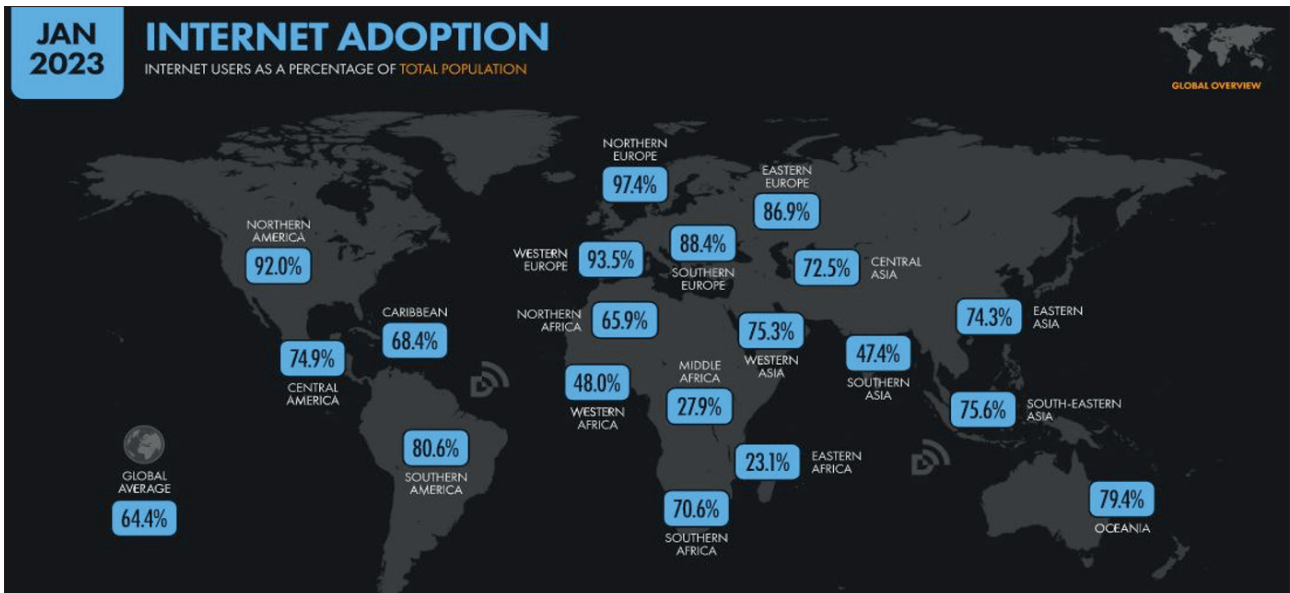


Figure 5. Internet adoption by country, updated to 2023. Source: *DataReportal* in collaboration with *Melwater* and *We Are Social*.

In reality, it can be understood that this is a final image that visualizes the status and evolution of countries and regions, and that ends up being the result of the confluence of all the factors previously indicated in the map regarding the death of paper newspapers.

As can be seen in Figure 5, the number of Internet users has grown again in the last year (1.9%) and now reaches a total of 5.16 billion people connected worldwide. Around the year 2000, the number of Internet users around the world was 396 million, a figure that had doubled by 3 years later (761 million in 2003), with exponential growth until 2021 when the curve began to flatten, a result of the end of the pandemic confinements and the gradual recovery of normality.

The two regions with the highest Internet penetration are Northern Europe (97.4%) and Western Europe (93.5%), followed by North America (92%), while the two regions with the lowest Internet penetration are Central Africa (27.9%) and East Africa (23.1%).

Connecting the map of Internet penetration and access, and looking at our contribution with the distribution of print and digital newspapers by region (Figure 6), it is precisely the European Nordic countries that occupy the first position overall (with representation based on the median excluding outliers), a position that speaks to the strength of its sample and the high average quality of the sites included. In this figure of geographical and political distribution, the position of the Asia-Pacific region also emerges as a developing focus of expansion in the economic, technological, and media fields.

From the point of view of regional analysis, we can render a more accurate reading by utilizing geographic zones, as seen in Figure 7. Apart from the reading of the medians, conditioned by the limitations and biases of the sample, as previously mentioned, it can be seen how the best media (with the best digital reputation) are located in the orbit of Western Europe and the United States (as can be seen in the figure, some media have above 90 for Overall). In fact, it is these two

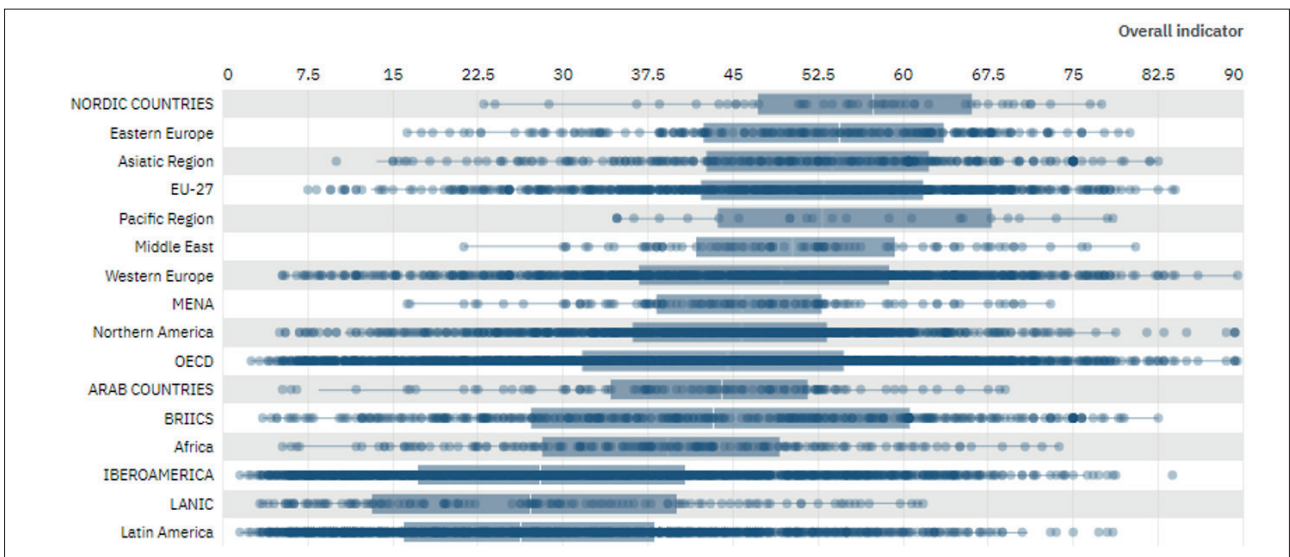


Figure 6. Distribution of media by region (geographic and political), ordered by median. Source: *SCImago Media Rankings*.

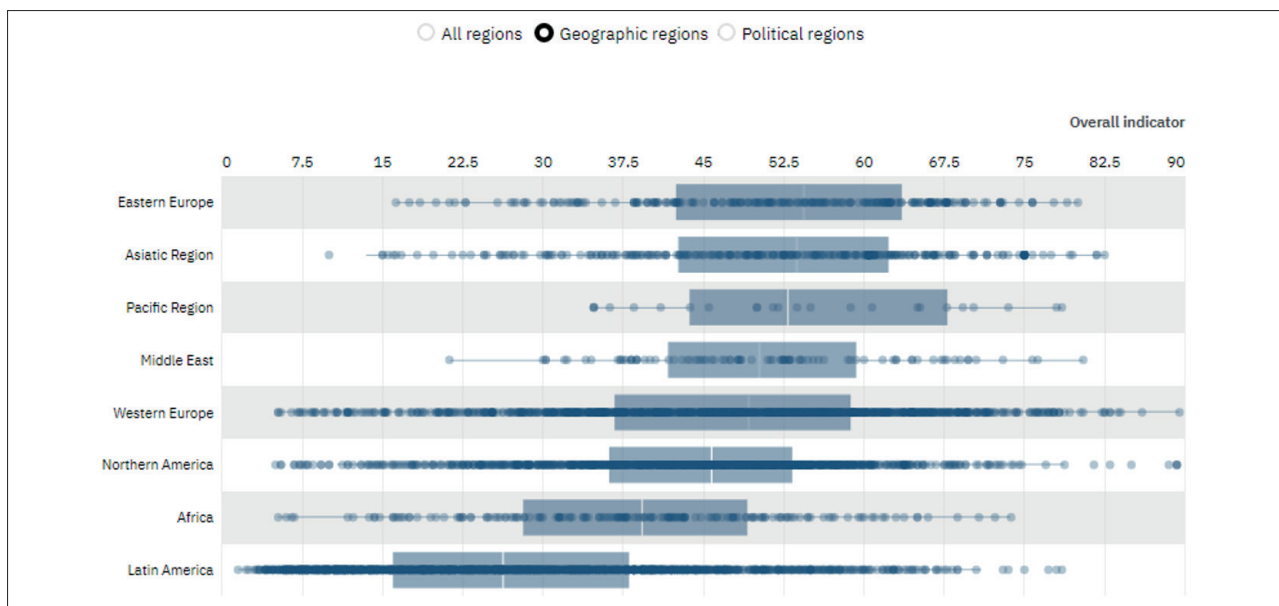


Figure 7. Distribution of media by geographic region, ordered by median. Source: SCImago Media Rankings.

regions that distribute media along the entire scale and that are in a more powerful relative position. In the case of Latin America, in last position in both geographic and political distribution, this is a circumstance conditioned by the strong presence of media from this region in the analysis sample and their heterogeneity.

The median ranking, regardless of the number of media in each region, should not be understood as a reflection of the quality of the media in each area, but rather as a reference that allows for comparability. In other words, it says more about the sample than about the quality of the media, and in this sense, the ontological argument about “what is seen at a great distance versus what is seen at close range” could be referenced. The limitation discussed in the research on the Western and Hispanic focus that characterizes this first selection of media can be observed here.

Along with regional readership, the distribution and positioning of the media according to language is very relevant. The general map of media included in this first sample includes newspapers published in more than 80 languages, dozens of them with bilingual and multilingual editions in response to their objective of global reach. There is a clear weight of the most widely spoken languages, such as English, Spanish, Portuguese, and Chinese, as well as unique journalistic projects that reflect the linguistic richness and cultural diversity of this first global ranking.

Figure 8, which compares the Spanish and English media, shows how the consolidation of English as the language not only of science but also of the web –of the digital ecosystem– has a direct impact on the media map. The same is true for the Spanish case when considering the world population that a Hispanic website can potentially reach due to quantitative population.

In this approach from the perspective of digital reputation, it can be observed how one of the characteristics of the media systems that is at the base of the large conglomerates with global projection is repeated: the socioeconomic and cultural characteristics of the country have an important weight, as well as the normative regulations and the historical tradition itself, but factors related to technology (greater or lesser penetration and access to the Internet) and language become determinants of their international trajectory and visibility. The *DRI* is based on access and link indicators, and thus elements such as the potentially objective Internet-connected population of the medium and the language filter are, evidently, fundamental.

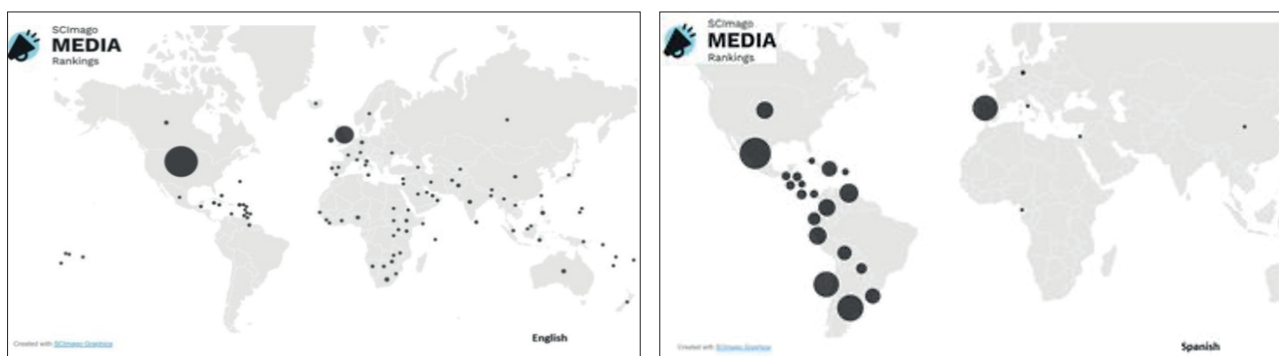


Figure 8. Distribution of media in English and Spanish. Source: SCImago Media Rankings.

A relevant datum that was extracted by analyzing the characteristics of access to the media outlets included in the sample refers to international traffic. *Semrush* was used as a source, with a reference sample of 4,500: 7% of the journalistic sites analyzed do not have a registry (317 media) and would therefore be left out (63 of them are Chinese, so the lack of data could be due to their restrictive regime regarding computer freedom).

A total of 4.6% (208 media) have more than half of their traffic outside their country of origin ($\leq 50\%$), and only 40 media publish a version in more than one language (within the same domain), which is less than 1%.

4.1. Comparison of the first three waves

From an evolutionary point of view, while waiting to reach the first year since the ranking was launched in autumn, the three analyses carried out allow for the observation of the general behavior of the media and for some initial comparisons to be made. As an overview for the analysis of the digital reputation of the media, Figure 9 shows how the bulk of newspapers (legacy and native) are located in the central part of the scale (out of 100), around a third of the sample occupies a very low position, while a small elite holds the top positions as global benchmarks in the journalistic field of general information.

In Figures 9 and 10, what can be observed is a behavior similar to what happens with other websites when facing the challenge of competition in the digital ecosystem: entry is relatively easy, as is climbing positions in the lower part of the ranking, but it is increasingly difficult to advance when in positions at the top. In the context of the 1–100 scale used by the *DRI* as a simple numerical matter, the differences in positions in the intermediate zones are very small, and become more demanding at the extremes, especially in the leadership zone.

In contrast, when analyzing global behavior among the first three waves used in the robustness test (Table 1), it can be seen how there are more than 1,000 media outlets that change more than 100 positions in the analysis interval. However, although such a figure might seem high, when looking at the Overall results, we see that only 100 of them change the value of the indicator by more than 10%. Most of them, moreover, are in the central zone of the ranking, where small changes can cause significant rises or falls in position.

Table 1. Overall position changes in the waves: Q3–22, Q4–22, and Q1–23.

Q	U S A	G B R	F R A	D E U	E S P	I N D	R U S	I T A	A U S	C H N	T U R	A R G	S W E	N L D	J P N	C A N	U K	I N	Z A	C O L	B R M	I S R	A U S	U K	B R	P R	F I N	B R A	C H I	I R	
Q3-22	20	10	8	8	6	4	7	4	4	7	1	3	2	2	3	2	4	1	0	1	0	0	0	0	0	0	0	0	1	0	2
Q4-22	20	11	7	8	6	7	8	4	5	7	1	2	2	2	3	2	0	0	1	0	0	0	1	0	0	0	0	1	0	0	
Q1-23	14	11	8	7	7	6	6	5	3	3	3	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	

From a comparative point of view, a general assessment can also be carried out as to how the languages in which the different media outlets publish affect their respective digital reputation rankings. The tentative tendency to incorporate

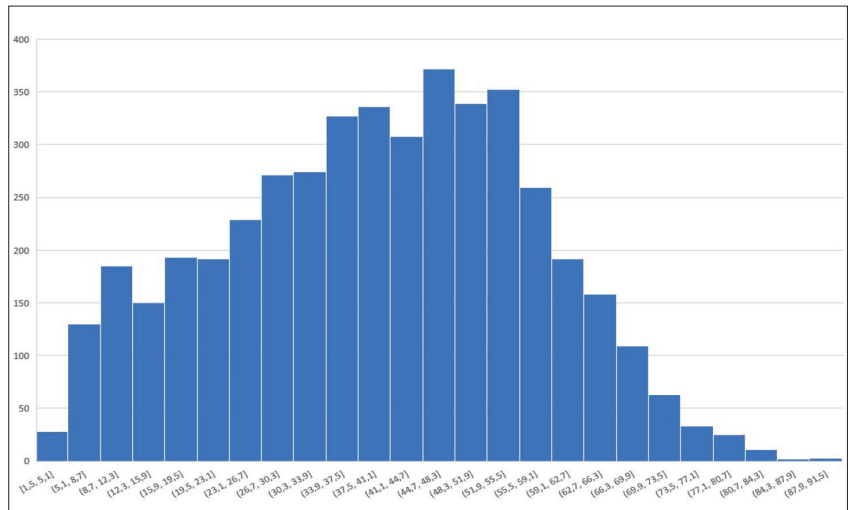


Figure 9. Distribution by rank of the 4,500 media (Q1 2023 Edition). Source: *SCImago Media Rankings*.

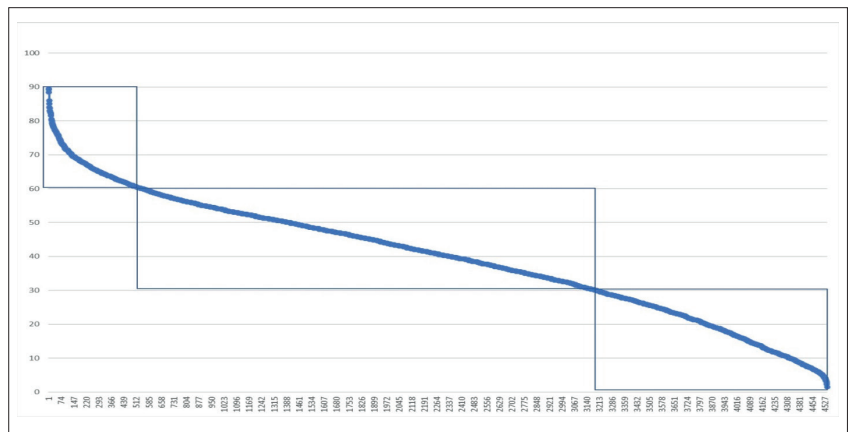


Figure 10. Distribution of media according to their Overall digital reputation. Source: *SCImago Media Rankings*.

more than one language (primarily English) is noticeable, in line with the observations already made above. This is another point of analysis to be corroborated and evaluated in future waves.

Table 2. Incidence of the language in the top 100 positions (+ indicates that, in addition to the main language, it coexists with other languages).

	Monolin- gual	More than one language	English	English +	Spanish	Spanish +	French	French +	German	German+
Q3 2022	94	6	45	51	9	11	8	10	7	9
Q4 2022	93	7	44	51	8	10	7	9	9	11
Q1 2023	90	10	38	47	11	14	8	10	8	10

The distribution of media by region shown in Figure 11 provides a global assessment of the volume of the selected samples (the point clouds show each of the sites analyzed according to their Overall position). From this point of view, the weight of the media environments in Latin America, Western Europe, and the United States is apparent, as these are the regions that make up the countries with the largest number of selected media. As can be seen, all the regions present a diamond shape that is less stylized (with a bulk of media in intermediate positions and a small group of leading sites), except in the Latin American case, which is presented in the form of a fir tree, showing a sample effect (media outlets in intermediate and low positions stand out).

“ The *Digital Reputation Indicator (DRI)*, used for the preparation of the media ranking presented in this paper, results from the triangulation of data obtained through *Semrush, Ahrefs, and Majestic*, three tools for analysis and data tracking for search engine optimization (SEO) that are consolidated and recognized by the web metrics sector at an international level ”

Altogether, this distribution is considered representative of the projection and global position that the media outlets occupy on the global map, incorporating socioeconomic factors, transparency policies, and indicators such as digitalization and literacy. The evolving analysis of this representation, and with the expansion of the sample being carried out in the regions of the East, Asia, the Pacific, and Africa, will also help in assessing in general terms how their media systems are developing.

The qualitative approach of this figure, which has to do with the Overall results of the different regions, is the one that already provides valuable information as to the quality of the media outlets according to their digital reputation when the boxplot that appears in gray on the point clouds of the different samples is analyzed. With this schematic representation, which will also facilitate longitudinal analysis, the distribution of means by quartiles and the average value in each region can both be seen. In the areas with a smaller number of websites included in the sample selection, a certain uniformity is observed in relatively high positions, which has to do with the above-mentioned sample effect as well as with the uniqueness of their media systems. It is noteworthy that the breadth for Pacific, Eastern Europe, and Asia is greater than that of the United States or Western Europe (media outlets with a medium-high Overall predominate), although

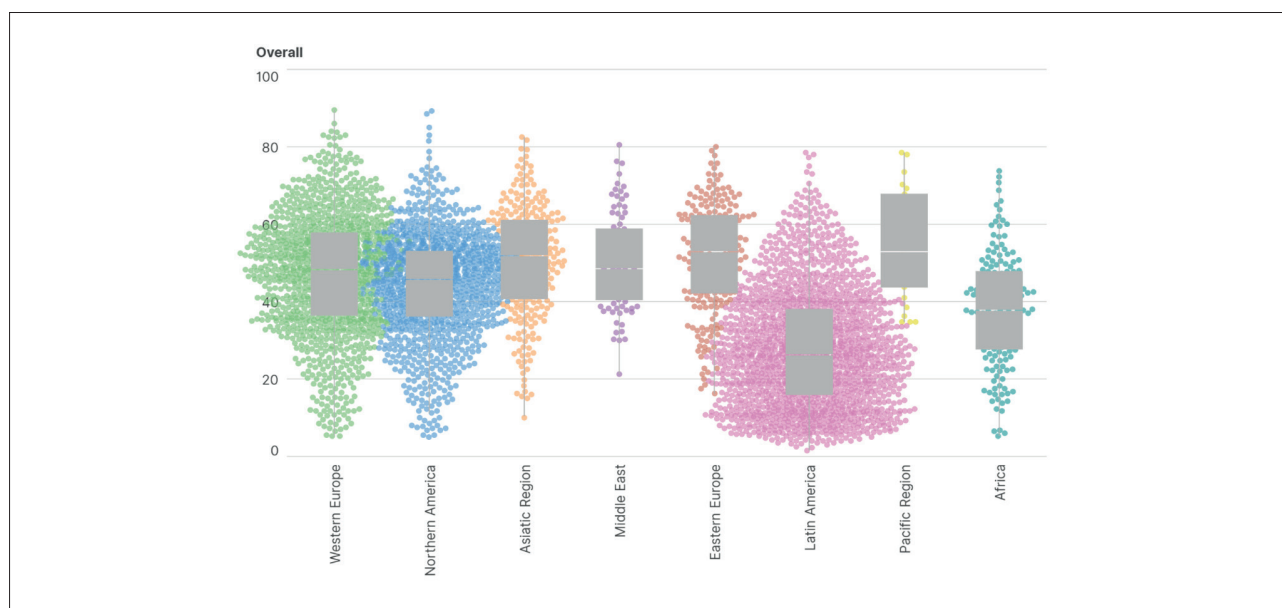


Figure 11. Comparison of media distribution by region according to Overall in the latest edition published. Source: SCImago Media Rankings.

the position of the North American and European presses becomes evident when analyzing which general information newspapers lead the global map.

In fact, when the changes in the positions of the media outlets are evaluated globally according to their digital reputation (Table 3), within the top 100 positions there are 69 journalistic brands that remained in the three quartiles within the top 100, mainly US and British media.

Likewise, if the analysis is broadened and the Overall result for each sample is averaged, we start from an overall position of 42.72 in the September wave, see a slight rise of 0.4 points in the December wave (43.12), and, breaking the trend, observe a drop of almost 3 points in the last wave carried out in March of this year (2.93 compared with the first one in September and 3.33 compared with the one in December).

However, when observing the leading positions, it is illustrative that, in contrast to the slight general worsening of the sample, there is a significant increase in the positions of the media outlets that are at the top: in the first ranking, *The New York Times* was in the leading position with an Overall of 87.25, followed by *USA Today* with 86.75 and *The Guardian* with 86.50. In the March edition, *The Guardian* ousted the US newspapers with 89.50, a quarter of a point ahead of *The New York Times* and ahead of *USA Today* (88.50) as well, which took first place in December but now slips to third.

The Overall analysis of the three waves allows for an important change to be distinguished in the maximum results when comparing the upper limit 87.25, which has remained constant in Q3 and Q4 of 2022, with 89.50 in Q1 of 2023. This generalized change in the overall ranking score may be due to the recent reformulation of *Semrush's Authority Score (AS)*, although this is an assessment that we will have to verify with new editions of the ranking.

In general, the dispute for the top positions is very fluid: all of them register advances in positions and improve their own records, but in the comparison, the positions are always interchanging. Pending new editions, what has been observed so far is that there is a prominent media elite at the international level, led by the most emblematic American and British newspapers (*The Washington Post*, *The Independent*, and the English tabloids join the aforementioned media outlets).

In the European case, the leadership tends to be quite stable with very consolidated positions and global reach, such as the Spanish *El País*, the French *Le Monde*, the German *Der Spiegel*, or the Italian *La Repubblica*. In the case of Spain, *Prisa's* newspaper *El País* is the first to appear in the global ranking, with a very significant rise between the first and last sample, from 19th to 7th place.

In Latin America, *Clarín*, *La Nación*, *Infobae*, and *O Globo* compete for the top positions, while in other regions such as Asia, two media outlets are clearly prominent, driven by the multilingual editions of their *sites*: the *Renmin Ribao/People Daily* newspaper and the *Xinhua* news agency. In the Pacific, there are emblematic newspapers from Australia such as *The Sydney Morning Herald* or *The Daily Telegraph*, of the *News Corporation* emporium, and from Japan, *Asahi Shimbun's* position (with English edition) is significant as one of the world's best-selling newspapers.

In general terms, in these first three waves, it is noteworthy that small changes in the positioning of the newspapers can already be seen, reflecting the movements that are taking place in the digital media landscape, largely as a result of the active processes of digital drive and transformation in which journalistic companies are immersed. Table 1 presents the world's top 50 presses according to their digital reputation.

5. Limitations and discussion

As mentioned above, *SCImago Media Rankings (SMR)* analyzes a relevant and qualified sample of the digital press at the global level. However, the ranking leaves out numerous digital publications that will gradually be incorporated into future updates. As the initial sample expands and *SMR* assigns greater prominence to countries with less developed journalistic ecosystems, the results obtained in the first waves will probably have to be put into context. The sample selection process has resulted in an over-representation of digital media based in English-, Spanish-, French-, and Portuguese-speaking countries, while other languages, such as Chinese, Hindi, or Arabic, have a relatively more modest presence in the ranking. The gradual inclusion of these less studied regions will undoubtedly lead to an enrichment of the results, as well as, in all likelihood, to a greater degree of nuance.

Another bias of the *SMR* in its initial versions is the preferential focus on digital publications originating from print media and, to a lesser extent, on digital native media. The project plans to include digital media linked to audiovisual media (radio and television) in future updates, which may also contextualize some of the results described in this study.

In addition, there are plans to expand the registry of specialized digital media, starting with sports and economy. These media outlets, which in some cases have a great

In this study, we present the results of three robustness tests of the *SMR* that have been carried out on the selected sample. The first corresponds to the pilot version of the ranking in September 2022 (Q3 edition 2022); the second was carried out in December of the same year (Q4 edition 2022); and the third is already in the first quarter of this year (Q1 edition 2023)

Table 3. Media outlets making up the global Top 50 for digital reputation. Q1 edition 2023.

Position	Media outlet	Domain	Country	Rank	Overall
1	THE GUARDIAN	theguardian.com	United Kingdom	1	89.50
2	THE NEW YORKTIMES	nytimes.com	United States	2	89.25
3	USA TODAY	usatoday.com	United States	3	88.50
4	THE INDEPENDENT	independent.co.uk	United Kingdom	4	86.00
5	THE WASHINGTON POST	washingtonpost.com	United States	5	85.00
6	LE MONDE	lemonde.fr	France	6	84.00
7	EL PAÍS	elpais.com	Spain	7	83.75
8	NEW YORK POST	nypost.com	United States	8	83.00
9	THE DAILY TELEGRAPH	telegraph.co.uk	United Kingdom	8	83.00
10	THE SUN	thesun.co.uk	United Kingdom	8	83.00
11	DAILY MIRROR	mirror.co.uk	United Kingdom	11	82.50
12	DER SPIEGEL	spiegel.de	Germany	11	82.50
13	THE TIMES OF INDIA	timesofindia.indiatimes.com	India	11	82.50
14	THE DAILY MAIL	dailymail.co.uk	United Kingdom	14	82.25
15	ASAHI SHIMBUN	asahi.com	Japan	15	81.75
16	THE WALL STREET JOURNAL	wsj.com	United States	16	81.50
17	BILD	bild.de	Germany	17	80.50
18	HÜRRİYET	hurriyet.com.tr	Turkey	17	80.50
19	DAILY EXPRESS	express.co.uk	United Kingdom	19	80.25
20	DNES	idnes.cz	Czech Republic	20	80.00
21	HINDUSTAN TIMES	hindustantimes.com	India	21	79.50
22	THE INDIAN EXPRESS	indianexpress.com	India	22	79.25
23	LA REPUBBLICA	repubblica.it	Italy	22	79.25
24	RIA NOVOSTI	ria.ru	Russia	24	79.00
25	ABC	abc.es	Spain	25	78.75
26	SAN FRANCISCO CHRONICLE	sfgate.com	United States	25	78.75
27	INFOBAE	infobae.com	Argentina	27	78.50
28	SYDNEY MORNING HERALD	smh.com.au	Australia	27	78.50
29	ANSA	ansa.it	Italy	29	78.25
30	CORRIERE DELLA SERA	corriere.it	Italy	29	78.25
31	CLARÍN	clarin.com	Argentina	31	78.00
32	NEWS.COM.AU	news.com.au	Australia	31	78.00
33	FOCUS	focus.de	Germany	33	77.75
34	KOMSOMOLSKAYA PRAVDA	kp.ru	Russia	33	77.75
35	LE PARISIEN	leparisien.fr	France	33	77.75
36	EXPRESSEN	expressen.se	Sweden	36	77.50
37	THE HINDU	thehindu.com	India	36	77.50
38	LA NACIÓN	lanacion.com.ar	Argentina	38	77.25
39	LE FIGARO	lefigaro.fr	France	38	77.25
40	OUEST FRANCE	ouest-france.fr	France	38	77.25
41	DE TELEGRAAF	telegraaf.nl	Netherlands	38	77.25
42	HOUSTON CHRONICLE	chron.com	United States	42	77.00
43	EVENING STANDARD	standard.co.uk	United Kingdom	42	77.00
44	EL ESPAÑOL	elespanol.com	Spain	44	76.75
45	KOMPAS	kompas.com	Indonesia	44	76.75
46	DIE ZEIT	zeit.de	Germany	44	76.75
47	AFTONBLADET	aftonbladet.se	Sweden	47	76.50
48	EL MUNDO	elmundo.es	Spain	48	76.25
49	NU	nu.nl	Netherlands	48	76.25
50	SABAH	sabah.com.tr	Turkey	48	76.25

reputation in and influence on the networks, are a strategic pillar in the transformation and evolution of the media to the new digital ecosystem and require specific analysis.

Critically speaking, presenting a digital reputation ranking is key, given that the media system is currently experiencing a reconfiguration after the impact and acceleration of the digitalization processes brought about by the Covid-19 pandemic (WHO, 2020; Salaverría et al., 2020; Pérez-Dasilva et al., 2020; Cerezo, 2020; López-García et al., 2021; Pérez-Escoda; Pedrero-Esteban, 2021). We are also in the midst of a debate on the viability and survival of media companies in a digital ecosystem where their “visibility” on the Internet is a determining factor in their sustainability strategies, and it is here where the concept of web reputation used for the evaluation of media and the construction of the ranking is crucial, because it has a direct connection with visibility.

Likewise, the value of connecting our study on digital reputation with a more intangible element such as the concept of branding is important in an environment increasingly saturated with actors, many of which are completely outside the media ecosystem and competing for the information business (Reig, 2011). Brand competitive advantage (Chan-Olmsted, 2011) and its impact on audience growth, thanks to the reputation effect (Pérez-Latre; Sánchez-Tabernero, 2012), constitutes a valuable line of development and innovation within the media, as well as of academic analysis.

Faced with the short-sighted vision of certain digital media, which continue to limit their strategies to simple monetization through the distribution of content on various platforms, the growing interest in cultivating a long-range “media branding” (Chan-Olmsted; Shay, 2015) as a way to strengthen the recognition and perception of quality by the public should not be ignored.

These final reflections on the future of the press and on the idea of branding as a factor of differentiation and value of the media, questions increasingly conditioned by actors outside the sector, are related to one of the most ambitious and long-term goals of the research project presented here: to position *SCImago Media Rankings* as a watchdog of analysis and evolution of the journalism industry and as a space for debate and discussion that has not only a dissemination aspect in academia, but also a practical aspect at the professional and business levels.

6. Conclusions and future avenues of research

The media industry, always at the center of the global stage, is one of the areas in which transformation processes are taking place in a more intense and dramatic way. The new *SCImago* project, presented here with this first assessment of the global press, was created with the aim of describing the current situation and contributing to the development of prospective studies: in the short term, to evaluate the research questions we have posed regarding the application of the *Digital Reputation Indicator (DRI)* to build a media ranking, and in the medium and long term, to become a new benchmark in the study of the sector.

After analyzing the results of the first *SMR*, we believe that it is possible to validate the objectives set out regarding the possibility of generating a world map of the media system on the basis of the big data provided by the Internet (with the limitations and future avenues of research already mentioned), as well as to lay the foundations for the development of longitudinal studies that will give us clues as to the processes of adaptation to the digital ecosystem.

From this perspective, in the “Results” section we have also confirmed the possibility of drawing conclusions about the territorial or regional distribution of the media outlets included in the map and connecting it to other global indicators such as the representation of Internet users and other socioeconomic factors related to, for example, the use of technology, digitalization, and literacy.

Even having presented and discussed the brand effect of the media in their positions in the *Digital Reputation Ranking*, the global assessment set forth in this paper shows the relative weight of communicational branding when it comes to competing for the attention of the digital ecosystem’s users: it can be concluded that, as has been argued, branding helps, but only based on the use of previous conditioning factors and context that have to do with the infrastructure of the countries and regions, socioeconomic conditions, and even cultural habits associated with the processes of digitization and globalization. Only from this perspective can it be explained, for example, that *The New York Times* is a leader (on its own merits, but also with a potential audience of millions of users world-

“ The *Digital Reputation Indicator (DRI)* combines the values of four metrics obtained from *Semrush*, *Ahrefs*, and *Majestic*, with which the composite indicator (Overall) is calculated. The combination of the number of inbound links (citations), their quality (relative weight of each link), and the estimation of traffic are the basis of the integral *DRI* indicator ”

“ In Latin America, *Clarín*, *La Nación*, *Infobae*, and *O Globo* compete for the top positions, while in other regions such as Asia, two media outlets are clearly prominent, driven by the multilingual editions of their sites: the *Renmin Ribao/People Daily* newspaper and the *Xinhua* news agency ”

wide), and that an equivalent media outlet in regions such as Asia or China with a very significant population weight is not found. The creation of high-value content is important, but so is the context.

In this sense, the *Digital Reputation Ranking* is based on the idea, established at an academic level on the key characteristics of media systems, that those who pull the strings at a global level (and in this case objectively underscored by their digital drive for global projection) are located in the North and speak English. The conviction that languages play a fundamental role in the world of communication to the point of clearly affecting the configuration of the map is strongly reinforced here, to the extent that they become a key factor in a media outlet's digital reputation, as has been observed in a singular way with websites that opt for English as a lingua franca and for multilingualism.

Finally, analyzing the Overall results of the analysis waves carried out between 2022 and 2023, we have seen how the media outlets with the best digital reputations are located in Western Europe and the United States, with a small group of leading media outlets competing for the top positions very closely, and who end up acting as an advance faction in their own digital innovation processes and strategies. In this leadership, a transfer of prestige from the historical paper newspapers to the digital environment can be observed (the brand effect previously mentioned), but the solid emergence of native projects capable of looking at the centuries-old brands from the perspective of digital reputation is already remarkable. The tension between legacy media and new media is a point of analysis that will be able to be evaluated more accurately as the sample expands and we respond to the dynamics of market evolution.

What has been observed so far is that there is a prominent media elite at the international level, led by the most emblematic American and British newspapers (*The Washington Post*, *The Independent*, and the English tabloids). In Europe, the leadership tends to be quite stable with very consolidated positions and global reach, such as the Spanish *El País*, the French *Le Monde*, the German *Der Spiegel*, or the Italian *La Repubblica*

7. Notes

1. We use the term “overall” in its methodological, neutral sense as a composite indicator (as stated on the project's website). Applied to our field of study and our proposed evaluation and media analysis, we refer to it as the *Digital Reputation Indicator (DRI)*.
2. In the summer of 2023, the study was completed with a fourth wave of analysis (Q2 edition 2023), and a new update is scheduled for release in October 2023.
3. All the identification and typification metadata are currently being completed and will be part of a new dissemination work.

8. References

Ahrefs (2023). *About us*, 28 February.
<https://ahrefs.com/about>

Ahrefs (2023). *Domain rating (DR)*, 28 February.
<https://ahrefs.com/blog/domain-rating>

Bustelo, Pablo (2022). *China en la economía mundial: fortalezas, debilidades y perspectivas*. Funcas.
https://www.funcas.es/wp-content/uploads/2022/04/CIE_186art8.pdf

Carlson, Matt (2020). *Measurable journalism. Digital platforms, news metrics and the quantified audience*. Routledge.
<https://doi.org/10.4324/9780429347108>

Cerezo, Pepe (2020). “El impacto de la pandemia en la prensa”. *Dossier Evoca*, Evoca Media.
<https://evocaimagen.com/dosieres/dossier-evoca-09-medios-y-coronavirus.pdf>

Chan-Olmsted, Sylvia M. (2011). “Media branding in a changing world: Challenges and opportunities 2.0”. *International journal on media management*, v. 13, n. 1, pp. 3-19.
<https://doi.org/10.1080/14241277.2011.568305>

Chan-Olmsted, Sylvia M.; Shay, Ronen (2015). “Media branding 3.0: From media brands to branded entertainment and information”. In: Siegert, G.; Förster, K.; Chan-Olmsted, S.; Ots, M. (eds.). *Handbook of media branding*, pp. 11-32. Springer.
https://doi.org/10.1007/978-3-319-18236-0_2

Costa-Sánchez, Carmen; López-García, Xosé (2020). “Communication and coronavirus crisis in Spain. First lessons”. *Profesional de la información*, v. 29, n. 3.
<https://doi.org/10.3145/epi.2020.may.04>

- Deuze, Mark; Witschge, Tamara** (2018). "Beyond journalism: Theorizing the transformation of journalism". *Journalism*, v. 19, n. 2, pp. 165-181.
<https://doi.org/10.1177/1464884916688550>
- Gómez-Mompert, Josep-Lluís** (2013). "Complejidad social y calidad informativa: hacia un periodismo 'glocal'". *Austral comunicación*, v. 2, n. 2, pp. 219-238.
<https://riu.austral.edu.ar/handle/123456789/269>
- Graves, Lucas; Kelly, John; Gluck, Marissa** (2010). *Confusion online: Faulty metrics and the future of digital journalism*. Tow Center for Digital Journalism, Columbia University.
<https://towcenter.columbia.edu/news/faulty-metrics-and-future-digital-journalism>
- Hernández-Pérez, Tony; Rodríguez-Mateos, David** (2016). "Medición integral de las audiencias: sobre los cambios en el consumo de información y la necesidad de nuevas métricas en medios digitales". *Hipertext.net*, n. 14.
<https://www.raco.cat/index.php/Hipertext/article/view/310968>
- Igartua, Juan-José; Ortega-Mohedano, Félix; Arcila-Calderón, Carlos** (2020). "The uses of communication in the time of coronavirus. A cross-cultural study". *Profesional de la información*, v. 29, n. 3.
<https://doi.org/10.3145/epi.2020.may.18>
- Kemp, Simon** (2023). *Digital 2023: Global overview report*, 11 March.
<https://datareportal.com/reports/digital-2023-global-overview-report>
- Lamas, Carlos** (2010). "Los medios interactivos y su publicidad: la medición de audiencias". *Telos: cuadernos de comunicación e innovación*, n. 82, 95-101.
<https://telos.fundaciontelefonica.com/archivo/numero082/los-medios-interactivos-y-su-publicidad-la-medicion-de-audiencias/?output=pdf>
- Livingstone, Sonia** (2019). "Audiences in an age of datafication: critical questions for media research". *Television & new media*, n. 202, pp. 170-183.
<https://doi.org/10.1177/1527476418811118>
- López-García, Xosé; Costa-Sánchez, Carmen; Vizoso, Ángel** (2021). "Journalistic fact-checking of information in pandemic: Stakeholders, hoaxes, and strategies to fight disinformation during the Covid-19 crisis in Spain". *International journal of environmental research and public health*, v. 18, n. 1227.
<https://doi.org/10.3390/ijerph18031227>
- Jarvis, Jeff** (2015). *El fin de los medios de comunicación de masas. ¿Cómo serán las noticias del futuro?*. Barcelona: Gestión 2000. ISBN: 8498754011
- Jenkins, Henry** (2008). *Convergence culture: la cultura de la convergencia de los medios de comunicación*. Barcelona, Paidós. ISBN: 978 84 493 2153 5
- Levinson, Paul** (2012). *New new media*. Boston: Pearson Higher Education. ISBN: 0205865577
- Maestro-Espínola, Lidia; García-Santamaría, José-Vicente; Pérez-Serrano, María-José** (2016). La 'prensa' digital generalista, como soporte publicitario y los cambios en su modelo de negocio. *Revista latina de comunicación social*, n. 71, pp. 1048-1068.
<https://doi.org/10.4185/RLCS-2016-1134>
- Magallón-Rosa, Raúl; Nachawati-Rego, Lella; Seoane-Pérez, Francisco** (2021). "Evaluación del riesgo de desinformación: el mercado de noticias online en España". *GDI - Global Disinformation Index*. Universidad Carlos III de Madrid.
<http://hdl.handle.net/10016/33587>
- Majestic** (2023). *About us*. February 28.
<https://majestic.com/company/about>
- Majestic** (2023). *Citation flow and trust flow*, February 28.
<https://majestic.com/help/glossary#CitationFlow>
- Masip, Pere** (2016). "Investigar el periodismo desde la perspectiva de las audiencias". *El profesional de la información*, v. 25, n. 3, pp. 323-330.
<https://doi.org/10.3145/epi.2016.may.01>
- Morales-Vargas, Alejandro; Pedraza-Jiménez, Rafael; Codina, Lluís** (2021). "Website quality in digital media: literature review on general evaluation methods and indicators and reliability attributes". *Revista latina de comunicación Social*, n. 80, pp. 39-60.
<https://doi.org/10.4185/rlcs-2022-1515>

- Newman, Nic; Fletcher, Richard; Robertson, Craig T.; Eddy, Kirsten; Kleis Nielsen, Rasmus-Kleis** (2022). *Digital news report 2022*. Reuters Institute for the Study of Journalism.
<https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2022>
- Nielsen (2020). *Nielsen launches next gen methodology for its flagship digital measurement products*. The Nielsen Company.
<https://bit.ly/3nzurMV>
- Olteanu, Alexandra; Castillo, Carlos; Díaz, Fernando; Kiciman, Emre** (2019). "Social data: Biases, methodological pitfalls, and ethical boundaries". *Frontiers in big data*, v. 2, n. 13.
<https://doi.org/10.3389/fdata.2019.00013>
- Organización Mundial de la Salud (2020). *Gestión de la infodemia sobre la Covid-19: Promover comportamientos saludables y mitigar los daños derivados de la información incorrecta y falsa*.
<https://bit.ly/3NDLQ44>
- Orihuela, José-Luis** (2021). *Culturas digitales. Textos breves para entender cómo y por qué internet nos cambió la vida*. Ediciones Eunate. ISBN: 8477684243
- Papí-Gálvez, Natalia; Perlado-Lamo-de-Espinosa, Marta** (2018). "Investigación de audiencias en las sociedades digitales: su medición desde la publicidad". *El profesional de la información*, v. 27, n. 2, pp. 383-393.
<https://doi.org/10.3145/epi.2018.mar.17>
- Pérez-Dasilva, Jesús-Ángel; Meso-Ayerdi, Koldobika; Mendiguren-Galdospín, Terese** (2020). "Fake news and coronavirus: Detecting key players and trends through analysis of Twitter conversations". *Profesional de la información*, v. 29, n. 3.
<https://doi.org/10.3145/epi.2020.may.08>
- Pérez-Escoda, Ana; Pedrero-Esteban, Luis-Miguel** (2021). "Challenges for journalism facing social networks, fake news and the distrust of Z generation". *Revista latina de comunicacion social*, v. 79, pp. 67-85.
<https://doi.org/10.4185/RLCS-2021-1519>
- Pérez-Latre, Francisco J.; Sánchez-Taberner, Alfonso** (2012). *Innovación en los medios: la ruta del cambio*. Eunsa. ISBN: 978 84 31328597
https://www.eunsa.es/libro/innovacion-en-los-medios_101841
- Peters, Kay; Chen, Yubo; Kaplan, Andreas M.; Ognibeni, Björn; Pauwels, Koen** (2013). "Social media metrics—A framework and guidelines for managing social media". *Journal of interactive marketing*, v. 27, n. 4, pp. 281-298.
<https://doi.org/10.1016/j.intmar.2013.09.007>
- Quintas-Froufe, Natalia; González-Neira, Ana** (2021). "Evolución de la medición digital de la audiencia en el mercado español: estado de la cuestión y retos de futuro". *Profesional de la información*, v. 30, n. 1, e300102.
<https://doi.org/10.3145/epi.2021.ene.02>
- Reig, Ramón** (2011). *Los dueños del periodismo*. Gedisa. ISBN: 978 84 9784 618 9
- Reig, Ramón; Labio-Bernal, Aurora** (2020). *El laberinto mundial de la información. Estructura mediática y poder*. Anthropos. ISBN: 8416421676
- Renó, Denis; Flores, Jesús** (2012). *Periodismo transmedia*. Fragua. ISBN: 978 84 70745041
- Reuters Institute (2021). *¿Cómo y por qué accede la audiencia a las noticias en redes sociales?* Reuters Institute for the Study of Journalism.
<https://bit.ly/3OTmXBB>
- Reyes-Lillo, Danilo; Morales-Vargas, Alejandro; Rovira, Cristòfol** (2023). "Reliability of domain authority scores calculated by Moz, Semrush, and Ahrefs". *Profesional de la información*, v. 32, n. 4, e320403.
<https://doi.org/10.3145/epi.2023.jul.03>
- Rodríguez-Vázquez, Ana-Isabel; Direito-Rebollal, Sabela; Silva-Rodríguez, Alba** (2018). "Audiencias crossmedia: nuevas métricas y perfiles profesionales en los medios españoles". *Profesional de la información*, v. 27, n. 4, pp. 793-800.
<https://doi.org/10.3145/epi.2018.jul.08>
- Salaverría, Ramón** (2019). "Digital journalism: 25 years of research. Review article". *Profesional de la información*, v. 28, n. 1, e280101.
<https://doi.org/10.3145/epi.2019.ene.01>
- Salaverría, Ramón; Buslón, Nataly; López-Pan, Fernando; León, Bienvenido; López-Goñi, Ignacio; Erviti, María-Carmen** (2020). "Disinformation in times of pandemic: Typology of hoaxes on Covid-19". *Profesional de la información*, v. 29, n. 3, e290315.
<https://doi.org/10.3145/epi.2020.may.15>

Salaverría, Ramón; Martínez-Costa, María-Pilar; Negro, Samuel; Paisana, Miguel; Crespo, Miguel (2022). *Iberifier Digital Media Dataset (1.1.0)* [Data set]. Zenodo.

<https://doi.org/10.5281/zenodo.7335551>

Sánchez, Álvaro (2023). "La India: un gigante demográfico que aún es una potencia incompleta". *El País*, 4 abril.

<https://elpais.com/economia/negocios/2023-02-11/la-india-un-gigante-demografico-que-aun-es-una-potencia-incompleta.html>

Scolari, Carlos A. (2021). "Adiós sociedad líquida, bienvenida sociedad gaseosa". *Hipermediaciones*.

<https://hipermediaciones.com/2021/08/13/adios-sociedad-liquida-bienvenida-sociedad-gaseosa>

Semrush (2023). *About Semrush*, 28 February.

<https://www.semrush.com/company>

Semrush (2023). *Crawlability score*, 28 February.

<https://www.semrush.com/blog/what-are-crawlability-and-indexability-of-a-website>

Solanas-García, Olga; Carreras-Margineda, Glòria (2011). "Nuevos medios publicitarios: la influencia del mundo digital en la planificación de medios". *Tripodos*, v. 28, pp. 63-81.

<https://raco.cat/index.php/Tripodos/article/view/247484>

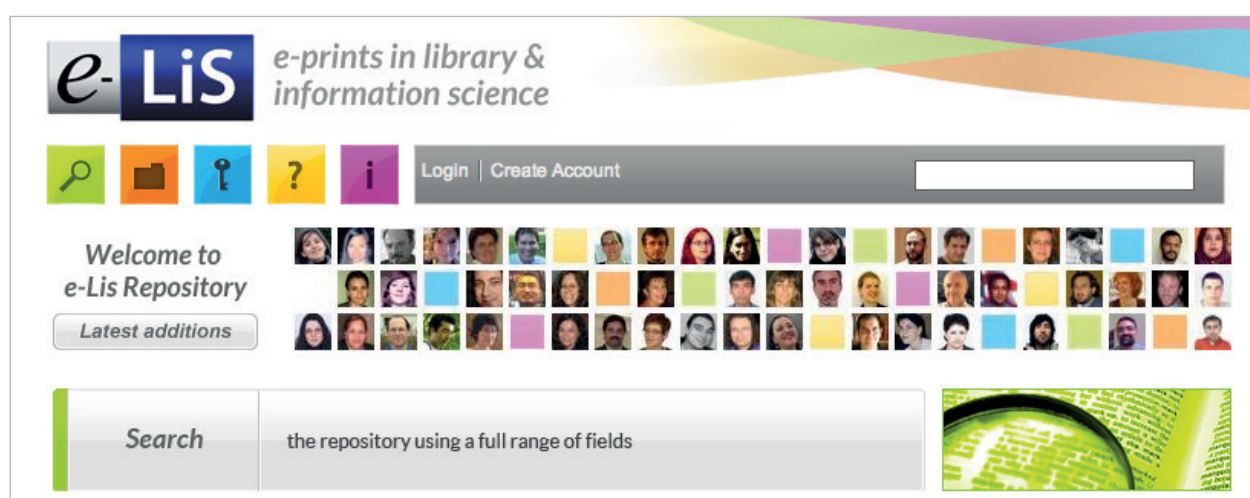
Toffler, Alvin (1980). *La tercera ola*. Plaza & Janés. ISBN: 84 01 37066 3

Tubella-Casadevall, Imma; Alberich-Pascual, Jordi (2012). *Comprender los media en la sociedad de la información*. Editorial UOC. ISBN: 978 84 9788 558 4

Vázquez-Herrero, Jorge; Negreira-Rey, María-Cruz; Silva-Rodríguez, Alba; Rodríguez-Vázquez, Ana-Isabel (2022). "The news media through the lens of the Spanish audience: A matter of trust". *Estudios sobre el mensaje periodístico*, v. 28, n. 2, pp. 447-459.

<https://doi.org/10.5209/esmp.77807>

Give visibility to your work by depositing it in e-LIS, the largest international repository on library & information science, and communication



<http://eprints.rclis.org>

Never mind predatory publishers... what about 'grey' publishers?

David Nicholas; Eti Herman; Abdullah Abrizah; Blanca Rodríguez-Bravo;
Chérifa Boukacem-Zeghmouri; Anthony Watkinson; Marzena Świgoń; Jie Xu;
Hamid R. Jamali; Carol Tenopir

Recommended citation:

Nicholas, David; Herman, Eti; Abrizah, Abdullah; Rodríguez-Bravo, Blanca; Boukacem-Zeghmouri, Chérifa; Xu, Jie; Watkinson, Anthony; Świgoń, Marzena; Jamali, Hamid R.; Tenopir, Carol (2023). "Never mind predatory publishers... what about 'grey' publishers?". *Profesional de la información*, v. 32, n. 5, e320509.

<https://doi.org/10.3145/epi.2023.sep.09>

Article received on August 26th 2023

Approved on September 8th 2023



David Nicholas

<https://orcid.org/0000-0001-8046-2835>

CIBER Research Ltd
Newbury, Berkshire
RG147RU, United Kingdom
dave.nicholas@ciber-research.com



Eti Herman

<https://orcid.org/0000-0001-8526-9081>

CIBER Research Ltd
Newbury, Berkshire
RG147RU, United Kingdom
eherman@univ.haifa.ac.il



Abdullah Abrizah

<https://orcid.org/0000-0002-8224-5268>

University of Malaya
Faculty of Arts and Social Sciences
Department of Library & Information Science
50603 Kuala Lumpur, Malaysia
abrizah@um.edu.my



Blanca Rodríguez-Bravo ✉

<https://orcid.org/0000-0002-9476-7602>

Universidad de León
Área de Biblioteconomía y Documentación
Campus de Vegazana
24071 León, Spain
blanca.rodriguez@unileon.es



Cherifa Boukacem-Zeghmouri

<https://orcid.org/0000-0002-0201-6159>

Université Claude Bernard Lyon 1
Département Informatique
69100 Villeurbanne, France
cherifa.boukacem-zeghmouri@univ-lyon1.fr



Anthony Watkinson

<https://orcid.org/0000-0002-2317-6557>

CIBER Research Ltd
Newbury, Berkshire
RG147RU, United Kingdom
anthony.watkinson@btinternet.com



Marzena Świgoń

<https://orcid.org/0000-0003-3600-8349>

Uniwersytet Warmińsko-Mazurski
Wydział Humanistyczny
Instytut Dziennikarstwa i Komunikacji
Społecznej
10-719 Olsztyn, Poland
marzena.swigon@uwm.edu.pl



Jie Xu

<https://orcid.org/0000-0002-9820-8066>

Wuhan University
School of Information Management
Hubei 430072, Wuhan, China
xuj@whu.edu.cn



Hamid R. Jamali

<https://orcid.org/0000-0003-1232-6473>

Charles Sturt University
School of Information Studies
Locked Bag 588, Wagga Wagga
NSW 2678, Australia
h.jamali@gmail.com



Carol Tenopir

<https://orcid.org/0000-0002-9056-8251>

University of Tennessee
School of Information Sciences
1345 Circle Park Drive, 451
Knoxville, TN 37996-0341, USA
ctenopir@utk.edu



Abstract

The *Harbingers* project, which studied the working lives and scholarly communication behaviour of early career researchers (ECRs) over 6 years, found evidence of changing attitudes to questionable (grey) publishing. Thus, whilst predatory publishers have come to be treated with equanimity, as a problem easily dealt with, there was growing concern with the high volume of low-grade research being generated, some of which by 'grey' open access publishers for want of a better name (questionable and non-standard have also been used). With the recent announcement (2023) that the government of Malaysia (a *Harbinger* case country) is not providing Article Processing Charges (APCs) for articles published by *MDPI*, *Frontiers* and *Hindawi* on quality and cost grounds, we set out to see what lay behind this decision and whether other countries exhibited similar concerns. Information was obtained by asking *Harbinger* country leads, mostly embedded in research universities, from Australia, China, France, Israel, Malaysia, Poland, Spain, UK, and the US to conduct desk research to establish what is happening. It was found that countries, like ECRs, appear to have formed into two different camps, with China, Poland, France, and Spain joining Malaysia in the camp of those who felt concerned about these publishers and the UK, US, Israel, and Australia belonging to the camp of the unconcerned. Explanations for the split are furnished and whether the Malaysian position will prevail elsewhere is considered. Finally, in this paper, we have aired issues/concerns, rather than provided robust, systematic data. For a systematic study we shall have to wait for the fuller study we are hoping to conduct.

Keywords

Predatory journals; Questionable research practices; Low-quality research; Policies against predatory journals; Publishing; Research; Scholarly communication; Early career researchers; ECRs; Pandemic consequences; *Harbingers project*; Interviews; Country differences; Australia; China; France; Israel; Malaysia; Poland; Spain; UK; United Kingdom; US; United States.

1. Introduction

*Harbingers*¹ is an international collaborative of researchers from nine countries (Australia, China, France, Israel, Malaysia, Poland, Spain, UK, and US²). Monitoring on a regular basis what early career researchers (ECRs) –tomorrow's leading professors, influencers, and decision makers– are up to, we, as a research group, seek to forecast what changes might occur to scholarly communications and research practices down the line. In this paper, we focus on questionable publishing practices, an undertaking prompted by the findings of the Sloan-funded *Harbingers-2* study (Nicholas *et al.*, 2023a). The study, which inspected every conceivable aspect of the scholarly communications system from the perspective of the millennial-generation early career researchers during the two demanding years of the pandemic, identified unethical/questionable practices as a crucially important 'crack' in the scholarly system. Indeed, with nearly two-thirds of ECRs saying that they were aware of questionable practices, and a third believing that the pandemic had fanned the flames, questionable practices were ranked third in importance among the vulnerabilities found, although in some ways they may be seen as the biggest overall threat because they undermine trust in the whole system. It is against the backdrop of these findings that we set out here to revisit a topic we recently reported on in this very journal: ECRs' views and actual practices of publishing in questionable open access (OA) journals (Nicholas *et al.*, 2023b).

2. Background and context

The problem of questionable publishers came to the fore in the scholarly world once a host of studies found predatory publishing to be a prevalent phenomenon among academics from all career levels (Alecci, 2018; Alrawadieh, 2018; Elliott *et al.*, 2022; Eykens *et al.*, 2019; Frandsen, 2022; Perlin; Imasato; Borenstein, 2018; Pyne, 2017; Shaghaei *et al.*, 2018; Wallace; Perri, 2018), and from all over the globe, inclusive of high- and upper-middle-income countries (IAP, 2022; Bagues; Sylos-Labini; Zinovyeva, 2017; Elliott *et al.*, 2022; Eykens *et al.*, 2019; Moher *et al.*, 2017; Shaghaei *et al.*, 2018), rather than limited to novice researchers from developing countries, as first explorations seemed to indicate (Alrawadieh, 2018; Eykens *et al.*, 2019; Frandsen, 2022; Perlin; Imasato; Borenstein, 2018; Pyne, 2017; Shaghaei *et al.*, 2018; Wallace; Perri, 2018; Demir, 2018; Frandsen, 2017; Kurt, 2018; McCann; Polacsek, 2018; Moher; Srivastava, 2015; Moher *et al.*, 2017; Nobes; Harris, 2019; Shen; Björk, 2015; Xia *et al.*, 2015). Still, over the years a change seemed to be taking place when it came to what Gallent-Torres (2022) aptly calls non-standard publishing, as the evidence emerging from the *Harbingers* project clearly indicated (Nicholas *et al.*, 2021; 2023b).

Indeed, running counter to the hue and cry raised about predatory publishing for quite some time, our ECR interviewees not only found predatory journals to be of no real concern, but even treated the phenomenon as a non-issue, something that barely merited a mention, certainly not a feature of their research world. Being well-aware of the risk to their careers that publishing in questionable journals posed, and informed enough to be able to identify such journals readily enough, they were not tempted to go for the predatory option. In any case, most published with senior colleagues who were knowledgeable on the topic and could have prevented them (if necessary) from falling into the predatory publishing trap. However, whilst predatory journals, thus, turned out to be 'old hat' to ECRs, not so when it came to what they identified as the real problem facing them, indeed, facing the entire research undertaking: a general lowering of scholarly standards that have been taking place as a result of increased opportunities to publish.

That their thinking should be along these lines is perhaps not very surprising, for the problem of questionable publishing has by now become much more nuanced and complex. In fact, predation in academic publishing now needs to be seen as a spectrum, with varying types and degrees of illegitimacy, with journals and publishers of multiple shades of grey

occupying borderline or ambiguous niches between predation and legitimacy (Siler, 2020). Thus, a journal may be increasingly/decreasingly predatory (or legitimate) in comparison to some standard or criteria, whether ideal or in relative contrast to one another (Teixeira-da-Silva *et al.*, 2021). Indeed, although our interviewees did not actually use the term 'grey publishers' in their discussions of questionable research practices, the conversation often switched (especially in France, Spain, and Poland) to their feeling that large volumes of low-quality research were being produced, with occasional blame for this situation being placed (unsolicited) at the door of MDPI³ and Frontiers⁴.

It must be kept in mind, of course, that these two publishers might have been singled out by our study participants because they have drawn particular attention to themselves by becoming bigger than most 'traditional' publishers, as MDPI's remarkable growth amply illustrates (Petrou, 2020). Indeed, the number of ECRs among the *Harbingers-2* interviewees, who thought that these publishers were simply fast, friendly, efficient, and receptive, was similar to the number of those who found fault with them. Admittedly, unlike predatory publishers, some 'grey' journals have respectable citation scores—the ultimate judge and jury on their worth. Also, unlike predatory publishers, 'grey' publishers do not deliberately and systematically set out to deceive authors or lie about their metrics and they run fair peer review processes (until it comes to special issues, it seems, a point we will discuss shortly). Nevertheless, in a recent study that analysed the behaviour of 53 MDPI-journals that were *Clarivate's Journal Citation Reports (JCR)* indexed in 2019, some of their characteristics were found to be equitable to those of predatory journals, such as their tendency to use very similar names to other journals with established reputations (Oviedo-García, 2023).

Still, some of the practices employed are likely, indeed, to make them suspect in the eyes of some. Thus, for example, the direct solicitation of potential authors raises quite a few eyebrows, although how otherwise can new journals start if the editor designate does not 'personally' invite people they know to submit articles or ask the publisher to do so on their behalf? The problem is that at times these publishers select potential authors to approach from a database, which is so badly put together that the recipient is not in the field which the journal covers. Also, authors and reviewers are sometimes offered discounts and incentives, which, given the accepted ways of doing things in the scholarly world, may be seen as commercial. According to Oviedo-García (2023), at least where JCR indexed MDPI-journals are concerned, they also have higher degrees of self-citation rates than found in 'traditional' journals and high citations rates from other MDPI-journals, the latter of which, of course, suggests citation cartels. However, perhaps above all, it is the practice of publishing huge quantities of special issues that gives these publishers truly bad publicity, for the sheer number of these publications obviously makes it impossible to guarantee minimum peer-review standards. Take just one example: according to an unpublished analysis, cited in Brainard (2023), in 2022 nearly 100 MDPI journals that have impact factors published more than 17,000 special issues, containing 187,000 articles.

3. Malaysia blows the whistle

Alarm bells truly started ringing for us when a member of the *Harbingers* collaborative, the Malaysian representative, informed the rest of the group of the Malaysian *Ministry of Higher Education's* decision to refrain from using public funding to cover APCs for publishing with MDPI, Frontiers, and also Hindawi⁵ (Japri, 2023). The decision is certainly indicative of a broader concern, which has at its heart the need for a more judicious allocation of resources that would focus on supporting research activities that contribute to the advancement of knowledge and societal development. Thus, plainly aiming at encouraging universities to exercise caution in their choice of publication venues, the *Ministry* seems to be advocating a more strategic approach to research dissemination, which would consider factors beyond the mere open-access nature of the journal and its indexation status (in *Web of Science* and/or *Scopus*). Among these factors the quality of peer review, the impact of the research, and the alignment of the publication with the university's research objectives are all seen as playing crucial roles in ensuring that research outputs contribute meaningfully to their respective fields.

However, beyond that, the *Ministry's* decision seems to have been sparked by the significant surge in the quantity of papers authored by Malaysian researchers that were published in MDPI's *Scopus*-indexed journals over consecutive years, possibly coupled with recent criticisms related to MDPI's publishing practices. Thus, as an unpublished analysis has shown, the number of Malaysian publications in MDPI *Scopus*-indexed journals has risen from 698 in 2018 to 1227 in 2019, 2281 in 2020, 3908 in 2021, and 4768 in 2022—plainly a substantial growth, amounting to nearly 580 percent during this time span. This, when fast-growing OA journals, of which MDPI, Frontiers and Hindawi are in all probability the best known, increasingly come under fire for their non-standard practices, some of which are seen as smacking of overly commercial considerations, if not predatory approaches, such as aggressive marketing, lenient peer review for the sake of achieving speed and quantity of publications, significant author fees, and the aforementioned overabundant publishing of special issues and manipulations of citations (Brockington, 2022; Crosetto, 2021; Gallent-Torres, 2022; Oviedo-García, 2023).

It is quite conceivable then that the straw that might have broken the camel's back in Malaysia was the recent delisting of 19 Hindawi and two MDPI journals from *Clarivate's World of Science (WoS)* database, entailing the journals' losing their impact factor, which carries so much weight with authors and institutions (Brainard, 2023). The move certainly cannot be dismissed lightly, especially as according to an expert opinion, cited in Brainard (2023), the expectation is that this initial delisting is only the tip of the iceberg.

Interestingly, despite the *Ministry's* decision, at least one first-line research university in Malaysia still permits accepting APCs from private or international funding sources for publishing in MDPI, Frontiers and Hindawi journals, as long as

they are indexed in *WoS* or *Scopus*, indeed, recognises such publications for assessment. Confusing signals then being transmitted, perhaps, further muddying the waters for Malaysian researchers, who (at least until recently) seemed oblivious to the problem of grey journals –our interviewees, unlike their counterparts in Poland, France, and Spain, were happy to publish in any journal, if it was indexed in *WoS/Scopus* (Nicholas *et al.*, 2022).

The question that naturally follows was whether these developments were unique to Malaysia. After all, publishing is a genuinely international business, with strong patterns of globalisation. Thus, we set out to see what the situation was in our other member countries: whether they exhibited the same concerns, and if they did, whether they were responding to the challenge similarly or differently and whether we have a trend on our hands.

4. So, what of other countries?

Data were obtained by asking country leads (mostly senior researchers embedded in research universities) from Australia, China, France, Israel, Poland, Spain, UK, and US to conduct desk research to establish what is going on. Interestingly, the picture that formed in result show that there are two different camps when it comes to ‘grey’ publishers, with China, France, Spain, and (less so) Poland joining Malaysia in the camp of those who felt concerned about these publishers and Australia, Israel, UK, and the US, belonging to the camp of the unconcerned.

5. Countries with concerns

China

Clearly, a lot of disquiet about some publishers has been long felt in China, entailing an assertive and direct governmental approach. Indeed, since 2020 the *Chinese Academy of Sciences* has been publishing a ‘warning journals list’ (blacklist) every year, with many of the journals –more than half in the 2023 list– published by *MDPI*, *Hindawi* or *Frontiers*. Inclusion of a journal in the list was determined based on several indicators: many factual and/or grammatical and/or spelling errors in the articles published; attempts to appoint editors without their authorisation; too many papers published in each issue; quick and questionable-quality peer review processes; production of large numbers of new titles.

Indeed, popular, and social media reports mention more than 20 universities and research institutions in China that have publicly stated their intention not to consider papers published in the ‘warning journals list’ for assessment purposes. It should be noted, however, that most of these universities and research hospitals are not first-line institutions, where the guidance of the list has been sorely needed, as they seem to face problems of academic misconduct, stemming from the requirement that clinicians seeking professional titles publish research papers.

However, whilst the approach to ‘grey’ publishers is thus not invariably defined as a formal policy in China, many major Chinese scientific research institutions are wary in their treatment of ‘grey’ journals, often excluding them from assessment procedures, as part and parcel of their reservations and prejudices about OA publishing in general. Not that this state of affairs has had much of an influence on institutions’ willingness to fund the payment of APCs for publishing in gold OA journals: unless the journals are high-level ones, such as journals with a high impact factor or a good reputation, researchers who want to do so need to pay APCs themselves. In any case, despite the unease surrounding publishing in ‘grey’ journals, there are many papers authored by Chinese academics, if often from second-level universities, to be found in them. Plainly, the short peer review procedure, the fast turnaround times of publication, the relatively more standardised/higher quality compared to local journals, and the lack of discrimination against young Chinese authors are strong incentives, especially as *MDPI*, with its huge offices in China, is a well-known company.

Spain

In Spain, too, things are far from straightforward, although journals from these publishers have been blacklisted. Thus, the central body in charge of the evaluations of research achievements for promotion purposes, *Aneca* (*Agencia Nacional de Evaluación de la Calidad y Acreditación*) issued as far back as 2021 a report (*Aneca*, 2021) on open-access publishers which questioned many journals, particularly journals from *MDPI* and *Frontiers*, said to exhibit ‘anomalous behaviour’.

<https://www.aneca.es>

Previously, many of the disciplinary committees, which evaluate the research accomplishments of universities every six years for budgetary reasons, stopped positively considering publication in certain journals, many of them *MDPI*.

The *Aneca* report was highly controversial because not all the measurement criteria used were considered adequate or logical. Indeed, ECRs feel confused, unable to understand why *JCR* indexed journals are not suitable for obtaining a good assessment. To complicate matters even further, two years later, *Aneca* withdrew its report, claiming to justify the withdrawal on the grounds of a more overarching need to reform the research evaluation models in Spain, towards more qualitative and less bibliometric evaluation models. More specifically, this need is for evaluating of the quality of scientific articles, rather than the quality of the journals in which they are published only, and, in addition, to consider the impact of the research from a broader perspective.

In any case, the funding of APCs is at the discretion of the individual university, and in general, it is allowed even where ‘grey’ publishers are concerned.

France

The picture is not that clear in France, either, although there is a much controversy in respect to the status of *Frontiers* and *MDPI*. On the one hand, more and more French papers are published in 'grey' journals because, as in the case of China, many researchers welcome the opportunity to publish their papers easily and quickly. This is especially so where the journals are indexed in *Scopus* and *WoS*, for then they are admissible in the research assessment process. On the other hand, it is said by some that researchers and more specifically early career researchers, risk their reputation by publishing in such journals, which, as rumour in France has it, are not well cited or, indeed, not cited at all, a rumour that is not sufficiently grounded. Still, as noted, given that the scientific value of some 'grey' journals have been contested, indeed, some have even been delisted from *WoS*, they are increasingly seen as questionable, with some institutions displaying their concerns by requesting researchers not pay the APCs for grey journals. There are also worries among French ECRs (and Polish ones too, as we shall see) that assessors will look less favourably at papers emanating from grey publishers. The case of *MDPI* has not been helped by the controversies associated with the publishing of thematic issues, with researchers denouncing them: <https://paolocrosetto.wordpress.com/blog>

Frontiers is more respected than *MDPI*, because of a rise in standards and their publication do attract more citations <https://www.frontiersin.org/about/impact#jif-and-citescore>

Poland

In Poland, it is the problem of *MDPI*, in particular, that has been discussed, indeed, still being discussed, but nothing has been resolved. Following the parameterization (measurement) of Polish universities, there has been a significant increase in publications by Polish authors in journals such as *MDPI* and other gold open access journals, which are regarded as venues that enable easy publishing, with a review process that takes a very short time. With the assembling of a copious publishing record of paramount importance for Polish academics, as their eligibility to get their doctorate and habilitation in each discipline at a given university largely depends on it, plainly the ease and speed of publishing their work becomes an important consideration.

As to the current situation –at least one university is said to have stopped funding APCs for publishing in *MDPI* journals. The Polish media reports that a great deal of Polish taxpayers' money, i.e., from the state budget, is spent on the publishing of articles in *MDPI* journals, but so far nothing has been heard on this from the *Ministry of Education and Science*. In fact, the *Ministry's* list of scored journals, assembled for assessment purposes, which includes more than 30,000 *WoS/Scopus* –indexed and Polish titles, even includes some journals published by *MDPI*. Nevertheless, all those who want to do a 'habilitation' or apply for a professorship try to make sure that they have a variety of journals on their publications list, and they certainly do not rely solely on *MDPI* because of the possibility that reviewers will criticise them for doing so.

Finally, Poland is looking to see how bigger and richer countries are dealing with the problem before they take any action.

6. Countries without concerns

Australia

Australia is different from countries, such as Malaysia, in that government/authorities usually do not interfere with publishing decisions, so that each university decides on its own when to fund a requested APC. However, not all universities have a specific OA policy, and those that have do not go beyond a generally felt strong opposition to paying for publishing OA in hybrid journals (*Wakeling et al.*, 2022). There is certainly no policy on the level of individual publishers –grey or otherwise.

Still, universities tend to favour white lists over black lists, so instead of banning publishing in certain journals, they favour others and encourage publishing in them. Universities might have different approaches to implementing this unwritten policy, but the basis for their decision-taking in these matters are usually the points awarded to a specific journal in their internal evaluations and promotions procedures. Thus, for example, if at a given university the achievements of academics are measured against *SCImago* (*Scopus* list), with Q1 and Q2 journals from that list accorded more points, researchers are encouraged to publish in journals listed in *Scopus*, be they 'traditional' or 'grey'. University libraries usually have *LibGuides* about publishing articles in which they warn researchers about predatory publishers. However, instead of blacklisting publishers they encourage researchers to follow best practices such as *Think, Check & Submit* protocols in their publishing decisions. Thus, the *LibGuide* for *Charles Sturt University* names *MDPI*, but it says it has a range of journals so authors need to decide case by case and does not say do not publish in *MDPI*.

<https://thinkchecksubmit.org>

<https://libguides.csu.edu.au/WhereToPublish/predatory#s-lg-box-22384022>

Israel

Publishing policies in Israel are not based on any nation-wide policy as to the legitimacy (or not) of journals, so that it is up to each faculty member to decide where to publish their work. University libraries do provide individual help, such as recommending *WoS* and *Scopus* indexed journals or warning against journals on *Cabells* list, and hold training sessions for those who want them.

Still, there is a list of journals widely used by universities and researchers to determine which journals are preferred venues for publication, originally assembled by the *Council for Higher Education*, through its *Planning and Budgeting Committee*, to serve institutional budgeting purposes. Thus, although the list was meant to serve solely as a means for

assessing the quality of a university's publication output, as measured by the impact factors of journals in which the publications appear, its use is customarily taken to vouch for the quality of the journals therein.

As to funding APCs, Israeli universities pay APCs for publishing in highly ranked journals (top ten in a subject area) and/or for publishing in journals that offer a discount obtained through deals with publishers.

UK

Grey publishers are not a big issue in the UK and, indeed, have become part of the scholarly infrastructure and welcomed to the top table by Jisc, who have conducted big deals with them on behalf of UK universities and have obtained APC discounts. ECRs indeed perceived journals particularly *Frontiers* as prestigious in their field and *MDPI* are admired because of their speed in processing. It is possible that their OA credentials protect them from criticism because the UK is very pro-open science and journals appear in the authoritative *DOAJ* directory. Librarians consulted were willing to pay APCs because the publishers were members of *OASPA* and the journals were in *DOAJ*, key marks of accessibility.

Nevertheless, there is still some unease about these publishers' outputs. The unease stems from direct and inappropriate solicitation of authors and editors annoying researchers by asking them to write papers, review papers or even edit special issues in disciplines which were not remotely related to their expertise. It is also believed that this aggressive targeting has increased in the last few years, presumably driven by internal targets for increases in submissions. *Frontiers* has admitted as such to some universities. This can lead to authors being often approached to write reviews and other non-research articles that universities may be unable to fund. Also, while not perhaps seeing them as promoting paper mills, although guest editors and salaried academic editors for *MDPI* were found to coordinate sale of authorship across four different *MDPI* journals, totalling over 20 papers (Grove, 2020), but as one university librarian said *some of their content does tend towards slicing and dicing, and amount to vanity publishing*.

US

In the US, there is no national rule, other than rules for federally funded research outputs. Each university sets its own standards for what journals count toward tenure and typically most universities leave that up to the academic unit to decide what are the best journals. Many universities now use a central metrics system to measure the impact and outputs of the faculty, but these are typically not used for tenure or retention decisions. The one unit on campus that pays most attention to research dissemination is usually the library. If the library pays for OA fees, they usually will not pay for hybrid journal fees and some may have a black or white list of ones they will pay. Libraries used to often use *Beall's* list years ago, but it is more individualized now. Thus, for example, in some universities eligible publications must be listed in the *Directory of Open Access Journals (DOAJ)* or the publisher must be a member of the *Open Access Scholarly Publishers Association (OASPA)*. Measured by these parameters, the three publishers discussed here would meet the criteria.

7. Conclusions

It seems that countries form a continuum in their approach to 'grey' publishers, from those who seem to have little concerns about them—Australia, Israel, UK, and US—to those growingly concerned about the phenomenon, with Malaysia the most concerned, but China, Poland, France, and Spain also exhibiting signs of unease. The level of concern is perhaps traceable to the assessment procedures in place in the different countries: the more a country relies on quantitative productivity coupled with citation-data based bibliometric indicators to measure research achievements, the more problematic the role of 'grey' publishers becomes. One wonders then whether the drip, drip news of retractions, special editions, government interventions and rumours might change the balance, with more countries moving into the 'concerned' camp, and, as a result, possibly changing their approach to research assessment policies. The greater presence of 'grey' publishers in the scholarly world may turn out to be then a blessing in disguise, which will spur on the long-hoped for change in the model for evaluating a scientist's work, whereby the currently prevalent 'bean-counting,' both where the number of publications and the citation score of publications are concerned, will be enriched by other measures. Indeed, in France the transition is already taking place, slowly but surely, and publishing papers in indexed journals is no longer considered as the only criterion of assessing research, whilst for Poland this is thought to be a pipe dream, at least for the time being. Thus, this paper, largely meant to be an opinion piece, has laid the ground for future systematic research on the part of the *Harbingers* group.

8. Notes

1. The original, 4-year (2016-2019), longitudinal *Harbingers-1* research project explored the working lives and scholarly communication behaviour of 116 junior science and social science researchers from eight countries (China, France, Malaysia, Poland, Russia, Spain, UK and US). Its 2-year extension (2020-2022), *Harbingers-2 –Early Career Researchers*, and the *Pandemic research project*.

<http://ciber-research.com/harbingers-2>

funded by the *Alfred P. Sloan Foundation*, enabled the investigation to continue to around 170 science and social science ECRs from the same countries, at a time when the pandemic brought about unprecedented change in scholars' circumstances.

2. While 8 countries have contributed data for this paper, Australia and Israel were not original *Harbinger* case studies.

3. *MDPI* is a publisher of 428 open access journals, including 419 peer-reviewed journals and 9 conference journals. Founded in 1996 and based in Basel, Switzerland, it has offices in Beijing, Wuhan, Tianjin and Nanjing (China), Barcelona

(Spain), Belgrade and Novi Sad (Serbia), Manchester (UK), Tokyo (Japan), Cluj, and Bucharest (Romania), Toronto (Canada), Kraków (Poland), Singapore (Singapore), and Bangkok (Thailand).

<https://en.wikipedia.org/wiki/MDPI>

4. *Frontiers Media SA* is a publisher of 185 peer reviewed, open access journals. Founded in 2007 it is based in Lausanne, Switzerland, with other offices in London, Madrid, Seattle and Brussels.

https://en.wikipedia.org/wiki/Frontiers_Media

5. *Hindawi* is a publisher of over 250 peer reviewed, open access, scientific, technical, and medical (STM) literature. It was founded in 1997 in Cairo, Egypt, but purchased in 2021 by *Wiley*, a publishing company based in the United States.

[https://en.wikipedia.org/wiki/Hindawi_\(publisher\)](https://en.wikipedia.org/wiki/Hindawi_(publisher))

9. References

Alecci, Scilla (2018). *New international investigation tackles 'fake science' and its poisonous effects*. Blog post, 20 July.

<https://www.icij.org/blog/2018/07/new-international-investigation-tackles-fake-science-and-its-poisonous-effects>

Alrawadieh, Zaid (2018). "Publishing in predatory tourism and hospitality journals: Mapping the academic market and identifying response strategies". *Tourism and hospitality research*, v. 20, n. 1, pp. 72-81

Aneca (2021). *Bibliometric analysis and impact of open-access publishers in Spain*.

<https://shorturl.at/kpyH4>

Bagues, Manuel F.; Sylos-Labini, Mauro; Zinovyeva, Natalia (2017). *A walk on the wild side: An investigation into the quantity and quality of 'predatory' publications in Italian academia* (N. 2017/01). LEM working paper series.

<http://hdl.handle.net/10419/174551>

Brockington, Dan (2022). *MDPI Journals: 2015-2021*. Blog post, Nov. 10.

<https://danbrockington.com/blog>

Crosetto, Paolo (2021). *Is MDPI a predatory publisher?*. Blog post, April 12.

<https://paolocrosetto.wordpress.com/2021/04/12/is-mdpi-a-predatory-publisher>

Demir, Selcuk-Besir (2018). "Predatory journals: Who publishes in them and why?". *Journal of informetrics*, v. 12, n. 4, pp. 1296-1311.

<https://doi.org/10.1016/j.joi.2018.10.008>

Elliott, Tracey; Fazeen, Bisma; Asrat, Asfawossen; Cetto, Ana-María; Eriksson, Stefan; Looi, Lai-Meng; Negra, Diane (2022). "Perceptions on the prevalence and impact of predatory academic journals and conferences: A global survey of researchers". *Learned publishing*, v. 35, n. 4, pp. 516-528.

<https://doi.org/10.1002/leap.1458>

Eykens, Joshua; Guns, Raf; Rahman, A. Jakaria; Engels, Tim C. (2019). "Identifying publications in questionable journals in the context of performance-based research funding". *PloS one*, v. 14, n. 11, e0224541.

<https://doi.org/10.1371/journal.pone.0224541>

Frandsen, Tove-Faber (2017). "Are predatory journals undermining the credibility of science? A bibliometric analysis of citers". *Scientometrics*, v. 113, n. 3, pp.1513-1528.

<https://link.springer.com/article/10.1007/s11192-017-2520-x>

Frandsen, Tove-Faber (2022). "Authors publishing repeatedly in predatory journals: An analysis of Scopus articles". *Learned publishing*, v. 35, n. 4, pp. 598-604.

<https://doi.org/10.1002/leap.1489>

Gallent-Torres, Cinta (2022). "Editorial misconduct: the case of online predatory journals". *Heliyon*, v. 8, n. 3, e08999.

<https://doi.org/10.1016/j.heliyon.2022.e08999>

Grove, Jack (2022). *Academic fraud factories are booming, warns plagiarism sleuth*. Times Higher Education, January 19.

<https://www.timeshighereducation.com/news/academic-fraud-factories-are-booming-warns-plagiarism-sleuth>

IAP - InterAcademy Partnership (2022). *Combatting predatory academic journals and conferences*. Report, March.

<https://www.interacademies.org/publication/predatory-practices-report-English>

Japri, Nurul-Ain (2023). *KPT larang universiti awam terbit artikel dalam Hindawi, Frontier dan MDPI [The Ministry of Higher Education prohibits universities from publishing articles in Hindawi, Frontier dan MDPI]*. Mediavariasi.

<https://mediavariasi.com/2023/07/kpt-larang-universiti-awam-terbit-artikel-dalam-hindawi-frontier-dan-mdpi>

Kurt, Serhat (2018). "Why do authors publish in predatory journals?". *Learned publishing*, v. 31, n. 2, pp. 141-147.

<https://doi.org/10.1002/leap.1150>

McCann, Terence V.; Polacsek, Meg (2018). "False gold: Safely navigating open access publishing to avoid predatory publishers and journals". *Journal of advanced nursing*, v. 74, n. 4, pp. 809-817.

<https://doi.org/10.1111/jan.13483>

Moher, David; Shamseer, Larissa; Cobey, Kelly D.; Lalu, Manoj M.; Galipeau, James; Avey, Marc T.; Ahmadzai, Nadera; Alabousi, Mostafa; Barbeau, Pauline; Beck, Andrew; Daniel, Raymond; Frank, Robert; Ghannad, Mona; Hamel, Candyce; Hersi, Mona; Hutton, Brian; Isupov, Inga; McGrath, Trevor A.; McInnes, Matthew D. F.; Page, Matthew J. et al. (2017). "Stop this waste of people, animals and money". *Nature*, v. 549, n. 7670, pp. 23-25.
<https://doi.org/10.1038/549023a>

Moher, David; Srivastava, Anubhav (2015). "You are invited to submit...". *BMC medicine*, v. 13, 180.
<https://doi.org/10.1186/s12916-015-0423-3>

Nicholas, David; Herman, Eti; Watkinson, Anthony; Xu, Jie; Abrizah, Abdullah; Rodríguez-Bravo, Blanca; Boukacem-Zeghmouri, Cherifa; Polezhaeva, Tatiana; Świgoń, Marzena (2021). "Early career researchers and predatory publishing: The views and behaviours of the millennials". *Foresight and STI governance*, v. 15, n. 1, pp. 56-65.
<https://doi.org/10.17323/2500-2597.2021.1.56.65>

Nicholas, David; Herman, Eti; Clark, David; Boukacem-Zeghmouri, Cherifa; Rodríguez-Bravo, Blanca; Abrizah, Abdullah; Watkinson, Anthony; Xu, Jie; Sims, David; Serbina, Galina; Świgoń, Marzena; Jamali, Hamid R.; Tenopir, Carol; Allard, Suzie (2022). "Choosing the 'right' journal for publication: Perceptions and practices of pandemic-era early career researchers". *Learned publishing*, v. 35, n. 4, pp. 605-616.
<https://doi.org/10.1002/leap.1488>

Nicholas, David; Rodríguez-Bravo, Blanca; Boukacem-Zeghmouri, Chérifa; Herman, Eti; Clark, David; Xu, Jie; Abrizah, Abdulah; Świgoń, Marzena; Watkinson, Anthony; Sims, David; Jamali, Hamid R.; Tenopir, Carol; Allard, Suzie (2023). "Early career researchers and predatory journals during the Covid-19 pandemic. An international analysis". *Profesional de la información*, v. 32, n. 1, e320117.
<https://doi.org/10.3145/epi.2023.ene.17>

Nobes, Andi; Harris, Siân (2019). "Open access in low-and middle-income countries: attitudes and experiences of researchers". *Emerald open research*, v. 1, n. 17.
<https://doi.org/10.35241/emeraldopenres.13325>

Oviedo-García, M. Ángeles (2021). "Journal citation reports and the definition of a predatory journal: The case of the Multidisciplinary Digital Publishing Institute (MDPI)". *Research evaluation*, v. 30, n. 3, pp. 405-419a.
<https://doi.org/10.1093/reseval/rvab020>

Perlin, Marcelo S.; Imasato, Takeyoshi; Borenstein, Denis (2018). "Is predatory publishing a real threat? Evidence from a large database study". *Scientometrics*, v. 116, n. 1, pp. 255-273.
<https://doi.org/10.1007/s11192-018-2750-6>

Petrou, Christos (2020). *Guest post – MDPI's remarkable growth*. The Scholarly Kitchen, Aug 10.
<https://scholarlykitchen.sspnet.org/2020/08/10/guest-post-mdpis-remarkable-growth>

Pyne, Derek (2017). "The rewards of predatory publications at a small business school". *Journal of scholarly publishing*, v. 48, n. 3, pp. 137-160.
<https://doi.org/10.3138/jsp.48.3.137>

Siler, Kyle (2020). "Demarcating spectrums of predatory publishing: Economic and institutional sources of academic legitimacy". *Journal of the Association for Information Science and Technology*, v. 71, n. 11, pp. 1386-1401.
<https://doi.org/10.1002/asi.24339>

Shaghaei, Najmeh; Wien, Charlotte; Holck, Jakob; Thiesen, Anita L.; Ellegaard, Ole; Vlachos, Evgenios; Drachen, Thea (2018). "Being a deliberate prey of a predator: Researchers' thoughts after having published in predatory journal". *Liber quarterly*, v. 28, n. 1.
<https://doi.org/10.18352/lq.10259>

Shen, Cenyu; Björk, Bo-Christer (2015). "'Predatory' open access: A longitudinal study of article volumes and market characteristics". *BMC medicine*, v. 13, 230.
<https://bmcmmedicine.biomedcentral.com/articles/10.1186/s12916-015-0469-2>

Teixeira-da-Silva, Jaime A.; Dunleavy, Daniel J.; Moradzadeh, Mina; Eykens, Joshua (2021). "A credit-like rating system to determine the legitimacy of scientific journals and publishers". *Scientometrics*, v. 126, n. 10, pp. 8589-8616.
<https://doi.org/10.1007/s11192-021-04118-3>

Wallace, Frederick H.; Perri, Timothy J. (2018). "Economists behaving badly: publications in predatory journals". *Scientometrics*, v. 115, n. 2, pp. 749-766.
<https://doi.org/10.1007/s11192-018-2690-1>

Xia, Jingfeng; Harmon, Jennifer L.; Connolly, Kevin G.; Donnelly, Ryan M.; Anderson, Mary R.; Howard, Heather A. (2015). "Who publishes in 'predatory' journals?". *Journal of the Association for Information Science & Technology*, v. 66, n. 7, pp. 1406-1417.
<https://doi.org/10.1002/asi.23265>

A search strategy for publications in interdisciplinary research

Wenjing Xiong; Ping Zhou

Recommended citation:

Xiong, Wenjing; Zhou, Ping (2023). "A search strategy for publications in interdisciplinary research". *Profesional de la información*, v. 32, n. 5, e320522.

<https://doi.org/10.3145/epi.2023.sep.22>

Article received on May 19th 2023
Approved on September 27th 2023



Wenjing Xiong

<https://orcid.org/0000-0002-8837-5178>

Zhejiang University
School of Public Affairs
Dept. of Information Resources Management
No. 866 Yuhangtang Road
310059 Hangzhou, China
wenjing_xiong@zju.edu.cn



Ping Zhou ✉

<https://orcid.org/0000-0001-5471-5261>

Zhejiang University
School of Public Affairs
Dept. of Information Resources Management
No. 866 Yuhangtang Road
310059 Hangzhou, China
pingzhou@zju.edu.cn

Abstract

To retrieve the right collection of publications in interdisciplinary research, we have developed a search strategy with four progressive steps and take the area of public affairs (PA) as a case study. A set of seed publications in PA is first identified, followed by the construction of a pool set of publications with wider coverage for refinement in the next step, which is critical and in which an expanded set of publications is established on the basis of the references and text semantic information, thus generating two respective subsets. One of these subsets is obtained on the basis of the number of references shared between each publication pair between the seed set and the pool set. To optimize the results, we construct two models, viz. a support vector machine (SVM) and a fully connected neural network (FCNN), and find that the FCNN model outperforms the SVM model. The second subset of publications are collected by selecting the publications with high topic similarity to the seed publications collected in the first step. The final step is to integrate the seed publications with the expanded publications collected in steps 1 and 3. The results show that PA research involves an extremely wide range of disciplines ($n = 45$), among which public administration, environmental sciences, economics, management, and health policy and services, among others, play the most significant roles.

Keywords

Search strategies; Information retrieval; Public affairs; Interdisciplinary research; Field delineation; Bibliometric analysis; Publications; Scientific research.

Funding

This study received financial support from the *National Natural Science Foundation of China* (NSFC, no. 71843012).

1. Introduction

Scientific research involving multiple disciplines/fields plays a significant role in the development of science and the social sciences, thus leading to numerous studies on interdisciplinary research from various perspectives (Lovelock; Margulis, 1974; Wagner *et al.*, 2011; Wang *et al.*, 2015; Glänzel; Debackere, 2021; Ledford, 2015). How to define the disciplinary attribution of interdisciplinary research remains a challenge because the involved knowledge is from two or more disciplines (Klein *et al.*, 1997; Repko *et al.*, 2007; OECD, 1972; Rotolo *et al.*, 2015) that may vary depending on the specific research topic (Glänzel *et al.*, 2016). It thus becomes difficult to retrieve interdisciplinary publications on the basis of existing classification schemes such as *Web of Science* (WoS) subject categories. Working at the level of scientific journals, the WoS scheme consists of approximately 250 research areas equivalent to subject categories. A multidisciplinary journal may be assigned to one or more research areas (Leydesdorff; Ràfols, 2009). To improve the WoS journal-based system, some science and technology organizations have developed their own, journal-level classification systems (Archambault *et al.*, 2011; Glänzel; Schubert, 2003; Boyack; Klavans, 2010). Such classification schemes, however, cannot be used directly to define interdisciplinary publications.

Classification schemes at the publication level may better reflect research subjects and have been explored at a relatively small scale in the early period (e.g., Griffith *et al.*, 1974; Small; Griffith, 1974; Small; Sweeney, 1985), until 2010 when a



large-scale publication-level classification scheme was proposed (Boyack; Klavans, 2010). On the basis of direct citation relationships between publications, the Leiden classification system covers scientific fields (Waltman; Van-Eck, 2012) and is composed of three hierarchies, with 20 research areas at the first level, 672 research areas at the second level, and 22,412 research areas at the third level. Working at publication level instead of journal level, the Leiden classification system can classify specific scientific areas in more detail and thus match the current structure of scientific research more closely. Using the Leiden methodology, interdisciplinary publications can be harvested by following simple steps. The requirements in terms of computing time and memory usage, however, are too demanding for individual researchers using a standard desktop computer. In addition, the Leiden methodology does not consider indirect citation relations, which may lead to an incomplete collection of interdisciplinary publications.

To collect interdisciplinary publications on specific topics such as innovation systems, big data, or public affairs, five methods can be considered. The first is to identify historical core contributions and community members in a specific area (Fagerberg *et al.*, 2012). With an implicit assumption that the historical core literature remains stable over time, such an approach is not able to reflect the dynamic changes in a specific field. The second relies on a set of keywords and has wider applications (e.g., Huang *et al.*, 2011; 2015; 2019; Ruiz-Navas; Miyazaki, 2018; Liu *et al.*, 2021). Variations exist in keyword approaches; For example, some researchers regard keywords as a core set that defines a specific field (Huang *et al.*, 2011), which requires strong knowledge of the field and the consideration of emerging terms. By making full use of the feature of semi-automatic iteration of keyword variants, some researchers first apply a keyword retrieval strategy to identify a core set of publications and then supplement or modify this core set by using high-frequency words (e.g., Liu *et al.*, 2021; Suominen *et al.*, 2016; Shapira *et al.*, 2017), consulting experts in related fields (Huang *et al.*, 2015), or using synonyms and subordinate words through a retrieval formula (Ruiz-Navas; Miyazaki, 2018). The third is to expand related publications through direct citation relations among publications or journals (e.g., Waltman; Van-Eck, 2012; Muñoz-Écija *et al.*, 2019; Bassecouard *et al.*, 2007), which requires high computational and memory capacity. To avoid such capacity demands, researchers start with a core set of publications and then collect publications linked by citing and cited relationships (Wang *et al.*, 2019). However, this may result in a large number of redundant documents with low relevance to the research topic being included in the core document set. In addition, starting from a core set of publications, the fourth takes advantage of co-citation relations and bibliographic coupling (Kostoff *et al.*, 2006; Zitt; Bassecouard, 2006; Alencar *et al.*, 2007; Soós; Kiss, 2020; Zhou *et al.*, 2019). This approach, however, is sensitive to the coverage of the starting corpus, in addition to the time lag of citations (Mogoutov; Kahane, 2007).

As a comprehensive retrieval strategy with multiple rounds of iterations (Glänzel, 2015), the fifth includes two critical tasks and is more accurate than the previous four. Step 1 is to determine a core set of documents (Zitt; Bassecouard, 2006) that can represent the subject in question well, and step 2 is to extend this core set with relevant documents on the basis of thematic similarity through different approaches. To ensure that the documents collected in the second step are relevant to the subject area, a bibliometric-added retrieval method was proposed, combining complex structures instead of individual search terms. All types of search fields, including keywords, terms, subject headings, journal titles, citations and references, and even organization addresses and author names/identifiers can be incorporated into this retrieval strategy. Taking the approach proposed by Rakas and Hain (2019), for example, they first collected a seed set of publications successively through keyword retrieval and highly cited publications, and then collected a set of “relevant documents” on the basis of overlapping bibliography (i.e., two articles with a high degree of overlap with a larger number of identical references). The two sets of publications are integrated as a final corpus for subsequent analysis. Duplicates and the publications that were not cited at least once are excluded. However, defects can exist in the process of publication collection because documents that are important but that have not yet attracted community attention may be missed. Moreover, extracting 500 publications with the highest bibliographical overlap with each seed publication may result in the inclusion of publications that are irrelevant or less relevant to the topic under consideration.

The diversity and variety of interdisciplinary research (Leydesdorff *et al.*, 2019) result in difficulty when it comes to the classification and retrieval of interdisciplinary publications. Most existing search strategies have pros and cons. The current paper proposes a general search strategy that can be applied to any interdisciplinary subject or field, and then subsequently applies this strategy to the discipline of public affairs (PA), a typical interdisciplinary area.

2. A general search strategy

The disciplines or fields involved in interdisciplinary research may vary significantly in terms of their number and type, resulting in various search strategies, such those introduced above. In the case where no single approach is widely accepted, it becomes necessary to propose a framework that is applicable to any interdisciplinary research field. Below, we first introduce such a general search strategy with four steps, and then apply the strategy to the publication collection of public affairs. Compared with other interdisciplinary subjects or fields, the interdisciplinary situation of public affairs is more complex in terms of both the variety and diversity of the disciplinary/field. A search strategy applicable to PA can thus be easily adjusted to other interdisciplinary subjects/fields with less diversity or variety.

A strategy should consider both recall and precision. According to Bradford’s law and by summarizing the pros and cons of existing methods, we propose a framework with four steps: Emphasizing precision, the first step is to construct a seed

set of publications. The second and third steps take recall into consideration but with different focuses: step 2 aims to construct a pool set of publications so as to enlarge the coverage, whereas step 3 aims to construct an expanded set of publications by extracting publications from the pool set that are similar to the target discipline or field. In the fourth step, a final corpus of publications is obtained by integrating the seed publications and expanded publications.

The seed publications collected in the first step should be the most representative of the target discipline or field, to ensure the relevance of the expanded publications collected in next steps. The methods to be chosen are dependent on the degree of interdisciplinarity, for instance, with or without a clear core field/discipline. For those with a clear core discipline or field (e.g., public affairs or digital finance), a list of representative journals—core journals according to Bradford’s law, subject category, or other combined methods—can be applied; for those without a definite core field or discipline (e.g., digital governance or innovation systems), a retrieval method using a set of topic words can be applied. To ensure the representativeness of the obtained collection, other bibliometric methods (e.g., top-high citations) can be applied.

To establish a pool set of publications with wider coverage, a well-accepted subject classification system with the targeted interdisciplinary category can be used. In addition to that of the *WoS*, national and organizational subject classification systems may also work. For example, the classification system of the *National Natural Science Foundation of China (NSFC)* is used in the current study. All the mentioned classification systems are based on journals.

Nevertheless, the problem with such types of classification is obvious, because not all publications in journals of the same interdisciplinary subject category are necessarily similar. Thus, publications with similarity in the pool set should be identified and extracted, which is the task of step 3: constructing an expanded set by assessing the subject similarity of publications in the pool with each of the seeds. An assessment of either reference similarity or topic similarity of the publications in the two sets (i.e., the pool set and the seed set) can serve this purpose. To assess reference similarity, publications in the two sets are paired to calculate dynamic thresholds for the number of common references by adopting models such as a support vector machine (SVM) or a fully connected neural network (FCNN). The operational results will indicate which model should be used. On the basis of the dynamic threshold, similar publication pairs can be identified. To evaluate topic similarity, a text semantic similarity algorithm (Jaccard, Word2vec, and TF-IDF) may be an option.

3. The case of public affairs

After more than 100 years of development, the discipline of public affairs has continuously absorbed knowledge from numerous disciplines/fields, such as political science, management, law, psychology, sociology, economics, and information science (Benz, 2005; Rodgers; Rodgers, 2000; Fleisher; Blair, 1999; Harris; Moss, 2001a; Harris; Fleisher, 2005), with the results of PA being regarded as a “borrowing discipline” (Stallings, 1986). With economic development and social progress, PA has been facing increasingly new problems and challenges (Raadschelders, 2011). Many sophisticated policy issues simply cannot be addressed within the narrow boundaries of traditional disciplines. Absorbing knowledge from a broader scope of subjects, including the natural sciences and engineering, has accompanied the development of PA as a discipline, and thus further increases its interdisciplinarity (Kettl; Milward, 1996). Some “distant” subjects, such as computer science, mathematics, and statistics, have become an inexhaustible resource for PA (Savage, 1974). Undoubtedly, the extensive assimilation of knowledge contributes significantly to the development of PA; in the meantime, however, there is difficulty in delineating the field composition of the discipline (Yang, 2019; Steckmest, 1982; Harris; Moss, 2001b). Failing to clarify this field/disciplinary attribution may limit researchers’ reference scope and thus affect the development of the discipline, weakening scholars’ sense of field belonging and even pushing PA into a crisis of discipline extinction over time (Ostrom, 1974; White et al., 1996; Seibel, 1996; Denhardt, 2004). In addition, the field variety within PA is significant owing to the broad inclusion of fields or disciplines such as public administration, environment protection, social security, digital governance, etc. Different public affairs universities may vary in field coverage, which results in enormous challenges to benchmarking practices (e.g., university ranking and performance evaluation) because of the dependence of the field on research evaluation (Hicks et al., 2015; Gómez-Núñez et al., 2014).

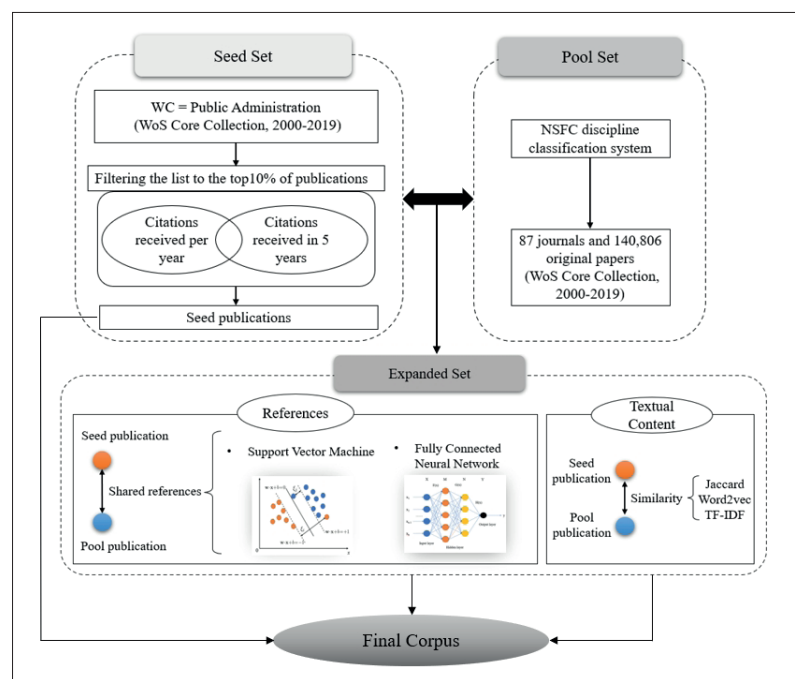


Figure 1. Search framework for publications in public affairs

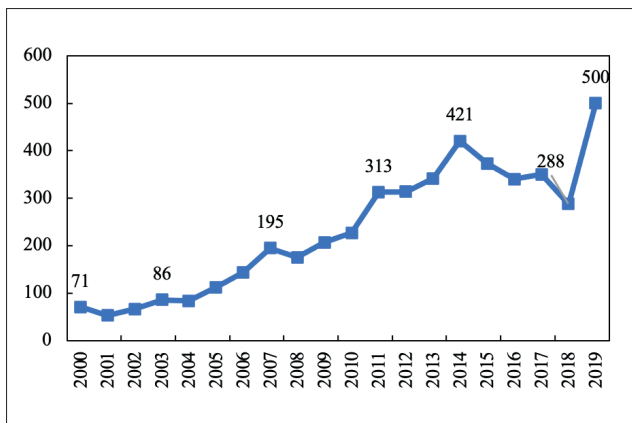


Figure 2. Annual development of the seed publications

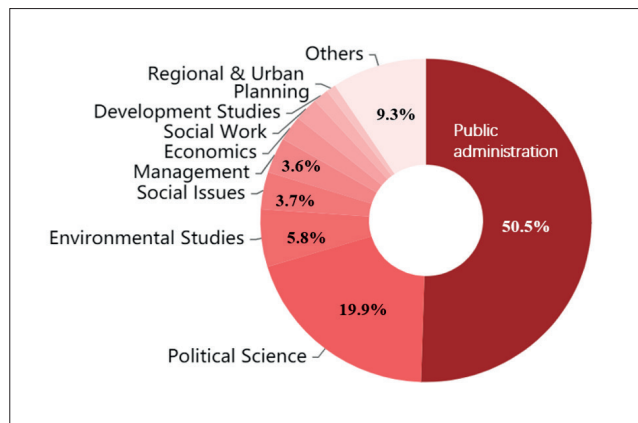


Figure 3. Discipline distribution of the seed publications

To collect publications in public affairs, we follow the general search strategy proposed above. The detailed framework is illustrated in Figure 1.

4. Step 1. Defining a seed set of publications

Public affairs involves various relationships surrounding governments (Fleisher, 2001; Harris; Fleisher, 2005; Lerbinger, 2006), with public administration as the core. Thus, public administration can be regarded as the core field of public affairs. Articles indexed in the *WoS* and published in the years 2000-2019 in the public administration subject category are first collected. Among the obtained publications, only those with top 10% citations in 5 years or per year are considered as the seed set of publications of public affairs. After removing duplicates, 4,664 seed publications are confirmed. The historical development and discipline distribution of the seed publications are shown in Figures 2 and 3. Over 10 years, the seed publications of PA have increased about seven fold. PA itself is multidisciplinary, involving more than 10 disciplines, and more than half of the seeds involve research in the public administration field. Political science is also one of the main fields of PA.

5. Step 2. Constructing a pool set of publications

Defining the seed set of publications by using highly cited articles may result in a lack of knowledge regarding other important contributions. It is necessary to construct a pool set of publications with wider coverage by searching the journals related to public affairs. The list of 91 journals in the current study is provided by the *Development Strategy for the Discipline of Public Affairs* (Xue et al., 2017), an output of a key project of the *National Natural Science Foundation of China*. On the basis of the journal list that is widely accepted by universities, as well as the practice of Yu (2019), four journals (energy, applied energy, agricultural water management, and energy economics) were eliminated owing to the keyword clustering analysis topics of the publications in the four journals mainly being related to energy, resource, and environmental technology, with little relevance to public affairs. Furthermore, the publication sizes of these four journals are unstable and have increased significantly from several hundred to more than 4,000 since 2010, which may affect the statistical results of PA research. The other 87 journals are indexed in the *Web of Science (WoS)*, and include a total of 140,806 articles published in the period 2000-2019. In other words, the pool set comprises 140,806 articles. The wide disciplinary coverage of publications in the pool set is presented in Table 1. The public administration field does not take first position anymore; instead, environment-related fields and economics take the lead. The purpose of enlarging the disciplinary coverage has been reached in a reasonable way.

Table 1. Distribution of top 20 disciplines in the pool

Discipline	N	Ratio
Environmental sciences	41,159	13.5%
Environmental studies	34,339	11.2%
Economics	26,401	8.6%
Public, environmental, and occupational health	16,023	5.2%
Public administration	13,665	4.5%
Health policy and services	13,208	4.3%
Healthcare sciences and services	12,559	4.1%
Ecology	10,297	3.4%
Energy and fuels	9,781	3.2%
Information science and library science	9,675	3.2%
Management	9,605	3.1%
Education and educational research	8,504	2.8%
Business	6,205	2.0%
Ergonomics	6,106	2.0%
Social sciences, interdisciplinary	5,872	1.9%
Transportation	5,798	1.9%
Geography	5,082	1.7%
Political science	4,969	1.6%
Computer science, interdisciplinary applications	4,926	1.6%
Agriculture, multidisciplinary	4,739	1.6%

6. Step 3. Establishing an expanded set of publications

The wider coverage of the pool set may include less or irrelevant documents. To ensure the publications are indeed relevant to public affairs, we adopt two methods to establishing an expanded set of publications, namely measuring the number of shared references and measuring the topic similarity between the seed publications and those in the pool set. Measuring shared references is more complicated and takes longer to illustrate.

Measurement of shared references. The references of the seed publications and the pool publications are compared in pairs. The number of shared references is used to determine whether two publications are similar. With regard to the appropriate threshold for the shared number of references, variation exists among different studies. The detailed operation is as follows: we first randomly select a sample of 500 publications from the seed set, and then select 500 publications from the pool set and pair them with each initial "seed." A dataset of 500 publication pairs (one-to-one matching) is obtained. Second, on the basis of the similarity of the title, abstract, keywords, and other information, we manually label whether each pair of publications is similar (1 for similar and 0 for not similar pairs). In the meantime, the number of references of each pair and the number of shared references are recorded. **Rakas and Hain (2019)** define the top 500 publications with a high overlap rate with the references in the pool set as the expanded publications of each seed publication. The disadvantages of this method are obvious: a lack of knowledge of citation disparities among document types, the time length of publication, and research topics. Such disadvantages would become severe when the number of shared references is small and where the number of references between publications varies greatly. The similarity between the target publications (a seed publication and a pool publication) in cases A and B in Table 2 would be considered equal on the basis of an absolute threshold, although it is clear that significant variation exists.

Table 2. Determination of similar publications based on an absolute threshold

Case	Seed R	Pool R	Shared R
A	34	294	1
B	4	7	1

Note: Calculated according to the method of **Rakas and Hain (2019)**.

Table 3. Determination of similar publications based on a relative threshold

Case	Seed R	Pool R	Shared R: 20%*Min (Num-seedR, Num-poolR)
C	100	100	20
D	150	140	28
E	100	10	2

Could a relative threshold be a better solution? We use three cases, presented in Table 3, for further understanding. When the difference in the number of references between the two compared papers is small, a relative threshold can be set to roughly determine the similarity on the basis of a smaller number of references [i.e., similarity = 20% × Min (Num-Seed R, Num-Pool R)]. Thus, in case D, the number of references shared between a seed publication and a pool publication is 28. When the difference is large, however, a relative threshold does not work, for instance, case E in Table 3.

Given that neither absolute nor relative thresholds can be used to reasonably determine publication similarity, this study tries to construct high-dimensional models for the determination of publication similarity. Two models, a support vector machine and a fully connected neural network, are constructed and compared to select the one that is optimal. The result of each model is compared with that of the method proposed by **Rakas and Hain (2019)**.

On the basis of sample records of 500 publications selected randomly and indexed manually, a formula is developed to simulate the relationship reflected by the number of references of each of the seed publications (X), the pool publication (Y), and the shared references (Z). The number of shared references (Z) of each pool publication and the seed publication are identified manually. The labeling variable is a dichotomous variable: whether each pair of publications is similar (1 for similar pairs and 0 for dissimilar pairs). It is well known that the number of references of individual publications varies significantly, from a few to tens or hundreds. To make it so that the selected publication is able to represent publications with different numbers of references, and thus to ensure the estimation accuracy and application scope of the model, we carried out five steps. Step 1 was to arrange the number of references of each seed and each pool publication in descending order and exclude those with fewer than two references. Step 2 was to classify publications on the basis of their number of references into subset 1 (2-100 references), subset 2 (101-200 references), etc. The determination of the interval of the number of references is empirical. Step 3 was to randomly select publications from each subset, and in the end, a total of 500 sample publications from the seed and pool publication set, respectively, were obtained. Step 4 was to pair-match each of the sample publications in the seed set with each of those in the pool set, whereby a set of 250,000 pairs was obtained ($500 \times 500 = 250,000$). Step 5 was to arrange the number of shared references of each pair in descending order after excluding those with fewer than two shared references, to classify publication pairs on the basis of their number of shared references into subset 1 (2-20 shared references), subset 2 (21-40 shared references), etc., and to select randomly a total of 500 publication pairs from the subsets.

Of the 500 publication pairs, the mean values of the number of references of the seed and the pool publications are 89.6 and 97.0, respectively, while the corresponding modes are 62 and 61, respectively (Table 4). In combination with the results of the analysis of extreme value and standard deviation, one can see that the mean values are significantly affected by extreme values, and the data sampling is relatively dispersed, which is conducive to improving the accuracy of the

subsequent model estimation. More specifically, the samples were divided into two groups according to similar pairs (22.0%) and dissimilar pairs (78.0%) of publication pairs. The mean values of X and Y of the similar publication pairs are greater than those of the dissimilar pairs, which implies that a publication pair with a larger number of references may have more shared references (similar pairs: $Z_{\text{mean}} = 20.5$) and thus be more likely to be judged as a similar publication pair.

Table 4. Descriptive statistics on the number of references of 500 publication pairs

Statistics	Total samples			Dissimilar pairs (78.0%)			Similar pairs (22.0%)		
	X	Y	Z	X	Y	Z	X	Y	Z
Mean	89.6	97.0	9.9	79.5	89.8	6.9	125.7	122.4	20.5
Median	53.0	56.5	3.0	53.0	55.0	3.0	67.0	58.0	12.0
Mode	62	61	2	62	61	2	11	34	5
Standard deviation	107.3	114.9	14.5	97.0	108.1	11.0	131.8	133.7	19.7
Minimum	2	4	2	2	5	2	3	4	2
Maximum	503	660	81	503	660	72	502	482	81

Significant correlations between different variables do exist but with low values (less than 0.6, see Table 5). The correlation between X and Y (0.55) is higher than that between Z and X, Y (0.514 and 0.453, respectively). The correlations of the dependent variable with X and Y (0.093 and 0.041, respectively) were much lower than that with Z (0.488).

Table 5. Spearman's correlations of variables

Variables	X	Y	Z	Similarity
X	1.000			
Y	0.550** (0.000)	1.000		
Z	0.514** (0.000)	0.453** (0.000)	1.000	
Similarity	0.093* (0.037)	0.041 (0.360)	0.488** (0.000)	1.000

Note: *5% and **1% significance levels. The numbers shown in parentheses correspond to p -values.

7. Model 1: support vector machine

The support vector machine (SVM) is a supervised machine learning model for data classification that works by identifying the optimal hyperplane on the basis of the principle of maximum interval in the feature space (Cortes; Vapnik, 1995). Compared with methods such as setting percentage or quantity thresholds, this model can capture high-dimensional data features effectively. Since its inception, the SVM has been widely used in areas such as facial recognition, text classification, data classification, etc. (Qin; He, 2005; Sun *et al.*, 2002; Chen *et al.*, 2001; Campbell *et al.*, 2006; Srivastava; Bhambhu, 2010). This study proposes to approach the data binary classification issue by using the SVM model. We divided the sample data into two parts, consisting of the training data (400 publication pairs) for estimating the model parameters, and the testing data (100 publication pairs) for evaluating the model accuracy. The training and test data were randomly selected in a ratio of 4:1 for model fitting for 10 times. The given training samples ($D = \{(x_1, y_1), (x_2, y_2), \dots, (x_s, y_s)\}$, where s is the number of samples) are put into either negative or positive categories as labeled [$y \in (-1, 1)$]. The model can identify the optimal hyperplane, which requires maximizing the sum of sample-plane distances by dynamically adjusting the hyperplane parameters after placing all the training samples onto both sides of the hyperplane. The expression for the hyperplane is

$$H = wx + b = 0, x = (x_1, x_2, \dots, x_s) \quad (1)$$

The normal vector of the hyperplane is denoted as $w = (w_1, w_2, \dots, w_s)$, and the displacement as b . Finding the optimal hyperplane can be transformed into an optimization problem:

$$\begin{aligned} \min & \frac{|w|^2}{2} \\ \text{s.t. } & y_i [(w^T x_i) + b] \geq 1 \quad i = 1, 2, \dots, s \end{aligned} \quad (2)$$

As completely linearly separable cases are comparatively rare in reality, we can lower the standard for the support vector machine from completely linearly separable to tolerate slight classification errors by adding a slack variable $\xi_i (\xi_i \geq 0)$ and a penalty term $C (C > 0)$ to the original linearly separable support vector machine. The optimization problem can be solved by:

$$\begin{aligned} \min & \frac{|w|^2}{2} + C \sum_{i=1}^l \xi_i \\ \text{s.t. } & y_i [(w^T x_i) + b] \geq 1 - \xi_i, i = 1, 2, \dots, s \\ & \xi_i \geq 0, i = 1, 2, \dots, s \end{aligned} \quad (3)$$

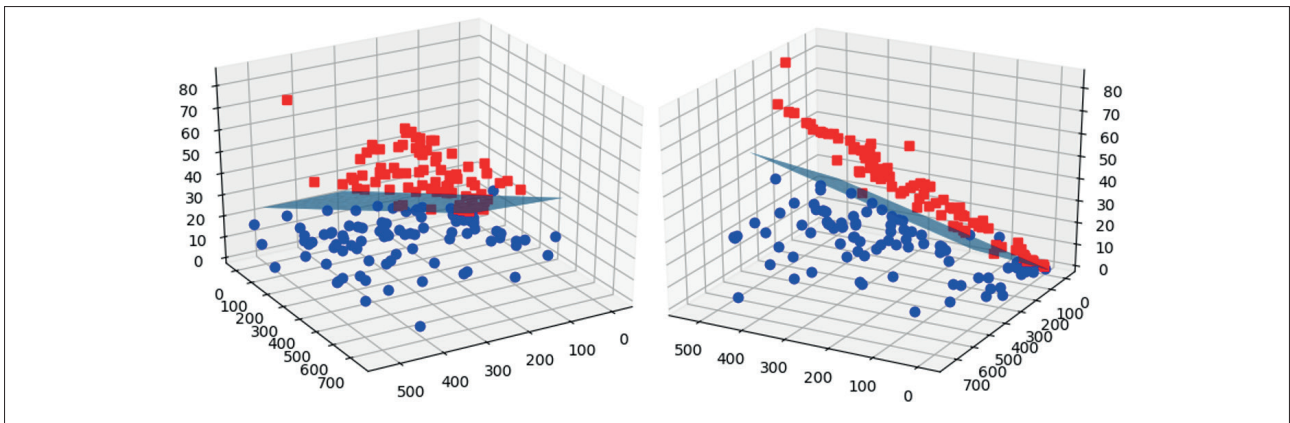


Figure 4. Support vector machine model

The model has a higher level of fault tolerance when $C = 1$. To solve Eq. (3), the Lagrange function can be constructed by adopting the Lagrange multiplier method, then the optimal solution to Eq. (3) can be obtained by solving its dual problem:

$$\min \frac{1}{2} \sum_{i=1}^s \sum_{j=1}^s \alpha_i \alpha_j y_i y_j x_i^T x_j - \sum_{i=1}^s \alpha_i \quad (4)$$

s.t. $\sum_{i=1}^n \alpha_i y_i = 0$
 $0 \leq \alpha_i \leq C, i = 1, 2, \dots, s$

The sequential minimal optimization (SMO) algorithm can then be used to solve for α . Subsequently, w is solved via the equation $w = \sum_{i=1}^n \alpha_i y_i x_i$. The final function is:

$$f(x) = w^T x + b = \sum_{i=1}^n \alpha_i y_i x_i^T x_j + b \quad (5)$$

The hyperplane fitting result is shown in Figure 4. (The data presented in Figure 4 are the training samples.)

8. Model 2: fully connected neural network

A fully connected neural network (FCNN) is often used in data classification, information recognition, and prediction because of its strong self-learning and fitting abilities. It can fully approximate any complex nonlinear relationship (Liu *et al.*, 2018; Raiyani *et al.*, 2018). The FCNN constructed in the current study contains an input layer, an output layer, and two hidden layers. The training and testing datasets are the same as for the SVM. The input of each layer can be regarded as the output of the previous layer. The process of network generation consists of two steps in opposite directions, that is, forward propagation to calculate the model loss and backward propagation to update the model parameters. Finally, the model parameter value under the condition of minimum loss is obtained (Figure 5).

The forward propagation process of the network is to take the vector $x = [x_1, x_2, x_3]^T$ as the input data and to combine it linearly with the weights w_1, w_2, w_3 , and the bias term b . The nonlinear transformation is achieved by a linear rectification linear unit (ReLU) function, and the output vector z^1 is obtained and then used as the input data for the first hidden layer:

$$z^1 = f_1(W^1 x + B^1) \quad (6)$$

where $W^1 \in R^{3 \times n}$, $B^1 \in R^n$, and n is the number of neurons (32 neurons in this study) in this layer. The output vector of the first hidden layer can be regarded as the input vector of the second hidden layer. The formula for the second hidden layer is as follows:

$$z^2 = f_2(W^2 z^1 + B^2) \quad (7)$$

where $W^2 \in R^{n \times m}$ is the output weight matrix, $B^2 \in R^m$ is the bias vector, and m is the number of neurons (16 neurons in this study) in this layer. To prevent the model from overfitting and to improve the generalization ability of the network, a dropout layer is added to randomly stop updating the weights of some neurons (20% of neurons). The output vector of the second hidden layer is then used as the input vector of the output layer:

$$\hat{y} = f_3(W^3 z^2 + B^3) \quad (8)$$

where $W^3 \in R^{m \times 2}$, $B^3 \in R^2$. The network can be regarded as a nonlinear composite function:

$$\hat{y} = f(x, W, B) = f_3(W^3 f_2(W^2 f_1(W^1 x + B^1) + B^2) + B^3) \quad (9)$$

Through a softmax function, the output vector is transformed into two probability values (between 0 and 1), and the larger one can be converted into the prediction value 0 or 1 for judging the similarity between the seed publications and

pool publications. The difference between the values of prediction (\hat{y}) and ground truth (y) is calculated by a loss function and can be used as the criterion for model convergence.

The process of backpropagation is applied to optimize the model through the gradient descent algorithm. Common gradient descent algorithms include full gradient descent (FGD) and batch gradient descent (BGD) with relatively stable parameter updating direction but slow convergence speed (Huo; Huang, 2017). Stochastic gradient descent (SDG) improves the convergence speed of the first two methods by selecting training samples randomly in each iteration, but has the disadvantage of unstable parameter updating (Bottou, 2012). The adaptive gradient algorithm (AdaGrad) can obtain efficient and relatively correct results by constantly adjusting the learning rate in the iterative process (η), although calculation errors caused by the sharp decrease of the gradient are hard to avoid (Liu et al., 2020). In the current study, an optimization algorithm using root-mean-square propagation (RMSprop) of AdaGrad is used to construct the prediction model with a learning rate of 0.001 (Duchi et al., 2011; Dauphin et al., 2015). The formula is as follows:

$$\begin{aligned}
 g &\leftarrow \partial f(x^i; \theta), y^i \\
 r &\leftarrow \rho r + (1 - \rho)g \odot g \\
 \Delta \theta &= -\frac{\eta}{\sqrt{\delta + r}} \odot g \\
 \theta &\leftarrow \theta + \Delta \theta
 \end{aligned}
 \tag{10}$$

where δ , ϑ , ρ , and r are the constant value, model parameter, the exponential decay rate, and the gradient accumulation, respectively. The final training model is obtained with 97.5% prediction accuracy on the basis of the test data.

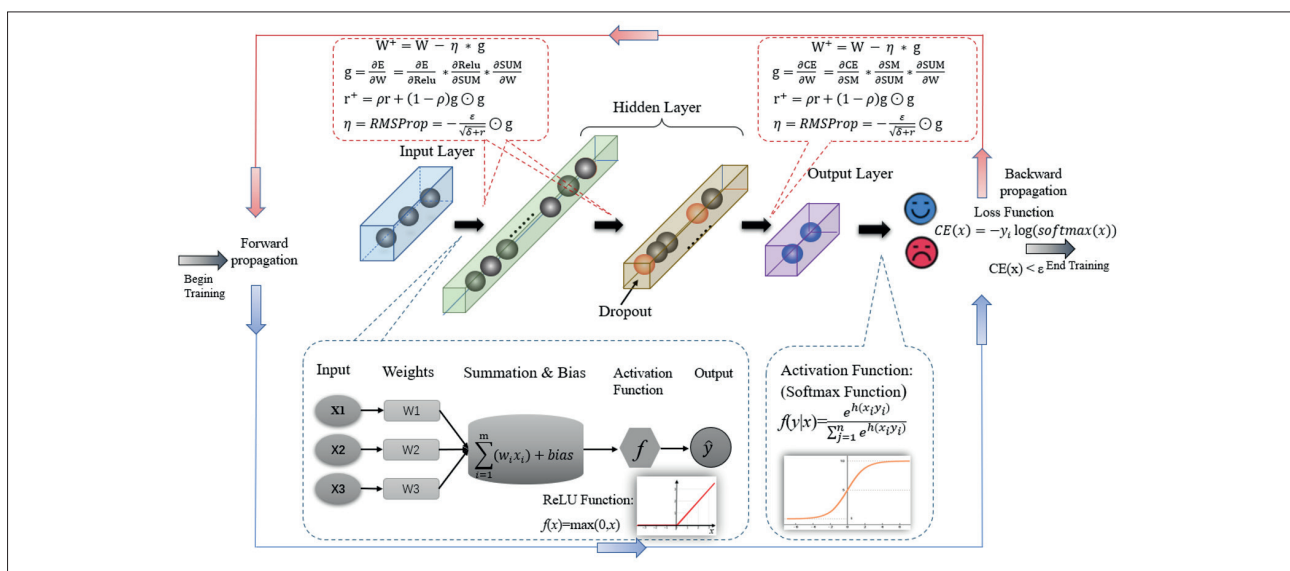


Figure 5. Fully connected neural network

9. Constructing two sets of expanded publications

Before running the above models, the follow steps are required to identify the shared references. The first step is to extract the shared references by comparing the DOIs of each reference pair between the seed publications and the pool publications. For references that cannot be matched by DOIs, other information in the CR field (i.e., author, year, journal, volume, and issue) can be used for fuzzy matching by adopting methods such as ignorance of punctuation differences, N-gram, and cosine (Abdulhayoglu et al., 2016; Glanzel; Czerwon, 1996; Sen; Gan, 1983). The current study adopts the method of ignorance of punctuation differences. By running the above models (SVM and FCNN), two sets of expanded publications are obtained (Figure 6). The training of the SVM or FCNN was performed on a NVIDIA GeForce GTX 1070 with an AMD Ryzen 1700. Sklearn 1.0.2 and Tensorflow 2.5.0 were used for the SVM and FCNN, respectively. For a computer with a memory of less than 500M, it takes less than 60 seconds to complete the training for each of the models.

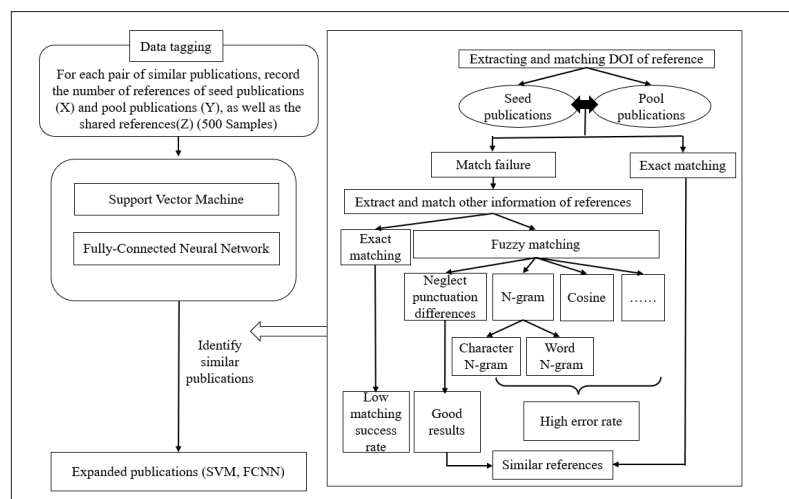


Figure 6. The process of constructing two sets of expanded publications

10. Comparison of the expanded collections obtained by the different methods

The three sets of expanded publications obtained by using the methods of Rakas and Hain (2019), SVM, and FCNN are shown in Figure 7. Overlaps exist between any two datasets, but with no full containment relationship. With 88,221 documents, the coverage of the method of Rakas and Hain is the widest, followed successively by that of SVM (22,138) and FCNN (15,925).

The accuracy and comprehensiveness of the three sets are presented in Table 6. The evaluation report of the models (e.g., accuracy, precision, recall, *f1*-score) and topic similarity (i.e., Jaccard, Word2vec, and TF-IDF) represent the fitting effect of each model and the semantic similarity between each seed document and its expanded documents, calculated according to the information in the title, abstract, keywords, and keywords plus. The results from the method of Rakas and Hain are unsatisfactory because of the large number of duplicate publications and the lowest average topic similarity (TF-IDF = 10.4%). The SVM exhibits relatively low accuracy (87.5%), recall (79.2%), and *f1*-score (82.7%). The FCNN performs best, with the highest model accuracy (97.5%), precision (96.2%), recall (98.2%), *f1*-score (97.5%), and average topic similarity (TF-IDF = 25.1%). Thus, the results of the FCNN model are incorporated into the expanded set.

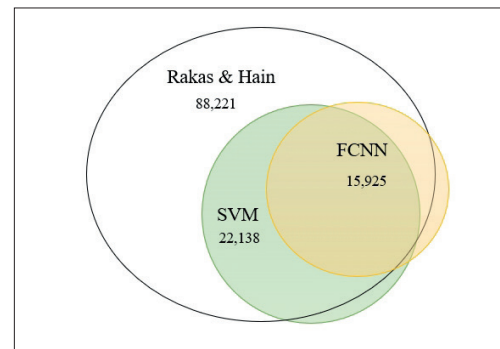


Figure 7. Coverage of the three sets of expanded publications

Table 6. Results of the three publication expansion models

Method	Expanded publications	Expanded publications (De-duplicate)	Model evaluation				Topic similarity		
			Accuracy	Precision	Recall	<i>f1</i> -score	Jaccard	Word2vec	TF-IDF
Rakas & Hain	1,668,408	88,221	-	-	-	-	4.0%	9.3%	10.4%
Support Vector Machine (Linear kernel function)	88,885	22,138	95.00%	94.05%	94.05%	94.05%	6.6%	18.4%	19.7%
Fully-Connected Neural Network	47,189	15,925	97.50%	96.15%	98.21%	97.53%	8.7%	21.9%	25.1%

11. Publication expansion based on topic similarity

Selecting expanded publications simply on the basis of shared references may neglect publication pairs that do not have enough shared references but do study similar topics. Therefore, we apply both shared references and topic similarity to screen out publications in the pool set. In other words, publications in the expanded set are screened out firstly by shared references and then by topic similarity. The topic similarity method is a two-step process. Step 1 is to calculate the topic similarity between the expanded publications obtained through shared references and their corresponding seed publications. The value of 25% (in line with Table 6) will be used to screen out publication pairs in step 3. Step 2 is to pair the rest of the publications (excluding those already screened out through shared references) in the pool set with those in the seed set, and calculate the TF-IDF value of each pair. Step 3 is to screen out those with topic similarity greater than 25% and add them to the expanded set.

By integrating the expanded publications on the basis of shared references and topic similarity, an expanded set (Table 7) containing 50,681 publications is obtained after removing duplicates, with 45 subject categories, among which environmental studies (12.6%), environmental sciences (10.4%), and economics (9.8%) account for a large proportion. The rest of the categories include health care sciences and services, education and educational research, political science,

Table 7. Distribution of the top 20 disciplines in the expanded set

Discipline	N	Ratio
Environmental studies	13,562	12.6%
Environmental sciences	11,218	10.4%
Economics	10,581	9.8%
Public administration	9,455	8.8%
Health policy and services	5,153	4.8%
Management	4,911	4.6%
Healthcare sciences and services	4,798	4.5%
Education and educational research	3,788	3.5%
Energy and fuels	3,767	3.5%
Political science	3,417	3.2%
Public, environmental, and occupational health	3,221	3.0%
Information science and library science	3,089	2.9%
Business	2,599	2.4%
Geography	2,150	2.0%
Ecology	1,775	1.6%
Urban studies	1,682	1.6%
Social issues	1,614	1.5%
Industrial relations and labor	1,475	1.4%
Computer science, interdisciplinary applications	1,402	1.3%
Regional and urban planning	1,287	1.2%

public, environmental and occupational health, information science and library science, etc. (Table 7). Compared with the pool set, the importance of some disciplines in this set has increased, for example, public administration (8.8%), health policy and services (4.8%), and management (4.6%).

12. Step 4. Construct a final corpus of publications

Finally, the seed publications and expanded publications are merged together by excluding duplicates, reviews, and uncited publications (Figure 8). The final corpus contains a total of 55,345 publications, including 45 WoS subject categories and 116 journals and research relating to scientific research management and innovation research, public health and health, safety and risk management, organizational management and human resource management, land resource management, water resource management, climate change, energy and environmental governance, governance research, education, social issues research, and other content connected to public affairs.

In short, the advantages of the multistep method are obvious. The seed set contains the core content of PA, while the expanded set considers its dynamic expansion. Thus, as a result of inclusiveness, it will not provide a specific boundary for PA, but rather a blurry one that tends to include publications from adjacent fields that are also substantially interrelated to PA.

13. Discussion and conclusions

To solve the problem of how to retrieve publications in interdisciplinary research, this study proposes a data-driven and multi-round iterative retrieval strategy that may ameliorate the deviation of retrieval results caused by single-way methods such as subject retrieval or journal retrieval. Selecting a reasonable method for expanding publication coverage is critical for the retrieval strategy, directly affecting the accuracy of the final retrieved results. Compared with the methods of setting an absolute number or relative threshold (e.g., Rakas; Hain, 2019) for expanding coverage, the FCNN model performs better because it addresses the difficulty of comparing the similarity of publications caused by large differences in the number of references. Furthermore, it improves the accuracy of determining similar publications. The proposed search strategy can be applied for the retrieval of any interdisciplinary publications. Adjustment, however, has to be done according to the characteristics of the target disciplines.

The third step of the search strategy—establishing an expanded set of publications—is somewhat complicated and requires knowledge of computer science and bibliometrics, which may, to some extent, limit the application of this methodology. To solve this problem, we plan to develop user-friendly guidelines and to make them freely accessible, so that researchers in fields with little knowledge of computer science and bibliometrics may apply the guidelines to collect an expanded set of publications in the interdisciplinary area of their choosing.

14. References

- Abdulhayoglu, Mehmet-Ali; Thijs, Bart; Jeuris, Wouter** (2016). "Using character n-grams to match a list of publications to references in bibliographic databases". *Scientometrics*, v. 109, n. 3, pp. 1525-1546.
<https://doi.org/10.1007/s11192-016-2066-3>
- Alencar, Maria-Simone-de-M.; Porter, Alan L.; Antunes, Adelaide-Maria-de-Souza** (2007). "Nanopatenting patterns in relation to product life cycle". *Technological forecasting and social change*, v. 74, n. 9, pp. 1661-1680.
<https://doi.org/10.1016/j.techfore.2007.04.002>
- Archambault, É.; Beauchesne, Olivier H.; Caruso, Julie** (2011). "Towards a multilingual, comprehensive and open scientific journal ontology". In: *Proceedings of the 13th International conference of the International Society for Scientometrics and Informetrics*, pp. 66-77. South Africa, Durban.
- Bassecouard, Elise; Lelu, Alain; Zitt, Michel** (2007). "Mapping nanosciences by citation flows: A preliminary analysis". *Scientometrics*, v. 70, n. 3, pp. 859-880.
<https://doi.org/10.1007/s11192-007-0315-1>
- Benz, Arthur** (2005). "Public administrative science in Germany: Problems and prospects of a composite discipline". *Public administration*, v. 83, n. 3, pp. 659-668.
<https://doi.org/10.1111/j.0033-3298.2005.00468.x>
- Bottou, Léon** (2012). *Stochastic gradient descent tricks*. Berlin, Heidelberg: Springer.
https://link.springer.com/chapter/10.1007/978-3-642-35289-8_25
- Boyack, Kevin W.; Klavans, Richard** (2010). "Co-citation analysis, bibliographic coupling, and direct citation: Which citation approach represents the research front most accurately?". *Journal of the American Society for Information Science and Technology*, v. 61, n. 12, pp. 2389-2404.
<https://doi.org/10.1002/asi.21419>

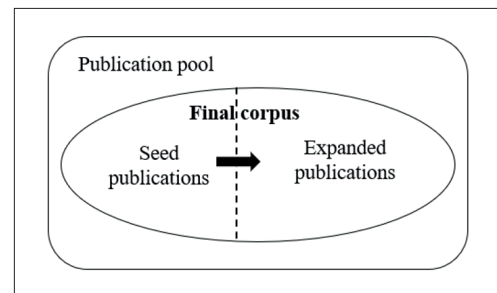


Figure 8. Relationship between different publication sets

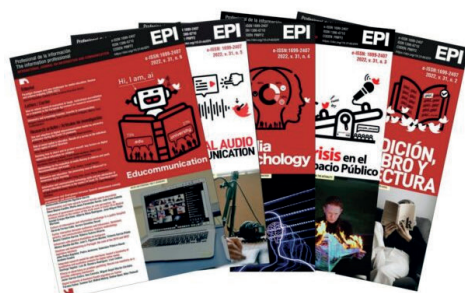
- Campbell, William M.; Sturim, Douglas E.; Reynolds, Douglas A.; Solomonoff, Alex** (2006). "SVM based speaker verification using a GMM supervector kernel and NAP variability compensation". In: *2006 IEEE International conference on acoustics speech and signal processing proceedings*, v. 1, pp. I-I.
- Chen, Yunqiang; Zhou, Xiang-Sean; Huang, Thomas S.** (2001). "One-class SVM for learning in image retrieval". In: *Proceedings 2001 international conference on image processing*, v. 1, pp. 34-37.
<https://doi.org/10.1109/ICIP.2001.958946>
- Cortes, Corinna; Vapnik, Vladimir** (1995). "Support-vector networks". *Machine learning*, v. 20, n. 3, pp. 273-297.
<https://doi.org/10.1007/BF00994018>
- Dauphin, Yan N.; De-Vries, Harm; Bengio, Yoshua** (2015). "Equilibrated adaptive learning rates for non-convex optimization". *Advances in neural information processing systems*.
- Denhardt, Robert B.** (2004). *Theories of public organization*. Belmont: Wadsworth.
- Duchi, John C.; Hazan, Elad; Singer, Yoram** (2011). "Adaptive subgradient methods for online learning and stochastic optimization". *Journal of machine learning research*, v. 12, n. 61, pp. 2121-2159.
<https://jmlr.org/papers/v12/duchi11a.html>
- Fagerberg, Jan; Fosaas, Morten; Sapprasert, Koson** (2012). "Innovation: Exploring the knowledge base". *Research policy*, v. 41, n. 7, pp. 1132-1153.
<https://doi.org/10.1016/j.respol.2012.03.008>
- Fleisher, Craig S.; Blair, Natasha M.** (1999). "Tracing the parallel evolution of public affairs and public relations: An examination of practice, scholarship and teaching". *Journal of communication management*, v. 3, n. 3, pp. 276-292.
<https://doi.org/10.1108/eb026052>
- Glänzel, Wolfgang** (2015). "Bibliometrics-aided retrieval: where information retrieval meets scientometrics". *Scientometrics*, v. 102, n. 3, pp. 2215-2222.
<https://doi.org/10.1007/s11192-014-1480-7>
- Glänzel, Wolfgang; Czerwon, Hans-Jürgen** (1996). "A new methodological approach to bibliographic coupling and its application to the national, regional and institutional level". *Scientometrics*, v. 37, n. 2, pp. 195-221.
<https://doi.org/10.1007/bf02093621>
- Glänzel, Wolfgang; Debackere, Koenraad** (2021). "Various aspects of interdisciplinarity in research and how to quantify and measure those". *Scientometrics*, v. 127, n. 9, pp. 5551-5569.
<https://doi.org/10.1007/s11192-021-04133-4>
- Glänzel, Wolfgang; Schubert, András** (2003). "A new classification scheme of science fields and subfields designed for scientometric evaluation purposes". *Scientometrics*, v. 56, n. 3, pp. 357-367.
<https://doi.org/10.1023/A:1022378804087>
- Gómez-Núñez, Antonio J.; Batagelj, Vladimir; Vargas-Quesada, Benjamín; De-Moya-Anegón, Félix; Chinchilla-Rodríguez, Zaida** (2014). "Optimizing SCImago Journal & Country Rank classification by community detection". *Journal of informetrics*, v. 8, n. 2, pp. 369-383.
<https://doi.org/10.1016/j.joi.2014.01.011>
- Griffith, Berver C.; Small, Henry G.; Stonehill, Judith A.; Dey, Sandra** (1974). "The structure of scientific literatures II: Toward a macro- and microstructure for science". *Science studies*, v. 4, n. 4, pp. 339-365.
<https://doi.org/10.1177/030631277400400402>
- Harris, Phil; Fleisher, Craig S.** (2005). *The handbook of public affairs*. Sage.
<https://doi.org/10.4135/9781848608108>
- Harris, Phil; Moss, Danny** (2001a). "In search of public affairs: a function in search of an identity". *Journal of public affairs*, v. 1, n. 2, pp. 102-110.
<https://doi.org/10.1002/pa.55>
- Harris, Phil; Moss, Danny** (2001b). "Understanding public affairs". *Journal of public affairs*, v. 1, n. 1, pp. 6-8.
<https://doi.org/10.1002/pa.45>
- Hicks, Diana; Wouters, Paul; Waltman, Ludo; De-Rijcke, Sarah; Ràfols, Ismael** (2015). "Bibliometrics: The Leiden Manifesto for research metrics". *Nature*, v. 520, n. 7548, pp. 429-431.
<https://doi.org/10.1038/520429a>
- Huang, Can; Notten, Ad; Rasters, Nico** (2010). "Nanoscience and technology publications and patents: a review of social science studies and search strategies". *The journal of technology transfer*, v. 36, n. 2, pp. 145-172.
<https://doi.org/10.1007/s10961-009-9149-8>

- Huang, Mu-Hsuan; Chang, Chia-Pin** (2013). "Detecting research fronts in OLED field using bibliographic coupling with sliding window". *Scientometrics*, v. 98, n. 3, pp. 1721-1744.
<https://doi.org/10.1007/s11192-013-1126-1>
- Huang, Ying; Schuehle, Jannik; Porter, Alan L.; Youtie, Jan** (2015). "A systematic method to create search strategies for emerging technologies based on the Web of Science: illustrated for 'Big Data'". *Scientometrics*, v. 105, n. 3, pp. 2005-2022.
<https://doi.org/10.1007/s11192-015-1638-y>
- Huo, Zhouyuan; Huang, Heng** (2017). "Asynchronous mini-batch gradient descent with variance reduction for non-convex optimization". In: *Proceedings of the AAAI Conference on artificial intelligence*, v. 31, n. 1.
<https://doi.org/10.1609/aaai.v31i1.10940>
- Kettl, Donald F.; Milward, H. Brinton** (1996). *The state of public management*. Baltimore: Johns Hopkins University Press.
- Klein, Julie-Thompson; Newell, William H.** (1996). "Advancing interdisciplinary studies". In: J. G. Gaff; J. L. Ratcliff; Associates (eds.). *Handbook of the undergraduate curriculum. A comprehensive guide to purposes, structures, practices, and change*, pp. 393-415. San Francisco, CA: Jossey-Bass.
- Kostoff, Ronald N.; Murday, James S.; Lau, Clifford G. Y.; Tolles, William M.** (2005). "The seminal literature of nanotechnology research". *Journal of nanoparticle research*, v. 8, n. 2, pp. 193-213.
<https://doi.org/10.1007/s11051-005-9034-9>
- Ledford, Heidi** (2015). "How to solve the world's biggest problems". *Nature*, v. 525, n. 7569, pp. 308-311.
<https://doi.org/10.1038/525308a>
- Lerbinger, Otto** (2006). *Corporate public affairs: Interacting with interest groups, media, and government*. Routledge. ISBN: 978 0 805856439
- Leydesdorff, Loet; Ràfols, Ismael** (2009). "A global map of science based on the ISI subject categories". *Journal of the American Society for Information Science and Technology*, v. 60, n. 2, pp. 348-362.
<https://doi.org/10.1002/asi.20967>
- Leydesdorff, Loet; Wagner, Caroline S.; Bornmann, Lutz** (2019). "Interdisciplinarity as diversity in citation patterns among journals: Rao-Stirling diversity, relative variety, and the Gini coefficient". *Journal of informetrics*, v. 13, n. 1, pp. 255-269.
<https://doi.org/10.1016/j.joi.2018.12.006>
- Liu, Kui; Kang, Guixia; Zhang, Ningbo; Hou, Beibei** (2018). "Breast cancer classification based on fully-connected layer first convolutional neural networks". *IEEE Access*, v. 6, pp. 23722-23732.
<https://doi.org/10.1109/ACCESS.2018.2817593>
- Liu, Mingrui; Mroueh, Youssef; Ross, Jerret; Zhang, Wei; Cui, Xiaodong; Das, Payel; Yang, Tianbao** (2020). "Towards better understanding of adaptive gradient algorithms in generative adversarial nets". *arXiv*, arXiv:1912.11940.
- Liu, Na; Shapira, Philip; Yue, Xiaoxu** (2021). "Tracking developments in artificial intelligence research: constructing and applying a new search strategy". *Scientometrics*, v. 126, n. 4, pp. 3153-3192.
<https://doi.org/10.1007/s11192-021-03868-4>
- Lovelock, James E.; Margulis, Lynn** (1974). "Atmospheric homeostasis by and for the biosphere: the gaia hypothesis". *Tellus*, v. 26, n. 1-2, pp. 2-10.
<https://doi.org/10.1111/j.2153-3490.1974.tb01946.x>
- Mogoutov, Andrei; Kahane, Bernard** (2007). "Data search strategy for science and technology emergence: A scalable and evolutionary query for nanotechnology tracking". *Research policy*, v. 36, n. 6, pp. 893-903.
<https://doi.org/10.1016/j.respol.2007.02.005>
- Muñoz-Écija, Teresa; Vargas-Quesada, Benjamín; Chinchilla-Rodríguez, Zaida** (2019). "Coping with methods for delineating emerging fields: Nanoscience and nanotechnology as a case study". *Journal of informetrics*, v. 13, n. 4, pp. 100976-100976.
<https://doi.org/10.1016/j.joi.2019.100976>
- Organisation for Economic Cooperation and Development* (1972). *Interdisciplinarity: Problems of teaching and research in universities*. Centre for Educational Research and Innovation. OECD Publications Center. Paris, France.
- Ostrom, Vincent** (1974). *The intellectual crisis in American public administration*. University, Ala., University of Alabama Press.
- Qin, Jun; He, Zhong-Shi** (2005). "A SVM face recognition method based on Gabor-featured key points". In: *2005 International conference on machine learning and cybernetics*, v. 8, pp. 5144-5149.

- Raadschelders, Jos C. N.** (2011). *Public administration: The interdisciplinary study of government*. Oxford/New York: Oxford University Press.
- Raiyani, Kashyap; Gonçalves, Teresa; Quaresma, Paulo; Nogueira, Vitor-Beires** (2018). "Fully connected neural network with advance preprocessor to identify aggression over Facebook and Twitter". In: *Proceedings of the first workshop on trolling, aggression and cyberbullying*, pp. 28-41.
<https://www.semanticscholar.org/paper/Fully-Connected-Neural-Network-with-Advance-to-over-Raiyani-Gon%C3%A7alves/ad0679c71836e32f3edf6f0234b9a2f486b132ed>
- Rakas, Marija; Hain, Daniel S.** (2019). "The state of innovation system research: What happens beneath the surface?". *Research policy*, v. 48, n. 9, 103787.
<https://doi.org/10.1016/j.respol.2019.04.011>
- Repko, A.; Nacakas, Francine; Fiscella, Joan B.** (2007). "Integrating Interdisciplinary: How the theories of common ground and cognitive interdisciplinarity are informing the debate on interdisciplinary integration". *Issues in interdisciplinary studies*, n. 25, 31 pp.
<https://www.semanticscholar.org/paper/Integrating-Interdisciplinary-Ity%3A-How-the-Theories-Repko-Navakas/b8c2f4e6d63c030068bbb93b65df408748a6b54d>
- Rodgers, Robert; Rodgers, Nanette** (2000). "Defining the boundaries of public administration: Undisciplined mongrels versus disciplined purists". *Public administration review*, v. 60, n. 5, pp. 435-445.
<https://doi.org/10.1111/0033-3352.00106>
- Rotolo, Daniele; Hicks, Diana; Martin, Ben R.** (2015). "What is an emerging technology?". *Research policy*, v. 44, n. 10, pp. 1827-1843.
<https://doi.org/10.1016/j.respol.2015.06.006>
- Ruiz-Navas, Santiago; Miyazaki, Kumiko** (2018). "A complement to lexical query's search-term selection for emerging technologies: the case of 'big data'". *Scientometrics*, v. 117, n. 1, pp. 141-162.
<https://doi.org/10.1007/s11192-018-2857-9>
- Savage, Peter** (1974). "Dismantling the administrative state: Paradigm reformulation in public administration". *Political studies*, v. 22, n. 2, pp. 147-157.
<https://doi.org/10.1111/j.1467-9248.1974.tb00014.x>
- Seibel, Wolfgang** (1996). "Administrative science as reform: German public administration". *Public administration review*, v. 56, n. 1, p. 74.
<https://doi.org/10.2307/3110057>
- Sen, Subir K.; Gan, Shymal K.** (1983). "A mathematical extension of the idea of bibliographic coupling and its applications". *Annals of library science and documentation*, v. 30, n. 2, pp. 78-82.
[https://nopr.niscpr.res.in/bitstream/123456789/28008/1/ALIS%2030\(2\)%2078-82.pdf](https://nopr.niscpr.res.in/bitstream/123456789/28008/1/ALIS%2030(2)%2078-82.pdf)
- Shapira, Philip; Kwon, Seokbeom; Youtie, Jan** (2017). "Tracking the emergence of synthetic biology". *Scientometrics*, v. 112, n. 3, pp. 1439-1469.
<https://doi.org/10.1007/s11192-017-2452-5>
- Small, Henry; Griffith, Belver C.** (1974). "The structure of scientific literatures I: Identifying and graphing specialties". *Science studies*, v. 4, n. 1, pp. 17-40.
<https://doi.org/10.1177/030631277400400102>
- Small, Henry; Sweeney, Ed; Greenlee, Edward** (1985). "Clustering the science citation index using co-citations. II. Mapping science". *Scientometrics*, v. 8, n. 5-6, pp. 321-340.
<https://doi.org/10.1007/bf02018057>
- Soós, Sándor; Kiss, Anna** (2020). "Informetrics and the study of science-society communications: a bibliometric scoping review". *Scientometrics*, v. 124, n. 2, pp. 825-842.
<https://doi.org/10.1007/s11192-020-03444-2>
- Srivastava, Durgesh, K.; Bhambhu, Lekha** (2010). "Data classification using support vector machine". *Journal of theoretical and applied information technology*, v. 12, n. 1.
- Stallings, Robert A.** (1986). "Doctoral programs in public administration: An outsider's perspective". *Public administration review*, v. 46, n. 3, p. 235.
<https://doi.org/10.2307/3110438>
- Steckmest, Francis W.; Business Roundtable** (1982). *Corporate performance: The key to public trust*. New York: McGraw-Hill.

- Sun, Aixin; Lim, Ee-Peng; Ng, Wee-Keong** (2002). "Web classification using support vector machine". In: *Proceedings of the fourth international workshop on Web information and data management - WIDM'02*.
<https://doi.org/10.1145/584931.584952>
- Suominen, Arho; Li, Yin; Youtie, Jan; Shapira, Philip** (2016). "A bibliometric analysis of the development of next generation active nanotechnologies". *Journal of nanoparticle research*, v. 18, n. 9.
<https://doi.org/10.1007/s11051-016-3578-8>
- Wagner, Caroline S.; Roessner, J. David; Bobb, Kamau; Klein, Julie-Thompson; Boyack, Kevin W.; Keyton, Joann; Ràfols, Ismael; Börner, Katy** (2011). "Approaches to understanding and measuring interdisciplinary scientific research (IDR): A review of the literature". *Journal of informetrics*, v. 5, n. 1, pp. 14-26.
<https://doi.org/10.1016/j.joi.2010.06.004>
- Waltman, Ludo; Van-Eck, Nees-Jan** (2012). "A new methodology for constructing a publication-level classification system of science". *Journal of the American Society for Information Science and Technology*, v. 63, n. 12, pp. 2378-2392.
<https://doi.org/10.1002/asi.22748>
- Wang, Jian; Thijs, Bart; Glänzel, Wolfgang** (2015). "Interdisciplinarity and impact: Distinct effects of variety, balance, and disparity". *PLoS one*, v. 10, n. 5, e0127298.
<https://doi.org/10.1371/journal.pone.0127298>
- Wang, Zhinan; Porter, Alan L.; Kwon, Seokbeom; Youtie, Jan; Shapira, Philip; Carley, Stephen F.; Liu, Xiaoyu** (2019). "Updating a search strategy to track emerging nanotechnologies". *Journal of nanoparticle research*, v. 21, n. 9.
<https://doi.org/10.1007/s11051-019-4627-x>
- White, Jay D.; Adams, Guy B.; Forrester, John P.** (1996). "Knowledge and theory development in public administration: The role of doctoral education and research". *Public administration review*, v. 56, n. 5, p. 441.
<https://doi.org/10.2307/977043>
- Xue Lan; Liang, Zheng; Yang, Liexun** (2017). *The development strategy of public management*. Science Press, Beijing.
- Yang, Lihua** (2019). "Public administration as a dynamic balance and integrative science across Politics, Management, and Law: Rosenbloom's framework and Chinese experiences". *The American review of public administration*, v. 49, n. 1, pp. 79-97.
<https://doi.org/10.1177/0275074018759337>
- Yu, Z.** (2019). *Research on the development and local construction of public administration discipline in China - a perspective of bibliometrics*. Unpublished doctoral dissertation, Zhejiang University, Hangzhou.
- Zhou, Xiao; Huang, Lu; Zhang, Yi; Yu, Miaomiao** (2019). "A hybrid approach to detecting technological recombination based on text mining and patent network analysis". *Scientometrics*, v. 121, n. 2, pp. 699-737.
<https://doi.org/10.1007/s11192-019-03218-5>
- Zitt, Michel; Bassecoulard, Elise** (2006). "Delineating complex scientific fields by an hybrid lexical-citation method: An application to nanosciences". *Information processing & management*, v. 42, n. 6, pp. 1513-1531.
<https://doi.org/10.1016/j.ipm.2006.03.016>

REVISTA EPI



<https://www.profesionaldelainformacion.com>

Profesional de la información es una revista científica sobre Comunicación, Documentación, Biblioteconomía, Informetría y Ciencias de la información. Se publica cada 2 meses (6 números al año).

e-ISSN: 1699-2407

<https://doi.org/10.3145/EPI>

SJR 2022 = 0,872 (Q1); JIF 2022 = 4,2 (Q1)

PROFESIONAL DE LA INFORMACIÓN PUBLICATIONS

PROFESIONAL DE LA INFORMACIÓN JOURNAL

Since 2015 it is published online only.

It is available in open access.

<https://revista.profesionaldelainformacion.com>

ANUARIO THINKEPI

It is available in open access.

<https://thinkepi.profesionaldelainformacion.com>

MANUAL SCIMAGO DE REVISTAS CIENTÍFICAS. CREACIÓN, GESTIÓN Y PUBLICACIÓN

Tomàs Baiget

a) Paperback: €29 +shipping costs

b) Free. Downlad here

<https://www.profesionaldelainformacion.com/producto/manual-scimago-de-revistas-cientificas-creacion-gestion-y-publicacion>

EPI SCHOLAR BOOK COLLECTION (Editorial UOC)

- | | | |
|--------------------------|---|--------|
| <input type="checkbox"/> | 1. Cibermetría. Midiendo el espacio red Enrique Orduña-Malea e Isidro F. Aguillo | 20,00€ |
| <input type="checkbox"/> | 2. La web social como nuevo medio de comunicación y evaluación científica Amalia Mas-Bleda e Isidro F. Aguillo | 21,00€ |
| <input type="checkbox"/> | 3. SEO. Introducción a la disciplina del posicionamiento en buscadores Jorge Serrano Cobos | 20,00€ |
| <input type="checkbox"/> | 4. Métricas de la web social para bibliotecas Nieves González-Fernández-Villavicencio | 26,00€ |
| <input type="checkbox"/> | 5. Calidad en sitios web. Métodos de análisis general, e-commerce, Imágenes, hemerotecas y turismo Rafael Pedraza-Jiménez, Lluís Codina, Javier Guallar | 19,50€ |
| <input type="checkbox"/> | 6. Gestión de datos de investigación Fabiano Couto Corrêa | 17,00€ |
| <input type="checkbox"/> | 7. Desafío a la investigación estándar en comunicación. Crítica y alternativas Manuel Goyanes | 20,00€ |
| <input type="checkbox"/> | 8. Bibliotecas integradas Pablo Parra-Valero | 15,00€ |
| <input type="checkbox"/> | 9. Gestión de información personal. Integración mediante el correo electrónico Jorge Franganillo | 19,00€ |
| <input type="checkbox"/> | 10. Datos abiertos enlazados de archivos, bibliotecas y museos en la web Carlos Henríquez Marcondes | 15,00€ |
| <input type="checkbox"/> | 11. Turismo rural online. Páginas web y redes sociales Gersón Beltrán | 18,00€ |
| <input type="checkbox"/> | 12. Documentación musical Cristina Martí-Martínez | 16,00€ |
| <input type="checkbox"/> | 13. Formatos digitales. Propiedades técnicas y contextos de uso Jorge Franganillo | 20,00€ |

<https://www.profesionaldelainformacion.com/libros-e-informes/?category=libros-epi-scholar>

Profesional de la
información

ANUARIO
Think
EPI

In the case of print publications, shipping costs have to be added

Information & orders

Isabel Olea

epi.iolea@gmail.com

Tel.: +34-608 491 521

LIBROS PROFESIONAL DE LA INFORMACIÓN

(Editorial UOC)

- | | | |
|--------------------------|---|---------|
| <input type="checkbox"/> | 56. Exposiciones virtuales en instituciones culturales. De la idea a la publicación Rubén Alcaraz-Martínez, Lluç Massaguer-Busqueta | 16,00 € |
| <input type="checkbox"/> | 55. Patentes Gema Velayos-Ortega, Rosana López Carreño | 13,00 € |
| <input type="checkbox"/> | 54. Gestión de la comunicación científica. 25 ideas para mejorar tu investigación José-Manuel Noguera-Vivo | 13,00 € |
| <input type="checkbox"/> | 53. Espacios de biblioteca: presente y futuro José-Pablo Gallo-León | 15,00 € |
| <input type="checkbox"/> | 52. Competencias mediáticas e informacionales: juego y aprendo Felicidad Campal y Ana Ordás-García | 15,00 € |
| <input type="checkbox"/> | 51. La biblioteca pública tras la pandemia del coronavirus Evelio Martínez-Cañadas | 12,00 € |
| <input type="checkbox"/> | 50. Vídeo para redes sociales. Guía de iniciación en la publicación y difusión Juan-José Boté | 14,00 € |
| <input type="checkbox"/> | 49. Biblioteconomía de guerrilla Evelio Martínez-Cañadas | 13,00 € |
| <input type="checkbox"/> | 48. Partituras electrónicas. Lectoescritura musical en el ámbito digital ... Javier Merchán-Sánchez-Jara | 14,00 € |
| <input type="checkbox"/> | 47. Makerspaces y bibliotecas Julio Alonso-Arévalo | 15,00 € |
| <input type="checkbox"/> | 46. Bibliotecas ciudadanas. Espacios de desarrollo y participación João De-Sousa-Guerreiro y María-Ramona Domínguez-Sanjurjo | 12,00 € |
| <input type="checkbox"/> | 45. Gamificación en bibliotecas. El juego como inspiración Ana Ordás | 15,00 € |
| <input type="checkbox"/> | 44. Biblioteca: el servicio público con corazón Magdalena Ortiz-Macías | 14,00 € |
| <input type="checkbox"/> | 43. La Hora del Cuento en las bibliotecas. Claves para su organización... Joan de Boer | 11,50 € |
| <input type="checkbox"/> | 42. Curación de contenidos para bibliotecas..... Evelio Martínez-Cañadas | 11,50 € |
| <input type="checkbox"/> | 41. Plan de contenidos para medios sociales Eva Sanagustín | 11,50 € |
| <input type="checkbox"/> | 40. Legislación de archivos Vicente Giménez-Chornet | 12,50 € |
| <input type="checkbox"/> | 39. Gestión de proyectos paso a paso Candela Ollé y Berta Cerezuela | 14,50 € |
| <input type="checkbox"/> | 38. Fuentes de información. Guía básica y nueva clasificación Rosana López-Carreño | 13,00 € |
| <input type="checkbox"/> | 37. La medición en plataformas sociales Adrián Segovia | 11,00 € |
| <input type="checkbox"/> | 36. Big data. Gestión y explotación de grandes volúmenes de datos Montserrat García-Alsina | 15,00 € |
| <input type="checkbox"/> | 35. Estudios de usuarios de información..... Isabel Villaseñor-Rodríguez | 12,00 € |
| <input type="checkbox"/> | 34. Autopublicar. Los nuevos circuitos para autores e investigadores.. Almudena Mangas-Vega | 12,00 € |
| <input type="checkbox"/> | 33. Lectura digital infantil. Dispositivos, aplicaciones y contenidos.... Araceli García-Rodríguez y Raquel Gómez Díaz | 13,00 € |
| <input type="checkbox"/> | 32. Geolocalización online. La importancia del dónde Gersón Beltrán | 12,00 € |
| <input type="checkbox"/> | 31. Biblioteca pública. Mientras llega el futuro Fernando Juárez-Urquijo | 14,50 € |
| <input type="checkbox"/> | 30. Geobibliotecas Estefanía Aguilar-Moreno y Carlos Granell-Canut | 11,50 € |
| <input type="checkbox"/> | 29. Los sexenios de investigación Álvaro Cabezas-Clavijo y Daniel Torres-Salinas | 12,00 € |
| <input type="checkbox"/> | 28. Los impresos antiguos Jon Zabala | 12,50 € |

<https://www.profesionaldelainformacion.com/libros-e-informes/?category=libros-epi-uoc>

<http://www.profesionaldelainformacion.com>

<http://www.thinkepi.net>

Profesional de la información (EPI) journal no longer accepts subscriptions.

Since January 1, 2023, EPI is a 100% Open Access journal, that is, all published content is available free.

If you wish to receive a summary when a new issue is published, please register here:

<https://revista.profesionaldelainformacion.com/index.php/EPI/user/register>

RULES FOR AUTHORS

Contributions must be original and unpublished, they cannot have been previously published on paper or electronically.

The text must be sent in Word or rtf. The tables must be pasted in the same document. All graphic materials (diagrams, photographs, screenshots, etc.) must be pasted into Word and also sent in separate files (in xls, jpg, pdf, etc. formats). Jpg images must have a resolution of at least 300 pp (about 200 KB each).

The text must be completely plane, without Word autoformats or automatism (subsections, bullets, linked citations, footers, indentations, tabs, colors, etc.), but must follow the EPI style in terms of bold (names of cited authors), italics (institutions, journal titles, brands) and capital letters. The urls must be without hyperlink.

Bibliographic citations in the text will be made in the form: (Surname, year) or (**SurnameAuthor1; SurnameAuthor2**, year).

The wording must be concise and precise, avoiding rhetoric.

Papers must include: a) title, b) abstract of 250-300 words, c) 14-15 keywords

In addition, the name of the authors, their place of work and address (postal and electronic), their passport type photo greater than 100 KB in jpg, and their orcid will be included.

PEER REVIEW

Papers are reviewed according to the traditional double-blind system by at least two experts on the subject, from the journal's Advisory Board and/or external experts. The journal undertakes to inform the authors of the result.

BIBLIOGRAPHIC REFERENCES

Sorted alphabetically by author, they will be limited to the works cited in the article.

No filler bibliography or "for more information" is accepted.

Journal article:

Lastname name1, First name1; Lastname2, First name2 (year). "Article title". *Journal name*, v., n., pp. xx-yy.

Url address starting in a new line, without period at the end
<https://doi.org/10.xxxx/doi>

Conference paper:

Lastname name1, First name1; Lastname2, First name2 (year). "Title of paper". In: Name of the congress, pp. xx-yy.

Url address starting in a new line, without period at the end
<https://doi.org/10.xxxx/doi>

Books and reports:

Lastname name1, First name1; Lastname2, First name2 (year). *Work title*. Place of publication: Publisher. ISBN: spaces, without hyphens and without period at the end

Url address starting in a new line, without period at the end
<https://doi.org/10.xxxx/doi>

Book chapter:

Lastname name1, First name1; Lastname2, First name2 (year). "Chapter title". In: Lastname name1, First name1; Lastname2, First name2. *Title of the book*. Place of publication: Publisher, pp. xx-yy. ISBN: spaces, without hyphens or end period

Url address starting in a new line, without period at the end
<https://doi.org/10.xxxx/doi>

Online resource:

Lastname name1, First name1; Lastname2, First name2 (year). *Title of the resource (without date of consultation)*.

Url address starting in a new line, without period at the end
<https://doi.org/10.xxxx/doi>

All contributions have to be sent to EPI's OJS platform

<https://revista.profesionaldelainformacion.com/index.php/EPI>

Previously, authors must register at:

<https://revista.profesionaldelainformacion.com/index.php/EPI/user/register>