



Editorial

Evolution of digital activism on social media: opportunities and challenges
Antonio Castillo-Esparcia; Lucía Caro-Castaño; Ana Almansa-Martínez

Research articles on Artificial intelligence

Fighting disinformation with artificial intelligence

A. Montoro-Montarroso; J. Cantón-Correa; P. Rosso; B. Chulvi; Á. Panizo-Lledot; J. Huertas-Tato; B. Calvo-Figueras; M. J. Rementería; J. Gómez-Romero

Designing personalisation of European public service media (PSM): trends on algorithms and artificial intelligence for content distribution

César Feiras-Ceide; Martín Vaz-Álvarez; Miguel Túñez-López

Semantic similarity models for automated fact-checking: *ClaimCheck* as a claim matching tool

Irene Larraz; Rubén Míguez; Francesca Sallicati

Research articles

Constructing counter-hegemony on *Twitter*: Discourse of Ibero-American political women of 'change' in the digital environment

Romina Pepe-Oliva; Andreu Casero-Ripollés

#Españalnvadida. Disinformation and hate speech towards refugees on *Twitter*: A challenge for critical thinking

Abdellah Essalhi-Rakrak; Ruth Pinedo-González

Donors wanted: narrative innovation on *TikTok* to enable mobilization

Raquel Martínez-Sanz; Amaia Arribas-Urrutia

Political conversations and regret: A qualitative evaluation on the aftermath of political discussions on social media

Manuel Goyanes; Porismita Borah; Homero Gil de Zúñiga

Using algorithms to identify social activism and climate skepticism in user-generated content on *Twitter*

N. Villagra; A. Reyes-Menéndez; J. Clemente-Mediavilla; D. J. Semova

Femitags in the networks and in the streets

Guiomar Rovira-Sancho; Jordi Morales-i-Gras

'As a rule, I don't have sex on the first date': Gender and age differences in motivations, perceptions, and experiences on *Tinder*

Carles Roca-Cuberes; Arnau Roig-Mora; Priscila Álvarez-Cueva

Spanish research on Communication in *WoS*

Francisco Segado-Boj; Valeriano Piñeiro-Naval; Tamara Antona-Jimeno

Public health communication and the Covid-19

A. Méndiz-Noguero; L. Wennberg-Capellades; E. Regadera-González; B. Goni-Fuste

Datadriven scientific research based on public statistics

J. E. Velasco-López; R. A. Carrasco; M. J. Cobo; G. Fernández-Avilés

Peer review: the attitudes and behaviours of Covid-19 pandemic-era ECR

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

The emerging phenomenon of L2 vlogging on *Bilibili*

Leticia-Tian Zhang; Boris Vázquez-Calvo; Daniel Cassany

Voice search optimization in digital media: challenges, use and training

Carlos Lopezosa; Lluís Codina; Javier Guallar; Mario Pérez-Montoro

Proposal for an index measuring the reputation of open data portals: The *Odapri*

Marta Ortiz-de-Urbina-Criado; Alberto Abella; Carmen De-Pablos-Herederó

Thirty years of research on high-growth entrepreneurship

M. P. Lechuga-Sancho; S. Martínez-Fierro; A. R. Ramos-Rodríguez

Body perception and frequency of exposure to advertising on social networks among adolescents

Beatriz Feijoo; Belén Cambroneró-Saiz; Begoña Miguel-San-Emeterio

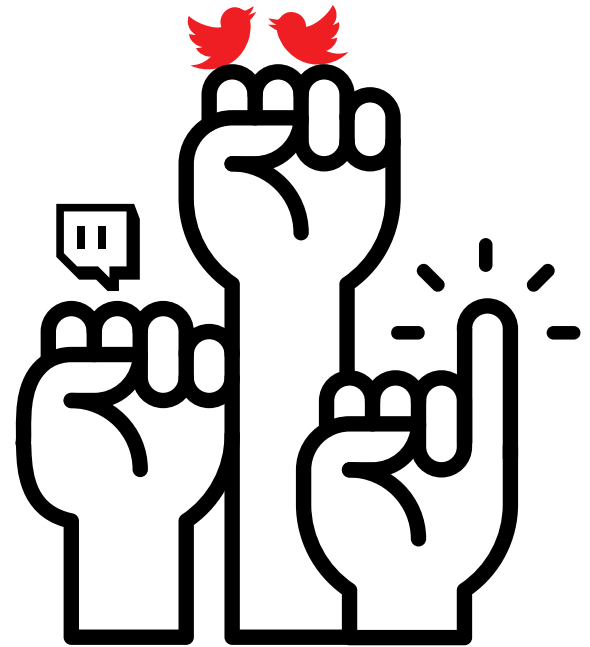
Political communication on social media in Latin America

Michael Haman

Report

The metaverse: updating the Internet (Web 3.0) or just a new development for immersive videogames?

Jean-Paul Simon



Activismos en redes

21:03 AM - PROTEST. USERS IN THE FIRST LINE



Profesional de la información es una revista de la editorial Ediciones

Profesionales de la Información S.L. NIF: B63664544

C/ Mistral, 36 (Cortijo del Aire)

18220 Albolote (Granada), España

Tel.: +34 - 608 491 521

<http://www.profesionaldelainformacion.com>

Cover design

Moisés Mañas

Universitat Politècnica de València

Foto de Moisés Mañas + DALL-E

Layout

Isabel Olea

isabel.iolea@gmail.com

Los trabajos publicados en EPI son aprobados según el sistema tradicional "peer review" en doble ciego: son revisados al menos por dos expertos en el tema, del Consejo Asesor de la revista y/o externos.

Para conseguir que los trabajos no pierdan actualidad, la dirección y los evaluadores de esta revista ponen especial esfuerzo en revisar los artículos con gran rapidez, consiguiendo un tiempo medio de aceptación o rechazo de los trabajos de sólo unas pocas semanas.

Postal address

Profesional de la información

Apartado 13

24080 León, España

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>

Depósito legal: B. 12.303-1997

EDITOR-IN-CHIEF

Tomàs Baiget

EPI SL

baiget@gmail.com

ASSOCIATE EDITOR-IN-CHIEF

Javier Guallar

Universitat de Barcelona

<http://sites.google.com/site/sitiodeljavierguallar>

EDITORIAL COORDINATION

Rodrigo Sánchez-Jiménez

Universidad Complutense de Madrid

rodsanch@uclm.es

Laura Alonso-Muñoz

Universitat Jaume I, Castellón

lalonso@uji.es

MANAGING EDITOR

Isabel Olea

EPI SL

isabel.iolea@gmail.com

GUEST EDITORS

Antonio Castillo-Esparcia

Universidad de Málaga

acastilloe@uma.es

Lucía Caro-Castaño

Universidad de Cádiz

lucia.caro@uca.es

Ana Almansa-Martínez

Universidad de Málaga

anaalmansa@uma.es

SOCIAL NETWORKS

Wileidys Artigas

High Rate Consulting, EUA

Wileartigas@gmail.com

ENGLISH LANGUAGE REVIEW

Colin Marsh

Cambridge Copy and Translation

colin.marsh@cambridgecopy.com

SCIENTIFIC COUNCIL

Ernest Abadal

Universitat de Barcelona.

Isidro F. Aguillo

Centro de CC Humanas y Sociales, CSIC, Madrid.

Rafael Aleixandre-Benavent

Universitat de València.

Laura Alonso-Muñoz

Universitat Jaume I, Castellón.

Lluís Anglada

Consorci de Serveis Universitaris de Catalunya (CSUC).

Natalia Arroyo-Vázquez

Universidad de La Laguna, Tenerife.

Wileidys Artigas

Universidad de León.

Tomàs Baiget

Universidad del Zulia, Venezuela.

Jesús Bustamante

Profesional de la Información.

Atilio Bustos-González

Cedefop, Salónica, Grecia.

Eva Campos-Domínguez

SCImago Research Group.

Andreu Casero-Ripollés

Universidad de Valladolid.

Lluís Codina

Universitat Jaume I, Castellón.

Carmen Costa-Sánchez

Universitat Pompeu Fabra, Barcelona.

Emilio Delgado-López-Cózar

Universidade da Coruña.

Javier Díaz-Noci

Universidad de Granada.

Roxana Dinu

Universitat Pompeu Fabra, Barcelona.

Ricardo Eíto-Brun

Biblioteca Națională a României.

Assumpció Estivill

Grupo GMV, Madrid.

Chiara Faggiolani

Universitat de Barcelona.

Andrés Fernández-Ramos

Università degli Studi di Roma "La Sapienza."

Antonia Ferrer-Sapena

Universidad de León.

António Fidalgo

Universitat Politècnica de València.

Francisco-Javier García-Marco

Universidade da Beira Interior, Portugal.

Homero Gil de Zúñiga

Universidad de Zaragoza.

Audilio Gonzales-Aguilar

Pennsylvania State University.

Manuel Goyanes

Université Paul-Valéry Montpellier 3. Francia.

Javier Guallar

Universidad Carlos III de Madrid.

Vicente P. Guerrero-Bote

Universitat de Barcelona.

Universidad de Extremadura.

Estefanía Herrán-Páez

SCImago Research Group.

Ying Huang

Wuhan University, China.

Johannes Keizer

Food and Agriculture Org. (FAO). Roma, Italia.

Jesús Lau

Universidad Veracruzana, Veracruz, México.

Javier Leiva-Aguilera

<http://www.javierleiva.com>, EUA.

Xosé López-García

Universidad de Santiago de Compostela, Santiago.

Carmen Marta-Lazo

Universidad de Zaragoza.

Maria-Jose Masanet-Jordà

Universitat de Barcelona.

Charles McCathieNevile

Yandex, Moscow, Rusia.

José-Antonio Moreiro-González

Universidad Carlos III de Madrid.

Isabel Olea

Profesional de la Información.

Enrique Orduña-Malea

Universitat Politècnica de València.

Juan-Antonio Pastor

Universidad de Murcia.

Fernanda Peset

Universitat Politècnica de València.

Rafael Repiso

Universidad Internacional de La Rioja.

Joan Roca

Minnesota State University, Mankato, EUA.

Blanca Rodríguez-Bravo

Universidad de León.

Luis Rodríguez-Yunta

Centro de CC Humanas y Sociales, CSIC, Madrid.

Ramón Salaverría

Universidad de Navarra, Pamplona.

Blanca San-José-Montano

Hospital Universitario de Móstoles, Madrid

Rodrigo Sánchez-Jiménez

Universidad Complutense de Madrid.

Tomás Saorín

Universidad de Murcia.

Robert Seal

Loyola Univ. Chicago, Evanston, Illinois, EUA.

Jorge Serrano-Cobos

Universitat Politècnica de València.

Ernesto Spinak

Consultor, Montevideo, Uruguay.

Emir-José Suaidén

Universidade de Brasília, Brasil.

Carlos-Miguel Tejada-Artigas

Universidad Complutense de Madrid.

Gerardo Tibaná

SCImago Research Group.

Daniel Torres-Salinas

Universidad de Granada, EC3metrics.

Jesús Tramullas

Universidad de Zaragoza.

Table of contents

EPI, 2023, v. 32, n. 3

Network activism Activismos en redes

EDITORIAL

- e320303 **Evolution of digital activism on social media: opportunities and challenges**
Antonio Castillo-Esparcia; Lucía Caro-Castaño; Ana Almansa-Martínez

RESEARCH ARTICLES ON ARTIFICIAL INTELLIGENCE

- e320322 **Fighting disinformation with artificial intelligence: fundamentals, advances and challenges**
Andrés Montoro-Montarroso; Javier Cantón-Correa; Paolo Rosso; Berta Chulvi; Ángel Panizo-Lledot; Javier Huertas-Tato; Blanca Calvo-Figueras; M. José Remente-ria; Juan Gómez-Romero
- e320311 **Designing personalisation of European public service media (PSM): trends on algorithms and artificial intelligence for content distribution**
César Feiras-Ceide; Martín Vaz-Álvarez; Miguel Túniz-López
- e320321 **Semantic similarity models for automated fact-checking: ClaimCheck as a claim matching tool**
Irene Larraz; Rubén Míguez; Francesca Sallicati

RESEARCH ARTICLES

- e320304 **Constructing counter-hegemony on Twitter: Discourse of Ibero-American political women of 'change' in the digital environment**
Romina Pepe-Oliva; Andreu Casero-Ripollés
- e320310 **#Españalnvadida. Disinformation and hate speech towards refugees on Twitter: A challenge for critical thinking**
Abdellah Essalhi-Rakrak; Ruth Pinedo-González
- e320305 **Donors wanted: narrative innovation on TikTok to enable mobilization**
Raquel Martínez-Sanz; Amaia Arribas-Urrutia
- e320302 **Political conversations and regret: A qualitative evaluation on the aftermath of political discussions on social media**
Manuel Goyanes; Porismita Borah; Homero Gil de Zúñiga
- e320315 **Using algorithms to identify social activism and climate skepticism in user-generated content on Twitter**
Nuria Villagra; Ana Reyes-Menéndez; Jorge Clemente-Mediavilla; Dimitrina J. Semova
- e320319 **Femitags in the networks and in the streets: 50 hashtags for feminist activism in Latin America**
Guiomar Rovira-Sancho; Jordi Morales-i-Gras
- e320308 **'As a rule, I don't have sex on the first date': Gender and age differences in motivations, perceptions, and experiences on Tinder**
Carles Roca-Cuberes; Arnau Roig-Mora; Priscila Álvarez-Cueva
- e320309 **Spanish research on Communication in WoS: thematic, methodological, and intellectual comparison between SSCI and ESCI**
Francisco Segado-Boj; Valeriano Piñeiro-Naval; Tamara Antona-Jimeno

- e320313 **Public health communication and the Covid-19: A review of the literature during the first wave**
Alfonso Méndiz-Noguero; Laia Wennberg-Capellades; Elisa Regadera-González; Blanca Goni-Fuste
- e320314 **Datadriven scientific research based on public statistics: a bibliometric perspective**
Jorge-Eusebio Velasco-López; Ramón-Alberto Carrasco; Manuel J. Cobo; Gema Fernández-Avilés
- e320306 **Peer review: the attitudes and behaviours of Covid-19 pandemic-era early career researchers**
David Nicholas; Eti Herman; Blanca Rodríguez-Bravo; Anthony Watkinson; Cherifa Boukacem-Zeghmouri; Marzena Świgoń; Abdullah Abrizah; David Sims; Jie Xu; David Clark; Galina Serbina; Hamid R. Jamali; Carol Tenopir; Suzie Allard
- e320301 **The emerging phenomenon of L2 vlogging on Bilibili: characteristics, engagement, and informal language learning**
Leticia-Tian Zhang; Boris Vázquez-Calvo; Daniel Cassany
- e320307 **Voice search optimization in digital media: challenges, use and training**
Carlos Lopezosa; Lluís Codina; Javier Guallar; Mario Pérez-Montoro
- e320312 **Proposal for an index measuring the reputation of open data portals: The Odapri**
Marta Ortiz-de-Urbina-Criado; Alberto Abella; Carmen De-Pablos-Heredero
- e320320 **Thirty years of research on high-growth entrepreneurship: bibliometric overview of its H-Classics**
María-Paula Lechuga-Sancho; Salustiano Martínez-Fierro; Antonio R. Ramos-Rodríguez
- e320318 **Body perception and frequency of exposure to advertising on social networks among adolescents**
Beatriz Feijoo; Belén Cambroner-Saiz; Begoña Miguel-San-Emeterio
- e320316 **Political communication on social media in Latin America: unequal use of Twitter by members of parliament**
Michael Haman

REPORT

- e320317 **The metaverse: updating the Internet (Web 3.0) or just a new development for immersive videogames?**
Jean-Paul Simon



La revista **Profesional de la información** está referenciada en:

DATABASES / BASES DE DATOS

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://scholar.cnki.net/journal/index/SJVD138667100001>

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)

<http://www.scopus.com>

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

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

FULL TEXTS / TEXTOS COMPLETOS

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=EI%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

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

Scipedia (2007-)

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

INDEXES / ÍNDICES

Dimensions

goo.gl/EMv87n

ERIH Plus

<https://dbh.nsd.uib.no/publiseringskanaler/erihplus/periodical/info.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 2021 = 3,596 Q2

Latindex (Catálogo)

<https://www.latindex.org/latindex/ficha/21379>

SCImago Journal & Country Rank (SJR)

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

Scimago Journal Rank 2022 = 0,872 Q1

SOCIAL MEDIA / REDES SOCIALES

http://twitter.com/revista_EPI

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

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

https://www.instagram.com/revista_ep

EDITORIAL MANAGER / GESTOR EDITORIAL (OJS / PKP)

Profesional de la información

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



Evolution of digital activism on social media: opportunities and challenges

Antonio Castillo-Esparcia; Lucía Caro-Castaño; Ana Almansa-Martínez

Nota: Este artículo se puede leer en español en:

<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87353>

Recommended citation:

Castillo-Esparcia, Antonio; Caro-Castaño, Lucía; Almansa-Martínez, Ana (2023). "Evolution of digital activism on social media: opportunities and challenges". *Profesional de la información*, v. 32, n. 3, e320303.

<https://doi.org/10.3145/epi.2023.may.03>

Article received on April 20th 2023



Antonio Castillo-Esparcia ✉

<https://orcid.org/0000-0002-9751-8628>

Universidad de Málaga
Facultad de Ciencias de la Comunicación,
León Tolstoi, s/n
29010 Málaga, Spain
acastilloe@uma.es



Lucía Caro-Castaño

<https://orcid.org/0000-0003-2720-1534>

Universidad de Cádiz
Facultad de Ciencias Sociales y de la
Comunicación
Avenida de la Universidad, s/n
11406 Jerez de la Frontera (Cádiz), Spain
lucia.caro@uca.es



Ana Almansa-Martínez

<https://orcid.org/0000-0003-0256-6369>

Universidad de Málaga
Facultad de Ciencias de la Comunicación,
León Tolstoi, s/n
29010 Málaga, Spain
anaalmansa@uma.es

Abstract

The social media revolution has affected all areas, including activism. However, there is scepticism about its emancipatory capacity, and it is considered by some to be a distorted form of activism. This article presents a review of the existing literature on digital activism, starting with the concept itself, and examines its impact on the organisation of activists and citizen participation. Likewise, based on the platformisation of these spaces and the evolution of their affordances, we observe a growth in individual, strategic and low-commitment participation, the accentuation of the role of emotions in media that promote virality, and the assumption of playful forms of politainment by activists –especially on *TikTok*–. On the other hand, the strategies carried out by organisations to promote activism in relation to their causes are investigated, such as the generation of arguments provided by the interest group and the involvement in an action of dissemination, co-creation and replication of activists. Finally, the main challenges of digital activism are indicated: the growing inequality in terms of access to algorithmic visibility between activists and brands, influencers and interest groups; the desirability of complementing playful activism with work in alternative digital media that allows for more stable forms of collaboration; and the need to protect activists against the growth of hate speech on these platforms.

Keywords

Digital activism; Social networking sites; Social media; Affordances; Affective publics; Influencers; Virality; Datafication; Connective action; Data activism; Interest groups; Hacktivism; Clicktivism; Slacktivism; Playful activism.

Funding

This study is a result of the project "Lobbying and communication in the European Union. Analysis of their communication strategies", funded by the *Ministry of Science and Innovation* of Spain National R+D+I Program, 2020. Project code: PID2020-118584RB-I00.



1. Introduction

Social media platforms have become key spaces for activism given the profound mediatisation of all human activities (Coudry; Hepp, 2017; Van-Dijck; Poell; Waall, 2018). This trend has been further exacerbated by the isolation measures of the COVID-19 pandemic biennium, which has led to a large part of the population shifting their information consumption habits to social media (Newman *et al.*, 2022). These spaces have become places from which to make visible support for social and political causes, to feel part of a larger movement, to organise to protest inside and outside the platforms, or as places from which to build and modify news frames and influence political and public agendas. Today, at any act of citizen protest or demonstration, a tide of mobile devices can be seen capturing the moment and broadcasting it, often live, through their *Instagram*, *Twitter*, *TikTok*, *Telegram* or *WhatsApp* group profiles.

The mobilisations of the Arab Spring, Occupy Wall Street or the 15M movement demonstrated the importance of new information technologies and networked communication to expand the forms of organisation, collective action, and social and political influence of citizens (Castells, 2009; 2015, Gerbaudo, 2012; 2017; Papacharissi; De-Oliveira, 2012). During the first era of digital activism these platforms enabled the formation of a collective identity through communicative processes (Gerbaudo, 2012; Treré, 2015; Toret *et al.*, 2013); however, despite the fact that digital activism has continued to develop online mobilisations of supranational following and impact, such as the #MeToo movements (Boyd; McEwan, 2022; Fileborn; Loney-Howes, 2019; Liao; Liqiu, 2022; Quan-Haase *et al.*, 2021), #BlackLivesMatter (Hockin-Boyers; Clifford-Astbury, 2021; Ince; Rojas; Rojas; Davis, 2017; Mundt; Ross; Burnett, 2018) or #FridaysForFuture (Díaz-Pérez; Soler-i-Martí; Ferrer-Fons, 2021; Herrmann; Rhein; Dorsch, 2022), the imaginary linked to these media points to a scepticism on the part of academics and activists regarding their emancipatory capacity today (Treré; Candón-Mena; Sola-Morales, 2021).

Among the causes of this scepticism is the process of platforming these digital spaces guided by corporate commercial interests (Butcher; Helmond, 2018; Helmond, 2015; Poell; Nieborg; Van-Dijck, 2019), which has undermined the structure of equal opportunity to be seen (Butcher, 2018; Hutchinson, 2021) and whose content moderation strategies often insufficiently protect activists from harassment and hate speech strategies (Gutiérrez, 2022; Van-Dijck; Lin, 2022; Vickery; Everbach, 2018; Weinmann; Nasri, 2020). These changes complicate the exercise of activism today, while facilitating the dissemination of content in these spaces by actors with greater economic resources such as governments, political parties, pressure groups, commercial brands and far-right movements (Chagas *et al.*, 2022; Freelon; Marwick; Kreiss, 2020; Schradie, 2019).

This article follows the following structure: First, it reviews the notion of digital activism, the empirical and theoretical studies that have reviewed its characteristics and impact in the last decade, as well as the way in which the platforming of the medium is affecting the exercise of activism and participation. Secondly, it analyses the activity of interest groups, corporations and brands as actors that advocate for social and political causes in the digital context as a way of generating consumer preferences by connecting the brand with the political and civic values of its consumers, thus giving rise, for example, to brand activism (Manfredi-Sánchez, 2019; Sarkar; Kotler, 2018; Vredenburg *et al.*, 2020). Finally, some of the challenges that have been identified in relation to digital activism on these platforms are discussed.

2. The hard definition of digital activism

Digital activism is a problematic concept on which there is no consensus among scholars and activists. Thus, Özkula describes it as

“a hazy and, as such, immanently problematic, if not dysfunctional, concept” (2021a, p. 61)

whose open debates include the question of which of the activities carried out through digital media should be considered activism (2021, p. 63), given the ambiguity and breadth of the term (Yang, 2016). In the same vein, Kaun and Uldam (2018) highlighted the growing interest of researchers from various academic disciplines in digital activism, almost a “hype” according to the authors, while noting the lack of

“it lacks a cohesive mode of inquiry that can leverage the advantages of different disciplinary insights and bring diverse perspectives and approaches into conversation” (2018, p. 2100).

Some authors have defined digital activism by focusing on the orientation of its actions towards its digital quality (Ben-nett; Segerberg, 2012; 2013; George; Leidner, 2019), which is problematic as it focuses exclusively on this condition, sometimes ignoring its contextual character as a socio-political act (Özkula, 2021a, p. 63). In this sense, several authors have insisted on the hybrid character between the digital and face-to-face planes in activist practices (Bustamante-Farias, 2019; Fischer, 2016; Gerbaudo, 2017; Greijdanus *et al.*, 2020; Karatzogianni, 2015; Nacher, 2021; Rodríguez-Suárez; Morán-Neches; Herrero-Olaizola, 2021; Theocharis *et al.*, 2023; Treré, 2019), although successful cases of digital native activism have been found, where

“the movement’s tactical repertoire is initiated, organised and coordinated online [...] with no pre-existing physical presence or offline campaign” (Li; Bernard; Luczak-Roesch, 2021, p. 1).

For Özkula (2021b) digital activism should be defined on the basis of its practices and not as a phenomenon originating from its relations with a specific technology, which would result in technological determinism, but also not as opposed to a romanticised view of traditional activism.

A remarkable definition because it emphasises its objectives and actors is the one proposed by **Karatzogianni** (2015), for whom digital activism describes:

“Political conduct aiming for reform or revolution by non-state actors and new socio-political formations such as social movements, protest organizations and individuals and groups from civil society, that is by social actors outside government and corporate influence” (**Karatzogianni**, 2015, p. 1).

The author thus coincides with **Diani's** (1992; 2000) classic definition of social movements where she clearly differentiates this type of formation from other collectives such as interest groups and political parties. From this perspective, the development of digital activist practices by actors such as commercial brands (**Manfredi-Sánchez**, 2019; **Sarkar**; **Kotler**, 2018; **Vredenburg et al.**, 2020) should not be considered digital activism. However, given that the consumption dimension is part of the areas of activism (**Teorell**; **Torcal**, 2007) and that this dimension is increasingly relevant for citizens in terms of social transformation (*CIS*, 2021, p. 4), it may not be a legitimate form of activism, but it is undoubtedly a type of initiative that brands will increasingly develop in order to connect with their audiences of interest and maintain themselves as culturally relevant (**Holt**, 2004; 2016) and socially responsible brands. Therefore, although it is understood that it is problematic to analyse these actions as forms of digital activism, they have been incorporated into this review because they have been conceptualised in this way by their authors, thus accounting for the transformation of the academic imaginary in relation to what is considered activism in the framework of social media today.

3. Impact of connective action

Some authors have analysed the emergence of digital activism in social media as a devalued mode of activism, ineffective for social change and whose main function is to make the practitioner feel good (**Gladwell**, 2010; **Miller**, 2017; **Morozov**, 2011). This critical view was expressed through neologisms such as slacktivism (**Christensen**, 2011) or lounge activism and clicktivism, terms that refer to the use of social media utilities that demand less effort (like, retweet, repost, sign petitions, etc.) and commitment to engage in civic and political campaigns (**Halupka**, 2018).

In contrast to this view, **Bennett** and **Segerberg** (2012; 2013) argued that Web 2.0 tools (blogs, forums, digital social networks, etc.) had given rise to a new form of personalised public engagement that they called the logic of connective action. Unlike traditional collective action, connective action is based on personalised content sharing through personal narratives rather than collective structures and identities. Thus, the effects of these online activities can lead to concrete policy changes through public pressure on governments (**Leong et al.**, 2021; **Uwalaka**, 2020) or mass boycott actions against companies (**Li**; **Bernard**; **Luczak-Roesch**, 2021), and extend political and civic outreach and participation by enabling new forms of mediated action (**Boulianne**, 2019; 2020; **Highfield**, 2016; **Kim**; **Hoewe**, 2020; **Lane et al.**, 2017; **Toret et al.**, 2013), especially attractive among young people (**Abbas et al.**, 2022; **Armstrong-Carter**; **Telzer**, 2021; **Boulianne**; **Theocharis**, 2020; **Cervi**; **Marín-Lladó**, 2022; **Jenkins et al.**, 2017; **Theocharis**; **De-Moor**; **Van-Deth**, 2019).

When assessing the role of clicktivism, it is interesting to incorporate the notion of subactivism to conceptualise the impact of these platforms on citizens' political and civic participation. **Bakardjieva** (2009) defined

“subactivism in my definition is a kind of politics that unfolds at the level of subjective experience and is submerging in the flow of everyday life” (2009, p. 92),

a form of empowering the citizen who finds on the internet the possibility of exploring his or her political agency from the everyday context. Subactivism thus points to

“[...] a major reservoir of civic energy” (2009, p. 103)

that finds in communication technologies the way to express itself in the public sphere. Subactivism thus breaks with the distinction between apolitical or anti-political stances and traditional activism (**Vihma**, 2016), and points towards the individualisation of social and political engagement characteristic of 21st century social movements (**Pleyers**, 2018), where for many activists it is not necessary to be part of formal structures and organisations to participate in campaigns and causes (**Sánchez-Duarte**; **Fernández-Romero**, 2017). From this perspective, connective action can be seen as a set of digital practices that have helped to bring to the surface that individual civic energy that was rarely expressed in the public sphere before the emergence of these media. Along these lines, **Nacher** (2021) analyses the use of social media platforms in feminist digital activism as a form of “weak opposition”, with weak being understood as those mundane acts, which take place in the context of everyday life, as opposed to “heroic” acts of high visibility and leadership.

As for the consequences of the expansion of connective action in the last decade in terms of participation, **Boulianne** (2015; 2019; 2020) has found in successive research a positive relationship between the use of these media and participation in civic and political life. According to the author, this relationship would have strengthened over time in connection with social media, the growth of the interactivity of websites and the emergence of new tools such as *Change.org* that would encourage the participation of social media users in offline actions such as street protests, boycotts of companies, etc. (**Boulianne**, 2020, pp. 961-962). Several studies establish a positive correlation between participation in online and offline activism activities (**Chae**; **Lee**; **Kim**, 2019; **Cardoso**; **Lapa**; **Di-Fatima**, 2016; **Gutiérrez**, 2022; **Slavina**; **Brym**, 2019). Furthermore, these platforms have broadened the basis of participation by encouraging the involvement of politically underrepresented groups such as women, ethnic minorities, youth and groups with lower levels of education (cf. **Ruess et al.**, 2021). Other studies have analysed the effects in terms of mobilisation among those already

sensitised to a cause. Thus, **Wilkins, Livingstone and Levine** (2019) found that participation in social media increased the involvement of subjects who were already aware if they perceived that their action made an effective contribution to the campaign. **Foster, Tassone and Matheson** (2021) found that women who responded to sexist attitudes on social media were enacting their social identity through these actions, which contributed to them becoming future activists.

For their part, **Pinazo-Calatayud, Nos-Aldás and Agut-Nieto** (2020) found that these media help to intensify motivation for the social cause by giving activists access to sources of information contrary to their convictions and disseminating negative messages that mobilise them against the cause and reinforce their involvement with it. These results coincide with recent work that questions the social isolation capacity of the algorithmic design of social media, pointing out that effects such as the bubble filter (**Pariser**, 2011) or the echo chamber (**Sunstein**, 2009) have been little demonstrated empirically (**Bruns**, 2019) and demonstrating that spaces allow the reception of accidental information contrary to one's own opinions more easily than other media (**Vaccari; Valeriani**, 2021). Other studies, on the contrary, have pointed out that these platforms tend to reinforce beliefs, although the personality of the subjects is relevant in relation to the development of increasingly active actions in pursuit of a cause (**Workman**, 2019).

In relation to the organisational impact of activism, **Leong et al.** (2018) argued that studies on the amplifying effects of connective action have focused too much on participation as the main outcome of grassroots involvement, thus neglecting other outcomes in terms of influence and control. In their theory of social media empowerment, the authors emphasise that social media enable the sustainability of activism over time, acting as a latent structure that retains the people needed to keep the movement alive, both in times of expansion and in times of suspension of the social movement. Similarly, **Leong et al.** (2021) empirically analysed the mechanisms underlying the shift from connective action to a more structured form of activism on social media, finding that platforms posed both opportunities and constraints for collective action. Against these positive views, several authors have analysed the constraints that the digitally mediated structure poses for activist organisation (**Dumitrica; Felt**, 2019), how they can lead to the disarticulation of collective action in non-profit organisations (**Özkula**, 2021c) and the weakening of alternative digital communication and information networks that were forged before the growth of these platforms (**Poell; Van-Dijck**, 2015; **Rendueles; Sádaba**, 2019).

In the following, we review the process of platforming these digital spaces and how the evolution of their affordances¹ design is affecting the practice of activism in recent years.

4. The evolution of activist practices in the context of social media platforms

The 2010s saw a key transformation for online citizenship communication: the shift from digital social networks (SNSs) to social media platforms (**Helmond**, 2015), something that reflects, for example, the strong presence of the label social media within the academic discourse around digital activism from 2010 onwards compared to previous terms (**Neumayer; Rossi**, 2016). While SNSs were conceived as a space structured around the user's profile links (**Boyd; Ellison**, 2008) and where visibility tended to be rather horizontal (**Hutchinson**, 2021), social media platforms are programmed to act as intermediaries with third parties, for which they

“process and reuse all the data generated therein and, through automated technical and social processes, provide guidance and behavioural guidelines to their users” (**Sued; Lugo**, 2022. pp. 1-2).

Thus, participating in online public discourse today requires adapting to certain communicative logics that the platform rewards through its algorithmic design (**Hutchinson**, 2021; **Van-Dijck; Poell**, 2014). This is potentially problematic not only because neoliberal interests and values are inscribed in the design of its algorithms (**Van-Dijck et al.**, 2018), but because its configuration is opaque to regulators and users (**Bishop**, 2019; **Cotter**, 2019; **Cardon**, 2018; **Miyazaki**, 2019) and algorithmic visibility marketing options have reduced the ability of activists to make themselves visible (**Hutchinson**, 2021; **Treré et al.**, 2021).

In response to this process of datification (**Van-Dijck**, 2017), different practices of resistance have emerged over the last decade that have been encompassed under the notion of data activism (**Gutiérrez**, 2018; **Milan; Van-der-Velden**, 2016; **Milan; Gutiérrez**, 2015), especially in response to events such as Snowden's revelations about US government spying (**Fuchs; Trottier**, 2017) or the *Cambridge Analytics* scandal (**Bennett; Lyon**, 2019), among others. This type of activism can take two approaches: reactive, focused on developing and disseminating practices of resistance to the control of data by governments and corporations; and proactive, focused on promoting uses of this data in favour of justice and social change (**Milan; Gutiérrez**, 2015). Connected to this reactive approach is the digital disconnection by some contemporary activists as a political practice, as well as data obfuscation to prevent its commercial and propagandistic use, and anonymous hacking practices (**Kaun; Treré** 2020).

4.1. Individual vs. collective identity

In the current media context, collaboration to promote political and civic initiatives and causes has become timelier and more strategic (**Highfield**, 2016; **Pecourt-Gracia**, 2015; **Pleyers**, 2018; **Zafra**, 2010) because connective action enables

“they connected individual viewpoints and in so doing organically assembled collaborative but not collective narratives” (**Papacharissi**, 2016, p. 314).

Participation in these campaigns also becomes useful material for expressing and self-validating individual identity online (Bennett; Segerberg, 2012; 2013; Lane *et al.*, 2017; Sánchez-Duarte; Fernández-Romero, 2017), which raises contradictions in the framework of collective action, as emphasising individual difference can hinder the development of stable alliances and fragment social movements (Caro-Castaño, 2015) within the quantitative logic of algorithmic visibility (Sued *et al.*, 2021). At the same time, the substitution of the collective by the aggregate of subjects is normalised in the imaginary of activism, which hides the complexity of collaborative processes (Dumitrica; Bakardjieva, 2018) and insists on a notion of the social and collective from the individual (Rendueles, 2013; Zajc, 2013).

In this sense, it is worth noting how individual participation in some digital activism campaigns has allowed users to gain visibility and engagement for their own personal brand, especially in the case of influencers (Abidin; Lee, 2022; Eriksson-Krutrök; Åkerlund, 2022; Riedl *et al.*, 2021) or minority celebrities (Abidin, 2019), online creators who use their position as part of a marginalised group to define their personal brand and seek online popularity. Thus, Eriksson-Krutrök and Åkerlund (2022) found that white influencers were highly visible when participating in the Black Lives Matter movement on *TikTok*, even appearing as leaders of the movement and generating positive exposure for their personal brand. According to the authors, while the narratives they constructed were positive for the movement, they shifted the focus away from those directly affected by racial injustice: the racialised people who had experienced police brutality. In the same vein, Boyd and McEwan (2022) observed that the virality achieved by the #MeToo hashtag by actress Alissa Milano on *Twitter* ultimately resulted in the erasure of the black women and women of colour community from public debate, despite the fact that the hashtag had been coined by activist Tarana Burke in 2006 to help black girls and women address their experiences of sexual violence.

4.2. Affective publics and playful activism

Emotions, especially positive ones (Berger; Milkman, 2012), are key to promoting mobilisation and sustaining activism insofar as it is these types of emotions that move the subject to repeat patterns of behaviour and involvement (Foster *et al.*, 2019), while they have been shown to have a great capacity to generate processes of emotional contagion with viral consequences online (Ahmed; Jaidka; Cho, 2017; Berger; Milkman, 2012; Castells, 2009; 2015; Gjerald; Eslen-Ziya, 2022). In this sense, Papacharissi (2016) proposed the convenience of analysing the affective processes that help to communicate and raise awareness in the online context and proposed the notion of “affective publics”: networked publics with a strong sense of agency that are mobilised by sentimental expressions. Affect thus contributes to establishing and maintaining feelings of community that can reflexively drive a movement in the context of social media (Papacharissi, 2016), constituting a starting point for social change (Castells, 2015). The author interprets the collaborative discourses built around hashtags in social media as “structures of feeling”, following Raymond Williams’ classic proposal, in her analysis, social experiences that reflect

“an organically developed pattern of impulses, constraints and tonality” (2016, p. 321)

that constitute meaning-making practices that can be revolutionary because they allow us to collaboratively imagine a common future (Papacharissi, 2016, p. 321).

The formation of these affective publics is possible thanks to certain affordances of the platforms that allow people who share interests and sensitivities to meet. One of the most widely used is the hashtag, which has become one of the central tools of digital activism (Bonilla; Rosa, 2015; Gjerald; Eslen-Ziya, 2022; Larrondo; Morales-i-Gras; Orbeago-Terradillos, 2019; Liao, 2019; Yang, 2016) as it allows the formation of affective audiences around an issue (Highfield, 2016; Zulli; Zulli, 2020) on a one-off basis or over time (Gjerald; Eslen-Ziya, 2022) and transnational in character, as has been the case with movements such as the anti-racist movement around #BlackLivesMatter (Hockin-Boyers; Clifford-Astbury, 2021; Ince; Rojas; Davis, 2017; Mundt; Ross; Burnett, 2018) or the feminist movement against sexual violence under the hashtag #MeToo (Boyd; McEwan, 2022; Fileborn; Loney-Howes, 2019; Liao; Liqiu, 2022; Quan-Haase *et al.*, 2021). It is worth noting that the hashtag remains the primary tool for rallying like-minded people and mobilising to seek visibility for progressive movements, however, neoconservative and far-right activist groups are developing strategies articulated around migration to alternative digital platforms without content moderation, dissemination through traditional like-minded media and manipulation and misinformation through digital media (Freelon *et al.*, 2020; Schradie, 2019), thanks to more resources and greater organisational stability (Schradie, 2019).

The platform that has become synonymous with online virality in recent years is *TikTok* (Abidin, 2021; Olivares-García; Méndez-Majuelos, 2020). Something common in the studies on digital activism reviewed on this platform is the need to learn and participate in the viral trends of the moment to remain visible through the use, not only of hashtags, but also of filters, sound and narrative memes, etc., seeking to be incorporated by the algorithm into the “For you” page and become popular (Abidin, 2021). Building on the notion of affective publics, Hautea *et al.* (2021) found in their study on climate activism on *TikTok* that the affordances of the platform facilitated a specific type of activism in which non-experts emerged as hyper-visible voices with the capacity to popularise certain ideas and consensus among younger people through the formation of an “atmosphere of connected concern” (2021, p. 12) about climate change. Likewise, Cervi and Marín-Lladó (2022) and Cervi and Divon (2023) found in their studies on the narratives of young Palestinian activists on *TikTok* that the platform’s potentialities favour a “playful activism” that helps these young people to approach audiences that had no previous interest and knowledge about the Palestinian-Israeli conflict. Abbas *et al.* (2022) also concluded

in their work on the use of *TikTok* by Palestinian activists how creative micro-videos could galvanise and shape opinions around the conflict through their ability to attract attention and awaken a sense of community through shared fun. In the same vein, **Vijay** and **Gekker** (2021) have argued that the platform functions as a space for playful political engagement in India.

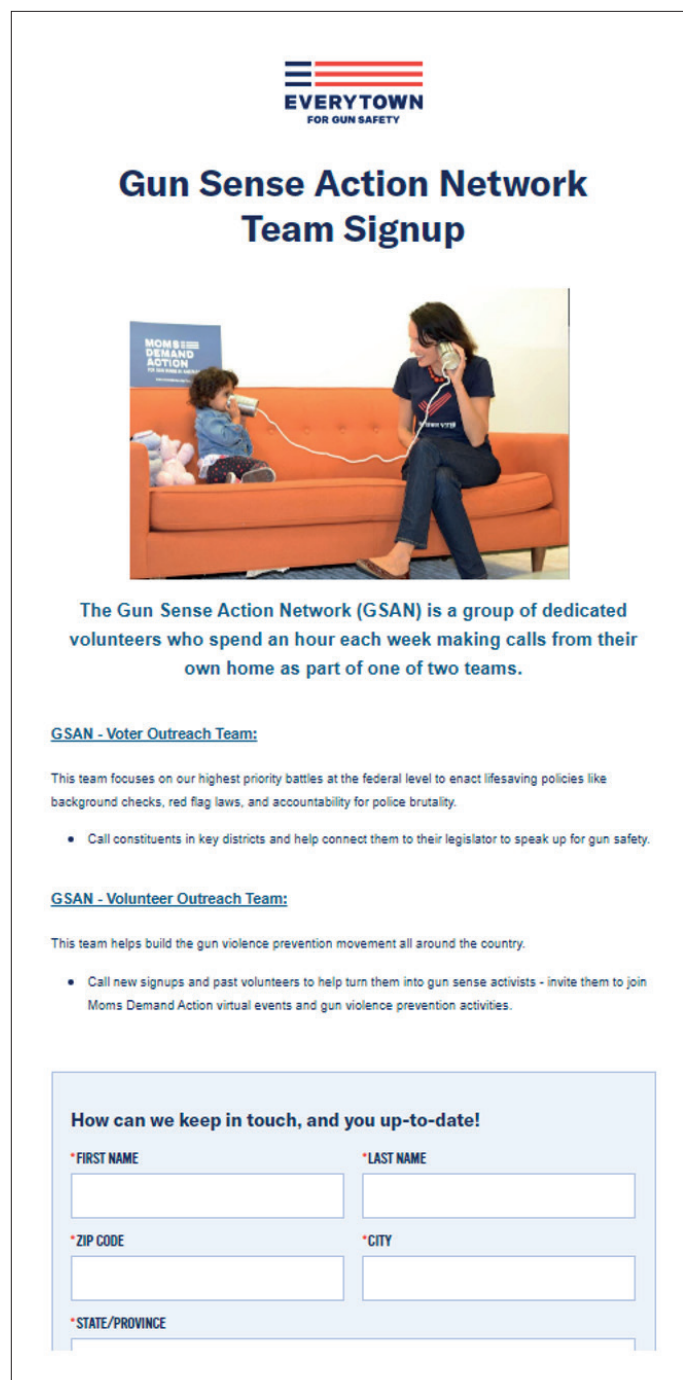
Recent studies on *TikTok* activism, on the other hand, have focused on how social media allows activism to create alternative narratives to hegemonic discourses (**Abidin; Zeng**, 2020; **Yang**, 2016; **Zhao; Abidin**, 2023) as well as access to algorithmic visibility for members of marginalised groups, such as migrants (**Jaramillo-Dent**, 2022; **Jaramillo-Dent; Contreras-Pulido; Pérez-Rodríguez**, 2022; **Sánchez-de-la-Vega**, 2021). Overall, the reviewed works point out that *TikTok* offers young people a space for political and civic engagement in an educational, entertaining, and interesting peer-to-peer format (**Abidin**, 2021, p. 84) where the confluence between influencers and the design of affordances is shaping a specific repertoire for social justice communication that **Abidin** and **Lee** (2022) have termed as the “popular cultures of social media”.

5. Advocacy and interest groups

Communication processes have undergone a significant modification since the emergence of communication 2.0, which is characterised, in general terms, by improved interaction and symmetry between audiences and organisations (**Ingenhoff; Koelling**, 2009; **Guillory; Sundar**, 2014). In this context of dialogic communication (**Kent; Taylor**, 1998) relationships are based on an increasing interactivity in which citizens take on more participatory roles and shared interactions (**McAllister-Spooner**, 2009; **Leiner; Quiring**, 2008). From this, five categories can be established according to the type of involvement and participation:

1. “Connect”, organised through a model of asymmetrical bidirectionality in which the activist maintains a passive action and which is essentially focused on the demand for information, such as subscribing to the organisation’s services (newsletter, downloading content, among others).
2. “Sharing”, organised as an activity in which the user provides information to other members of the network or the organisation, becoming a disseminator of information.
3. “Commenting”, which involves an increase in the activist’s activity and consists of comments, added to the communicative spaces generated by the organisation or as a result of a specific campaign on the network. It takes the form of comments and opinions on the campaign, whether in the news generated in the media or on social networks.
4. “Participate”, structured as an action that involves getting involved in the mobilisation actions of a campaign, such as writing to the media, writing to public representatives, signing petitions or intervening in the streets.
5. “Collaborate”, at this stage people increase their involvement and become co-creators of content with their own contributions or with documentation provided by the organisation.

The possibility of generating new supporters can be done by the organisation itself, but also by activists directly contacting other people to become activists. The latter approach can be more effective (**Almansa-Martínez; Quintana-Pujalte; Castillo-Esparcia**, 2023). One of the reasons for this is the



EVERYTOWN
FOR GUN SAFETY

Gun Sense Action Network Team Signup

MOMS DEMAND ACTION

The Gun Sense Action Network (GSAN) is a group of dedicated volunteers who spend an hour each week making calls from their own home as part of one of two teams.

GSAN - Voter Outreach Team:

This team focuses on our highest priority battles at the federal level to enact lifesaving policies like background checks, red flag laws, and accountability for police brutality.

- Call constituents in key districts and help connect them to their legislator to speak up for gun safety.

GSAN - Volunteer Outreach Team:

This team helps build the gun violence prevention movement all around the country.

- Call new signups and past volunteers to help turn them into gun sense activists - invite them to join Moms Demand Action virtual events and gun violence prevention activities.

How can we keep in touch, and you up-to-date!

*FIRST NAME

*LAST NAME

*ZIP CODE

*CITY

*STATE/PROVINCE

Image 1. The Gun Sense Action Network. Source: Everytown (2023). <https://secure.everyaction.com/RvQymTQk-E-Eyw-jFksSHg2>

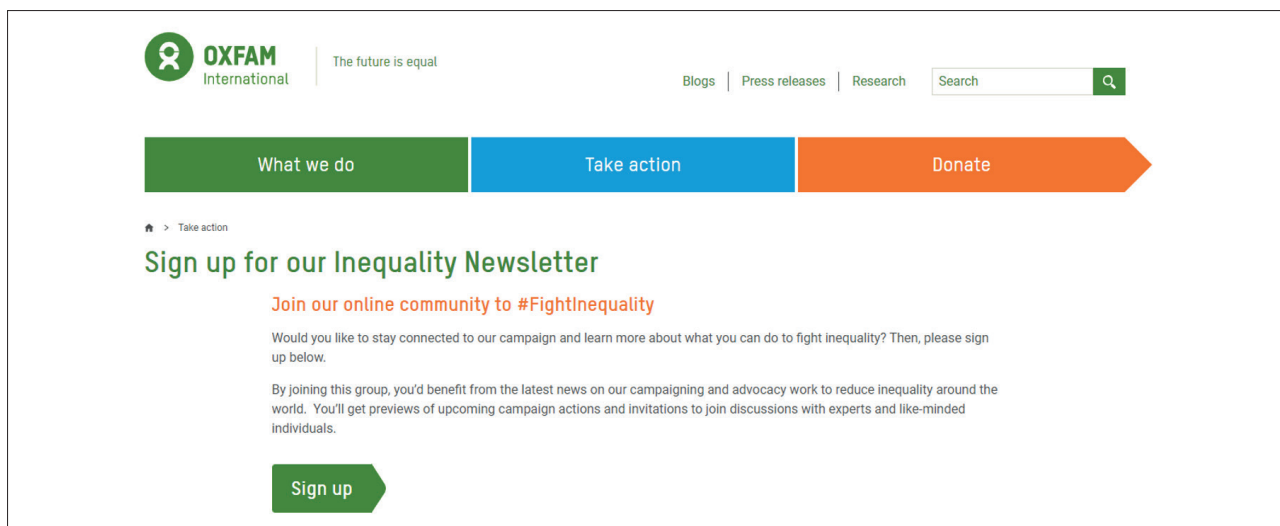


Image 2. Oxfam International newsletter subscriptions. Source: Oxfam International (2023)
<https://www.oxfam.org/en/take-action/join-global-movement-change>

greater involvement, because personal connections between activists create stronger and closer ties. This is the case of the lobby *Mom Demand Action. For Gun Sense in America*, as reflected in Image 1.

The different parameters for developing a digital activism strategy are modulated by two strategies. The first involves the generation of arguments, sources or documents provided by the organisation that has planned the campaign. The second involves an action of dissemination, co-creation and replication by the activists who are part of the campaign. In this sense, a digital strategy involves a minimum number of people who are the ones who initiate the viralisation process and who, through personal and collective networks, extend the echo of the messages. Thus, any digital activism campaign brings together a series of characteristics such as:

- Generate a strategy of what we want to transmit, based on a planning that modulates the audiences we want to involve as disseminators and recipients, the objectives of the campaign, the timing of the communication process and the availability of technological, instrumental and personnel resources.
- Participatory concept of the process in which there must be different modes that generate communicative spaces in the form of a network and that allow information to be sent to the different audiences (Image 2).
- Implementing clear, transparent and concrete messages so that activists can disseminate, participate and co-create new communication scenarios.
- Generate documentation and messages for activists. Thus, it is necessary to have a series of arguments that can be disseminated not with closed phrases, but ideas that can produce messages specific to each activist in order to show an image of private and unorganised initiative. At the same time, it is necessary to share documentation (texts, photographs, videos, posters, etc.) that can serve as support for these activists in the action of disseminating the campaign.
- Knowing the most appropriate time to carry out the activist campaign, which will allow a better visibility of the cause to be disseminated.

In this sense, in terms of the degree of activists' involvement, they can be organised into the following typologies:

- Hacktivism: involves a high involvement of activists with actions to influence and lobby for the achievement of objectives. It is based on the term hacking, which involves intrusive access to systems and organisations.
- Slacktivism, with minimal action on the part of the individual. The commitment to the cause or demand is minimal and focuses on passive feed-back, such as "liking" or sharing content.
- Clicktivism: focused on getting a large number of people to click on a petition, send online messages or letters to certain politicians or join platforms such as *Change.org*. The most important thing is to get a high number of clicks that allow them to justify metrics with many people supporting a demand or petition.

A significant part of activist campaigns is strengthened when they move from social media to traditional media (either in the off or on format), as they are still media of high social influence on the political class. One of the recent cases is the *Bud Light* beer controversy

"because it sponsored two *Instagram* posts from a transgender woman" (Wiener-Bronner, 2023).

This event has been followed by all kinds of pro and anti-transgender aftershocks in different social networks and digital spaces.

One of the aspects to consider in digital activism is that it is part of an organised communicative process that is located in the digital environment and online platforms. This interactive process can involve organisational aspects, but

also inter-subjective aspects, and everyone participates in communication strategies aimed at achieving certain ends. Therefore, the need to measure and evaluate digital campaigns is inherent to any strategy. It is necessary to start from the axiom that measuring is not evaluating, because evaluation is a complex process in which the degree of fulfilment of objectives is verified based on the results obtained in the measurement. It is therefore essential to establish a series of levels of impact and evaluation of a digital campaign in any digital activity. In this sense, *International Association for the Measurement and Evaluation of Communication (AMEC)* establishes a taxonomy of evaluation based on the following concepts:

Outputs: as the production of communication campaigns and the reception of a given audience. It focuses on measuring the impact or presence of our tools in different communication scenarios. It can be tracking coverage rates, number of news items published in the media, the media spaces in which they have been published, whether they are spaces in newspaper pages or time on radio and television, the number of views of videos, posts, tweets, number of blog entries, audience of the different media in which we appear, number of comments on the news, attendees at an event, click throughs. It is a taxonomy in which the impact on the population that our activity has had is counted.

Outtakes: refers to the degree of reception of the public and their attention to what has been communicated, whether it is attention, awareness, engagement, or participation. Examples of this type of reception are subscriptions to newsletters, recollection of knowing or receiving the information, retweets, comments, among others.

Outcomes: focused on knowing the effects of the communication on the audience. These are medium and long-term effects and address cognitive aspects on individuals, such as whether their behaviour towards the organisation has changed, whether there has been a change in attitude, whether audience preferences have changed, whether there has been an increase or decrease in support for an organisation, or whether knowledge about a specific topic has improved.

It is these three taxonomies that help us determine the type of reception, awareness, and uptake of campaigns.

6. In conclusion: key challenges

1. The commercialisation of social media visibility. The transformation of SNSs into social media platforms developed according to technocommercial interests and mechanisms during the 2010s has meant that corporations, stakeholders, governments, and neoconservative movements (with more economic resources and organisational stability) are gaining presence in these spaces while citizens and activists are increasingly forced to adapt to the affordances of the platform if they do not want to be punished with algorithmic invisibility. Given that a large part of citizens' lives is mediated by these platforms, it is critical for the proper functioning of deliberative democracy to advance transparency around the functioning of these algorithms. Likewise, it seems appropriate to study limitations by supranational governments on the commercialisation of visibility in these spaces, as this visibility tangibly affects citizens' participation in the public space. However, recent moves such as *Twitter's* decision to transform its old identity verification system (blue check) into a pay-for-visibility system for user tweets point to a worsening of this logic (Bécares, 2023).

2. The growth of playful activism and its limits for raising awareness. The importance of adapting to the affordances of the medium has led to the development of forms of activism that base their messages on entertainment formats, thus incorporating themselves into the media logic of traditional media politainment. This strategy, while useful for raising awareness and broadening the basis for social mobilisation, especially among young people, must be complemented with less superficial forms of activism on alternative platforms that offer more space and time for reflection and where users are more able to control their relationships and access to information. In this sense, it is interesting to note the movements of *Twitter* users towards *Mastodon* and other decentralised platforms following Elon Musk's purchase of the service (Chan, 2023), as a way in which users seek to regain control over their experience of the medium.

3. The need to protect activists. The governance of platforms and the problematic ways in which they exercise content moderation has allowed the proliferation of hate speech that has resulted in many activists -especially women- leaving these spaces, as well as the censorship of those who denounce abuses (Jaramillo *et al.*, 2021). Governments and supranational institutions must address this issue beyond the promises of self-regulation by these corporations, whose actions do not seem to be producing the necessary results.

7. Note

Given the diversity of approaches to the notion of affordances, this paper assumes the definition proposed by Ronzhyn, Cardenal and Batlle-Rubio (2022, p. 14):

“Social media affordances are the perceived properties, real or imagined, of social media, which arise from the relationship between the technological, the social and the contextual, and which enable and constrain specific uses of the platforms.”

8. References

Abbas, Laila; Fahmy, Shahira S.; Ayad, Sherry; Ibrahim, Mirna; Ali, Abdelmoneim H. (2022). “TikTok intifada: Analyzing social media activism among youth”. *Online media and global communication*, v. 1, n. 2, pp. 287-314. <https://doi.org/10.1515/omgc-2022-0014>

- Abidin, Crystal** (2019). "Minahs and minority celebrity: Parody YouTube influencers and minority politics in Singapore". *Celebrity studies*, v. 12, n. 4, pp. 598-617.
<https://doi.org/10.1080/19392397.2019.1698816>
- Abidin, Crystal** (2021). "Mapping internet celebrity on TikTok: Exploring attention economies and visibility labours". *Cultural science journal*, v. 12, n. 1, pp. 77-103.
<https://doi.org/10.5334/csci.140>
- Abidin, Crystal; Lee, Jin** (2022). *Social Justice through social media pop cultures: Case studies and reading resources on influencers and TikTok*. TikTok cultures research network (TCRN) & Social media pop cultures programme, Centre for Culture and Technology (CCAT), Curtin University.
<https://TikTokCultures.com/socialjustice2022>
- Abidin, Crystal; Zeng, Jing** (2020). "<? covid19?> Feeling Asian together: Coping With# COVIDRacism on subtle Asian traits". *Social media + society*, v. 6, n. 3.
<https://doi.org/10.1177/2056305120948223>
- Ahmed, Saifuddin; Jaidka, Kokil; Cho, Jaeho** (2017). "Tweeting India's Nirbhaya protest: a study of emotional dynamics in an online social movement". *Social movement studies*, v. 16, n. 4, pp. 447-465.
<https://doi.org/10.1080/14742837.2016.1192457>
- Almansa-Martínez, Ana; Quintana-Pujalte, Leticia; Castillo-Esparcia, Antonio** (2023). "Social media and citizen mobilization: The case of NGOs in European Union". In: López-López, Paulo, C. *et al. Smart innovation, systems and technologies*, v. 318. Singapore: Springer, pp. 347-356. ISBN: 978 981 19 6346 9
- Armstrong-Carter, Emma; Telzer, Eva H.** (2021). "Advancing measurement and research on youths' prosocial behavior in the digital age". *Child development perspectives*, v. 15, n. 1, pp. 31-36.
<https://doi.org/10.1111/cdep.12396>
- Bakardjieva, Maria** (2009). "Subactivism: Lifeworld and politics in the age of the internet". *The information society*, v. 25, n. 2, pp. 91-104.
<https://doi.org/10.1080/01972240802701627>
- Bécares, Bárbara** (2023). "Elon Musk apuesta por un Twitter de pago: quien no lo haga perderá mucha visibilidad en la red social". *www.GenBeta.com*, 29 marzo.
<https://www.genbeta.com/actualidad/nuevo-plan-elon-musk-que-twitter-gratis-sea-incluso-peor-no-pagas-no-podras-votar-perderas-visibilidad>
- Bennett, Colin J.; Lyon, David** (2019). "Data-driven elections: implications and challenges for democratic societies". *Internet policy review*, v. 8, n. 4.
<https://doi.org/10.14763/2019.4.1433>
- Bennett, W. Lance; Segerberg, Alexandra** (2012). "The logic of connective action. Digital media and the personalization of contentious politics". *Information, communication & society*, v. 15, n. 5, pp. 739-768.
<https://doi.org/10.1080/1369118X.2012.670661>
- Bennett, W. Lance; Segerberg, Alexandra** (2013). *The logic of connective action. Digital media and the personalization of contentious politics*. Cambridge University Press. ISBN: 978 1 107025745
- Berger, Jonah; Milkman, Katherine L.** (2012). "What makes online content viral?". *Journal of marketing research*, v. 49, n. 2, pp. 192-205.
<https://doi.org/10.1509/jmr.10.0353>
- Bishop, Sophie** (2019). "Managing visibility on YouTube through algorithmic gossip". *New media & society*, v. 21, n. 11-12, pp. 2589-2606.
<https://doi.org/10.1177/1461444819854731>
- Bonilla, Yarimar; Rosa, Jonathan** (2015). "#Ferguson: Digital protest, hashtag ethnography, and the racial politics of social media in the United States". *American ethnologist*, v. 42, n. 1, pp. 4-17.
<https://doi.org/10.1111/amet.12112>
- Boulianne, Shelley** (2015). "Social media use and participation: a meta-analysis of current research". *Information, communication & society*, v. 18, n. 5, pp. 524-538.
<https://doi.org/10.1080/1369118X.2015.1008542>
- Boulianne, Shelley** (2019). "Revolution in the making? Social media effects across the globe". *Information, communication & society*, v. 22, n. 1, pp. 39-54.
<https://doi.org/10.1080/1369118X.2017.1353641>

- Boulianne, Shelley** (2020). "Twenty years of digital media effects on civic and political participation". *Communication research*, v. 47, n. 7, pp. 947-966.
<https://doi.org/10.1177/0093650218808186>
- Boulianne, Shelley; Theocharis, Yannis** (2020). "Young people, digital media, and engagement: A meta-analysis of research". *Social science computer review*, v. 38, n. 2, pp. 111-127.
<https://doi.org/10.1177/0894439318814190>
- Boyd, Alicia; McEwan, Bree** (2022). "Viral paradox: The intersection of 'me too' and #MeToo". *New media & society*, online first.
<https://doi.org/10.1177/14614448221099187>
- Brown, Yanni; Pini, Barbara; Pavlidis, Adele** (2022). "Affective design and memetic qualities: Generating affect and political engagement through bushfire TikToks". *Journal of sociology*, online first.
<https://doi.org/10.1177/14407833221110267>
- Bruns, Axel** (2019). *Are filter bubbles real?*. John Wiley & Sons. ISBN: 978 1 509536450
- Bucher, Taina** (2018). *If... then: Algorithmic power and politics*. Oxford University Press. ISBN: 978 0 190493059
- Bustamante-Farías, Óscar** (2019). "Teoría de la activación digital. Propuesta teórico-metodológica de las prácticas comunicativas de los movimientos sociales en red". *Correspondencias & análisis*, n. 10.
<https://doi.org/10.24265/cian.2019.n10.05>
- Butcher, Taina; Helmond, Anne** (2018). "The affordances of social media platforms". In: Burgess, Jane; Marwick, Alice; Poell, Thomas (eds.). *The SAGE handbook of social media*. SAGE Publications, pp. 223-253. ISBN: 978 1 473995802
- Cardon, Dominique** (2018). *Con qué sueñan los algoritmos. Nuestras vidas en el tiempo del big data*. Dado. ISBN: 978 84 945072 8 1
- Cardoso, Gustavo; Lapa, Tiago; Di-Fatima, Branco** (2016). "People are the message? Social mobilization and social media in Brazil". *International journal of communication*, v. 10, n. 22, pp. 3909-3930.
<https://ijoc.org/index.php/ijoc/article/view/3301>
- Caro-Castaño, Lucía** (2015). "Construir y comunicar un 'nosotras' feminista desde los medios sociales. Una reflexión acerca del 'feminismo del hashtag'". *Commons. Revista de comunicación y ciudadanía digital*, v. 4, n. 2 pp. 124-154.
<http://doi.org/10.25267/commons.2015.v4.i2.06>
- Castells, Manuel** (2009). *Comunicación y poder*. Alianza Editorial. ISBN: 978 84 206 8499 4
- Castells, Manuel** (2015). *Networks of outrage and hope: Social movements in the internet age*. Polity. ISBN: 978 0 745695761
- Centro de Investigaciones Sociológicas (CIS)* (2021). *Encuesta sobre tendencias sociales (I)*. Estudio n. 3343, noviembre.
- Cervi, Laura; Divon, Tom** (2023). "Playful activism: Memetic performances of Palestinian resistance in TikTok #Challenges". *Social media + society*, n. 9, n. 1.
<https://doi.org/10.1177/20563051231157607>
- Cervi, Laura; Marín-Lladó, Carles** (2022). "Freepalestine on TikTok: from performative activism to (meaningful) playful activism". *Journal of international and intercultural communication*, v. 15, n. 4, pp. 414-434.
<https://doi.org/10.1080/17513057.2022.2131883>
- Chae, Younggil; Lee, Sookjung; Kim, Yeolib** (2019). "Meta-analysis of the relationship between Internet use and political participation: examining main and moderating effects". *Asian journal of communication*, v. 29, n. 1, pp. 35-54.
<https://doi.org/10.1080/01292986.2018.1499121>
- Chagas, Viktor; Carreiro, Rodrigo; Santos, Nina; Popolin, Guilherme** (2022). "Far-right digital activism in polarized contexts: A comparative analysis of engagement in hashtag wars". *Media and communication*, v. 10, n. 4, pp. 42-55.
<https://doi.org/10.17645/mac.v10i4.5622>
- Chan, Wilfred** (2023). "Thousands fled to Mastodon after Musk bought Twitter. Are they still 'tooting'?". *www.TheGuardian.com*, April 18.
<https://www.theguardian.com/technology/2023/apr/18/mastodon-users-twitter-elon-musk-social-media>
- Christensen, Henrik-Serup** (2011). "Political activities on the Internet: Slacktivism or political participation by other means?". *First Monday*, v. 16, n. 2.
<https://doi.org/10.5210/fm.v16i2.3336>
- 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>

- Couldry, Nick; Hepp, Andreas** (2017). *The mediated construction of reality*. John Wiley & Sons. ISBN: 978 0 7456 8131 3
- Denisova, Anastasia; O'Brien, Michaela** (2019). "From high visibility to high vulnerability: Feminist, postcolonial and anti-gentrification activism at risk". *Westminster papers in communication and culture*, v. 14, n. 1, pp. 94-98.
<https://doi.org/10.16997/wppc.323>
- Diani, Mario** (1992). "The concept of social movement". *The sociological review*, v. 40, v. 1.
<https://doi.org/10.1111/j.1467-954X.1992.tb02943.x>
- Diani, Mario** (2000). "Social movement networks virtual and real". *Information, communication & society*, v. 3, n. 3, pp. 386-401.
<https://doi.org/10.1080/13691180051033333>
- Díaz-Pérez, Silvia; Soler-i-Martí, Roger; Ferrer-Fons, Mariona** (2021). "From the global myth to local mobilization: Creation and resonance of Greta Thunberg's frame". *Comunicar*, n. 68, pp. 33-43.
<https://doi.org/10.3916/C68-2021-03>
- Dumitrica, Delia; Bakardjieva, Maria** (2018). "The personalization of engagement: the symbolic construction of social media and grassroots mobilization in Canadian newspapers". *Media, culture & society*, v. 40, n. 6, pp. 817-837.
<https://doi.org/10.1177/0163443717734406>
- Dumitrica, Delia; Felt, Mylynn** (2019). "Mediated grassroots collective action: negotiating barriers of digital activism". *Information, communication & society*, v. 23, n. 13, pp. 1821-1837.
<https://doi.org/10.1080/1369118X.2019.1618891>
- Eriksson-Krutrök, Moa; Åkerlund, Mathilda** (2022). "Through a white lens: Black victimhood, visibility, and whiteness in the Black Lives Matter movement on TikTok". *Information, communication & society*.
<https://doi.org/10.1080/1369118X.2022.2065211>
- Fileborn, Bianca; Loney-Howes, Rachel** (2019). *#MeToo and the politics of social change*. Cham: Palgrave Macmillan. ISBN: 978 3 0301 5213 0
- Fischer, Mia** (2016). "#Free_CeCe: the material convergence of social media activism". *Feminist media studies*, v. 16, n. 5, pp. 755-771.
<https://doi.org/10.1080/14680777.2016.1140668>
- Foster, Mindi D.; Hennessey, Eden; Blankenship, Benjamin T.; Stewart, Abigail** (2019). "Can 'slacktivism' work? Perceived power differences moderate the relationship between social media activism and collective action intentions through positive affect". *Cyberpsychology: Journal of psychosocial research on cyberspace*, v. 13, n. 4, art. 6.
<https://doi.org/10.5817/CP2019-4-6>
- Foster, Mindi D.; Tassone, Adrianna; Matheson, Kimberly** (2021). "Tweeting about sexism motivates further activism: A social identity perspective". *British journal of social psychology*, v. 60, n. 3, pp. 741-764.
<https://doi.org/10.1111/bjso.12431>
- Freelon, Deen; Marwick, Alice; Kreiss, Daniel** (2020). "False equivalencies: Online activism from left to right". *Science*, v. 369, n. 6508, pp. 1197-1201.
<https://doi.org/10.1126/science.abb2428>
- Fuchs, Christian; Trottier, Daniel** (2017). "Internet surveillance after Snowden: A critical empirical study of computer experts' attitudes on commercial and state surveillance of the Internet and social media post-Edward Snowden". *Journal of information, communication and ethics in society*, v. 15, n. 4, pp. 412-444.
<https://doi.org/10.1108/JICES-01-2016-0004>
- George, Jordana J.; Leidner, Dorothy E.** (2019). "From clicktivism to hacktivism: Understanding digital activism". *Information and organization*, v. 29, n. 3, 100249.
<https://doi.org/10.1016/j.infoandorg.2019.04.001>
- Gerbaudo, Paolo** (2012). *Tweets and the streets: Social media and contemporary activism*. Pluto Press. ISBN: 978 0 7453 3248 2
- Gerbaudo, Paolo** (2017). "From cyber-autonomism to cyber-populism: An ideological history of digital activism". *TripleC: Communication, capitalism & critique*, v. 15, n. 2, pp. 477-489.
<http://www.triple-c.at/index.php/tripleC/article/view/773/1012>
- Gjerald, Olga; Eslen-Ziya, Hande** (2022). "From discontent to action: #quarantinehotel as not just a hashtag". *Cogent social sciences*, v. 8, n. 1.
<https://doi.org/10.1080/23311886.2022.2051806>
- Gladwell, Malcom** (2010). "Small change: Why the revolution will not be tweeted". *The New Yorker*, September 27.
<https://www.newyorker.com/magazine/2010/10/04/small-change-malcolm-gladwell>

- Greijdanus, Hedy; De-Matos-Fernandes, Carlos A.; Turner-Zwinkels, Felicity; Honari, Ali; Roos, Carla A.; Rosenbusch, Hannes; Postmes, Tom** (2020). "The psychology of online activism and social movements: Relations between online and offline collective action". *Current opinion in psychology*, n. 35, pp. 49-54.
<https://doi.org/10.1016/j.copsyc.2020.03.003>
- Guillory, Jamie E.; Sundar, Shyam S.** (2014). "How does web site interactivity affect our perceptions of an organization?". *Journal of public relations research*, v. 26, n. 1, pp. 44-61.
<https://doi.org/10.1080/1062726X.2013.795866>
- Gutiérrez, Miren** (2018). *Data activism and social change*. Palgrave Macmillan. ISBN: 978 3 319 78318 5
- Gutiérrez, Miren** (2022). "Democratic practice in the era of platforms: From clicktivism to open-source intelligence". In: Gómez-Gutiérrez, Juan-José; Abdelnour-Nocera, José; Anchústegui-Igartua, Esteban (eds.). *Democratic institutions and practices: A debate on governments, parties, theories and movements in today's world*. Springer International Publishing, pp. 197-215.
https://doi.org/10.1007/978-3-031-10808-2_13
- Halupka, Max** (2018). "The legitimisation of clicktivism". *Australian journal of political science*, v. 53, n. 1, pp. 130-141.
<https://doi.org/10.1080/10361146.2017.1416586>
- Hautea, Samantha; Parks, Perry; Takahashi, Bruno; Zeng, Jing** (2021). "Showing they care (or don't): Affective publics and ambivalent climate activism on TikTok". *Social media + society*, v.7, n. 2.
<https://doi.org/10.1177/20563051211012344>
- Helmond, Anne** (2015). "The platformization of the web: Making web data platform ready". *Social media+ society*, v. 1, n. 2.
<https://doi.org/10.1177/2056305115603080>
- Herrmann, Christoph; Rhein, Sebastian; Dorsch, Isabelle** (2022). "#fridaysforfuture-What does Instagram tell us about a social movement?". *Journal of information science*, online first.
<https://doi.org/10.1177/0165551521106362>
- Highfield, Tim** (2016). *Social media and everyday politics*. Polity. ISBN: 978 0 7456 9134 3
- Hockin-Boyers, Hester; Clifford-Astbury, Chloe** (2021). "The politics of # diversifyyourfeed in the context of Black Lives Matter". *Feminist media studies*, v. 21, v. 3, pp. 504-509.
<https://doi.org/10.1080/14680777.2021.1925727>
- Holt, Douglas** (2004). *How brands become icons: The principles of cultural branding*. Harvard Business Press. ISBN: 1 57851 774 5
- Holt, Douglas** (2016). "Branding in the age of social media". *Harvard business review*, v. 94, n. 3, pp. 40-50.
<https://hbr.org/2016/03/branding-in-the-age-of-social-media>
- Hutchinson, Jonathon** (2021). "Micro-platformization for digital activism on social media". *Information, communication & society*, v. 24, n. 1, pp. 35-51.
<https://doi.org/10.1080/1369118X.2019.1629612>
- Ince, Jelani; Rojas, Fabio; Davis, Clayton A.** (2017). "The social media response to Black Lives Matter: How Twitter users interact with Black Lives Matter through hashtag use". *Ethnic and racial studies*, v. 40, n. 11, pp. 1814-1830.
<https://doi.org/10.1080/01419870.2017.1334931>
- Ingenhoff, Diana; Koelling, A. Martina** (2009). "The potential of web sites as a relationship building tool for charitable fundraising NPOs". *Public relations review*, v. 35, n. 1, pp. 66-73.
<https://doi.org/10.1016/j.pubrev.2008.09.023>
- Jaramillo-Dent, Daniela** (2022). "Algorithmic (in)visibility tactics among immigrant tiktokers". *Mediální studia*, v. 16, n. 2, pp. 215-235.
- Jaramillo-Dent, Daniela; Contreras-Pulido, Paloma; Pérez-Rodríguez, Amor** (2022). "Immigrant influencers on TikTok: Diverse microcelebrity profiles and algorithmic (in)visibility". *Media and communication*, v. 10, n. 1, pp. 208-221.
<https://doi.org/10.17645/mac.v10i1.4743>
- Jenkins, Henry; Billard, Thomas J.; Close, Samantha; Elsayed, Yomna; Forelle, Michelle C.; López, Rogelio; Yang, Emilia** (2017). "Participatory politics". In: Navas, Eduardo; Gallagher, Owen; Burrough, Xtine (eds.). *Keywords in remix studies*. Routledge, pp. 230-245. ISBN: 978 1 315516417
- Karatzogianni, Athina** (2015). *Firebrand waves of digital activism 1994-2014: The rise and spread of hacktivism and cyberconflict*. Palgrave Macmillan. ISBN: 978 0 230 24246 3

- Kaun, Anne; Treré, Emiliano** (2020). "Repression, resistance and lifestyle: charting (dis)connection and activism in times of accelerated capitalism". *Social movement studies*, v. 19, n. 5-6, pp. 697-715.
<https://doi.org/10.1080/14742837.2018.1555752>
- Kaun, Anne; Uldam, Julie** (2018). "Digital activism: After the hype". *New media & society*, v. 20, n. 6, pp. 2099-2106.
<https://doi.org/10.1177/1461444817731924>
- Kent, Michael L.; Taylor, Maureen** (1998). "Building dialogic relationships through the World Wide Web". *Public relations review*, v. 24, n. 3, pp. 321-334.
- Kim, Bumsoo; Hoewe, Jennifer** (2020). "Developing contemporary factors of political participation". *The social science journal*.
<https://doi.org/10.1080/03623319.2020.1782641>
- Lane, Daniel S.; Kim, Dam H.; Lee, Slgi S.; Weeks, Brian E.; Kwak, Nohin** (2017). "From online disagreement to offline action: How diverse motivations for using social media can increase political information sharing and catalyze offline political participation". *Social media + society*, v. 3, n. 3.
<https://doi.org/10.1177/2056305117716274>
- Larrondo, Ainara; Morales-i-Gras, Jordi; Orbegozo-Terradillos, Julen** (2019). "Feminist hashtag activism in Spain: Measuring the degree of politicisation of online discourse on #YoSíTeCreo, #HermanaYoSíTeCreo, #Cuéntalo y #NoEstásSola". *Communication & society*, v. 32, v. 4, pp. 207-221.
<https://doi.org/10.15581/003.32.34734>
- Leiner, Dominik J.; Quiring, Oliver** (2008). "What interactivity means to the user essential insights into and a scale for perceived interactivity". *Journal of computer-mediated communication*, v. 14, n. 1, pp. 127-155.
<https://doi.org/10.1111/j.1083-6101.2008.01434.x>
- Leong, Carmen; Faik, Isam; Tan, Felix T.; Tan, Barney; Khoo, Ying-Hooi** (2021). "Digital organizing of a global social movement: From connective to collective action". *Information and organization*, v. 30, n. 4, 100324.
<https://doi.org/10.1016/j.infoandorg.2020.100324>
- Leong, Carmen; Pan, Shan L.; Bahri, Shamshul; Fauzi, Ali** (2018). "Social media empowerment in social movements: power activation and power accrual in digital activism". *European journal of information systems*, v. 28, n. 2, pp. 173-204.
<https://doi.org/10.1080/0960085X.2018.1512944>
- Li, Yevgeniya; Bernard, Jean G.; Luczak-Roesch, Markus** (2021). "Beyond clicktivism: What makes digitally native activism effective? An exploration of the sleeping giants movement". *Social media+ society*, v. 7, n. 3.
<https://doi.org/10.1177/205630512111035357>
- Liao, Sara** (2019). "'# IAmGay# What about you?': Storytelling, discursive politics, and the affective dimension of social media activism against censorship in China". *International journal of communication*, v. 13, n. 21, pp. 2314-2333.
<https://ijoc.org/index.php/ijoc/article/view/10376/2658>
- Liao, Sara; Luqiu, Luwei R.** (2022). "#MeToo in China: The dynamic of digital activism against sexual assault and harassment in higher education". *Signs: Journal of women in culture and society*, v. 47, n. 3, pp. 741-764.
<https://www.journals.uchicago.edu/doi/abs/10.1086/717712>
- Manfredi-Sánchez, Juan-Luis** (2019). "Brand activism". *Communication & society*, v. 32, n. 4, pp. 343-359.
<https://doi.org/10.15581/003.32.37294>
- McAllister-Spooner, Sheila M.** (2009). "Fulfilling the dialogic promise: A ten-year reflective survey on dialogic Internet principles". *Public relations review*, v. 35, n. 3, pp. 320-322.
<https://doi.org/10.1016/j.pubrev.2009.03.008>
- Milan, Stefania; Gutiérrez, Miren** (2015). "Medios ciudadanos y big data: La emergencia del activismo de datos". *Mediaciones*, v. 11, n. 14, pp. 10-26.
- Milan, Stefania; Van-der-Velden, Lonneke** (2016). "The alternative epistemologies of data activism". *Digital culture & society*, v. 2, n. 2, pp. 57-74.
<https://doi.org/10.14361/dcs-2016-0205>
- Miller, Vincent** (2017). "Phatic culture and the status quo: reconsidering the purpose of social media activism". *Convergence*, v. 23, n. 3, pp. 251-269.
<https://doi.org/10.1177/1354856515592512>
- Miyazaki, Shintaro** (2019). "Take back the algorithms! A media theory of commonistic affordance". *Media theory*, v. 3, n. 1, pp. 269-286.
<https://hal.science/hal-02470429>

- Morozov, Evgeny** (2011). *The net delusion: How not to liberate the world*. Allen Lane. ISBN: 978 1 846143533
- Mundt, Marcia; Ross, Karen; Burnett, Charla M.** (2018). "Scaling social movements through social media: The case of Black Lives Matter". *Social media + society*, v. 4, n. 4.
<https://doi.org/10.1177/2056305118807911>
- Nacher, Anna** (2021). "#BlackProtest from the web to the streets and back: Feminist digital activism in Poland and narrative potential of the hashtag". *European journal of women's studies*, v. 28, n. 2, pp. 260-273.
<https://doi.org/10.1177/1350506820976900>
- Neumayer, Christina; Rossi, Luca** (2016). "15 years of protest and media technologies scholarship: A sociotechnical timeline". *Social media+ society*, v. 2, n. 3.
<https://doi.org/10.1177/2056305116662180>
- Newman, Nic; Fletcher, Richard; Robertson, Craig T.; Eddy, Kirsten; Nielsen, Rasmus-Kleis** (2022). *Reuters Institute digital news report 2022*. Reuters Institute for the Study of Journalism.
<https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2022>
- Olivares-García, Francisco-José; Méndez-Majuelos, Inés** (2020). "Análisis de las principales tendencias aparecidas en TikTok durante el periodo de cuarentena por la COVID-19". *Revista española de comunicación en salud*, Extra 1, pp. 243-252.
<https://doi.org/10.20318/recs.2020.5422>
- Özkula, Suay-Melissa** (2021a). "What is digital activism anyway? Social constructions of the 'digital' in contemporary activism". *Journal of digital social research*, v. 3, n. 3, pp. 60-84.
<https://doi.org/10.33621/jdsr.v3i3.44>
- Özkula, Suay-Melissa** (2021b). "The problem of history in digital activism: Ideological narratives in digital activism literature". *First Monday*, v. 26, n. 8.
<https://doi.org/10.5210/fm.v26i8.10597>
- Özkula, Suay-Melissa** (2021c). "The unmaking of collective action: changing organizing logics in civil society organizations through social media activism culture". *International journal of communication*, v. 15, n. 19, pp. 1984-2002.
<https://ijoc.org/index.php/ijoc/article/view/15930/3425>
- Papacharissi, Zizi** (2016). "Affective publics and structures of storytelling: sentiment, events and mediality". *Information, communication & society*, v.19, n. 3, pp. 307-324.
<https://doi.org/10.1080/1369118x.2015.1109697>
- Papacharissi, Zizi; Oliveira, Maria-de-Fatima** (2012). "Affective news and networked publics: The rhythms of news storytelling on #Egypt". *Journal of communication*, v. 62, n. 2, pp. 66-282.
<https://doi.org/10.1111/j.1460-2466.2012.01630.x>
- Pariser, Eli** (2011). *The filter bubble: How the new personalized web is changing what we read and how we think*. Penguin. ISBN: 978 1 594 20300 8
- Pecourt-Gracia, Juan** (2015). "La esfera pública digital y el activismo político", *Política y sociedad*, v. 52, n. 1, pp. 75-98.
- Pinazo-Calatayud, Daniel; Nos-Aldás, Eloísa; Agut-Nieto, Sonia** (2020). "Comunicar en positivo o negativo en el activismo social". *Comunicar*, n. 62, pp. 69-78.
<https://doi.org/10.3916/C62-2020-06>
- Pleyers, Geoffrey** (2018). *Movimientos sociales en el siglo XXI. Perspectivas y herramientas analíticas*. Clacso. ISBN: 978 987 722 373 6
- Poell, Thomas; Nieborg, David; Van-Dijck, José** (2019). "Platformisation". *Internet policy review*, v. 8, n. 4.
<https://doi.org/10.14763/2019.4.1425>
- Poell, Thomas; Van-Dijck, José** (2015). "Social media and activist communication". In: Atton, Chris (ed.). *The Routledge companion to alternative and community media*. Routledge, pp. 527-537. ISBN: 978 1 1383 9183 3
- Prada, María J.** (2022). "La pandemia dispara el activismo en Internet", *Elpais.com*, 20 septiembre.
<https://elpais.com/sociedad/2020-09-20/la-pandemia-dispara-el-activismo-en-internet.html>
- Quan-Haase, Anabel; Mendes, Kaitlynn; Ho, Dennis; Lake, Olivia; Nau, Charlotte; Pieber, Darryl** (2021). "Mapping#MeToo: A synthesis review of digital feminist research across social media platforms". *New media & society*, v. 23, n. 6, pp. 1700-1720.
<https://doi.org/10.1177/1461444820984457>
- Rendueles, César** (2013). *Sociofobia: el cambio político en la era de la utopía digital*. Capitán Swing. ISBN: 978 84 941690 0 7

- Rendueles, César; Sádaba, Igor** (2019). "Digitalización y cambio social. De las expectativas apocalípticas a la tecnopolítica del presente". *Cuadernos de relaciones laborales*, v. 37, n. 2, pp. 331-349.
<https://doi.org/10.5209/crla.66041>
- Riedl, Magdalena; Schwemmer, Carsten; Ziewiecki, Sandra; Ross, Lisa M.** (2021). "The rise of political influencers - perspectives on a trend towards meaningful content". *Frontiers in communication*, n. 247.
<https://doi.org/10.3389/fcomm.2021.752656>
- Rodríguez-Suárez, Julio; Morán-Neches, Lorena; Herrero-Olaizola, Juan B.** (2021). "Online research, new languages and symbolism of digital activism: A systematic review". *Comunicar*, n. 68, pp. 47-58.
<https://doi.org/10.3916/C68-2021-04>
- Ruess, Christina; Hoffmann, Christian-Pieter; Boulianne, Shelley; Heger, Katharina** (2021). "Online political participation: the evolution of a concept". *Information, communication & society*.
<https://doi.org/10.1080/1369118X.2021.2013919>
- Sánchez-de-la-Vega-González, Fania** (2021). "Las nuevas TICs como herramientas tecnopolíticas para el activismo digital: el caso del movimiento unDocuyouth". In: Sierra, Francisco; Leetoy, Salvador; Gravante, Tommaso (coords.). *Democracia inconclusa: Movimientos sociales, esfera, pública y redes digitales*. Universidad Autónoma de México, pp. 179-200. ISBN: 978 607 30 4517 9
- Sánchez-Duarte, José-Manuel; Fernández-Romero, Diana** (2017). "Subactivismo feminista y repertorios de acción colectiva digitales: prácticas ciberfeministas en Twitter". *Profesional de la información*, v. 26, n. 5, pp. 894-902.
<https://doi.org/10.3145/epi.2017.sep.11>
- Sarkar, Christian; Kotler, Philip** (2018). *Brand activism: From purpose to action*. Idea Bite Press. ISBN: 978 0 990576792
- Schradie, Jen** (2019). *The revolution that wasn't: How digital activism favors conservatives*. Harvard University Press. ISBN: 978 0 6742 4044 5
- Slavina, Anna; Brym, Robert** (2020). "Demonstrating in the internet age: a test of Castells' theory". *Social movement studies*, v. 19, n. 2, pp. 201-221.
<https://doi.org/10.1080/14742837.2019.1627866>
- Sued, Gabriela-Elisa; Castillo-González, María-Concepción; Pedraza, Claudia; Flores-Márquez, Dorismilda; Álamo, Sofía; Ortiz, María; Lugo, Nohemí; Arroyo, Rosa-Elba** (2021). "Vernacular visibility and algorithmic resistance in the public expression of Latin American feminism". *Media international Australia*, v. 183, n. 1, pp. 60-76.
<https://doi.org/10.1177/1329878x2111067571>
- Sued, Gabriela-Elisa; Lugo, Nohemí** (2022). "Innovación metodológica para el estudio de la plataformización". *Virtualis*, v. 13, n. 24.
<https://doi.org/10.46530/virtualis.v13i24.407>
- Sunstein, Cass R.** (2009). *Going to extremes: How like minds unite and divide*. Oxford University Press. ISBN: 978 0 19 537801 6
- Teorell, Jan; Torcal, Mariano** (2007). "Political participation: Mapping the terrain". In: Van-Deth, Jan W.; Montero, José-Ramón; Westholm, Anders (eds). *Citizenship and involvement in European democracies: A comparative analysis*. Routledge, pp. 334-357. ISBN: 978 0 4154 7960 8
- Theocharis, Yannis; Boulianne, Shelley; Koc-Michalska, Karolina; Bimber, Bruce** (2023). "Platform affordances and political participation: how social media reshape political engagement". *West European politics*, v. 46, n. 4, pp. 788-811.
<https://doi.org/10.1080/01402382.2022.2087410>
- Theocharis, Yannis; De-Moor, Joost; Van-Deth, Jan W.** (2021). "Digitally networked participation and lifestyle politics as new modes of political participation". *Policy & internet*, v. 13, n. 1, pp. 30-53.
<https://doi.org/10.1002/poi3.231>
- Toret-Medina, Javier; Calleja-López, Antonio; Marín-Miró, Óscar; Aragón, Pablo; Aguilera, Miguel; Lumbreras, Alberto** (2015). *Tecnopolítica: La potencia de las multitudes conectadas*. Editorial UOC. ISBN: 978 84 90644584
- Treré, Emiliano; Candón-Mena, José; Sola-Morales, Salomé** (2021). "Imaginario activistas sobre Internet: Del mito tecno-utópico al desencanto digital". *CIC. Cuadernos de información y comunicación*, n. 26, pp. 33-53.
<https://doi.org/10.5209/ciyc.76147>
- Uwalaka, Temple** (2020). "Clicktivism and political engagement in Nigeria". *The Nigerian journal of communication*, v. 17, n. 1.
<http://tnjc.org.ng/wp-content/uploads/Vol17N1Full/Clicktivism-and-Political-Engagement-in-Nigeria.pdf>

- Vaccari, Cristian; Valeriani, Augusto** (2021). *Outside the bubble: Social media and political participation in Western democracies*. Oxford University Press. ISBN: 978 0 1908 585 06
- Van-Dijck, José** (2017). "In data we trust? The implications of datafication for social monitoring". *Matrizes*, v. 11, n. 1, pp. 39-59.
<https://doi.org/10.11606/issn.1982-8160.v11i1p39-59>
- Van-Dijck, José; Lin, Jian** (2022). "Deplatformization, platform governance and global geopolitics: Interview with José van Dijck". *Communication and the public*, v. 7, n. 2, pp. 59-66.
<https://doi.org/10.1177/20570473221102199>
- Van-Dijck, José; Poell, Thomas** (2014). "Understanding social media logic". *Media and communication*, v. 1, n. 1, pp. 2-14.
<https://doi.org/10.17645/mac.v1i1.70>
- Van-Dijck, José; Poell, Thomas; De-Waal, Martijn** (2018). *The platform society: Public values in a connective world*. Oxford University Press. ISBN: 978 0 190889777
- Vickery, Jacqueline-Ryan; Everbach, Tracy** (2018). *Mediating misogyny: Gender, technology, and harassment*. ISBN: 978 3 319 72917 6
- Vihma, Peeter** (2016). "Internet, activism and politics. The repertoires and rhetoric of Estonian internet activists". *Studies of transition states and societies*, v. 8, n. 2, pp. 64-80.
<https://www.ceeol.com/search/article-detail?id=426601>
- Vijay, Darsana; Gekker, Alex** (2021). "Playing politics: How Sabarimala played out on TikTok". *American behavioral scientist*, v. 65, n. 5, pp. 712-734.
<https://doi.org/10.1177/0002764221989769>
- Vredenburg, Jessica; Kapitan, Sommer; Spry, Amanda; Kemper, Joya A.** (2020). "Brands taking a stand: Authentic brand activism or woke washing?". *Journal of public policy & marketing*, v. 39, n. 4, pp. 444-460.
<https://doi.org/10.1177/0743915620947359>
- Weimann, Gabriel; Masri, Natalie** (2020). "Research note: Spreading hate on TikTok". *Studies in conflict & terrorism*, v. 46, n. 5, pp. 752-765.
<https://doi.org/10.1080/1057610X.2020.1780027>
- Wiener-Bronner, Danielle** (2023). "Anheuser-Busch facilities face threats after Bud Light backlash", *www.CNN.com*.
<https://edition.cnn.com/2023/04/20/business/bud-light-threats/index.html>
- Wilkins, Denise J.; Livingstone, Andrew G.; Levine, Mark** (2019). "All click, no action? Online action, efficacy perceptions, and prior experience combine to affect future collective action". *Computers in human behavior*, n. 91, pp. 97-105.
<https://doi.org/10.1016/j.chb.2018.09.007>
- Workman, Michael D.** (2021). "Antecedents of social media activism: an analytical study of a controversial topic". *Behaviour & information technology*, v. 40, n. 3, pp. 300-310.
<https://doi.org/10.1080/0144929X.2019.1687755>
- Yang, Guobin** (2016). "Narrative agency in hashtag activism: The case of #BlackLivesMatter". *Media and communication*, v. 4, n. 4, pp. 13-17.
<https://doi.org/10.17645/mac.v4i4.692>
- Zafra, Remedios** (2010). *Un cuarto propio conectado. (Ciber)espacio y (auto)gestión del y (auto)gestión del yo*. Fórcola Ediciones. ISBN: 978 84 15174 01 1
- Zajc, Melita** (2015). "Social media, prosumption, and dispositives: New mechanisms of the construction of subjectivity". *Journal of consumer culture*, v. 15, n. 1, pp. 28-47.
<https://doi.org/10.1177/1469540513493201>
- Zhao, Xinyu; Abidin, Crystal** (2023). "The "Fox Eye" challenge trend: Anti-racism work, platform affordances, and the vernacular of gesticular activism on TikTok". *Social media + society*, v. 9, n. 1.
<https://doi.org/10.1177/20563051231157590>
- Zulli, Diana; Zulli, David-James** (2020). "Extending the Internet meme: Conceptualizing technological mimesis and imitation publics on the TikTok platform". *New media & society*. v. 24, n. 8, pp. 1872-1890.
<https://doi.org/10.1177/1461444820983603>

Fighting disinformation with artificial intelligence: fundamentals, advances and challenges

Andrés Montoro-Montarroso; Javier Cantón-Correa; Paolo Rosso; Berta Chulvi; Ángel Panizo-Lledot; Javier Huertas-Tato; Blanca Calvo-Figueras; M. José Rementeria; Juan Gómez-Romero

Note: Este artículo se puede leer en español en:
<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87328>

Recommended citation:

Montoro-Montarroso, Andrés; Cantón-Correa, Javier; Rosso, Paolo; Chulvi, Berta; Panizo-Lledot, Ángel; Huertas-Tato, Javier; Calvo-Figueras, Blanca; Rementeria, M. José; Gómez-Romero, Juan (2023). "Fighting disinformation with artificial intelligence: fundamentals, advances and challenges". *Profesional de la información*, v. 32, n. 3, e320322.
<https://doi.org/10.3145/epi.2023.may.22>

Manuscript received on 27th March 2023
Accepted on 17th May 2023



Andrés Montoro-Montarroso ✉
<https://orcid.org/0000-0003-1893-3346>
Universidad de Granada
Decsai
Citic-UGR
Periodista Rafael Gómez Montero, 2
18014 Granada, Spain
andres.montoro@ugr.es



Javier Cantón-Correa
<https://orcid.org/0000-0002-8466-1679>
Universidad Internacional de La Rioja
Fac. de Ciencias Sociales y Humanidades
Universidad de Granada
Decsai
Citic-UGR, Spain
javicanton@ugr.es



Paolo Rosso
<https://orcid.org/0000-0002-8922-1242>
Universitat Politècnica de València
Pattern Recognition and Human Language
Technologies (PRHLT) Research Center
Camí de Vera, s/n
46022 Valencia, Spain
proso@dsic.upv.es



Berta Chulvi
<https://orcid.org/0000-0003-1169-0978>
Universitat Politècnica de València
Pattern Recognition and Human Language
Technologies (PRHLT) Research Center
Camí de Vera, s/n
46022 Valencia, Spain
berta.chulvi@upv.es



Ángel Panizo-Lledot
<https://orcid.org/0000-0002-2195-3527>
Universidad Politécnica de Madrid
Escuela Técnica Superior de Ingeniería de
Sistemas Informáticos
Alan Turing, s/n
28031 Madrid, Spain
angel.panizo@upm.es



Javier Huertas-Tato
<https://orcid.org/0000-0003-4127-5505>
Universidad Politécnica de Madrid
Escuela Técnica Superior de Ingeniería de
Sistemas Informáticos
Alan Turing, s/n
28031 Madrid, Spain
javier.huertas.tato@upm.es



Blanca Calvo-Figueras
<https://orcid.org/0000-0001-6939-3576>
Barcelona Supercomputing Center (BSC)
Language Technologies Unit
Plaça Eusebi Güell, 1-3
08034 Barcelona, Spain
blanca.calvo@bsc.es



M. José Rementeria
<https://orcid.org/0000-0002-3140-1160>
Barcelona Supercomputing Center (BSC)
Social and Media Impact Evaluation
Plaça Eusebi Güell, 1-3
08034 Barcelona, Spain
maria.rementeria@bsc.es





Juan Gómez-Romero

<https://orcid.org/0000-0003-0439-3692>

Universidad de Granada

Decsai

Citic-UGR

Periodista Rafael Gómez Montero, 2

18014 Granada, Spain

jgomez@decsai.ugr.es

Abstract

Internet and social media have revolutionised the way news is distributed and consumed. However, the constant flow of massive amounts of content has made it difficult to discern between truth and falsehood, especially in online platforms plagued with malicious actors who create and spread harmful stories. Debunking disinformation is costly, which has put artificial intelligence (AI) and, more specifically, machine learning (ML) in the spotlight as a solution to this problem. This work revises recent literature on AI and ML techniques to combat disinformation, ranging from automatic classification to feature extraction, as well as their role in creating realistic synthetic content. We conclude that ML advances have been mainly focused on automatic classification and scarcely adopted outside research labs due to their dependence on limited-scope datasets. Therefore, research efforts should be redirected towards developing AI-based systems that are reliable and trustworthy in supporting humans in early disinformation detection instead of fully automated solutions.

Keywords

Journalism; Disinformation; Computing; Artificial intelligence; AI; Machine learning; Fact-checking; Datasets; Natural language processing; NLP; Social network analysis; Deepfakes; Large language models.

Funding

This work was funded by the *European Commission*, project *Iberifier (Iberian Digital Media Research and Fact-Checking Hub)*, under the call CEF-TC-2020–2 (*European Digital Media Observatory*), grant number 2020-EU-IA-0252.

1. Introduction

Amidst the prevailing post-truth era, people are overwhelmed with an enormous and uninterrupted flow of information, making it difficult to discern reliable material from content that seeks to mislead, whether intentionally (i.e., disinformation) or unintentionally (i.e., misinformation) (Wardle; Derakhshan, 2017). As a result, disinformation poses a significant and wide-ranging threat that can potentially transform any society's political, economic, and cultural fabric, thus eroding the fundamental principles of democratic nations.

While domain experts and fact-checkers may find it relatively easy to disprove hoaxes, more resources are necessary to drive and speed up their work and empower non-specialised citizens and organisations. Hence, the interest in developing technological tools for automatic information verification has grown, particularly in the ever-changing social media environment. Machine learning (ML), a subfield of Artificial Intelligence (AI), has significantly contributed to combating disinformation in recent years. Essentially, ML algorithms can be trained with data to automatically detect patterns indicative of disinformation and then apply these patterns to discern the likely truth or falsehood of unseen content. Deep Learning (DL), a subset of ML algorithms based on neural networks, has proved very useful in multiple domains (LeCun; Bengio; Hinton, 2015) and currently completely dominates the AI landscape. ML is also the predominant approach to fight disinformation (Xu; Sheng; Wang, 2023), but at the same time it can be used to generate synthetic content, increasing the impact of disinformation (Masood *et al.*, 2022).

ML is a very active, technical, and complex subject, making it difficult for non-specialists to understand and incorporate solutions arising in this field. At the same time, ML researchers must be aware of the multiple facets of a social problem like disinformation. Consequently, the research objective of this paper is to provide a brief and multidisciplinary guide to navigate the recent literature on AI to combat disinformation focusing on ML. This paper discusses the effectiveness of AI and ML techniques in detecting and counter-fighting disinformation and identifies the challenges and limitations of current approaches. We also suggest research directions for developing trustworthy AI-based systems that can assist humans in the early detection of disinformation.

Disinformation in social and digital media has prevalently spread through text. Therefore, when training ML algorithms, the primary characteristics considered are related to the syntax and content of the messages, including aspects such as syntactic, lexical, stylistic, and semantic features, which fall into the field of natural language processing (NLP). Furthermore, social network analysis (SNA) has researched the topology of disinformation networks. By

Machine learning (ML), a subfield of Artificial Intelligence (AI), has significantly contributed to combating disinformation in recent years

analysing the network structure and identifying communities, it is possible to identify groups of users who are likely to generate and disseminate harmful content, whether in a coordinated or uncoordinated way. Accordingly, we centre our work on NLP and SNA as the areas of AI more often related to disinformation analysis.

Automated disinformation analysis has been addressed from multiple perspectives. Here we propose an organisation into three overlapping approaches:

- disinformation identification by automated classification;
- feature extraction to characterise disinformation; and
- providing support to fact-checking tasks.

This organization is consistent with the approaches of the revised research works and reflects the historical development of the area:

- Disinformation classification. Automated classification is the most straightforward way of disinformation analysis –given a labelled dataset, we can train an ML classification model to distinguish legit contents. However, this methodology has the drawback that trained models on one domain are hardly extensive to others.
- Feature-based disinformation identification. Feature extraction, in turn, focuses on finding characteristics of disinformation that can be used manually or automatically to detect content and communities of interest afterwards.
- Hybrid-based fact-checking. Detecting misleading content by specialized journalists has proved very effective for disinformation analysis but also bottleneck in the process. This limitation has led to the emergence of a third type of approach known as semi-automated fact-checking.

The remainder of this manuscript is accordingly divided into three parts. The first describes AI techniques and methods used to detect disinformative content. The second focuses on the AI methods proposed in the literature to combat disinformation, including the features used to train these models and how fact-checkers can take advantage of these technological advances. The third one describes the increasing use of AI to generate disinformative content automatically. Finally, we end the paper with a summary of the main findings and the most promising research lines for future work.

2. Background

ML is a powerful tool within AI that can help to address the growing problem of disinformation by automating the detection and analysis of untrustworthy content. This section provides the reader with a background on ML and an overview of the fundamentals of Natural Language Processing and Social Network Analysis. Readers familiar with AI and ML can skip this section; otherwise, more information can be found in the classical books by **Russell and Norvig (2020)** and **Bishop (2006)**.

2.1. Machine Learning

Machine Learning is a field of AI that encompasses a range of methods, techniques, and tools for building intelligent systems by exploiting large volumes of data related to a specific problem. Specifically, ML falls under the pattern recognition paradigm, i.e., it identifies repeating characteristics in a data sample using statistical and computational processes. These patterns serve two primary functions: making predictions about future events (predictive analysis) and uncovering insights from the data (descriptive analysis). Depending on the learning mode and the process of obtaining patterns, there are three main families of ML techniques: Supervised, Unsupervised, and Reinforcement Learning. Based on artificial neural networks, Deep Learning mainly falls into Supervised Learning, but it can also be applied in Unsupervised and Reinforcement Learning setups. This subsection focuses on Supervised and Unsupervised techniques (including Deep Learning), the most representative ML techniques to combat disinformation.

Supervised Learning seeks to develop models from labelled training data that allows predicting the labels of unseen or future data. Supervised Learning can be classified into two basic categories, depending on the nature of the target variable: classification and regression. In classification, the target variable has a limited number of discrete values. Archetypical methods within this category are Decision Trees, Logistic Regression, Support Vector Machines, and the K-Nearest-Neighbour algorithm. In regression, the target variable is a real number. Some regression algorithms are Linear Regression, Polynomial Regression, Regression Splines, and Regression Trees. Supervised Learning methods are often combined to increase accuracy, yielding ensemble models such as Bagging, Boosting, and Random Forest.

Unsupervised Learning refers to techniques that deal with unlabelled or unstructured data. The most prevalent technique is clustering, utilised to identify hidden groups within a dataset for descriptive analysis. We have partitional clustering, where clusters are disjoint and typically encompass the entire item set (e.g., the *DbSCAN* and k-means algorithms), and hierarchical clustering, where groups are organised into a hierarchy. Another notable technique within Unsupervised Learning is association rules, which aim to discover dependencies between a set of items in a database.

The current dominant trend in ML is Deep Learning (**Goodfellow; Bengio; Courville, 2016**), which was first applied in Supervised Learning setups and has since been extended to other paradigms. Deep Learning enhances traditional neural networks, which are computational models that, inspired by the form of neuronal synapses, can learn intricate decision boundaries from data. Because deep neural networks possess more intermediate layers and neurons in each layer, they can capture complex relationships in large datasets. Different types of algorithms fall within Deep Learning, such as

- Convolutional Neural Networks (CNNs), which are specialised neural networks that process data with a regular structure, like images;
- Recurrent Neural Networks (RNNs), which process sequential data allowing feedback loops in the networks and work well with time series; and
- Transformers, which learn to identify relevant sections of sequences by applying attention models and are very useful with textual data.

2.2. Natural Language Processing

Natural Language Processing (NLP) involves using computational linguistics techniques to analyse text in a specific language, whether written or spoken (**Manning; Schütze**, 1999). Before developing a ML model for natural language processing (e.g., a language model), it is crucial to tackling three critical challenges: text preprocessing, feature extraction, and representation.

1) Text preprocessing involves cleaning the text and eliminating unimportant elements so that only useful information remains. The fundamental steps of text preprocessing are tokenisation (the division of the raw text into units), stopword removal (elimination of common words not significant for the analysis) and stemming (heuristic-type rules for cutting off the ends of words or affix removal) or lemmatisation (transformation of words into their base form or lemma).

2) Feature extraction involves identifying and selecting basic features from raw text data suitable to the task. Some of the most widely used techniques for feature extraction are Part-Of-Speech tagging (POS) to identify lexical categories, Named-Entity Recognition (NER) for identifying entities within the text, and bag-of-words to represent linguistic units based on their frequency of occurrence.

Another more advanced feature extraction technique is Sentiment Analysis (SA), also called Opinion Mining, which aims to automatically grasp a text's sentiments, opinions, emotions, or attitudes (**Serrano-Guerrero et al.**, 2015). It can also include eliciting the author's psychological traits through specific-purpose annotated lexicons (**John; Srivastava**, 1999; **Pennebaker et al.**, 2015).

3) Representation involves creating a numerical encoding of the text so that other ML algorithms can use it. Many techniques exist, but word embeddings are the most widely used today. They are representations of text units in the form of numerical vectors that capture their semantics. *Word2Vec* (**Mikolov et al.**, 2013) and *GloVe* (**Pennington; Socher; Manning**, 2014) are the most used techniques for obtaining embeddings. There are also available public embeddings for common terms precalculated from massive text sources, like *Wikipedia*, that can be reused in other applications. Once a document is represented as numbers, ML techniques (and particularly Deep Learning methods) can be applied to solve a downstream task (e.g., text classification or text prediction).

In this regard, Transformer networks with attention mechanisms aim to overcome the limitations of previous methods by learning to hold on to essential parts of the input text (**Vaswani et al.**, 2017). Particularly, Large Language Models (LLMs) are neural network base systems specialised in predicting the next word in a sequence that can be used for text generation and translation between sequences. A particularly noteworthy LLM is the Generative Pre-trained Transformer (GPT) (**Brown et al.**, 2020). Its current incarnation *GPT-3* and *GPT-4* can generate natural language and perform a wide range of NLP tasks, such as text generation, machine translation and question-answering (**Zhu; Luo**, 2022). More recently, a variant of *GPT-3* named *ChatGPT* has been successfully trained through human interaction to engage in realistic conversations (**Megahed et al.**, 2023).

2.3. Social Network Analysis (SNA)

Social Network Analysis is the computational field that explores social entities' relationships, patterns, and structures to understand the system, position, and linkage between these actors (**Barabási**, 2016). SNA uses mathematical and computational methods to analyse data from social media through two different approaches (**Aggarwal**, 2011; **Camacho et al.**, 2020):

- structural analysis (topology of the network, communities, and important nodes); and
- content-based analysis (information about social media users, shared content).

Structural analysis focuses on studying the topology of a network by applying graph theory. Often-used structural metrics include local measures like centrality, degree, closeness or betweenness –used for identifying the importance of certain nodes (users) within the network, and global measures such as density, diameter, radius, or transitivity –used to study the global structure of the network. An essential problem in SNA is community detection, which aims to identify sets of more tightly connected nodes (**Bedi; Sharma**, 2016). The task of community detection is closely related to the clustering problem, so most techniques belong to this broad family of algorithms (**Fortunato**, 2010). Other approaches are based on the maximisation of modularity, a measure that balances the number of internal and external connections of a community. Some algorithms based on modularity are Newman's greedy method (**Newman**, 2004) and the Blondel method (**Blondel et al.**, 2008).

Content-based analysis examines both the content and the connections between nodes, for example, by incorporating text to provide additional context to the network (**Cambria; Wang; White**, 2014). Content analysis is commonly applied in the following ways:

- user profiling, which gathers extra information about the human actors –e.g. behaviour or physical features– in a network (**Harrigan et al.**, 2021);
- topic extraction, which identifies the main themes of discussion among a group of nodes (**Yin et al.**, 2012), or the interests of users through their social connections (**Wang et al.**, 2013);
- sentiment analysis, which examines the tone of the messages exchanges among the nodes (**Camacho et al.**, 2020).

3. Disinformation classification with Machine Learning

Supervised Learning is the most widely employed approach for the automatic identification of disinformation. Thereby, the identification of disinformation is usually modelled as a binary classification problem. Given a set of representative features of an information item I , the task is to predict whether I is truthful or not, i.e.:

$$f(I) = \begin{cases} 1, & \text{if } I \text{ is a disinformation item} \\ 0, & \text{if } I \text{ is not disinformative} \end{cases}$$

where f is the function we want to learn from the available data. The combination of the features to obtain f can be done manually or automatically. In the first case, Multi-Criteria Decision Making (MCDM) has been applied to define criteria and probability weights to calculate an information credibility score and rank the candidate solutions (**Pasi; De-Grandis; Viviani**, 2020). In the second case, DL has been applied to learn the features and the combination weights (**Amador; Molina-Solana; Gómez-Romero**, 2019; **Molina-Solana; Amador; Gómez-Romero**, 2018).

Nevertheless, disinformation flows in shades of grey, not black and white, rendering a binary classification insufficient. In the literature, we can find more precise definitions of labels to capture the more subtle nuances of disinformation. For example, **Wang** (2017) proposed

a manually labelled dataset with six fine-grained labels where the degree of truthfulness (pants-fire, false, barely true, half-true, mostly true, and true) of thousands of statements was evaluated. **Nakamura, Levy and Wang** (2020) used a labelling hierarchy of two, three, and six categories for each sample of their multimodal dataset enabling the implementation of classification models at different levels of granularity.

“ The performance of Machine Learning depends directly on the quality of the data ”

The performance of Supervised Learning depends directly on the quality of the labelled data, which usually represents situations, making it difficult to extend the models to other similar domains. This limitation is even more noticeable when applied to the automatic detection of disinformation since it is challenging to build datasets with enough quality to cover the nuances of disinformation in heterogeneous contexts (**Shu et al.**, 2017). Dataset construction involves

- data extraction, either through APIs provided by platform owners or web scraping methods, and
- annotation, which is a manual time-consuming and error-prone task with little automatic support (**Simko et al.**, 2021).

Annex includes datasets used in the literature for testing ML disinformation classification models.

As reported several times (**Guo et al.**, 2020; **Meel; Vishwakarma**, 2020; **Zhang; Ghorbani**, 2020), studies that work directly on automatic disinformation detection with Unsupervised Learning are scarce. Some works formulate automatic identification of disinformation as an anomaly detection problem on social networks, employing an autoencoder as an Unsupervised Learning method (**Li et al.**, 2021), another uses Bayesian statistics to compute the veracity of news and the credibility of their authors (**Yang et al.**, 2019). Nevertheless, most studies use Unsupervised Learning in a complementary way to Supervised Learning; that is, they use a Semi-supervised approach (**De-Souza et al.**, 2022; **Dong; Victor; Qian**, 2020; **Li; Lu et al.**, 2022; **Meel; Vishwakarma**, 2021; **Paka et al.**, 2021).

4. Feature-based automated disinformation detection

As explained, methods for disinformation detection need relevant features representative of the news items. Classically, they have been classified into content-based and context-based features (**Bondielli; Marcelloni**, 2019).

- Content-based features are relevant attributes extracted directly from the data item, usually a text stating or supporting the potential hoax and often associated with several images or videos that reinforce it.
- Context-based features refer to data or metadata surrounding the piece of information. This section focuses on various features that can be extracted and used to detect false information.

4.1. Natural language processing for stylistic characterisation of messages

Content-based methods use the linguistic features of false information, including syntactic and semantic characteristics (**Zhou et al.**, 2020) that can be obtained by applying NLP techniques (**Ruffo et al.**, 2023). Among syntactic features, we can find POS tags and relevant groups of words (bigrams, trigrams, or n-grams). Semantic features can be obtained through sentiment analysis, opinion mining, topic detection, or encodings with word embeddings.

A specific kind of linguistic feature is style-based features. The rationale behind methods based on them is that ML can capture the distinctive style that malicious actors use to increase the diffusion and acceptance of their content (**Zhou; Zafarani**, 2020). The style of news text has been formalised and measured in terms of the frequency of morphological

patterns (Castelo *et al.*, 2019; Vogel; Meghana, 2020), the presence of structural elements (Bonet-Jover *et al.*, 2021), the lexical variety and the use of punctuation symbols (Azevedo *et al.*, 2021), the complexity and level of readability of the text (Castelo *et al.*, 2019) and the emotional tone (Giachanou; Rosso; Crestani, 2019).

“ The features employed for disinformation classification can be categorized into two groups: content-based features and context-based features ”

Regarding morphological patterns, in an early study, Afroz, Brennan and Greenstadt(2012) were able to identify false information by analysing the number of syllables and words, vocabulary, grammatical complexity, and POS tags. Misleading content spreaders were also found to use more informal language (Giachanou *et al.*, 2022), e.g., certain patterns in the use of personal pronouns and swear words (Rashkin *et al.*, 2017). Regarding the emotional tone of the discourse, Del-Vicario *et al.* (2016) showed that the emotional state of social media users is linked to their level of engagement in the community –more activity leads to more negative emotions and vice versa. Accordingly, the use of polarised language patterns is often seen as a sign of message engineering to increase impact by provoking negative emotions in the receiver, such as anger, disgust, or fear (Giachanou; Rosso; Crestani, 2021), and therefore, an indicator of low credibility (Ghanem *et al.*, 2021; Stella; Ferrara; De-Domenico, 2018).

Conversely, disinformers can learn style-based features to replicate the writing styles of trustworthy information sources and disguise their actions. This is particularly problematic if language models are used to generate disinformation, which is currently a trend and a challenge. For example, Schuster *et al.* (2020) showed that NLP models for disinformation identification based on stylistic features work well with human writing. Still, they tend to fail when confronted with synthetic text created by language models trained to replicate trusted media.

4.2. Contextual aspects of disinformation in social networks

Contextual features are extracted by considering the relevant data related to an information item, including metadata or other external elements. This information is primarily available in social networks, where context can be connected to the users, their posted messages, or the network (Guo *et al.*, 2020).

4.2.1. Features based on the context of the users

User-based features include the number of posts, number of followers, demographics, whether the account is verified, or the age of the account on the platform. A usual metric built from such profile data is user credibility, which can indicate the likelihood of sharing false information (Shu; Wang; Liu, 2019). Credibility can be obtained from network metadata to analyse whether there is a correlation between a user profile and the publication of false information (Shu *et al.*, 2019). Furthermore, user engagement (likes, retweets, and replies) with tweets written by verified users can also be used to assess credibility (Yang *et al.*, 2019).

A very interesting type of social network user is bots. Bots are computer programs that carry out autonomous actions, including automatically generating false information and amplifying disinformation during the initial dissemination stages (Shao; Ciampaglia *et al.*, 2018). Bots tend to have particular profiles on social networks, e.g., they are usually recent accounts (Davis *et al.*, 2016) with lengthy usernames using weird characters (Oehmichen *et al.*, 2019). Their behaviour is also different from humans' (Ruffo *et al.*, 2023); e.g., they retweet more, get fewer retweets, receive fewer replies and mentions, and publish fewer original tweets (Ferrara *et al.*, 2016). All these features can be obtained from the public profiles and the graph of retweets for automatic bot identification alone (Des-Mesnards *et al.*, 2022) or combined with message data (Kudugunta; Ferrara, 2018).

Disinformation is closely related to the user's personality and mental processes. Given that psychological characteristics regulate behaviour and interaction in the physical world, it is logical to assume that they also impact virtual communities. Psychological traits can influence how individuals interpret and engage with information, increasing the likelihood of spreading false information and toxic narratives. For example, inherently human cognitive biases such as limited reality perception and confirmation bias can increase the likelihood of perceiving fake news as real and thus encourage its dissemination (Shu *et al.*, 2017). Unlike disseminators of accurate information, disinformers have been found to be extroverted, less neurotic and present more stress in their tweets (Shrestha; Spezzano, 2022). In contrast, Srinivas, Das and Pulabaigari (2022) suggest that users who spread false political information are neurotic, conservative and have psychopathic traits. The difference in the conclusions of these works is mainly due to the way of detecting and measuring these psychological traits.

4.2.2. Features based on the context of the messages

Contextual user and message-based features are often not clearly distinguished (Guo *et al.*, 2020) and even merged (Yang *et al.*, 2019). Still, for clarity, we consider the context of the posted messages separately, which are different, more dynamic, and specific than the users' features (Tacchini *et al.*, 2017). Thus, metadata about posts in social networks has been mainly used to increase the effectiveness of another principal feature (Della-Vedova *et al.*, 2018). Likewise, multimedia resources associated with messages have been used to complement ML models, yielding multimodal disinformation analysis (Hangloo; Arora, 2022).

Multimodal analysis has been focused to date on images and addressed in three main forms: forensic –evaluates whether an image has been subjected to modification or manipulation (Qi *et al.*, 2019)–, contextual –the image and the text are consistent (Kang; Hwang; Yu, 2020; Xiong *et al.*, 2023)–, and hybrid –the image is processed to extract additional information to be used in (Giachanou; Zhang; Rosso, 2020; Jing *et al.*, 2023; Khattar *et al.*, 2019; Li; Yao *et al.*, 2022; Singh *et al.*, 2023; Wang *et al.*, 2018). For example, Zhang, Giachanou y Rosso (2022) combined textual, visual, and contextual information to build the “scene” depicted in the post, obtaining statistically significant differences in the appearance of specific places, weather, and seasons in false and truthful content.

4.2.3. Features based on the network structure

Network-based features refer to the static structure of the social network, such as central nodes and communities based on users’ connections, and the more dynamic propagation of (dis)information, including critical actors, dissemination paths, and infiltration from one community to another (Bondielli; Marcelloni, 2019; Zhou; Zafarani, 2020).

Most works in the literature focus on detecting false information by modelling the information dissemination network, assuming that true and false information have different propagation patterns (De-Souza *et al.*, 2022; Liu; Wu, 2018; Liu; Xu, 2016; Song *et al.*, 2022). Other works have combined the analysis of propagation paths with spreaders’ characteristics for disinformation classification (Grinberg *et al.*, 2019; Shao; Ciampaglia *et al.*, 2018; Shao; Hui; *et al.*, 2018). This approach is highly effective for stopping the propagation of disinformation, as it prioritises identifying (and removing) disinformative over the more costly analysis of individual publications. Specifically, the networking characteristics of users involved in disseminating false information have been investigated through initiatives such as the PAN challenges (Buda; Bolonyai, 2020; Vogel; Meghana, 2020). In addition, modern ML techniques have been recently applied to this topic, e.g., Rath, Salecha y Srivastava (2022) proposed a graph neural network model to identify nodes prone to disseminate false information using network topology and historical user activity data.

5. AI-supported fact-checking

Fact-checking is journalism focused on checking public assertions (Graves; Nyhan; Reifler, 2016). While verifying information is a foundational part of journalism, fact-checking emphasises the relevance of the checking process and the development of methods and tools to do so effectively and transparently. The first proposals to automate online fact-checking appeared more than 15 years ago (Graves, 2018), already highlighting that full automation is practically impossible because of the critical judgment, sensitivity, and experience required to make a decision that is not binary (Arnold, 2020). The fact-checking community acknowledges that the rapid dissemination of false information presents scalability issues –i.e., spreading a lie is way faster than debunking it (Vosoughi; Roy; Aral, 2018)– but this should not undermine the rigour of the fact-checking process.

Accordingly, the approaches in the literature tend to AI-supported fact-checking rather than automated fact-checking, which is often labelled as human-in-the-loop systems (La-Barbera; Roitero; Mizzaro, 2022; Shabani *et al.*, 2021; Yang *et al.*, 2021). AI can support fact-checking at different stages of the verification workflow (Guo; Schlichtkrull; Vlachos 2022; Nakov; Corney *et al.*, 2021):

- (1) monitoring, recognition, and prioritisation of content susceptible to verification;
- (2) evaluating whether claims are verifiable or not and topic prioritisation;
- (3) searching for previous verifications that apply to the same case;
- (4) retrieval of evidence for further investigation;
- (5) semi-automated classification in categories (hoax, misleading content, false context, etc.);
- (6) dissemination of the verifications; and
- (7) speeding-up writing and documenting the fact-checks.

Proposals in the literature have primarily focused on stages 1-4. For stage 5, the contributions described in Section 3 could be applied, although they show limitations in their applicability to multiple domains, as already described.

Various methods have been proposed for the check-worthiness of claims (stages 1 and 2), either based on the ranking of claims by score prediction (Kartal; Kutlu, 2023; Nakov; Da-San-Martino *et al.*, 2021) or the classification of the claim using specific annotations (Konstantinovskiy *et al.*, 2021). Since automated systems can introduce biases in claim selection, research has pivoted towards tools like news alerts, speech recognition, and translation models to filter claims more effectively (Rashkin *et al.*, 2017).

Detecting previously fact-checked claims, including those verified in other languages or countries, has been addressed with NLP and information retrieval techniques (stages 3 and 4). In the first case, semantic textual similarity has been applied to match new claims with already-verified ones in English (Thorne; Vlachos, 2018) and Spanish (Martín *et al.*, 2022). In the second case, software tools with different levels of intelligence have been developed for evidence retrieval, including structured data extraction, speech recognition, reverse image search, video forensics, or natural language search (Das *et al.*, 2023).

“ The current trend leans towards fact-checking assisted by Artificial Intelligence rather than relying solely on fully automated fact-checking ”

One remarkable tool covering different stages is *InVid*, a free platform that hosts tools to detect, authenticate and check the reliability and authenticity of images and videos.

<https://www.invid-project.eu>

The *vera.ai* project is expected to continue and expand AI-supported verification tools and services in Europe.

<https://www.veraai.eu>

6. The challenge of the automatic generation of disinformation

Large Language Models (LLMs), introduced in Section 2.2, are one of the most challenging technologies for the massive generation of textual disinformative content. For example, *GPT-3* and *ChatGPT* can produce synthetic text that can be exploited to spread disinformation in many ways (Solaiman *et al.*, 2019):

- to camouflage false content under the guise of real information;
- to create bots and web pages amplifying a disinformative discourse;
- to elude stylistic checkers, etc.

Additionally, since there is no control over the sources used to train LLMs, much of the content they learn and produce is false and biased (Marcus, 2022). Therefore, it is crucial to develop effective methods for detecting and mitigating the impact of LLM-generated disinformation; unfortunately, attempts to date are still ineffective (Mitchell *et al.*, 2023).

Disinformation is not limited to text format; images, videos, and audio can also be spread, often even more harmful than text. The term deepfake denotes very realistic content automatically generated or altered with DL techniques like Generative Adversarial Networks (GANs) (Goodfellow *et al.*, 2014), e.g., fake avatar generation, face and speech manipulation, and person and background substitution. Not surprisingly, deepfakes have been applied for disinformative purposes like damaging an individual's reputation or manipulating elections (Greengard, 2019; Masood *et al.*, 2022). Therefore, fighting deepfakes entails researching how they can be generated and detected (Dagar; Vishwakarma, 2022; Mirsky; Lee, 2022; Saif; Tehseen, 2022).

Face manipulation in images and videos, either partial or total, has been one of the most active areas of research to date and one that poses a significant challenge to fighting disinformation. Full face generation refers to the creation of a completely fake face (Serengil; Ozpinar, 2021) using architectures such as *ProGAN* (Karras *et al.*, 2018) or *StyleGAN* (Karras; Laine; Aila, 2019). Partial manipulation, in turn, refers to modifications like face swapping, attribute manipulation (hair, skin tone, eyes, etc.), face re-enactment, and lip-syncing (Tolosana *et al.*, 2020). Conversely, there is a wide range of ML-based techniques for detecting deepfakes. Particularly, convolutional neural networks (CNNs) with attention mechanisms have been recently used (Dagar; Vishwakarma, 2022; Rana *et al.*, 2022; Tolosana *et al.*, 2020), but their effectiveness lags behind the advances in deepfake generation and the possibility of manually refine the deepfakes in post-production.

7. Conclusions and future work

The speed and the amount of available data make it challenging to distinguish trustworthy information from disinformative content that is often disguised as legit and appeals to emotions and beliefs. Computational technologies have arisen as suitable tools to address disinformation but also have exceptional capabilities to exacerbate the problem through invention and falsification. In this manuscript, we have described the current trends in AI and ML applied to disinformation detection and characterisation, as well as the challenges posed by synthetic text and media generation.

Most of the reviewed proposals perform an a posteriori analysis of disinformative content once it has become impactful, focusing on different features that can be used in automatic classification. While the approaches assume that solutions in specific problems and domains can be extended to others, they strongly depend on the datasets used and the processes to create them. Therefore, there is a need for new, high-quality, and unbiased datasets, particularly in languages other than English. Furthermore, more efforts are required to transfer and evaluate trained models from one domain to another. Similarly, AI-based disinformation analysis tools are not widely available or lack the maturity that non-technological users need.

Early detection of disinformation is crucial to limit the impact of a phenomenon that otherwise is impossible to deter. Therefore, we identify two future research directions for fighting disinformation with AI and ML. The first one is the study of patterns of creation and propagation, including paths and ecosystems, to understand better and anticipate the spread of harmful propaganda and conspiracy theories. The second one is the application of intelligent technologies to amplify the scope of fact-checks and media literacy, similarly as disinformers engineer their messages to reach wider audiences.

These initiatives will require creating explainable AI methods to provide results and justify them and facilitating the interplay between technological tools and practitioners with deep domain knowledge, including fact-checkers, experts, and decision-makers. By addressing these challenges, we will progress towards AI-based systems that can detect and combat disinformation more effectively, ultimately contributing to a better-informed society.

“ Early detection of disinformation is crucial to limit the impact of a phenomenon that otherwise is impossible to deter ”

8. References

- Afroz, Sadia; Brennan, Michael; Greenstadt, Rachel** (2012). "Detecting hoaxes, frauds, and deception in writing style online". In: *IEEE symposium on security and privacy*, pp. 461-475.
<https://doi.org/10.1109/SP.2012.34>
- Aggarwal, Charu C.** (2011). "An introduction to social network data analytics". In: Aggarwal, Charu C. (ed.). *Social network data analytics*. Springer.
<https://doi.org/10.1007/978-1-4419-8462-3>
- Amador, Julio; Molina-Solana, Miguel; Gómez-Romero, Juan** (2019). "Towards easy-to-implement misinformation automatic detection for online social media". In: *Proceedings of the conference for truth and trust online 2019*.
<https://doi.org/10.36370/tto.2019.4>
- Arnold, Phoebe** (2020). "The challenges of online fact checking". *Full fact*, 17 December.
<https://fullfact.org/blog/2020/dec/the-challenges-of-online-fact-checking-how-technology-can-and-cant-help>
- Azevedo, Lucas; D'Aquin, Mathieu; Davis, Brian; Zarrouk, Manel** (2021). "LUX (linguistic aspects under examination): discourse analysis for automatic fake news classification". In: *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pp. 41-56.
<https://doi.org/10.18653/v1/2021.findings-acl.4>
- Barabási, Albert-László** (2016). *Network science*. Cambridge University Press. ISBN: 978 1 107 07626 6
<http://networksciencebook.com>
- Bedi, Punam; Sharma, Chhavi** (2016). "Community detection in social networks". *Wiley interdisciplinary reviews: Data mining and knowledge discovery*, v. 6, n. 3, pp. 115-135.
<https://doi.org/10.1002/widm.1178>
- Bishop, Christopher M.** (2006). *Pattern recognition and machine learning*. Springer. ISBN: 978 0 387 31073 2
<https://www.microsoft.com/en-us/research/uploads/prod/2006/01/Bishop-Pattern-Recognition-and-Machine-Learning-2006.pdf>
- Blondel, Vincent D.; Guillaume, Jean-Loup; Lambiotte, Renaud; Lefebvre, Etienne** (2008). "Fast unfolding of communities in large networks". *Journal of statistical mechanics: theory and experiment*, n. 10, pp. P10008.
<https://doi.org/10.1088/1742-5468/2008/10/P10008>
- Bondielli, Alessandro; Marcelloni, Francesco** (2019). "A survey on fake news and rumour detection techniques". *Information sciences*, v. 497, pp. 38-55.
<https://doi.org/10.1016/j.ins.2019.05.035>
- Bonet-Jover, Alba; Piad-Morffis, Alejandro; Saquete, Estela; Martínez-Barco, Patricio; García-Cumbreras, Miguel-Ángel** (2021). "Exploiting discourse structure of traditional digital media to enhance automatic fake news detection". *Expert systems with applications*, v. 169, 114340.
<https://doi.org/10.1016/j.eswa.2020.114340>
- 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://papers.nips.cc/paper/2020/hash/1457c0d6bfc4967418bfb8ac142f64a-Abstract.html>
- Buda, Jakab; Bolonyai, Flora** (2020). "An ensemble model using n-grams and statistical features to identify fake news spreaders on Twitter". In: *Working notes of CLEF 2020 - Conference and labs of the evaluation forum*, v. 2696.
https://ceur-ws.org/Vol-2696/paper_189.pdf
- Camacho, David; Panizo-Lledot, Ángel; Bello-Organ, Gema; González-Pardo, Antonio; Cambria, Erik** (2020). "The four dimensions of social network analysis: an overview of research methods, applications, and software tools". *Information fusion*, v. 63, pp. 88-120.
<https://doi.org/10.1016/j.inffus.2020.05.009>
- Cambria, Erik; Wang, Haixun; White, Bebo** (2014). "Guest editorial: big social data analysis". *Knowledge-based systems*, v. 69.
<https://doi.org/10.1016/j.knosys.2014.07.002>
- Castelo, Sonia; Almeida, Thais; Elghafari, Anas; Santos, Aécio; Pham, Kien; Nakamura, Eduardo; Freire, Juliana** (2019). "A topic-agnostic approach for identifying fake news pages". In: *Companion proceedings of the 2019 World Wide Web conference*, pp. 975-980.
<https://doi.org/10.1145/3308560.3316739>

- Dagar, Deepak; Vishwakarma, Dinesh K.** (2022). "A literature review and perspectives in deepfakes: generation, detection, and applications". *International journal of multimedia information retrieval*, v. 11, n. 3, pp. 219-289.
<https://doi.org/10.1007/s13735-022-00241-w>
- Das, Anubrata; Liu, Houjiang; Kovatchev, Venelin; Lease, Matthew** (2023). "The state of human-centered NLP technology for fact-checking". *Information processing & management*, v. 60, n. 2, 103219.
<https://doi.org/10.1016/j.ipm.2022.103219>
- Davis, Clayton-Allen; Varol, Onur; Ferrara, Emilio; Flammini, Alessandro; Menczer, Filippo** (2016). "BotOrNot: a system to evaluate social bots". In: *Proceedings of the 25th International conference companion on World Wide Web*, pp. 273-274.
<https://doi.org/10.1145/2872518.2889302>
- Della-Vedova, Marco L.; Tacchini, Eugenio; Moret, Stefano; Ballarin, Gabriele; DiPierro, Massimo; De-Alfaro, Luca** (2018). "Automatic online fake news detection combining content and social signals". In: *22nd Conference of open innovations association (Fruct)*, pp. 272-279.
<https://doi.org/10.23919/FRUCT.2018.8468301>
- De-Souza, Mariana C.; Nogueira, Bruno-Magalhães; Rossi, Rafael-Geraldeli; Marcacini, Ricardo-Marcondes; Dos-Santos, Bruce-Neves; Rezende, Solange-Oliveira** (2022). "A network-based positive and unlabeled learning approach for fake news detection". *Machine learning*, v. 111, n. 10, pp. 3549-3592.
<https://doi.org/10.1007/s10994-021-06111-6>
- Del-Vicario, Michela; Vivaldo, Gianna; Bessi, Alessandro; Zollo, Fabiana; Scala, Antonio; Caldarelli, Guido; Quattrociocchi, Walter** (2016). "Echo chambers: emotional contagion and group polarization on facebook". *Scientific reports*, v. 6, 37825.
<https://doi.org/10.1038/srep37825>
- Des-Mesnards, Nicolas-Guenon; Hunter, David-Scott; El-Hjouji, Zakaria; Zaman, Tauhid** (2022). "Detecting bots and assessing their impact in social networks". *Operations research*, v. 70, n. 1.
<https://doi.org/10.1287/opre.2021.2118>
- Dong, Xishuang; Victor, Uboho; Qian, Lijun** (2020). "Two-path deep semisupervised learning for timely fake news detection". *IEEE transactions on computational social systems*, v. 7, n. 6, pp. 1386-1398.
<https://doi.org/10.1109/TCSS.2020.3027639>
- Ferrara, Emilio; Varol, Onur; Davis, Clayton; Menczer, Filippo; Flammini, Alessandro** (2016). "The rise of social bots". *Communications of the ACM*, v. 59, n. 7, pp. 96-104.
<https://doi.org/10.1145/2818717>
- Fortunato, Santo** (2010). "Community detection in graphs". *Physics reports*, v. 486, n. 3-5, pp. 75-174.
<https://doi.org/10.1016/j.physrep.2009.11.002>
- Ghanem, Bilal; Ponzetto, Simone P.; Rosso, Paolo; Rangel, Francisco** (2021). "FakeFlow: fake news detection by modeling the flow of affective information". In: *Proceedings of the 16th Conference of the European chapter of the Association for Computational Linguistics*, pp. 679-689.
<https://doi.org/10.18653/v1/2021.eacl-main.56>
- Giachanou, Anastasia; Ghanem, Bilal; Rísola, Esteban A.; Rosso, Paolo; Crestani, Fabio; Oberski, Daniel** (2022). "The impact of psycholinguistic patterns in discriminating between fake news spreaders and fact checkers". *Data & knowledge engineering*, v. 138, 101960.
<https://doi.org/10.1016/j.datak.2021.101960>
- Giachanou, Anastasia; Rosso, Paolo; Crestani, Fabio** (2019). "Leveraging emotional signals for credibility detection". In: *Proceedings of the 42nd international ACM SIGIR conference on research and development in information retrieval*, pp. 877-880.
<https://doi.org/10.1145/3331184.3331285>
- Giachanou, Anastasia; Rosso, Paolo; Crestani, Fabio** (2021). "The impact of emotional signals on credibility assessment". *Journal of the Association for Information Science and Technology*, v. 72, n. 9, pp. 1117-1132.
<https://doi.org/10.1002/asi.24480>
- Giachanou, Anastasia; Zhang, Guobiao; Rosso, Paolo** (2020). "Multimodal multi-image fake news detection". In: *IEEE 7th International conference on data science and advanced analytics (DSAA)*, pp. 647-654.
<https://doi.org/10.1109/DSAA49011.2020.00091>
- Goodfellow, Ian; Bengio, Yoshua; Courville, Aaron** (2016). *Deep learning*. MIT Press. ISBN: 978 0 262 035613
- Goodfellow, Ian J.; Pouget-Abadie, Jean; Mirza, Mehdi; Xu, Bing; Warde-Farley, David; Ozair, Sherjil; Courville, Aaron; Bengio, Yoshua** (2014). "Generative adversarial nets". *Advances in neural information processing systems*, v. 27.
<https://papers.nips.cc/paper/5423-generative-adversarial-nets>

- Graves, Lucas** (2018). *Understanding the promise and limits of automated fact-checking*. Reuters Institute, University of Oxford. https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2018-02/graves_factsheet_180226%20FINAL.pdf
- Graves, Lucas; Nyhan, Brendan; Reifler, Jason** (2016). "Understanding innovations in journalistic practice: a field experiment examining motivations for fact-checking". *Journal of communication*, v. 66, n. 1, pp. 102-138. <https://doi.org/10.1111/jcom.12198>
- Greengard, Samuel** (2019). "Will deepfakes do deep damage?". *Communications of the ACM*, v. 63, n. 1, pp. 17-19. <https://doi.org/10.1145/3371409>
- Grinberg, Nir; Joseph, Kenneth; Friedland, Lisa; Swire-Thompson, Briony; Lazer, David** (2019). "Fake news on Twitter during the 2016 U.S. presidential election". *Science*, v. 363, n. 6425, pp. 374-378. <https://doi.org/10.1126/science.aau2706>
- Guo, Bin; Ding, Yasan; Yao, Lina; Liang, Yunji; Yu, Zhiwen** (2020). "The future of false information detection on social media: new perspectives and trends". *ACM computing surveys*, v. 53, n. 4. <https://doi.org/10.1145/3393880>
- 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
- Hangloo, Sakshini; Arora, Bhavna** (2022). "Combating multimodal fake news on social media: methods, datasets, and future perspective". *Multimedia systems*, v. 28, n. 6, pp. 2391-2422. <https://doi.org/10.1007/s00530-022-00966-y>
- Harrigan, Paul; Daly, Timothy M.; Coussement, Kristof; Lee, Julie A.; Soutar, Geoffrey N.; Evers, Uwana** (2021). "Identifying influencers on social media". *International journal of information management*, v. 56, 102246. <https://doi.org/10.1016/j.ijinfomgt.2020.102246>
- Jing, Jing; Wu, Hongchen; Sun, Jie; Fang, Xiaochang; Zhang, Huaxiang** (2023). "Multimodal fake news detection via progressive fusion networks". *Information processing & management*, v. 60, n. 1, 103120. <https://doi.org/10.1016/j.ipm.2022.103120>
- John, Oliver P.; Srivastava, Sanjay** (1999). "The big five trait taxonomy: history, measurement, and theoretical perspectives". In: Pervin, Lawrence A.; John, Oliver P. (eds.). *Handbook of personality: Theory and research*, pp. 102-138. <https://pages.uoregon.edu/sanjay/pubs/bigfive.pdf>
- Kang, SeongKu; Hwang, Junyoung; Yu, Hwanjo** (2020). "Multi-modal component embedding for fake news detection". In: *14th international conference on ubiquitous information management and communication (Imcom)*. <https://doi.org/10.1109/IMCOM48794.2020.9001800>
- Karras, Tero; Aila, Timo; Laine, Samuli; Lehtinen, Jaakko** (2018). "Progressive growing of GANs for improved quality, stability, and variation". In: *6th International conference on learning representations*. https://research.nvidia.com/sites/default/files/pubs/2017-10_Progressive-Growing-of/karras2018iclr-paper.pdf
- Karras, Tero; Laine, Samuli; Aila, Timo** (2019). "A style-based generator architecture for generative adversarial networks". In: *Proceedings of the IEEE computer society conference on computer vision and pattern recognition*, pp. 4396-4405. <https://doi.org/10.1109/CVPR.2019.00453>
- Kartal, Yavuz-Selim; Kutlu, Mucahid** (2023). "Re-think before you share: a comprehensive study on prioritizing check-worthy claims". *IEEE transactions on computational social systems*, v. 10, n. 1, pp. 362-375. <https://doi.org/10.1109/TCSS.2021.3138642>
- Khattar, Dhruv; Goud, Jaipal-Singh; Gupta, Manish; Varma, Vasudeva** (2019). "MVAE: multimodal variational autoencoder for fake news detection". In: *The World Wide Web conference*, pp. 2915-2921. <https://doi.org/10.1145/3308558.3313552>
- Konstantinovskiy, Lev; Price, Oliver; Babakar, Mevan; Zubiaga, Arkaitz** (2021). "Toward automated factchecking: developing an annotation schema and benchmark for consistent automated claim detection". *Digital threats: research and practice*, v. 2, n. 2. <https://doi.org/10.1145/3412869>
- Kudugunta, Sneha; Ferrara, Emilio** (2018). "Deep neural networks for bot detection". *Information sciences*, v. 467, pp. 312-322. <https://doi.org/10.1016/j.ins.2018.08.019>
- La-Barbera, David; Roitero, Kevin; Mizzaro, Stefano** (2022). "A hybrid human-in-the-loop framework for fact checking". In: *Proceedings of the 6th Workshop on natural language for artificial intelligence (NL4AI 2022)*, v. 3287. <https://ceur-ws.org/Vol-3287/paper4.pdf>

- LeCun, Yann; Bengio, Yoshua; Hinton, Geoffrey** (2015). "Deep learning". *Nature*, v. 521, n. 7553, pp. 436-444.
<https://doi.org/10.1038/nature14539>
- Li, Dun; Guo, Haimei; Wang, Zhenfei; Zheng, Zhiyun** (2021). "Unsupervised fake news detection based on autoencoder". *IEEE access*, v. 9, pp. 29356-29365.
<https://doi.org/10.1109/ACCESS.2021.3058809>
- Li, Shuo; Yao, Tao; Li, Saifei; Yan, Lianshan** (2022). "Semantic-enhanced multimodal fusion network for fake news detection". *International journal of intelligent systems*, v. 37, n. 12, pp. 12235-12251.
<https://doi.org/10.1002/int.23084>
- Li, Xin; Lu, Peixin; Hu, Lianting; Wang, Xiao-Guang; Lu, Long** (2022). "A novel self-learning semi-supervised deep learning network to detect fake news on social media". *Multimedia tools and applications*, v. 81, n. 14, pp. 19341-19349.
<https://doi.org/10.1007/s11042-021-11065-x>
- Liu, Yang; Wu, Yi-Fang** (2018). "Early detection of fake news on social media through propagation path classification with recurrent and convolutional networks". *Proceedings of the AAAI conference on artificial intelligence*, v. 32, n. 1, pp. 354-361.
<https://doi.org/10.1609/aaai.v32i1.11268>
- Liu, Yang; Xu, Songhua** (2016). "Detecting rumors through modeling information propagation networks in a social media environment". *IEEE transactions on computational social systems*, v. 3, n. 2, pp. 46-62.
<https://doi.org/10.1109/TCSS.2016.2612980>
- Manning, Christopher D.; Schütze, Hinrich** (1999). *Foundations of statistical natural language processing*. MIT Press. ISBN: 978 0 262 133609
- Marcus, Gary** (2022). "AI platforms like *chatGPT* are easy to use but also potentially dangerous". *Scientific American*, 19 December.
<https://www.scientificamerican.com/article/ai-platforms-like-chatgpt-are-easy-to-use-but-also-potentially-dangerous>
- Martín, Alejandro; Huertas-Tato, Javier; Huertas-García, Álvaro; Villar-Rodríguez, Guillermo; Camacho, David** (2022). "FacTeR-check: semi-automated fact-checking through semantic similarity and natural language inference". *Knowledge-based systems*, v. 251, 109265.
<https://doi.org/10.1016/j.knsys.2022.109265>
- Masood, Momina; Nawaz, Mariam; Malik, Khalid M.; Javed, Ali; Irtaza, Aun; Malik, Hafiz** (2022). "Deepfakes generation and detection: state-of-the-art, open challenges, countermeasures, and way forward". *Applied intelligence*, v. 54, pp. 3974-4026.
<https://doi.org/10.1007/s10489-022-03766-z>
- Meel, Priyanka; Vishwakarma, Dinesh K.** (2020). "Fake news, rumor, information pollution in social media and web: a contemporary survey of state-of-the-arts, challenges and opportunities". *Expert systems with applications*, v. 153, 112986.
<https://doi.org/10.1016/j.eswa.2019.112986>
- Meel, Priyanka; Vishwakarma, Dinesh K.** (2021). "A temporal ensembling based semi-supervised convnet for the detection of fake news articles". *Expert systems with applications*, v. 177, 115002.
<https://doi.org/10.1016/j.eswa.2021.115002>
- Megahed, Fadel M.; Chen, Ying-Ju; Ferris, Joshua A.; Knoth, Sven; Jones-Farmer, L. Allison** (2023). "How generative AI models such as *chatGPT* can be (mis)used in SPC practice, education, and research? An exploratory study". *ArXiv*.
<https://doi.org/10.48550/arXiv.2302.10916>
- Mikolov, Tomas; Chen, Kai; Corrado, Greg; Dean, Jeffrey** (2013). "Efficient estimation of word representations in vector space". In: *1st International conference on learning representations (ICLR)*.
<https://arxiv.org/abs/1301.3781>
- Mirsky, Yisroel; Lee, Wenke** (2022). "The creation and detection of deepfakes". *ACM computing surveys*, v. 54, n. 1.
<https://doi.org/10.1145/3425780>
- Mitchell, Eric; Lee, Yoonho; Khazatsky, Alexander; Manning, Christopher D.; Finn, Chelsea** (2023). "DetectGPT: zero-shot machine-generated text detection using probability curvature". *ArXiv*.
<https://doi.org/10.48550/arXiv.2301.11305>
- Molina-Solana, Miguel; Amador, Julio; Gómez-Romero, Juan** (2018). "Deep learning for fake news classification". In: *Workshop on deep learning*, pp. 1197-1201.
https://sci2s.ugr.es/caepia18/proceedings/docs/CAEPIA2018_paper_207.pdf
- Nakamura, Kai; Levy, Sharon; Wang, William Y.** (2020). "Fakeddit: a new multimodal benchmark dataset for fine-grained fake news detection". In: *Proceedings of the 12th International conference on language resources and evaluation*, pp. 6149-6157.
<https://aclanthology.org/2020.lrec-1.755.pdf>

- Nakov, Preslav; Corney, David; Hasanain, Maram; Alam, Firoj; Elsayed, Tamer; Barrón-Cedeño, Alberto; Papotti, Paolo; Shaar, Shaden; Da-San-Martino, Giovanni** (2021). "Automated fact-checking for assisting human fact-checkers". In: *Proceedings of the Thirtieth international joint conference on artificial intelligence (IJCAI)*, pp. 4551-4558.
<https://doi.org/10.24963/ijcai.2021/619>
- Nakov, Preslav; Da-San-Martino, Giovanni; Elsayed, Tamer; Barrón-Cedeño, Alberto; Míguez, Rubén; Shaar, Shaden; Alam, Firoj; Haouari, Fatima; Hasanain, Maram; Mansour, Watheq; Hamdan, Bayan; Ali, Zien-Sheikh; Babulkov, Nikolay; Nikolov, Alex; Shahi, Gautam-Kishore; Struß, Julia-Maria; Mandl, Thomas; Kutlu, Mucahid; Kartal, Yavuz-Selim** (2021). "Overview of the clef-2021 checkthat! Lab on detecting check-worthy claims, previously fact-checked claims, and fake news". In: *International conference of the cross-language evaluation forum for European languages. Experimental IR meets multilinguality, multimodality, and interaction*, pp. 264-291.
https://doi.org/10.1007/978-3-030-85251-1_19
- Newman, Mark E. J.** (2004). "Fast algorithm for detecting community structure in networks". *Physical review E*, v. 69, n. 6, 066133.
<https://doi.org/10.1103/PhysRevE.69.066133>
- Oehmichen, Axel; Hua, Kevin; Amador, Julio; Molina-Solana, Miguel; Gómez-Romero, Juan; Guo, Yi-ke** (2019). "Not all lies are equal. A study into the engineering of political misinformation in the 2016 US presidential election". *IEEE access*, v. 7, pp. 126305-126314.
<https://doi.org/10.1109/ACCESS.2019.2938389>
- Paka, William-Scott; Bansal, Rachit; Kaushik, Abhay; Sengupta, Shubhashis; Chakraborty, Tanmoy** (2021). "Cross-sean: a cross-stitch semi-supervised neural attention model for Covid-19 fake news detection". *Applied soft computing*, v. 107.
<https://doi.org/10.1016/j.asoc.2021.107393>
- Pasi, Gabriella; De-Grandis, Marco; Viviani, Marco** (2020). "Decision making over multiple criteria to assess news credibility in microblogging sites". In: *IEEE International conference on fuzzy systems (FUZZ-IEEE)*.
<https://doi.org/10.1109/FUZZ48607.2020.9177751>
- Pennebaker, James W.; Boyd, Ryan L.; Jordan, Kayla; Blackburn, Kate** (2015). *The development and psychometric properties of LIWC2015*. Austin, TX: University of Texas at Austin.
<https://repositories.lib.utexas.edu/handle/2152/31333>
- Pennington, Jeffrey; Socher, Richard; Manning, Christopher** (2014). "GloVe: global vectors for word representation". In: *Proceedings of the 2014 Conference on empirical methods in natural language processing (Emnlp)*, pp. 1532-1543.
<https://doi.org/10.3115/v1/D14-1162>
- Qi, Peng; Cao, Juan; Yang, Tianyun; Guo, Junbo; Li, Jintao** (2019). "Exploiting multi-domain visual information for fake news detection". In: *IEEE International conference on data mining (ICDM)*, pp. 518-527.
<https://doi.org/10.1109/ICDM.2019.00062>
- Rana, Md-Shohel; Nobi, Mohammad-Nur; Murali, Beddhu; Sung, Andrew H.** (2022). "Deepfake detection: a systematic literature review". *IEEE access*, v. 10, pp. 25494-25513.
<https://doi.org/10.1109/ACCESS.2022.3154404>
- Rashkin, Hannah; Choi, Eunsol; Jang, Jin Y.; Volkova, Svitlana; Choi, Yejin** (2017). "Truth of varying shades: analyzing language in fake news and political fact-checking". In: *Proceedings of the 2017 Conference on empirical methods in natural language processing*, pp. 2931-2937.
<https://doi.org/10.18653/v1/D17-1317>
- Rath, Bhavtosh; Salecha, Aadesh; Srivastava, Jaideep** (2022). "Fake news spreader detection using trust-based strategies in social networks with bot filtration". *Social network analysis and mining*, v. 12, n. 66.
<https://doi.org/10.1007/s13278-022-00890-z>
- Ruffo, Giancarlo; Semeraro, Alfonso; Giachanou, Anastasia; Rosso, Paolo** (2023). "Studying fake news spreading, polarisation dynamics, and manipulation by bots: a tale of networks and language". *Computer science review*, v. 47, 100531.
<https://doi.org/10.1016/j.cosrev.2022.100531>
- Russell, Stuart; Norvig, Peter** (2020). *Artificial intelligence: a modern approach*. Pearson Series. ISBN: 978 0 134 610993
- Saif, Shahela; Tehseen, Samabia** (2022). "Deepfake videos: synthesis and detection techniques - a survey". *Journal of intelligent and fuzzy systems*, v. 42, n. 4, pp. 2989-3009.
<https://doi.org/10.3233/JIFS-210625>
- Schuster, Tal; Schuster, Roei; Shah, Darsh J.; Barzilay, Regina** (2020). "The limitations of stylometry for detecting machine-generated fake news". *Computational linguistics*, v. 46, n. 2, pp. 499-510.
https://doi.org/10.1162/coli_a_00380

- Serengil, Sefik I.; Ozpinar, Alper** (2021). "HyperExtended lightface: a facial attribute analysis framework". In: *International conference on engineering and emerging technologies (Iceet)*.
<https://doi.org/10.1109/ICEET53442.2021.9659697>
- Serrano-Guerrero, Jesús; Olivas, José A.; Romero, Francisco P.; Herrera-Viedma, Enrique** (2015). "Sentiment analysis: a review and comparative analysis of web services". *Information sciences*, v. 311, pp. 18-38.
<https://doi.org/10.1016/j.ins.2015.03.040>
- Shabani, Shaban; Charlesworth, Zarina; Sokhn, Maria; Schuldt, Heiko** (2021). "SAMS: human-in-the-loop approach to combat the sharing of digital misinformation". *CEUR workshop proceedings*, v. 2846.
<https://ceur-ws.org/Vol-2846/paper27.pdf>
- Shao, Chengcheng; Ciampaglia, Giovanni-Luca; Varol, Onur; Yang, Kai-Cheng; Flammini, Alessandro; Menczer, Filippo** (2018). "The spread of low-credibility content by social bots". *Nature communications*, v. 9, n. 1, pp. 4787.
<https://doi.org/10.1038/s41467-018-06930-7>
- Shao, Chengcheng; Hui, Pik-Mai; Wang, Lei; Jiang, Xinwen; Flammini, Alessandro; Menczer, Filippo; Ciampaglia, Giovanni-Luca** (2018). "Anatomy of an online misinformation network". *Plos one*, v. 13, n. 4, e0196087.
<https://doi.org/10.1371/journal.pone.0196087>
- Shrestha, Anu; Spezzano, Francesca** (2022). "Characterizing and predicting fake news spreaders in social networks". *International journal of data science and analytics*, v. 13, n. 4, pp. 385-398.
<https://doi.org/10.1007/s41060-021-00291-z>
- Shu, Kai; Sliva, Amy; Wang, Suhang; Tang, Jiliang; Liu, Huan** (2017). "Fake news detection on social media: a data mining perspective". *ACM SIGKDD explorations newsletter*, v. 19, n. 1, pp. 22-36.
<https://doi.org/10.1145/3137597.3137600>
- Shu, Kai; Wang, Suhang; Liu, Huan** (2019). "Beyond news contents: the role of social context for fake news detection". In: *Proceedings of the 12th ACM International conference on web search and data mining*, pp. 312-320.
<https://doi.org/10.1145/3289600.3290994>
- Shu, Kai; Zhou, Xinyi; Wang, Suhang; Zafarani, Reza; Liu, Huan** (2019). "The role of user profiles for fake news detection". In: *Proceedings of the 2019 IEEE/ACM International conference on advances in social networks analysis and mining*, pp. 436-439.
<https://doi.org/10.1145/3341161.3342927>
- Simko, Jakub; Racsko, Patrik; Tomlein, Matus; Hanakova, Martina; Moro, Robert; Bielikova, Maria** (2021). "A study of fake news reading and annotating in social media context". *New review of hypermedia and multimedia*, v. 27, n. 1-2, pp. 97-127.
<https://doi.org/10.1080/13614568.2021.1889691>
- Singh, Prabhav; Srivastava, Ridam; Rana, K. P. S.; Kumar, Vineet** (2023). "SEMI-fnd: stacked ensemble based multimodal inferencing framework for faster fake news detection". *Expert systems with applications*, v. 215, 119302.
<https://doi.org/10.1016/j.eswa.2022.119302>
- Solaiman, Irene; Brundage, Miles; Clark, Jack; Askill, Amanda; Herbert-Voss, Ariel; Wu, Jeff; Radford, Alec; Krueger, Gretchen; Kim, Jong-Wook; Kreps, Sarah; McCain, Miles; Newhouse, Alex; Blazakis, Jason; McGuffie, Kris; Wang, Jasmine** (2019). "Release strategies and the social impacts of language models". *ArXiv*.
<https://doi.org/10.48550/arXiv.1908.09203>
- Song, Chenguang; Teng, Yiyang; Zhu, Yangfu; Wei, Siqi; Wu, Bin** (2022). "Dynamic graph neural network for fake news detection". *Neurocomputing*, v. 505, pp. 362-374.
<https://doi.org/10.1016/j.neucom.2022.07.057>
- Srinivas, P. Y. K. L.; Das, Amitava; Pulabaigari, Viswanath** (2022). "Fake spreader is narcissist; real spreader is Machiavellian prediction of fake news diffusion using psycho-sociological facets". *Expert systems with applications*, v. 207, 117952.
<https://doi.org/10.1016/j.eswa.2022.117952>
- Stella, Massimo; Ferrara, Emilio; De-Domenico, Manlio** (2018). "Bots increase exposure to negative and inflammatory content in online social systems". *Proceedings of the National Academy of Sciences*, v. 115, n. 49, pp. 12435-12440.
<https://doi.org/10.1073/pnas.1803470115>
- Tacchini, Eugenio; Ballarin, Gabriele; Della-Vedova, Marco L.; Moret, Stefano; De-Alfaro, Luca** (2017). "Some like it hoax: automated fake news detection in social networks". In: *CEUR Workshop proceedings*, v. 1960.
<https://arxiv.org/abs/1704.07506>
- 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>

- Tolosana, Rubén; Vera-Rodríguez, Rubén; Fierrez, Julián; Morales, Aythami; Ortega-García, Javier** (2020). "Deepfakes and beyond: a survey of face manipulation and fake detection". *Information fusion*, v. 64, pp. 131-148.
<https://doi.org/10.1016/j.inffus.2020.06.014>
- Vaswani, Ashish; Shazeer, Noam; Parmar, Niki; Uszkoreit, Jakob; Jones, Llion; Gomez, Aidan N.; Kaiser, Łukasz; Polosukhin, Illia** (2017). "Attention is all you need". In: *31st Conference on neural information processing systems*.
<https://papers.nips.cc/paper/7181-attention-is-all-you-need.pdf>
- Vogel, Inna; Meghana, Meghana** (2020). "Fake news spreader detection on Twitter using character n-grams". In: *CEUR Workshop proceedings*, v. 2696.
https://ceur-ws.org/Vol-2696/paper_59.pdf
- 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, Tingting; Liu, Hongyan; He, Jun; Du, Xiaoyong** (2013). "Mining user interests from information sharing behaviors in social media". In: *Pacific-Asia conference on knowledge discovery and data mining*, pp. 85-98.
https://doi.org/10.1007/978-3-642-37456-2_8
- Wang, William Y.** (2017). "'Liar, liar pants on fire': a new benchmark dataset for fake news detection". In: *55th Annual meeting of the Association for Computational Linguistics*, v. 2, pp. 422-426.
<https://doi.org/10.18653/v1/P17-2067>
- Wang, Yaqing; Ma, Fenglong; Jin, Zhiwei; Yuan, Ye; Xun, Guangxu; Jha, Kishlay; Su, Lu; Gao, Jing** (2018). "EANN: event adversarial neural networks for multi-modal fake news detection". In: *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*, pp. 849-857.
<https://doi.org/10.1145/3219819.3219903>
- Wardle, Claire; Derakhshan, Hossein** (2017). *Information disorder: toward an interdisciplinary framework for research and policy making*. Council of Europe report.
<https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-research/168076277c>
- Xiong, Shufeng; Zhang, Gupei; Batra, Vishwash; Xi, Lei; Shi, Lei; Liu, Liangliang** (2023). "Trimoon: two-round inconsistency-based multi-modal fusion network for fake news detection". *Information fusion*, v. 93, pp. 150-158.
<https://doi.org/10.1016/j.inffus.2022.12.016>
- Xu, Fan; Sheng, Victor S.; Wang, Mingwen** (2023). "A unified perspective for disinformation detection and truth discovery in social sensing: a survey". *ACM computing surveys*, v. 55, n. 1.
<https://doi.org/10.1145/3477138>
- Yang, Jing; Vega-Oliveros, Didier; Seibt, Tais; Rocha, Anderson** (2021). "Scalable fact-checking with human-in-the-loop". In: *IEEE International workshop on information forensics and security (WIFS)*.
<https://doi.org/10.1109/WIFS53200.2021.9648388>
- Yang, Shuo; Shu, Kai; Wang, Suhang; Gu, Renjie; Wu, Fan; Liu, Huan** (2019). "Unsupervised fake news detection on social media: a generative approach". *Proceedings of the AAAI Conference on artificial intelligence*, v. 33, n. 1, pp. 5644-5651.
<https://doi.org/10.1609/aaai.v33i01.33015644>
- Yin, Zhijun; Cao, Liangliang; Gu, Quanquan; Han, Jiawei** (2012). "Latent community topic analysis". *ACM transactions on intelligent systems and technology*, v. 3, n. 4.
<https://doi.org/10.1145/2337542.2337548>
- Zhang, Guobiao; Giachanou, Anastasia; Rosso, Paolo** (2022). "SceneFND: multimodal fake news detection by modelling scene context information". *Journal of information science*, Online first.
<https://doi.org/10.1177/01655515221087683>
- Zhang, Xichen; Ghorbani, Ali A.** (2020). "An overview of online fake news: characterization, detection, and discussion". *Information processing and management*, v. 57, n. 2.
<https://doi.org/10.1016/j.ipm.2019.03.004>
- Zhou, Xinyi; Jain, Atishay; Phoha, Vir V.; Zafarani, Reza** (2020). "Fake news early detection". *Digital threats: research and practice*, v. 1, n. 2.
<https://doi.org/10.1145/3377478>
- Zhou, Xinyi; Zafarani, Reza** (2020). "A survey of fake news: fundamental theories, detection methods, and opportunities". *ACM computing surveys*, v. 53, n. 5.
<https://doi.org/10.1145/3395046>
- Zhu, Q.; Luo, J.** (2022). "Generative pre-trained transformer for design concept generation: an exploration". *Proceedings of the design society*, v. 2, pp. 1825-1834.
<https://doi.org/10.1017/pds.2022.185>

9. Annex

Table 1. Datasets created to train models for disinformation classification

Datasets	Application	Source	Size	Source information	Labels	Annotations	Feature coverage	Language	Publicly available	URL
CREDBANK	Credibility assessment	Twitter	> 60 million	Social media posts about 1,049 events	Tuple <degree (certainly, probably, uncertain), polarity (accurate, inaccurate, uncertain)>	Mechanical Turk	Content and context features	English	Yes	https://compsocial.github.io/CREDBANK-data/
PHEME	Rumour detection	Twitter	5,802	Social media posts about 1,049 events	Rumour (1,972), Non-rumour (3,830)	Expert annotation	Content features	English	Yes	https://figshare.com/articles/dataset/PHEME_dataset_of_rumours_and_non_rumours/4010619
LIAR	False information detection	PolitiFact.com	12,800	Political statements	pants-fire (1,047), false (2,507), barely-true (2,103), half-true (2,627), mostly-true (2,454), and true (2,053)	Expert annotation	Content and context features	English	Yes	https://www.cs.ucsb.edu/~william/data/liar_dataset.zip
FakeNews-Net	Study false information on social media	BuzzFeed.com and PolitiFact.com	422	News content	Fake (211), Real (211)	Expert annotation	Content and context features	English	Yes	https://github.com/KaiDMML/FakeNewsNet
MuMiN	Misinformation detection	Twitter and 115 fact-checking organisations	12,914 fact-checked claims and 21,565,018 tweets	Social media post and fact-checked claims	Misinformation, factual	Semi-automatic	Content and context features	Multi-lingual	Yes	https://mumin-dataset.github.io/gettingstarted/
MediaEval	Misinformation and conspiracies detection	Twitter	3,389	Social media posts	Promotes/Supports Conspiracy, Discusses Conspiracy and Non-Conspiracy	Expert annotation	Content and context features	English	Under request	https://multimediaeval.github.io
Buzz-FeedNews dataset	False information detection	Facebook	2,282	Social media posts from 9 sources (3 right-wing bias, 3 left-wing bias and 3 credible)	Most true (1,669), No factual content (264), Mixture of true and false (245), Mostly false (104)	Expert annotation	Content and context features	English	Yes	https://webis.de/data/buzzfeed-webis-fake-news-16.html
BuzzFace dataset	False information detection and bots detection	Facebook	> 1.6 millions	Social media posts verified by BuzzFeed and comments and reactions about this posts	Only source data (BuzzFeedNews dataset) are labelled	Expert annotation	Content and context features	English	Yes	https://github.com/gsantia/BuzzFace
FacebookHoax	Hoax detection	Facebook	15,500	Social media posts 32 pages (14 conspiracy and 18 scientific)	Hoax (8,923), Non-Hoax (6,577)	Pages assumptions	Content and context features	English	Yes	https://github.com/gabll/some-like-it-hoax
FACTOID	False information spreaders detection	Reddit	4,150	3,354,450 social media posts authored by 4,150 users	Real news spreader (3,071), Fake news spreader (1,079)	Expert-based automatic annotation	Content and context features	English	Yes	https://github.com/caisa-lab/FAC-TOID-dataset
Spanish Fake News Corpus	False information detection	News media websites	971	News from 9 different topics	Fake (480), Real (491)	Expert annotation	Content features	Spanish	Yes	https://github.com/jpposadas/Fake-NewsCorpusSpanish
Spanish Fake News Corpus 2.0	False information detection	News media websites and social networks	1,543	News and social media post from 12 different topics	Fake (766), Real (777)	Expert annotation	Content features	Spanish	Yes	https://github.com/jpposadas/Fake-NewsCorpusSpanish
NLI19-SP	Misinformation detection	Twitter	46,919	Social media posts related with a pool of 61 hoaxes identified by fact-checker organisations	Contradiction (406), Entailment (2,521), Neutral (43,992)	Automatic annotation	Content and context features	Spanish and English	Under request	https://aida.etsisi.upm.es/download/nli19-sp-dataset-facter-check
PAN-AP-2020 corpus	False information spreaders detection	Twitter	500	Social media users from news posted on Twitter.	Real news spreader (250), fake news spreader (250)	Expert annotation	Content and context features	Spanish and English	Under request	https://zenodo.org/record/4039435#.Y2z2f8ryRs

Designing personalisation of European public service media (PSM): trends on algorithms and artificial intelligence for content distribution

César Feiras-Ceide; Martín Vaz-Álvarez; Miguel Túñez-López

Nota: Este artículo se puede leer en español en:
<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87298>

Recommended citation:

Feiras-Ceide, César; Vaz-Álvarez, Martín; Túñez-López, Miguel (2023). "Designing personalisation of European public service media (PSM): trends on algorithms and artificial intelligence for content distribution". *Profesional de la información*, v. 32, n. 3, e320311.

<https://doi.org/10.3145/epi.2023.may.11>

Manuscript received on 1st March 2023
Accepted on 09th May 2023



César Feiras-Ceide ✉
<https://orcid.org/0000-0001-5606-3236>

Universidade de Santiago de Compostela
Facultad de Ciencias de la Comunicación
Av. de Castelao, s/n. Campus Norte
15782 Santiago de Compostela, Spain
cesar.feiras@rai.usc.es



Martín Vaz-Álvarez
<https://orcid.org/0000-0002-4848-9795>

Universidade de Santiago de Compostela
Facultad de Ciencias de la Comunicación
Av. de Castelao, s/n. Campus Norte
15782 Santiago de Compostela, Spain
martin.vaz.alvarez@usc.es



Miguel Túñez-López ✉
<https://orcid.org/0000-0002-5036-9143>

Universidade de Santiago de Compostela
Facultad de Ciencias de la Comunicación
Av. de Castelao, s/n. Campus Norte
15782 Santiago de Compostela, Spain
miguel.tunez@usc.es

Abstract

The migration of audiences to digital environments has motivated the media to develop a content distribution strategy that has a presence in these new spaces. In the case of European public broadcasters, they have strengthened their digital news services and have built video-on-demand platforms where they organise and screen their products. Even so, the overload of information and content reaching users forces corporations to look for new mechanisms to present an adequate, interesting and diverse offering to each of their followers. This research project analyses the use of artificial intelligence in the recommendation systems implemented by 14 European public broadcasters in Germany (ARD and ZDF), Belgium (VRT and RTBF), Denmark (DR), Spain (RTVE), Finland (YLE), France (France TV), Great Britain (BBC), the Netherlands (NPO), Ireland (RTÉ), Italy (RAI), Sweden (SVT) and Switzerland (RTS). The results reveal that there is no unanimity among the corporations with regard to the operation and origin of these systems, which vary between home-made developments, acquired from third parties, or collaborative solutions. Operators differentiate between news recommendation processes and those executed on their VoD platforms and aim to distance their systems from those of commercial media, for which they have already started working on a public service media (PSM) algorithm that includes traditional public media values, avoids filter bubbles, and pays special attention to the *European General Data Protection Regulation (GDPR)*.

Keywords

Public service broadcasting; PSM; Recommendation systems; Artificial intelligence; AI; Public service media; Algorithms; Bubble filters; Video on demand; VoD; Journalism; Audiovisual communication; Content distribution; News; Trends; Europe.



Funding

This article is part of the activities of the project “Public audiovisual media in the face of the platform ecosystem: management models and evaluation of the public value of reference for Spain” (PID2021-122386OB-I00), funded by the *Ministry of Science and Innovation (Micin, Spain)*, *Agencia Estatal de Investigación (AEI, Spain)*, and *European Regional Development Fund (ERDF, EU)*. In addition to being integrated in the activities of the *RTVE USC Chair on Public Service Media in Europe*.

1. Introduction

In the age of digitisation, public service media (PSM) have been forced to look for new alternatives to remain relevant to their audiences, improve their experience and ultimately fulfil their public service mission. Competition with the private media and the services they offer have also prompted European public broadcasters to implement disruptive technologies to close the competitive gap with them. This is why they have chosen to implement content recommendation systems in their corporations, with which they achieve an effective distribution where users and their interests converge at a common point: Video on Demand (VoD) platforms or news apps.

The algorithms that feed these systems are mathematical models that work with huge amounts of data and have already penetrated the different spheres of everyday life (Thurman; Lewis; Kunert, 2019). However, the opacity of their operation sometimes limits the credibility of their internal processes, which, together with the accelerated implementation of resources based on artificial intelligence, points to three major challenges for democracy in relation to these technologies (García-Orosa, 2022):

- polarisation;
- fake news, deepfakes and astroturfing;
- echo chambers and bubble filters.

The *Reuters Institute* believes that AI will be the technology with the greatest impact on journalism in the coming years (Newman, 2021, p. 30). Internationally, these solutions have already been implemented and accommodated in newsrooms at all stages of creation and distribution, from the search for newsworthy elements and content analysis, to automated production, distribution and verification (Sanahuja-Sanahuja; López-Rabadán, 2022), which means that journalists inevitably need to become more technologically literate (EBU, 2019).

In the case of Public Service Media, they have traditionally worked with the aim of offering relevant content to the whole population (Sørensen, 2019), and although traditional algorithmic recommendation (also called collaborative filtering) optimises content exposure and protects the user from information overflow (Herlocker *et al.*, 1999), the application of intelligent tools that select, filter and distribute content is particularly delicate (Feiras-Ceide; Vaz-Álvarez; Túñez-López, 2022), as these mechanisms have to be aligned with the core values of these media (Aslama-Horowitz; Nieminen, 2017): universality, diversity, independence and distinctiveness (Unesco, 2001).

For effective use of recommender systems, PSMs must find a way to provide a critical mass of information sufficient to provide these personalisation engines with material to generate relevant product recommendations for each user.

This is expected to lead to changes in production routines and newsrooms, and some corporations may even decide to incorporate content from other media into their offerings (EBU, 2019).

Regarding news and media, there has always been a concern about the potential loss of diversity that algorithmic recommender systems could lead to (Napoli, 2011). Subsequently, the debate has centred on filter bubbles (Pariser, 2011) that would be caused by this collaborative filtering by omitting content of specific interest in favour of the most popular (Bozdog; Van-den-Hoven, 2015). However, it has been shown that algorithms are just as good at providing diversity as human editors (Möller *et al.*, 2018) and that the uncertainty of filter bubbles may have been overestimated (Zuiderveen-Borgesius *et al.*, 2016).

The use of recommender systems in PSMs has been discussed by Fields, Jones, and Cowlshaw from the *BBC* (2018) and also by Van-den-Bulck and Moe (2017), comparing the organisational strategies of European public corporations with respect to personalisation, and focusing their study on *VRT* in Belgium and *NRK* in Norway. Pöchhacker *et al.* (2017) analyse the incorporation of a recommender system in the German media ecosystem, discussed in depth by Schmidt *et al.* (2018). Sørensen, in 2013, analysed the first tests of personalisation in online services of PSMs, while Bodo (2018) and Kunert and Thurman (2019) studied the use and typologies of personalisation in both private and public media.

The fact that PSMs can offer personalisation options is an important value proposition, as well as a competitive alternative to other media. Even so, Sørensen (2019) argues that data analysts, programmers and data curators insist on the complexity of explaining these systems to editors and journalists, who struggle to understand the importance of generating quality metadata, which directly affects the recommendation outcome. Public service organisations also have a duty to explain to users the procedure of personalisation and use of their personal data, despite the complexity involved (Sørensen, 2019).

Regarding the provenance of these systems, most of the PSMs studied by **Sørensen** (2019) rely on open-source tools to develop them. Only the *DR* in Denmark opted for a proprietary solution offered by an external provider, as they argue that they should move away from the technology side and focus on content creation. Corporations, in general, opt to build their own systems in order to have more internal control over customisation, and for the tool to be tailored specifically to their environment, ensuring the technical independence of PSMs in the future (**Sørensen; Van-den-Bulck**, 2018).

YLE Finland was one of the pioneering PSMs (2014) to experiment with personalisation through its personalised news application, *YLE NewsWatch*. This tool collects data from three sources: “the user’s active choices, other users’ behaviour and editorial decisions.” In addition, it incorporated the first intelligent personal news assistant (*Voitto*) that displays news recommendations directly on the lock screen. *Voitto* has also helped editors in the coverage of events such as municipal elections or *National Hockey League (NHL)* games (*YLEisradio*, 2018).

For its part, the *BBC* has set itself the challenge of incorporating “public service” status into the algorithm. They are already segmenting their services for different audiences, so that the homepage of their website looks different depending on age and region of origin (*EBU*, 2019). The *BBC*’s news bots have been basic, do not use machine learning, and have rarely been integrated into news production. At the same time, the corporation is laying the groundwork for developing more interactive news formats, with a more conversational tone that emphasises greater personalisation (**Jones; Jones**, 2019).

In terms of international cooperation, European public broadcasters joined forces in building a platform through which to share their data science and personalisation tools. This *EBU* project known as *Peach* (personalisation for everyone) operates diversified algorithms, through which it attempts to expand the user’s catalogue of content (*EBU*, 2019). This initiative was inaugurated by the *Bayerische Rundfunk (BR)* in Germany, and later continued as a project of the *European Broadcasting Union* (**Sørensen**, 2019).

In the *Peach* project, the recommendation system is complemented by the work of editors in the news section. Editors can choose content by searching by category, news section, source, or date, and once the selection is complete, the final part of the process begins: publication on websites and apps (**Canavilhas**, 2022). The *EBU* (2021) summarises this initiative as ‘the right content at the right time for the right person on the right device’.

Through AI, this collaboratively developed, open-source technology seeks to adapt the information offer to the user’s interests, combining audience preferences with current news trends (**Canavilhas**, 2022). The *Peach* project includes algorithms for all types of content (text, sound, video, text), with some 20 million elements collected daily. Each content is transformed into a vector representation by means of previously trained machine learning models, aiming to associate similar pieces, even if their descriptors use different words (**Canavilhas**, 2022).

In parallel, the Belgian public service corporation *RTBF* developed another software project aimed at creating generic recommendation solutions for PSM, which, like the *Peach* project, addresses the much-needed diversity requirement in public media (**Sørensen**, 2019). With regard to the quality control of its content, the *Österreichischer Rundfunk (ORF)* identifies the most structured and comprehensive quality measurement and control system in Europe. The logic applied by *ORF* is simple: if the content offered matches the individual interests of a user, it will be more relevant to them. For public service media, this means that they must listen to the feedback of the algorithm in order to modify the offer, both in terms of topics and approach (*EBU*, 2019).

2. Methodology

This research is a continuation of a previous study by the authors analysing artificial intelligence strategies in European public broadcasters, which detects the incipient interest of these corporations in integrating, enhancing, and personalising their content recommendation systems (**Fieiras; Vaz; Túnñez**, 2022). Despite this growing interest, no previous projects have been identified that provide an overview of the use of these technologies in these broadcasters, the procedures followed for their implementation, or the perceptions of the professionals linked to them in relation to their operation.

This is why the central objective that marks the beginning of this research is the following: to extract a complete radiography of the current situation of the use of recommender systems in European public broadcasters, paying special attention to the functioning of these systems, their origin, the teams in charge of their maintenance and implementation, and the concept of public service algorithm. To this end, the following research questions are posed, which are progressively addressed in the results sections:

- Q1. What is the perception and intention of corporations in relation to these algorithmic recommender systems?
- Q2. What is the concept of a public service algorithm, and what is the procedure proposed by the corporations for its construction and implementation?
- Q3. Do the recommendation systems used by these broadcasters come from external companies, collaborative projects, or are they developed by their own internal teams?
- Q4. How is content recommendation executed in each of the systems of the corporations studied?

In order to achieve the central objective of the research and to answer the questions posed, qualitative methods were used and personal in-depth interviews were conducted with the innovation managers, strategy and artificial intelligence of the public broadcasters of Germany (*ARD* and *ZDF*), Belgium (*VRT* and *RTBF*), Denmark (*DR*); Spain (*RTVE*); Finland (*YLE*); France (*France TV*); Great Britain (*BBC*); the Netherlands (*NPO*); Ireland (*RTÉ*); Italy (*RAI*); Sweden (*SVT*) and Switzerland (*RTS*); in Delphi mode and in two rounds.

A purposive convenience sample was selected, which was expanded in snowball mode based on the contributions of the interviewees, and a second consultation was carried out to clarify details of the information collected. We worked with a semi-structured questionnaire that never exceeded 8 questions and in which specific questions were included depending on the corporation interviewed.

Even so, the main blocks in all of them were: digital strategy of content distribution; application and origin of recommendation systems; public service algorithm; and data privacy and regulations. Thirty-three contacts were made and a convenience sample was finally validated in which PSMs from the three media models described by **Hallin** and **Mancini** (2004) were represented. The final panel of 14 respondents consisted of:

Table 1. Purposive convenience sample of the study

Abbreviation	Name	Corporation	Position
(AM/RAI)	Alberto Messina	RAI	R&D Director
(BV/NPO/EBU)	Bob Van de Velde	NPO/EBU	Head of Data and Archives at NPO; Co-Director of IA at EBU
(DC/RTVE)	David Corral	RTVE	Head of content innovation
(JF/DR)	Jakob Faarvang	DR	Digital Product Manager
(JK/YLE)	Jarno Koponen	YLE	Head of AI and personalisation
(JA/BBC)	Jatin Aythora	BBC	Chief Architect, responsible for IA
(JL/SVT)	Johan Linden	SVT	Strategy Director
(JB/BR/ARD)	Jonas Bedford	BR/ARD	Innovation Leader at <i>Bayerischer Rundfunk</i>
(KB/FranceTV)	Kati Bremme	France TV	Artificial Intelligence Vision Director
(LB/RTS/EBU)	Léonard Bouchet	RTS/EBU	Data and Archives Lead at RTS; Co-Director of IA at EBU
(LV/RTBF)	Loïc de Visscher	RTBF	Innovation Director
(MM/VRT)	Mike Matton	VRT	Responsible for international innovation partnerships
(RW/RTÉ)	Richard Waghorn	RTÉ	Director of Operations, Technology and Transformation
(RA/ZDF)	Robert Amlung	ZDF	Digital Strategy Director

Thus, by models, the polarised pluralist model is represented by *RTVE*, *France TV* and *RAI*; the corporatist-democratic model by *VRT*, *RTBF*, *DR*, *RTS*, *SVT*, *ZDF*, *ARD*, *NPO*, and *YLE*; and finally the liberal model with the *BBC* and *RTÉ*.

Table 2. Correspondence between corporations/Hallin Mancini model (2004)

Polarised pluralist	Corporatist democratic	Liberal
<i>France TV</i> (France) <i>RAI</i> (Italy) <i>RTVE</i> (Spain)	<i>ARD</i> (Germany) <i>DR</i> (Denmark) <i>NPO</i> (Netherlands) <i>RTBF</i> (Belgium) <i>RTS</i> (Switzerland) <i>SVT</i> (Sweden) <i>VRT</i> (Belgium) <i>YLE</i> (Finland) <i>ZDF</i> (Germany)	<i>BBC</i> (United Kingdom) <i>RTÉ</i> (Ireland)

Interviews were conducted between 11 December 2021 and 25 November 2022 via *Skype*, *Zoom* and *Google Meetings*, with an average duration of over 30 minutes. The transcription and translation of the material was carried out in parallel to its collection, while the analysis of the results began after the end of this period.

3. Results

3.1. Recommendation systems in European public broadcasters

The corporations studied agree that the linear services they offer will be maintained in the future and will be complemented by their growing digital offer, which will gain a progressive relevance that will end up imposing itself as reference content in their structures. These media prioritise the establishment of a differentiation with the usual recommendation systems of private operators, in which revenue generation, keeping the user in front of the screen, and where, according to the professionals contacted, “it does not give the sensation of being assisted by people, but by machines”, take precedence.

Regarding the construction of public service algorithms, there are discrepancies among corporations. Some, such as the *NPO*, explain that, in the creation process, they try to relate the PSM values to numerical codes, so that the system can interpret it and apply it in the subsequent recommendation. However, the *BBC* or the *DR* state that they do not see it possible to define what public service really is, so they emphasise giving diversity, plurality, accountability and transparency to the algorithm to fit its mission and avoid filter bubbles.

“The lack of data and engineering skills, forces, on most occasions, corporations to turn to suppliers to provide them with the algorithms”

The lack of data and engineering skills, and of professional profiles that can carry out these systems, forces, on most occasions, corporations to turn to suppliers to provide them with the algorithms and then adapt them to their structure. However, at *YLE*, *VRT*, *RTBF*, *NPO* or *SVT*, they practically develop all this technology in house with the contribution of their employees.

In cases such as *ARD* or *RTS*, they use the common open solution proposed by the *EBU*, even though the results, at least in the last of these corporations, have not been entirely favourable. In the short term, the *BBC* intends to replace the third-party algorithms it owns with its own algorithms over which it has full control and which fully meet its needs.

In evaluating these algorithms, corporations claim to employ quantitative methods such as A/B testing, alongside qualitative methods such as real user testing and interviews. The most common systems are based on collaborative filtering, which is based on user characteristics and offers content based on what similar profiles have consumed; and content-based filtering, which has the product at the heart of the recommendation. European PSMs insist on differentiating between the procedures for generating suggestions in VoD platforms and in news apps, where they stress the need to return to a journalistic emphasis.

Striking is the case of *YLE*, with its intelligent assistant for the *Voitto* news app, which recommends content directly to the user's lock screen if the user allows it. Multimodal audio, video and text recommendations are a growing trend that corporations have not yet mastered but which is identified in their roadmap for the coming years.

On user consent on notifications, data usage and other actions that may infringe on user privacy, there is considerable interest from PSMs. In addition, there is concern about the attitude of part of their audience, who are indifferent to sharing their personal information with any type of platform. In this regard, the importance of maintaining ethics and transparency in the recommendations is emphasised. Specifically, the *YLE* publishes several annual reports in which it explains exactly how it uses data, the type of data it collects and the benefits that these systems bring to users.

In terms of current regulations, corporations agree that they are not pressured by their respective governments, but that they follow the guidelines carefully. At the European level, they highlight the *General Data Protection Regulation (GDPR)*, whose reinterpretation recently forced a change in the functioning of cookie-based systems and led the affected corporations to look for alternatives in subscription portals and log in formulas.

3.2. Concept, construction, and implementation of the public service algorithm

VoD services offered by European PSMs are gaining relevance by leaps and bounds. Even so, as *ZDF* points out, in their case, they are suffering significant losses in linear services that are not compensated by the *Mediathek* (online services). This decline is mainly among young audiences, so it is essential to renew services to capture their attention. One issue on which there is no unanimity among corporations is their position on third party platforms such as *YouTube* or *Facebook* to distribute their content.

On these social media, *SVT* Sweden limits itself to reporting on its programmes and maintaining an ongoing dialogue with the audience. They choose to publish products only on their own sites, want to have full control of their signals and shy away from any payment schemes that may arise in these spaces. However, other broadcasters do not hesitate to upload their pieces to these platforms as they understand that part of their users browse them, which they see as a strategic opportunity.

(*JL/SVT*) “One of our decisions is not to publish full content on *YouTube* or *Facebook*. We use third platforms to report on programmes and dialogue with young audiences. We have to be free and free for all, if you have to pay for a *YouTube* subscription to watch us without ads, we are no longer fulfilling that. We also want to be 100 percent in control of our signals and publishing.”

(*LV/RTBF*) “We made the decision to be on *YouTube*, *Facebook* or *Instagram*. We have a mission to bring public service values to all citizens, so we must have a presence on *YouTube* to bring information to young people. We want to be relevant in the future.”

Corporations agree that the PSM algorithm is necessary to appropriately target content to the population. The media have an offer that is impossible to concentrate on their homepage, so they need to incorporate the recommendations for a satisfactory user experience. Moreover, the audience is routinely adapting to this new presentation of individualised content, so that rather than a differential value, it has already become an indispensable requirement.

The *EBU* points to the complexity of “estimating what news is important to people.” In the exceptional context of the coronavirus, there was a central event that was newsworthy on a global scale, but in everyday life there is not usually news that is relevant for everyone.

(RA/ZDF) “The idea of personalisation in PSM is to target content. We have a lot of content, it is impossible to put everything on the front page and we have to offer a better user experience. Users have *Netflix* and *Amazon* accounts, they are getting used to this kind of content presentation, we have to react.”

“The most common systems are based on collaborative filtering, which is based on user characteristics, and content-based filtering, which has the product at the heart of the recommendation”

To build this public service algorithmic recommendation system, it must be as unbiased as possible so that it improves accessibility issues and allows corporations to reach those populations they do not usually find. *SVT* uses the algorithm to assist publishers. They explain that it should reflect “all parts of the country, the diversity of geography and perspectives.” Meanwhile, *DR* is working on a new service that allows users to create lists of content and then run recommendations based on them.

(JL/SVT) “It’s hard to know how the public service algorithm manages to offer personalisation without being annoying or creating filter bubbles. We don’t operate with it now; we use it to help publishers.”

(JF/DR) “We have been working on a new proposition, *FlexyBlocks*, where we allow the user to create lists of content. We analyse your decisions and recommend you around it. This is our best guess for using recommendation in a public service.”

In the process of creating the PSM algorithm, Bon Van de Velde, head of data and archives at *NPO* and co-director of AI at the *EBU*, points out that the first step is to represent the content input, determine the independent variables and dummies, this is called feature space. Then, the algorithm maps this area, so the input-output set must be quantitative.

The last part includes the evaluation criteria, “whether or not people see this content on recommendation, if they click on it to search for it”, where there must be some kind of score to conclude if it works. In the first step, values can be included, and this is where the introduction of the public service fits in.

(BV/NPO/EBU) “The utility must be introduced in the first step, where features such as values can be included. This algorithm can only optimise one numerical feature. If you do survey research you know that there are many things that can be put in the numbers, so that’s how we represent this public service value.”

3.3. Provenance and evaluation of recommendation algorithms in European public broadcasters

European public broadcasters have not followed the same roadmap when building or acquiring their algorithmic recommender systems. In the cases of *YLE*, *VRT*, *RTBF*, *NPO* or *SVT*, they develop their own concepts and core technologies, although they combine them with resources from suppliers to finalise their operation. *YLE* says that in selecting these one-off collaborations, they choose the company carefully, in the same way they would a worker, as they need them to understand the fundamentals of delivering on their mission and values. The *BBC* is looking at and working on migrating from third-party systems to in-house solutions, as they need better “control, understanding and transparency.”

(JK/YLE) “In all our systems we combine resources, but the concept and the basic principles all come from in-house, we need to understand the systems thoroughly, it’s how they really work. There are always two things we try to do: to deliver and respect our values, mission, and objectives; and to carefully select our technology partners so that they do so as well.”

(JA/BBC) “Our top innovation priority is algorithm-driven recommendation engines. We want to build them all in-house and move away from third-party systems, we are in that transition to have better control, understanding and improve transparency.”

Other corporations such as *RAI* decided to go directly to suppliers that provide a solution that only needs to be customised. The Italian organisation did not go as far as to develop a build strategy, as they needed to have the technology immediately and did not want to spend “human resources, time or money” on the development team. The RD explains that many of his colleagues at PSM think that off-the-shelf solutions do not work and are difficult to control. Yet, they claim that they have always been successful.

(AM/RAI) “We go to suppliers. They basically sell their solution, and you customise it. We decided to do it this way because we wanted to have something ready as soon as possible and not to have to spend human resources, time or money on setting up a development team.”

(JF/DR) “We used machine learning for personalisation. We created a platform for news where you can customise whatever you want. In conversations with colleagues from other PSMs they think we buy solutions that we cannot control, yet it has worked well for us. The news where the editorial factor is important is sorted by humans, while the rest is sorted by the algorithm.”

Corporations such as *RTVE* or *France TV* complain about the lack of professionals specialised in data or engineering and architecture to be able to create these intelligent solutions. This is why *Radiotelevisión Española* follows the trend

of contracting services through public tenders, as they did at the beginning of 2020 with the *EFE* agency and the company *Narrativa* in the area of automatic news creation.

Regarding *France TV*, they are working with startups and are starting to hire staff with more technological profiles. They are in the strategic design phase, cleaning their data to optimise results and deciding what kind of algorithm they want for their system. Another alternative way of working is the one followed by *ZDF*, which takes advantage of synergies in its relationship with other corporations or with specialists from universities.

(DC/RTVE) “We don’t have a team of engineers to design the algorithms, our profiles are what they are. We do not have a data analyst to make user profiles for us. As we have done with *EFE* and *Narrativa* with respect to news, the trend is for this service to be contracted through a public tender.”

(RA/ZDF) “We have some internal knowledge. But we do not know enough to create the algorithm completely in-house. We are partnering. We have a good network formed by other European PSMs, specialised people from universities and with commercial companies through public offerings.”

Corporations such as *RTS* or *BR* use the customisation engine offered by the *EBU* to all public broadcasters, the *Peach* project. However, in the particular experience of the Swiss broadcaster, they are not fully satisfied with the results and are trying to define a new alternative.

For them, the most important thing is “the overall experience, perception, quality and relevance of the experience”, which they will try to improve in relation to the results obtained with the *EBU* technology. The *ARD*’s *BR* created its own system from this project, which is still operational and is not the same as the one used in the central corporation.

(LB/RTS/EBU) “We are part of the *Peach* project. It is a customisation engine for all *EBU* projects. We are not completely satisfied with it, so we are taking another route now that offers more individualised results.”

(JB/BR/ARD) “In *BR* we now have our own recommendation system that we built within the *EBU Peach* project. We still have it implemented today.”

To assess the correct functioning of the recommendation algorithms, the PSMs use both quantitative and qualitative methods. Regarding the former, they are doing A/B testing, which consists of running different alternatives for the same use and measuring the results. On the qualitative level, the *YLE* conducts “real user tests and interviews” to get feedback and improve the system based on the information collected.

An important question is the opinion of corporations on their preference to include editorial decisions in the algorithm or, on the contrary, to teach the algorithm to make editorial decisions. The conclusion is that these are complementary options that are being implemented jointly by European public broadcasters.

(RA/ZDF) “You have to train the algorithm to make editorial decisions and include these decisions in the algorithm. You really need to do the training at the machine level. Feed it data that is given in a public service context. And then, as with any recommender system, you need to interpret the results. This is human work.”

(MM/VRT) “As a researcher I would say we want to explore including editorial decisions in the algorithm and teaching the algorithm to make editorial decisions. Still, I think we will be more successful if we create a hybrid approach where you feed the recommender system, but of course you give more personalisation of the recommender system to the user.”

3.4. Description of the functioning of the recommendation systems of European PSMs

The recommender systems implemented by European PSMs have different strategic origins. One of them is technologies based on collaborative filtering, in which contributors (users) are analysed to establish profiles and recommend products based on the similarity between these profiles. These are the most common and understand that if one user ‘x’ is interested in ‘x’ content, another with similar behaviour will also like it.

On the other hand, there are systems that are based on content and that implement recommendations based on content. Below are two tables of results from the interviews carried out, in which each corporation is listed with a brief description of the recommendation system it applies. In the first table, broadcasters with a provider algorithm are grouped together, while in the second one, media with home-made systems or from collaborative solutions are grouped together.

VoD services offered by European MSPs are gaining in importance by leaps and bounds, but sometimes this increase does not cover the losses suffered in linear services

European public broadcasters have not followed the same roadmap when building or acquiring their algorithmic recommender systems

Table 3. Description of the recommendation system of corporations with supplier algorithms

Corporation	Recommender system (purchased)	Source
RAI	Content-based filtering. Mix between editorial packaging and automatic recommendation. Editorial line still takes precedence.	3 rd party
RTVE	Collaborative filtering. Algorithm that is fed by the information provided by the user through the subscription portal.	3 rd party
France TV	Recommendation by algorithm built by startup.	3 rd party
BBC	Multimodal audio, video, and text recommendations through algorithms. They work to provide a “universal recommendation.”	Currently third-party algorithm, but they intend to switch to home production.
DR	Content curation. They mainly make recommendations with algorithms, but also include editorial decisions on news stories.	3 rd party
	They bought analytics systems for the streaming platform; about news, they bought in the company <i>Cxense</i> . Their service does not require registration, it is powered by cookies. <i>FlexyBlocks</i> : new personalisation proposal, the user creates lists from which they select the most relevant content for each moment.	
RTÉ	Recommendation algorithms together with editorial decisions, no personalisation. They do not have mandatory registration in VoD, an element of personalisation “which is not meaningful” acts on users accessing the log in.	3 rd party

Table 4. Description of the recommendation system of the corporations with a home-made algorithm or from common solutions

Corporation	Recommender system (in-house or collaborative)	Source
VRT	Collaborative filtering, content curation	In-house
RTBF	Collaborative filtering and content-based filtering. Mix between “home-made PSM algorithm” and editorial decisions.	In-house
SVT	Collaborative filtering, filtered playlists. Editorial decisions assisted by a simple home-made algorithm fed by cookies/user IP number. No log in.	In-house
NPO	Content-based filtering. They use several simple algorithms for personalisation. Work with educational research institutes. They have subscription on the VoD platform.	In-house
RTS	They are part of the <i>EBU's Peach</i> project.	<i>Peach</i> project (EBU)
YLE	Different recommendation methods. Both in VoD and in their news system they use algorithms.	They combine resources, but the concept and ideas come from their in-house team.
	In the news, they use their <i>Voitto</i> intelligent assistant. This offers personalisation from the mobile screen based on interests and location. Editorial decisions overwrite the algorithm, they are above it. The streaming service acts differently, with more weight given to the user's story and the action of the algorithm. Even so, they want to give diversity to the recommendations.	
BR/ARD	At <i>BR</i> they use a shared system through the <i>Peach</i> project of the <i>EBU</i> . In the <i>ARD</i> they use their own independent system.	<i>Peach</i> project (EBU)
ZDF	Mix between algorithm recommendation and editorial decisions. The algorithm analyses users and collects information about them, but it is humans who interpret.	They have some in-house expertise, but mainly draw on partnerships with other entities.

Corporations differentiate between the processes of building recommendations for VoD platforms and those for news apps. The former follows a similar pattern to companies like *Netflix* or *Amazon* where personalisation is completed based on the IP number, while news pieces have higher production rates and the procedure is different. PSMs avoid mixing suggestions for one and the other space.

(JL/SVT) “It is important to differentiate between recommendation in the VoD service, which is very similar to *Netflix*, and news. News has much higher production rates, thousands of stories are created. In VoD services we personalise through the IP number, we try to do that as much as we can without going over that integrity threshold. That is done transparently, and the user can decide to turn it off if they want to.”

(LV/RTBF) “The recommendations we make on the player are not breaking news, that's what we have our news site for. They are usually about new fiction that we co-produce and we want them to gain relevance. If you look at our platform, it is less about breaking news than entertainment content.

Relevant to the study of personalisation processes in news apps is the analysis of *YLE's* platform. The Finnish broadcaster uses “different algorithmic systems for both VoD recommendation and news piece services.” There are two use cases

that act completely differently. On the news side, they stress the importance of returning to the journalistic emphasis. Users can use the app without personalisation or with the use of their *Voitto* smart assistant, which delivers two different types of notifications directly to the locked screen: some based on location and some based on their interests.

(JK/YLE) “We are very human-centred, so the user can always decide the level of personalisation. At the end of the day there must be a really big story for the newsroom to think it’s so important that as many people as possible should know about it. That overwrites the algorithm. So, we’re focused on making sure that these journalistic principles are also built into the core of our algorithmic recommendation engine.”

PSMs aim to “open the minds” of users with their recommendations. Filter bubbles are one of the cornerstone problems of public service media, which is why they should focus their efforts on implementing systems that advocate diversity and plurality of content. This is also one of the premises of the public service algorithm. However, this is not about recommending content that is totally opposed to the user’s interests, but about doing so in a gradual, logistically sensitive way, and with results that allow the audience to reach relevant content that they would not go to in their usual behavioural patterns.

(LV/RTBF) “On the VoD platform we have a recommendation engine that was built in-house, we created it to establish an algorithm that we call the PSM algorithm. The goal was to generate a filter that didn’t do the bubble effect.”

(BV/NPO) “You shouldn’t recommend the opposite of what people are seeing because people are not interested in that, so you have to do it very gradually. But if you come up with something that is not in their normal viewing pattern, you still get a little bit of additional diversity.”

Not all corporations include user registration procedures in their systems. In the case of *SVT* they decided to dispense with log in, as they understand that public services should not monitor audiences, so they focus exclusively on cookies to build recommendations. The misuse that *Facebook* and other large companies make of these logs is what makes people believe that they are a threat, when in the opinion of the Swedish public broadcaster this should not be the case. At *RTÉ*, the service offered after registration is improved, although it is not a mandatory requirement for using the platform. *YLE* even personalises the cover images of content that users see on its streaming site.

(JK/YLE) “We are experimenting with the automatic creation of images that are used in recommendation systems. For example, in our audio-visual services or in our streaming service we are using different images for different people when we recommend audio-visual content.”

European PSMs are already working on the next generation of recommendations, multimodal recommendations, which combine text, audio, and video. The *BBC* calls them “universal” and says there is still “a long way to go” to master them, but they understand that they will be of great value to users. For their part, the *DR* says they have seen poor results in their attempts, but they are not giving up on trying in the future. They point out that their metadata structure does not work correctly for all types of content, so they will have to work on them to polish this new form of recommendation.

(JF/DR) “We couldn’t get our multimodal recommendations to work. It was like telling the user to read your economists’ magazine, then watch an episode of *Peppa Pig* and then listen to a podcast. We couldn’t make it work, but we haven’t given up on it.”

(JA/BBC) “The future is in universal recommendations, in multimodal, multi-modal, audio, video and text recommendation experiences. We still have a long way to go, we’re still pretty much driven by editorial value and manual curation.”

4. Conclusions

The role of public media has historically occupied a special position in the media landscape, since in exchange for public funding they have had the arduous task of contributing to social inclusion and cohesion, of reinforcing local culture, and of fostering democratic processes with plural and diverse content (Sørensen, 2019).

Recommender systems have established themselves in this sense as a fundamental tool for individualising attention to each user, providing them with content that best suits their tastes and interests. Even so, the socio-political contexts of each country make it impossible to establish a common model for all of them, even though, in order to comply with the premise of offering each viewer what they need, all public media must guarantee the principles of diversity of exposure and surprise in their proposals (Fields; Jones; Cowlshaw, 2018).

In the case of European public broadcasters, corporations agree that their linear services will be maintained in the future, but insist on the growing relative weight of VoD platforms and digital news applications. They also insist on the need to implement effective recommendation systems to offer viewers the content they want, without producing the isolation of filter bubbles. It will be interesting to monitor the coming years to identify whether this upward forecast is consolidated in the next media stage, and whether the recommendation systems of public media will definitely manage to broaden the field of vision of society without incurring in the isolation typical of private operators’ recommenders.

Corporations agree that the PSM algorithm is necessary to appropriately target content to the population

In this study, corporations from the three media models described by **Hallin** and **Mancini** (2004) have been analysed, without obtaining coinciding results between corporations belonging to the same blocks, so it is concluded that their particular characteristics do not affect the approach, construction and implementation of their recommendation systems. The main results collected from the research questions formulated at the beginning of this research are presented in Table 5 below.

Corporations differentiate between the processes of building recommendations for VoD platforms and those for news apps

Table 5. Results obtained for each research question in this study

Research question	Results of the study
Q1. What is the perception and intention of corporations in relation to these algorithmic recommender systems?	They see them as a prerequisite for optimising the user experience and offering an individualised and tailored service.
	They seek differentiation from private operators' systems.
	They intend to establish recommendations with different guidelines on their VoD platforms and news sites.
	They maintain a rigorous control of GDPR compliance.
Q2. What is the concept of the public service algorithm, and what is the procedure proposed by the corporations for its construction and implementation?	Algorithm integrating traditional public media values.
	Necessary to appropriately target content to the population.
	It should be a pluralistic algorithm reflecting all parts of the country of origin, geographical diversity, and diversity of perspectives.
	The utility values must be represented by numerical codes in order to enter them in the construction of the system.
Q3. Do the recommendation systems used by these broadcasters come from external companies, collaborative projects, or are they developed by their own internal teams?	It is essential to establish valid evaluation criteria that analyse their functioning from the outset.
	There is no unanimity on the origin of the systems. Most often they are purchased from third parties.
	The EBU's <i>Peach</i> project is the most widely used collaborative solution.
	The lack of specialised data or computer engineering professionals in corporate teams hampers home-grown developments.
	The long development times and cost also motivate corporations to turn in the first instance to off-the-shelf solutions to adapt them to their infrastructures.
Q4. How is content recommendation executed in each of the systems of the corporations studied?	Corporations are putting a lot of effort into cleaning and organising their metadata, as it is the central input to the recommendation.
	Collaborative filtering is the most used method, although content-based recommendations are also common.
	The system employed (collaborative or content-based) is always supported by the editorial decisions of the professionals involved.
	Avoiding filter bubbles is one of the main objectives of broadcasters.
	Multimodal recommendations (text, audio and video) mark the new line of work of these corporations.

The PSM's efforts to translate its traditional values into numerical scoring codes to be introduced in the creation phases of the systems will form the basis of the public service algorithm models that will begin to be implemented in these corporations. For this reason, the testing of these solutions to check their validity and effectiveness, in comparison with the results obtained from commercial proposals, is consolidated as a line of study. It will also be interesting to monitor whether these media decide to hire more professionals with technological profiles to dedicate themselves to these processes internally, or whether they will continue to rely on external companies for most of the tasks.

As the most promising initiative, multimodal text, audio, and video recommendations represent a new dimension in the task of personalisation, understanding the particular situation of each user at the specific moment when the viewer requests content, in order to offer the piece that is most needed at that moment. This is why the most technologically advanced corporations will be the first to polish and establish these solutions on a permanent basis, which will guide the rest of the operators to propose a model that is expected to be the protagonist in the coming years.

5. References

- Aslama-Horowitz, Minna; Nieminen, Hannu** (2017). "Diversity and rights. Connecting media reform and public service media". *IC. Revista científica de información y comunicación*, v. 14, pp. 99-119.
<http://icjournal-ojs.org/index.php/IC-Journal/article/view/385/341>
- Bodo, Bolázs** (2018). "Means, not an end (of the world) - The customization of news personalization by European news media". In: *Algorithms, automation, and news conference, Munich*, pp. 22-23.
<https://doi.org/10.2139/ssrn.3141810>

- Bozdag, Engin; Van-den-Hoven, Jeroen** (2015). "Breaking the filter bubble: democracy and design". *Ethics and information technology*, v. 17, pp. 249-265.
<https://doi.org/10.1007/s10676-015-9380-y>
- Canavilhas, João** (2022). "Inteligencia artificial aplicada al periodismo: estudio de caso del proyecto "A European perspective" (UER)". *Revista latina de comunicación social*, n. 80.
<https://www.doi.org/10.4185/RLCS-2022-1534>
- EBU** (2019). *News report 2019. The next newsroom*. Geneva: European Broadcasting Union.
https://www.ebu.ch/publications/strategic/login_only/report/news-report-2019
- EBU** (2021). *Personalisation and recommendation ecosystem developed by broadcasters for broadcasters. Media Intelligence Service*. Geneva: European Broadcasting Union.
<https://peach.ebu.io/docs>
- Fieiras-Ceide, César; Vaz-Álvarez, Martín; Túnñez-López, Miguel** (2022). "Artificial intelligence strategies in European public broadcasters: Uses, forecasts and future challenges". *Profesional de la información*, v. 31, n. 5, e310518.
<https://doi.org/10.3145/epi.2022.sep.18>
- Fields, Ben; Jones, Rhianne; Cowlishaw, Tim** (2018). *The case for public service recommender algorithms*. In: *Fatrec Workshop*. London: BBC, pp. 22-24.
<https://pirt.gitlab.io/fatrec2018/program/fatrec2018-fields.pdf>
- García-Orosa, Berta** (2022). "Digital political communication: Hybrid intelligence, algorithms, automation and disinformation in the fourth wave". In: García-Orosa, Berta (ed.). *Digital political communication strategies*, Palgrave, pp. 3-23. ISBN: 978 3 030 81568 4
https://doi.org/10.1007/978-3-030-81568-4_1
- Hallin, Daniel C.; Mancini, Paolo** (2004). *Comparing media systems: Three models of media and politics*. Cambridge, UK: Cambridge University Press. ISBN: 978 0 511790867
<https://doi.org/10.1017/CBO9780511790867>
- Herlocker, Jonathan L.; Konstan, Joseph A.; Borchers, Al; Riedl, John** (1999). "An algorithmic framework for performing collaborative filtering". *ACM Sigir*, v. 22, pp. 230-237.
<https://doi.org/10.1145/312624.312682>
- Jones, Bronwyn; Jones, Rhianne** (2019). "Public service chatbots: Automating conversation with *BBC News*". *Digital journalism*, v. 7, n. 8.
<https://doi.org/10.1080/21670811.2019.1609371>
- Kunert, Jessica; Thurman, Neil** (2019). "The form of content personalisation at mainstream, transatlantic news outlets". *Journalism practice*, v. 13, n. 7, pp. 759-780.
<https://doi.org/10.1080/17512786.2019.1567271>
- Möller, Judith; Trilling, Damian; Helberger, Natali; Van-Es, Bram** (2018). "Do not blame it on the algorithm: an empirical assessment of multiple recommender systems and their impact on content diversity". *Information, communication & society*, v. 21, n. 7, pp. 959-977.
<https://doi.org/10.1080/1369118X.2018.1444076>
- Napoli, Philip M.** (2011). "Exposure diversity reconsidered". *Journal of information policy*, v. 1, pp. 246-259.
<https://doi.org/10.5325/jinfopoli.1.2011.0246>
- Newman, Nic** (2021). *Journalism, media, and technology trends and predictions 2021*. UK: The Reuters Institute for the Study of Journalism.
https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2021-01/Newman_Predictions_2021_FINAL.pdf
- Pariser, Eli** (2011). *The filter bubble: what the Internet is hiding from you*. London: Penguin Books. ISBN: 978 0 241954522
- Pöchhacker, Nikolaus; Burkhardt, Marcus; Geipel, Andrea; Passoth, Jan-Hendrik** (2017). "Interventionen in die produktion algorithmischer öffentlichkeiten: Recommender systeme als herausforderung für öffentlich-rechtliche sendeanstalten". *Kommunikation & gesellschaft*, v. 18.
<https://nbn-resolving.org/urn:nbn:de:0168-ssoa-51500-9>
- 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>

Schmidt, Jan-Hinrik; Sørensen, Jannick-Kirk; Dreyer, Stephan; Hasebrink, Uwe (2018). “Wie können empfehlungssysteme zur vielfalt von medieninhalten beitragen? Perspektiven für öffentlich-rechtliche rundfunkanstalten”. *Media Perspektiven*, v. 11, pp. 522-531.
https://www.ard-werbung.de/fileadmin/user_upload/media-perspektiven/pdf/2018/1118_Schmidt_Soerensen_Dreyer_Hasebrink.pdf

Sørensen, Jannick-Kirk (2013). “PSB goes personal: The failure of personalised PSB web pages”. *MedieKultur*, v. 29, n. 55, pp. 43-71.
<https://doi.org/10.7146/mediekultur.v29i55.7993>

Sørensen, Jannick-Kirk (2019). *Public service media, diversity and algorithmic recommendation: Tensions between editorial principles and algorithms in European PSM Organizations*. INRA.
<https://pdfs.semanticscholar.org/0d63/5ba73cf3a0ad7a036692f9bb94242499248f.pdf>

Sørensen, Jannick-Kirk; Van-den-Bulck, Hilde (2018). “Public service media online, advertising and the third-party user data business: A trade versus trust dilemma?”. *Convergence: The international journal of research into new media technologies*, v. 26, n. 2, pp. 421-447.
<https://doi.org/10.1177/1354856518790203>

Thurman, Neil; Lewis, Seth C.; Kunert, Jessica (2019). “Algorithms, automation, and news”. *Digital journalism*, v. 7, n. 8, pp. 980-992.
<https://doi.org/10.1080/21670811.2019.1685395>

Unesco (2001). *Public broadcasting why? How? Technical report*. Paris: Unesco.
<http://unesdoc.unesco.org/images/0012/001240/124058eo.pdf>

Van-den-Bulck, Hilde; Moe, Hallvard (2017). “Public service media, universality and personalisation through algorithms: mapping strategies and exploring dilemmas”. *Media, culture & society*, v. 40, n. 6.
<https://doi.org/10.1177/0163443717734407>

YLEisradio (2018). *The first of its kind in the world: YLE NewsWatch’s smart Voitto assistant shows recommendations directly on the lock screen*. Helsinki: YLE.
<https://yle.fi/aihe/artikkeli/2018/10/12/the-first-of-its-kind-in-the-world-yle-newswatchs-smart-voitto-assistant-shows>

Zuiderveen-Borgesius, Frederik; Trilling, Damian; Moeller, Judith; Bodó, Balázs; De-Vreese, Claes H.; Helberger, Natali (2016). “Should we worry about filter bubbles?”. *Internet policy review. Journal on internet regulation*, v. 5, n. 1.
<https://ssrn.com/abstract=2758126>

Anuario ThinkEPI

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



Semantic similarity models for automated fact-checking: *ClaimCheck* as a claim matching tool

Irene Larraz; Rubén Míguez; Francesca Sallicati

Nota: Este artículo se puede leer en español en:
<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87284>

Recommended citation:

Larraz, Irene; Míguez, Rubén; Sallicati, Francesca (2023). "Semantic similarity models for automated fact-checking: *ClaimCheck* as a claim matching tool". *Profesional de la información*, v. 32, n. 3, e320321.

<https://doi.org/10.3145/epi.2023.may.21>

Manuscript received on 15th February 2023

Accepted on 18th May 2023



Irene Larraz ✉
<https://orcid.org/0000-0002-9164-6610>

Universidad de Navarra
Campus Universitario
31009 Pamplona, Spain
ilarraz@alumni.unav.es



Rubén Míguez
<https://orcid.org/0000-0001-9395-1849>

Newtral
Vandergoten, 1
28014 Madrid, Spain
ruben.miguez@newtral.es



Francesca Sallicati
<https://orcid.org/0000-0001-9981-8360>

Newtral
Vandergoten, 1
28014 Madrid, Spain
francesca.sallicati@gmail.com

Abstract

This article presents the experimental design of *ClaimCheck*, an artificial intelligence tool for detecting repeated falsehoods in political discourse using a semantic similarity model developed by the fact-checking organization *Newtral* in collaboration with *ABC Australia*. The study reviews the state of the art in algorithmic fact-checking and proposes a definition of claim matching. Additionally, it outlines the scheme for annotating similar sentences and presents the results of experiments conducted with the tool.

Keywords

Verification; Automated fact-checking; Claim matching; Semantic similarity; Paraphrase models; Disinformation; Artificial intelligence; AI; Algorithms; Software.

Funding

This publication is part of the R&D&I project "Digital native media in Spain: typologies, audiences, building trust and keys to journalistic sustainability" - *Diginativemedia*. Reference PID2021-122534OB-C22, funded by the *Ministry of Science and Innovation/State Research Agency/10.13039/501100011033/* and the *European Regional Development Fund "A way of doing Europe"*.

1. Introduction

In 2020, *The Washington Post* identified more than 50 lies that former US President Donald Trump had repeated at least 20 times during his term in office. Some of them had been uttered more than 200 times (Kessler; Fox, 2021). Two years earlier, the media had to add a new category to rate such lies repeated more than 20 times. The head of fact-checking at the media outlet, Glenn Kessler, explained that this figure is sufficient to demonstrate that, far from being a mistake, there is an intention to deceive, which ultimately turns into propaganda (*The Washington Post*, 2018).



Fact-checkers dedicate considerable time and effort to combat misinformation, providing verified information to the public discourse and increasing the political cost of lying. Therefore, identifying repeated lies has become a priority task in political discourse fact-checking.

In recent years, advancements in the field of artificial intelligence and language models through deep learning have propelled the development of automated solutions to detect these repetitions. The aim is to assist journalists in enhancing their ability to monitor politicians' statements, increase the reach of fact-checking, and provide faster responses to audiences. However, the use of natural language processing (NLP) techniques has encountered various challenges in identifying similar phrases, particularly when pronounced with different words or expressions. Consequently, most models designed for this purpose are not yet operational in newsrooms.

Fact-checking organizations have also begun developing their own tools to automate claim matching, which refers to the task of identifying phrases with the same meaning. In the context of fact-checking, this concept entails the process of comparing a claim with previously verified ones to determine if there is any match or similarity.

The UK-based fact-checking organization *Full Fact* is working on a tool to review whether a politician's claim resembles another that they have already verified, comparing it against their archive of publications and alerting journalists when a repetition is detected (Corney, 2021a; Floodpage, 2021). Similarly, the *Duke Reporters' Lab* at *Duke University* is experimenting with *Squash*, a tool that allows for finding similar claims among fact-checks from various US media outlets to provide live alerts during electoral debates (Adair, 2021).

The purpose of this study is to address the challenge of detecting repeated false information by designing an experiment to validate a potential solution: the reuse of previous fact-checks and the improvement of semantic similarity systems to search for linguistic repetitions in political discourse. The central hypothesis is that the use of semantic similarity systems can significantly enhance the automation of fact-checking by identifying false claims that utilize paraphrasing to express the same idea. This article provides a detailed description of the experimental methodology and the obtained results, as well as their implications for future research on detecting false information in the political domain.

Intending to improve the precision of detecting similar phrases, this study conducts a comparative analysis of three architectures by evaluating their performance in creating *ClaimCheck*. This tool was developed jointly by the fact-checking organization *Newtral* and *ABC Australia*, as part of the *JournalismAI* program coordinated by the *London School of Economics*. The system was designed to meet the specific needs of journalists and is the result of collaboration between engineers and fact-checkers. Furthermore, the model has been tested in a newsroom environment, where it has been used to continuously improve its results.

Additionally, the study reviews the state of the art in the application of such algorithms for fact-checking (Section 2), proposes a definition of claim matching (Section 3), details the methodology of the experimental design, including an analysis of the challenges faced by these models (Section 4), and conducts an experiment to compare three potential architectures (Section 5). Finally, it presents an analysis of the tool's implementation in the newsroom and suggests future research directions to enhance claim matching and contribute to the automation of fact-checking through artificial intelligence (Section 6).

2. State of the art

The academic literature has extensively documented the need to address repeated lies and the problems generated by the repetition of misinformation, as well as the negative effects of amplifying false messages (Phillips, 2018; Wardle, 2018).

Babakar and Moy (2016) emphasize that public discourse heavily relies on repetition, making it a crucial aspect of the roadmap for automated fact-checking and a challenge that researchers still need to address. **Graves** (2018) also highlights the detection of repeated falsehoods as a fundamental element of automated fact-checking and suggests that the most effective method for these tools is to compare statements with a library of previously verified claims by one or multiple fact-checking organizations to enhance their reach and responsiveness to false assertions.

Thorne and Vlachos (2018) also refer to claim matching as a popular model employed by fact-checking organizations, wherein a claim is compared with previously verified ones. Similarly, **Nakov et al.** (2021) provide an overview of the technological advancements available for automated fact-checking and the detection of previously verified claims as one of the primary approaches. They also emphasize the role of context in this task, including the use of neighboring phrases, coreference resolution, and reasoning about the target text.

The effects of repeated lies have also been examined in other fields, such as Psychology, which has investigated what is known as the illusory truth effect. This effect occurs when the repetition of a deception creates familiarity in the listener, reducing doubts about its truthfulness and leading to the belief that what is being said is

“ We understand claim matching as the task of identifying statements that share a common meaning, even if they are expressed in different ways ”

true (Hassan; Barber, 2021; Murray *et al.*, 2020; Agadjanian *et al.*, 2019). Some of these studies have also attempted to measure the likelihood of politicians repeating a claim after it has been fact-checked. For instance, one study found that the chances of a claim being repeated in the five days following a fact-check publication decreased by 9.5 percentage points (Lim, 2018).

Scientific evidence has also analyzed the various political responses that arise from fact-checking. In some cases, it has been found that politicians continue to cling to falsehoods even after being debunked with compelling evidence against their integrity. This can be attributed to reasons ranging from disagreement to demagogic strategies (Sippitt, 2020; Porter; Wood, 2021).

On the other hand, in the field of computer science, claim matching has emerged as one of the primary benchmarks for NLP (Nakov *et al.*, 2022). Although paraphrasing and semantic similarity models still pose unresolved challenges, advancements in language models, specifically large language models (LLMs), have brought new perspectives to enhance the accuracy of these methods. Specifically, the domain of Semantic Textual Similarity (STS) focuses on the task of measuring the similarity in meaning between sentences. In this regard, models assess the extent to which two claims are semantically, textually, lexically, and referentially similar (Hövelmeyer; Boland; Dietze, 2022; Martín *et al.*, 2021; Sheng *et al.*, 2021).

To train these models, researchers work with datasets and pre-trained models such as *SemEval-PIT2015* (en), the *STS Benchmark*, or the *Microsoft Research Paraphrase Corpus*, among others (Dolan; Brockett, 2005; Lan *et al.*, 2019).

Researchers have attempted to address some of the challenges faced by the technology, such as the ambiguity of claims and the need for more contextual information (Shaar *et al.*, 2021a; 2021b; Reimers; Gurevych, 2019; Nguyen; Karimi; Xing, 2021). Others have taken different approaches, for example, framing it as a classification problem over a collection of previously verified claims (Mansour; Elsayed; Al-Ali, 2022). Kazemi *et al.* (2022) also propose a model based on a binary classifier using *XLM-RoBERTa* (*XLM-R*), a popular multilingual language model, and a semantic similarity search system utilizing sentence embeddings from *LaBSE*, *SBERT*, and pairwise cosine similarity.

To achieve this, various paths have been explored. For instance, the *University of Texas at Arlington* explores structured and semantic models that can capture multiple aspects of a factual claim, such as the topic, the type of expressed data, the involved entities and their relationships, quantities, times, intervals, comparisons, and aggregated structures (Arslan, 2021). Another example is the *Berkeley FrameNet* system (Baker; Fillmore; Lowe, 1998), which relies on frame semantics, a branch of meaning theory, with 13 categories of factual claims. Currently, the tools are designed as hybrid systems in which the judgment of the journalist is necessary, and the technology is considered a support to enhance search capabilities and improve response speed when a falsehood resurfaces.

Within the broader framework of studies at the intersection of Journalism and artificial intelligence, fact-checking has been the subject of extensive analysis due to its potential for automation (*Stanford Institute for Human-Centered Artificial Intelligence*, 2023). Despite persistent challenges in achieving full automation of fact-checking, such as the need to improve the accuracy of natural language processing tools and a more comprehensive database of verified facts, there is a growing number of initiatives that leverage AI to perform this task.

In the field of automated fact-checking, claim matching is intertwined with other tasks, such as claim detection, which involves identifying factual claims within political discourse; check-worthiness, which assesses the relevance of a claim to prioritize verifications; and claim validation, which involves corroborating the data to confirm the veracity of a claim by seeking evidence from open data sources (Hassan *et al.*, 2015; 2017; Graves, 2018; Zeng; Abumansour; Zubiaga, 2021). These tasks have driven the development of artificial intelligence applied to fact-checking, an area that has attracted significant attention due to its distinctive features, such as the structure of fact-checks or the methodology employed (Nakov *et al.*, 2021).

3. Towards a definition of claim matching

Revised definitions identify claim matching as a task aimed at detecting pairs of statements whose meaning coincides, even if the words or grammatical structures used to convey that meaning differ. However, although there is a general consensus on the nature of this task, there is no universal definition that precisely delineates the concept.

Full Fact describes claim matching for fact-checking purposes as a task to identify statements with a shared truth condition (Corney, 2021b). In other words, similar phrases are considered if both contain a variable that makes them true. This means that there cannot be one certain statement and another that is not. For example, they cite the statement “It is raining in London,” which is true under the condition that it is raining in that city. A similar phrase could be “It is damp outside” or “It is pouring.”

On the other hand, Kazemi *et al.* (2021) define claim matching as the task of identifying pairs of textual messages containing statements that can be addressed with the same fact-check. From a more technical perspective, Shaar *et al.* (2020) reflect a similar idea, framing the task as a classification problem to find verifications that can help debunk the initial claim. This implies that the fact-check contains the new statement. The work of Adair *et al.* (2018) follows the same line of thought.

Jiang *et al.* (2021) provide a broader definition that includes

“messages containing falsehoods, inaccuracies, rumors, decontextualized truths, or deceptive logical leaps that convey similar information/themes as the claim but, for example, with a different name of the person or event (...).”

For these authors, it is not only about identifying phrases with a common meaning but also grouping statements that share a theme or information, even if they are presented differently or use different names or events. This expansion of the definition highlights the importance of considering the broader context in which they are presented, such as when a politician repeats the same data but changes the reference to the city in which it is being uttered.

In summary, our proposal is based on the definitions provided by the aforementioned authors and aims to encompass the common factors among them. Thus, we understand claim matching as the task of identifying statements that share a common meaning, even if they are expressed in different ways.

4. Experimental design methodology

The experimental design of the claim matching model was based on an iterative process that involved the evaluation and comparison of various architectures to determine their effectiveness. To carry out this task, a comprehensive analysis of previously published scientific literature in related studies was conducted, and consultations were made with experts in the field of data verification. Based on this knowledge, empirical evaluation of the models was performed through testing using a training dataset that contained phrases with varying degrees of similarity.

In doing so, the work of Dolan and Brockett (2005) on the *Microsoft Research Paraphrase Corpus* and the study by Kazemi *et al.* (2021) served as references. From this corpus, an annotation guideline was developed, which defined the criteria for labeling pairs of statements.

Table 1. Examples of annotated statements and reasoning for the decision

Claim 1	Claim 2	Tag	Reason
More than 4,100 arrested for gender-based violence since the start of the lockdown	Nearly 9,000 arrested for gender-based violence during the state of emergency	Similar	Both messages contain a figure regarding the number of arrests during the state of emergency.
The government inflates the data of tests conducted to boast that we are in the 'top ten' worldwide.	The government falsifies the data it sends to the OECD to be in the top 10 of tests conducted.	Similar	Both statements point to manipulated data by the government regarding the number of tests conducted to appear "in the top 10".
20% of the infections in Spain are among healthcare personnel.	For the first time, more than 20 percent of the infected population consists of healthcare workers.	Similar	There are three shared elements: the percentage, and references to infections and healthcare personnel.
"Gender-based violence is the most common crime during the state of alarm."	The pandemic of sexism: Aggressions against minors or people close to women increase as a form of sexist violence, according to the <i>Prosecutor's Office</i> .	Dissimilar	Both refer to gender-based or sexist violence during the lockdown, however, one focuses on the frequency of the crime while the other addresses where it occurs.
282,891 more unemployed, and we must add the million self-employed who have ceased their activity and the 3.3 million Spaniards in ERTE*.	To boast about management with over 70,000 deaths, nearly four million unemployed, 750,000 in ERTE*, and a million self-employed on the brink can only be done by someone living in La Moncloa, oblivious to the problems of the Spanish people.	Similar	Although both statements provide multiple data points, there is one that coincides in both regarding the number of self-employed individuals who have lost their activity.

* Employment Regulation File

The training data was obtained from a database of 200,000 factual statements provided by *Newtral*. For each sentence, the three most similar candidates were selected based on cosine similarity scores, a measure used to assess the similarity between two entities in a vector space. From these candidates, sentence pairs were extracted for annotation based on threshold values predefined in a similar dataset (Dolan; Brockett, 2005). According to these thresholds, the following were selected:

- 50% of pairs with cosine similarity ≥ 0.8
- 25% of pairs with cosine similarity < 0.8 and ≥ 0.7
- 25% of pairs with cosine similarity < 0.7 and ≥ 0.4

The objective was to obtain a training dataset consisting of 10,000 annotated pairs of statements, including both manually annotated pairs and pairs with weak labels obtained through heuristics. The heuristics involved applying rules to identify common features in factual statements, such as the presence of numbers, percentages, geographic names, and other indicators of factuality. While labels obtained through heuristics are less precise than manually obtained labels, they allow for a significant expansion of the sample size and, consequently, improve the effectiveness of the training process.

Once the dataset was obtained, labels or similarity parameters were assigned to the sentences, and annotation criteria were defined. Different approaches were considered in this process, including assigning numerical values for measuring similarity or classifying similarity into distinct categories.

“ The experimental design of the claim matching model was based on an iterative process that involved the evaluation and comparison of various architectures to determine their effectiveness ”

Firstly, an evaluation was conducted using a preliminary definition of criteria on 100 pairs of statements, with three different annotators labeling the pairs of sentences into three categories: similar, dissimilar, or related. We also included a “N/A” mark in case the proposed sentence was not factual. The results revealed a high discrepancy, with a 50% difference among the annotators. Out of those sentences, half showed a complete disagreement among the annotators, as each one applied a different label. Subsequently, the three annotators conducted a joint review to align the annotation criteria. A second trial was prepared, and this time an improvement in the alignment of the annotations was observed. Although a 30% discrepancy persisted, only one annotator marked a different label on this occasion. Finally, the annotation criteria were reviewed, and it was agreed to eliminate one of the labels (‘related’) to avoid noise. With this new approach, the data were reannotated.

4.1. Confining degrees of similarity: a proposed framework for data annotation

Confining the definition of claim matching entails identifying what is considered a similar sentence. Several factors come into play in this step. The first one arises from language usage issues, such as the use of pronouns like ‘he’ or temporal and spatial references like ‘yesterday’ or ‘here’. An extreme example of this is how one can refer to the same person using their name, surname, initials, position, or former position, among many others, which makes it challenging for automated systems to identify them. This problem falls within a broader issue, that of the ambiguity of some statements.

In addition to the mentioned factors, another essential component for deciphering the meaning of a statement and determining its similarity to others is context. This element is crucial in understanding short statements where the context is difficult to extrapolate (Shaar *et al.*, 2021b). Particularly in oral language statements, necessary details are often omitted to understand what is being referred to. The conversation’s context or the speaker’s context may contain relevant information that algorithms find difficult to deduce otherwise. For instance, if it is uttered by a regional politician, it likely refers to that specific region, whereas if the same statement is made by a national politician, it extends to a national scope.

The data itself poses another obstacle to retrieving previously verified sentences, as the same data can be presented in different ways. In some cases, politicians may use rounding or switch from absolute numbers to relative values, which can lead to interpretation errors. Fact-checkers have pointed out that in many cases, the idea politicians seek to support is more relevant than the data itself. An example of this is the statement made by some members of the *Popular Party* in Spain regarding the number of unemployed individuals in the country, where a variety of data points were used, exceeding the actual figure. In total, different party members had repeated at least 19 times that there were more millions of unemployed individuals than those actually recorded by the *National Institute of Statistics* (Real, 2021). However, the figures mentioned by the political representatives ranged from four to six million unemployed, making it difficult for a system that only identified the repetition of figures to detect this discrepancy.

Taking the above into consideration, for the development of *ClaimCheck*, two sentences were considered similar if they met the following criteria:

- They refer to the same thing, even if the data varies or is contradictory.
- One of the statements includes specific details that do not invalidate the other.
- One of the statements refers to the same reality from a different perspective than the other; for example, one provides the total quantity while the other refers to the percentage variation, but both convey the same message.

According to the established criteria, the data annotation stage used to train a claim matching model is a critical aspect for the success of its implementation, as the selected parameters and labels must accurately reflect the similarity between the statements. In the case of *ClaimCheck*, a binary classification with two categories was chosen in order to simplify the classification task and obtain more precise results. It is important to emphasize that the selection of categorization should be carefully evaluated to avoid classification errors and ensure the quality of the annotations made.

4.2. Issues in fact-checks retrieval

Beyond what is considered similar in the annotation phase, there are problems arising from its specific application for fact-checking in order to retrieve previously published fact-checks. In these cases, other factors also come into play, such as temporality. A statement that was false at a certain moment may become true over time, or vice versa. This problem arises when retrieving fact-checks whose data may have been modified over time due to updated statistics and fluctuating values, as well as variations in the context in which the data is used. In other words, Statement A and Statement B are similar, and the similarity between them persists over time, but it may not be possible to use Statement A to retrieve a fact-check with Statement B if that verification has “expired,” even though it does not affect the underlying concept or definition of similarity between them.

Nuances are another crucial element in identifying two similar statements in the field of fact-checking, as the introduction of a single word can completely change the meaning of the statement, altering the rating or classification given to it before, shifting from true to false or other categories. For example, the statement “50% of new contracts are indefinite” was true at a certain moment if it referred to those contracts generated from a specific date, in this case, following labor reform. However, if the word “new” is removed and the totality of contracts is analyzed, it would be false to claim that “50% of contracts are indefinite.” Again, in this case, both statements are similar according to the criteria established in the experimental design, but it would create problems to retrieve a fact-check to assess the other.

“ Finding a balance between system precision and its ability to retrieve all similar phrases is a challenging task on claim matching ”

The same occurs with the context in the specific case of fact-checking. Referring to the previous example, it is not the same for the statement “50% of new contracts are indefinite” to be made by a minister, who is referring to the national level, as another statement indicating that “50% of the contracts we have made are indefinite” said by a regional president, who generally refers to the regional scope of their community. Both statements are semantically similar, but the context implies that they are referring to different data related to different locations. These types of problems present challenges when interpreting the results generated by the algorithm.

5. Method and experiments on *ClaimCheck*

ClaimCheck is built on the foundation of *ClaimHunter*, a claim detection tool that identifies factual statements in *Twitter* messages (Beltrán; Míguez; Larraz, 2019). These are different algorithms applied in different phases of the process. While the latter focuses on the detection of statements to be verified, the former works once that filtering is done and seeks similar cases that have been previously verified. Thus, to process each statement, political discourse first goes through the *ClaimHunter* algorithm to confirm its factual nature and then through the *ClaimCheck* algorithm, which is responsible for retrieving candidate phrases to be labeled as similar.

The first step is to identify what is used as the archive, whether it includes only the dataset of previous fact-checks or also other statements from politicians on social media or in the media. *ClaimCheck* utilizes both sources to achieve two objectives: on one hand, it allows the reuse of previously conducted verifications for faster action, and on the other hand, it helps identify disinformation campaigns when multiple statements articulated in different ways convey the same false message. To do this, *ClaimCheck* collects fact-checks from the *Fact Check Explorer*, a *Google* tool that stores fact-checkers’ publications using the *ClaimReview* markup. The *ClaimReview* markup is a structured data system that facilitates the extraction of categories within a verification, such as the claim being verified, the author of the statement, and the rating indicating the degree of falsehood. Additionally, *ClaimCheck* also utilizes verifications from other fact-checkers not included in the *Google* tool. As a result, it has a database of 300,000 fact-checks in over 20 different languages, which in the future could expand its scope beyond political discourse fact-checking to include the debunking of disinformation circulating on social media.

Next, we detail each of the phases: the construction of the training and test datasets, the design of the architectures, and the analysis of the experimental results.

5.1. Construction of the training and test datasets

To train the model, *ClaimCheck* required identifying valid candidates, i.e., detecting phrases in the fact-checking database that had the potential to convey the same meaning as the input phrase. In this process, the aforementioned problems of semantic similarity arise (see Section 4.2), as the system searches for similar phrases. Additionally, the *ClaimCheck* system needed to evaluate whether the two phrases refer exactly to the same thing, which poses a problem of semantic meaning.

The *ClaimCheck* system, therefore, consists of two components: a first component that performs a search for similar phrases, and a classifier that evaluates the similarity between the original claim and the proposed candidates.

To construct the training dataset, annotations from the *Fact Check Explorer* were used. On the other hand, pairs were selected from the *ClaimHunter* dataset to construct the test dataset (test benchmark), which was used to evaluate the classification model.

This led to the adoption of two strategies. On one hand, phrases were extracted from tweets that could contain more than one sentence, as well as complete tweets that included additional information. This extraction of phrases was performed by applying tokenization functions to obtain the different phrases that compose each tweet, and they were then passed through *ClaimHunter* to discard non-factual phrases. Specifically, three types of pairs were generated from the *ClaimHunter* database:

- Tweet-to-tweet pairs
- Phrase-to-tweet pairs
- Phrase-to-phrase pairs

All these possible combinations of phrase pairs underwent cosine similarity filters using the sentence-transformer library and the *sentence-transformers/paraphrase-multilingual-MiniLM-L12-v2* model. Specifically, pairs with a cosine similarity below 0.7 were discarded, and a subset was selected for annotation. In total, the training dataset consisted of 7,762 pairs, with 66% being similar claims, while the test (*ClaimHunter*) benchmark comprised 2,246 pairs with a 45% similarity. Additionally, the *Prodigy* tool was used for annotation purposes.

Table 2. Datasets used for the development of *ClaimCheck*

Dataset	Source	Annotation	Similar pairs	Manually annotated pairs	Total pairs
Training	<i>Fact Check Explorer</i> and others	Similar	66%	46% *	7,760
Test	<i>ClaimHunter Benchmark</i>	Similar	45%	100%	2,246
	<i>Prodigy Benchmark</i>	Similar	31%	100%	7,443

*The other 54% are weak labels.

5.2. Experimental design

The experimentation with the design of the architectures took place between July and October 2022, carried out by a team of three researchers who implemented the prototyping proposals. For this purpose, a protocol was developed for data collection, and three potential experimental scenarios with their respective prototypes were designed. Each architecture presented below represents a different experiment, approaching the steps of experimentation in its own manner.

5.2.1. Classifier

The first approach involved retrieving information using standard *ElasticSearch* methods, a system that utilizes keywords and rules to search for the top K suitable candidates, and then filtering them using a classifier. Specifically, a *BERT*-like classifier was employed to detect pairs of similar sentences.

5.2.2. Semantic search + threshold

The second strategy also relied on information retrieval, but this time using a semantic search approach with *K-Nearest Neighbors (KNN)* implemented with *OpenSearch*. *KNN* is a type of supervised learning algorithm used for both regression and classification. Regular search with *ElasticSearch* to find matches is useful for general queries, but *KNN*-based search is more suitable for specific searches. This is because in *KNN*-based search, language features known as embeddings or ‘vector representations’ are extracted. The term ‘embedding’ is used in natural language processing to refer to the technique of representing words or phrases as numerical vectors, enabling measurement of proximity between them. Similarity is established based on how close these values are in the vector space.

Mukherjee, Sela and Al-Saadoon (2020) provide the following example: when searching for a wedding dress using a *KNN*-based search application, it produces similar results if you type “wedding dress” or “bridal gown.” Thus, “summer dress” and “floral summer dress” are considered similar due to the proximity between their embeddings, unlike “summer dress” and “wedding dress.”

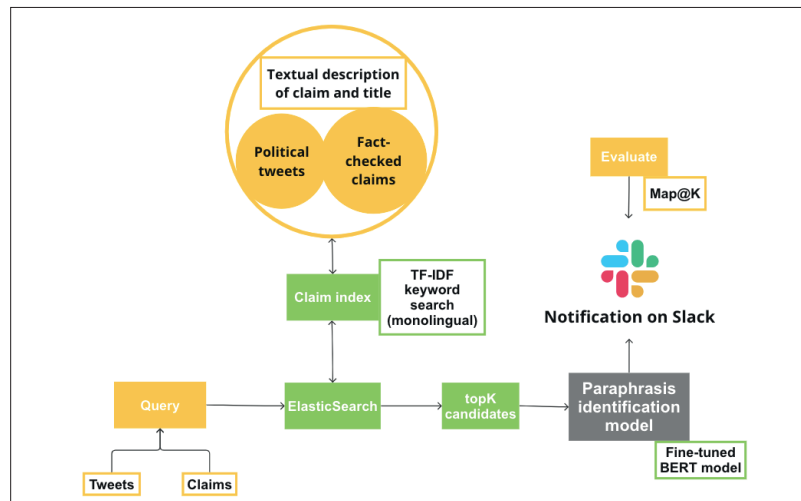


Figure 1. The architecture of the model with a classifier

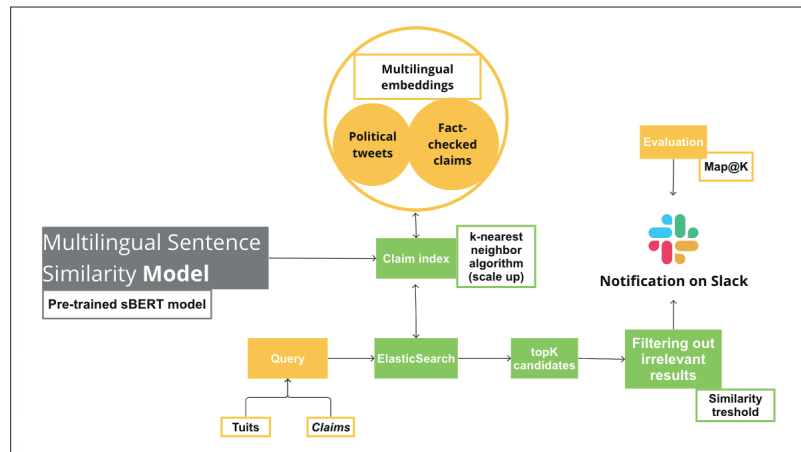


Figure 2. The architecture of the model with a semantic search system and a threshold

To retrieve the most similar or vectorially closest candidate phrases, a cosine similarity threshold is set to filter out irrelevant candidates. This concept is used to establish a critical value or cutoff point from which a decision or classification is made. The challenge lies in defining an appropriate similarity threshold for our solution. In this case, instead of using an AI-based classifier like in point 1, we simply used the similarity threshold as the classifier. With this approach, we consider the retrieval model to be good enough so that anything exceeding the threshold would be classified as a similar phrase.

5.2.3. Semantic search + classifier

The third proposal involved combining two artificial intelligence models:

- one for generating vector representations in the pre-trained semantic search process to retrieve candidates, and
- another model that acts as a binary classifier to identify paraphrases among the retrieved candidates, i.e., to determine whether the two phrases actually convey the same meaning or not.

Similar to the previous experiment, the first step was to train a pre-trained semantic search model, and in the second step, train a binary paraphrase classifier model using the labeled training dataset specifically created for this purpose. Both models were integrated into a prototype that allows for semantic search first and, in a subsequent phase, identifies paraphrases using the classification system.

As shown in Figure 3, the prototype is based on an embedding generator using politicians' statements on *Twitter* and the constructed fact-checking archive from previously published fact-checks (*Claim index*). From new tweets and statements (*Query*), a search is performed (*ElasticSearch*) on the embeddings, and the most accurate results (*topK* candidates) are retrieved. To filter the *topK* candidates, the paraphrase model is utilized, which selects the results that yield better matching (Paraphrase identification model), and sends an alert to the *Slack* messaging program.

The feedback from journalists on *Slack*, where they label whether the candidates selected by the tool are similar or not, enables an evaluation of the model in the real world using the mean average precision at different values of *K* (*MAP@K*).

To determine the best option, we conducted an evaluation using *MAP@K* for each of the three architectures. Specifically, this involves quantitatively assessing the percentage of *topK* candidates retrieved, which represents the system's precision. On an individual level, in the first and third cases, the evaluation focuses on the classifier's precision (percentage of correctly classified records), recall (percentage of similar records that are retrieved), and F1 score (a metric that combines the two previous metrics).

5.3. Results of the experiments

The third strategy (semantic search + classifier) proved to be the most successful for our use case. Regarding the similarity classification model, we conducted several tests to choose one of the pre-trained versions of *BERT*-like models offered by *Huggingface*, which is one of the most popular and widely used libraries in the field of natural language processing. Among these tests, the models with the best performance were *microsoft/Multilingual-MiniLM-L12-v2* and *xlm-roberta-base*.

Both models performed well in terms of precision and recall, but we chose the *microsoft/Multilingual-MiniLM-L12-v2* model because it is lighter, making it more suitable for use in the editorial workflow.

The *sentence-transformers/paraphrase-multilingual-MiniLM-L12-v2* model was used in the experiments to test candidate retrieval. The effectiveness of the model was evaluated by adding an additional filter using the *microsoft/Multilingual-MiniLM-L12-H384* model. Three commonly used metrics in information retrieval systems were used to measure performance: *MAP@K*, *Recall@K*, and *Mean Reciprocal Ranking*. The goal was to evaluate the ability of each model to retrieve the most relevant results according to the user's search criteria.

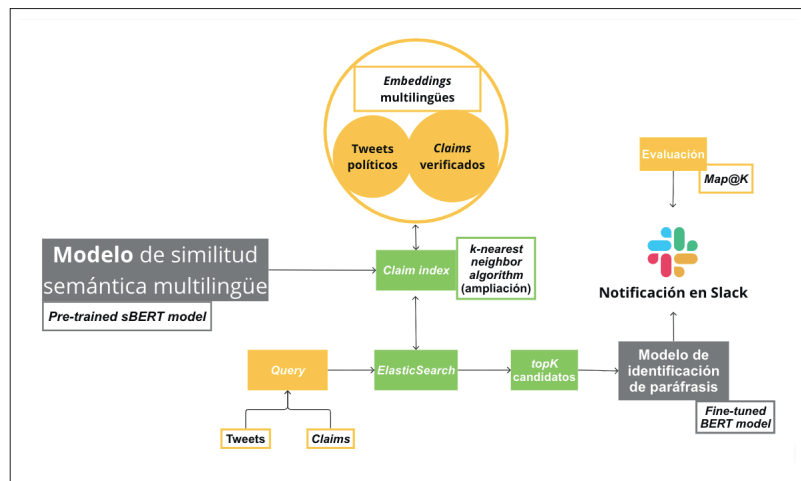


Figure 3. The architecture of the model with a semantic search system and classifier

“ The combination of semantic search systems and classifiers for retrieving previous fact-checks can enhance the efficiency and effectiveness of claim matching, enabling to detect previously verified claims more quickly and accurately ”

Table 3. Preliminary results of the trial with different pre-trained models

Models - 2checks Training	Train set	Thres-hold	Set	Similarity %	Precision	Recall	F1-score	Accuracy
<i>microsoft/Multilingual-MiniLM-L12-H384</i>	2checks (66% similar)	0.5	Test MRPC	66%	67.32%	98.86%	80.10%	67.35%
			Test Benchmark - ClaimHunter	45%	79.45%	71.32%	75.16%	78.05%
			Test PAWSX-es	44%	44.79%	99.89%	61.84%	44.82%
<i>xlm-roberta-base (bce - preprocess tweets)</i>	2checks (66% similar)	0.5	Test MRPC	66%	67.36%	99.21%	80.24%	67.52%
			Test Benchmark - ClaimHunter	45%	80.88%	69.98%	75.04%	78.32%
			Test Benchmark - ClaimHunter - only Sentences (Set6)	60%	92.14%	50.96%	65.62%	67.58%
			Test PAWSX-es	44%	44.76%	99.89%	61.82%	44.77%
<i>xlm-roberta-base (bce - preprocess assign-labels - current ml commons version)</i>	2checks (66% similar)	0.5	Test MRPC	66%	67.46%	98.77%	80.17%	96.18%
			Test Benchmark - ClaimHunter	45%	79.59%	67.11%	72.82%	76.67%
			Test PAWSX-es	44%	44.79%	100.00%	61.87%	44.82%
<i>xlm-roberta-base (bce - preprocess assign-labels - local)</i>	2checks (66% similar)	0.5	Test MRPC	66%	66.90%	98.77%	79.77%	66.71%
			Test Benchmark - ClaimHunter	45%	78.38%	66.54%	71.98%	75.87%
			Test PAWSX-es	44%	44.79%	100.00%	61.87%	44.82%
<i>xlm-roberta-base (bce - preprocess tweets)</i>	2checks (66% similar)	ROC optimal = 0.331424	Test MRPC	66%	67.35%	98.25%	79.91%	67.17%
			Test Benchmark - ClaimHunter	45%	75.52%	75.81%	75.67%	77.29%
			Test PAWSX-es	44%	44.82%	100.00%	61.89%	44.87%
<i>microsoft/mdeberta-v3-base (bce - preprocess tweets)</i>	2checks (66% similar)	0.5	Test MRPC	66%	66.98%	99.82%	80.17%	67.17%
			Test Benchmark - ClaimHunter	45%	66.44%	92.73%	77.41%	74.80%
			Test PAWSX-es	44%	44.77%	100.00%	61.85%	44.77%
<i>microsoft/mdeberta-v3-base (bce - preprocess-tweets) adjustment of deberta specific parameters (warmup steps, epsilon, weight decay)</i>	2checks (66% similar)	0.5	Test MRPC	66%	67.24%	99.74%	99.69%	67.52%
			Test Benchmark - ClaimHunter	45%	66.17%	92.73%	77.23%	74.53%
			Test PAWSX-es	44%	44.77%	100.00%	61.85%	44.77%
<i>microsoft/mdeberta-v3-base (bce - preprocess-tweets) adjustment of deberta specific parameters (warmup steps, epsilon, weight decay)</i>	2checks (66% similar)	ROC optimal = 0.998581	Test MRPC	66%	67.80%	99.21%	80.55%	68.18%
			Test Benchmark - ClaimHunter	45%	78.55%	80.88%	79.70%	80.81%
			Test PAWSX-es	44%	44.79%	100.00%	61.87%	44.82%
<i>bert-base-multilingual-cased (bce - preprocess-tweets)</i>	2checks (66% similar)	0.5	Test MRPC	66%	66.71%	99.82%	79.97%	66.76%
			Test Benchmark - ClaimHunter	45%	71.76%	76.29%	73.96%	74.98%
			Test PAWSX-es	44%	44.79%	100.00%	61.87%	44.82%
<i>sentence-transformers/paraphrase-multilingual-mpnet-base-v2 (bce - preprocess-tweets)</i>	2checks (66% similar)	0.5	Test MRPC	66%	66.39%	99.30%	79.58%	66.12%
			Test Benchmark - ClaimHunter	45%	56.22%	82.89%	67.03%	61.98%
			Test PAWSX-es	44%	44.79%	100.00%	61.87%	44.82%
<i>sentence-transformers/stsb-xlm-r-multilingual (bce - preprocess-tweets)</i>	2checks (66% similar)	0.5	Test MRPC	66%	67.53%	93.43%	78.40%	65.77%
			Test Benchmark - ClaimHunter	45%	66.97%	69.79%	68.35%	69.90%
			Test PAWSX-es	44%	44.79%	99.89%	61.84%	44.82%
<i>sentence-transformers/paraphrase-xlm-r-multilingual-v1 (bce - preprocess-tweets)</i>	2checks (66% similar)	0.5	Test MRPC	66%	66.82%	90.46%	76.86%	63.80%
			Test Benchmark - ClaimHunter	45%	62.49%	69.12%	65.64%	66.30%
			Test PAWSX-es	44%	44.81%	99.89%	61.87%	44.87%

Tabla 4. Resultados de los ensayos del modelo de clasificación

Semantic Similarity - Classifier Model													
Modelo	Training dataset	Training distribution	Size	Test dataset	Test size	Test distribution	Precision		Recall		F1		Accuracy
							class 0	class 1	class 0	class 1	class 0	class 1	
<i>xlm-roberta-base</i>	Fact Check Explorer y otros	66% similar	7,762	Test Benchmark	2,246	45% similar	77.2%	80%	84.5%	71.3%	80.7%	75.4%	78.4%
				Prodigy Benchmark	7,443	31% similar	85.4%	67%	84.4%	68.8%	84.9%	67.9%	79.5%
<i>microsoft/Multilingual-MiniLM-L12-H384</i>	Fact Check Explorer y otros	66% similar	7,762	Test Benchmark	2,246	45% similar	77%	79.4%	83.9%	71.3%	80.3%	75.2%	78%
				Prodigy Benchmark	7,443	31% similar	86.3%	64.7%	82%	71.8%	84.1%	68.1%	78.8%

Table 5. Average results of the three experiments

Evaluation of candidate retrieval (average of the three architectures)					Evaluation of candidate retrieval (+ filtering model) (average of the three architectures)				
	K = 1	K = 3	K = 5	K = 10		K = 1	K = 3	K = 5	K = 10
MAP@K	0.7471	0.6971	0.6842	0.6864	MAP@K	0.8237	0.7122	0.6837	0.6646
Recall@K	0.3383	0.5615	0.6646	0.7836	Recall@K	0.3638	0.5625	0.6392	0.7148
MRR	0.8528				MRR	0.9169			

By comparing the results of the experiment, it was confirmed that the filtering of the model with the third architecture contributes to retrieving more suitable candidates in the top positions (Table 5). This translates into higher precision of the first retrieved candidates, as can be observed in the higher values of the *MAP@K* metric for smaller values of *K*.

6. Discussion and future research directions

The first finding of this research indicates that finding a balance between system precision and its ability to retrieve all similar phrases is a challenging task. As the retrieval rate increases, the precision in the recommendations provided to fact-checkers decreases, which negatively affects their confidence in the algorithm and the effort required to review the information. Given that, in general, the volume of unrelated phrases is significantly higher than that of similar phrases, the goal is to optimize the system's precision in order to avoid a high incidence of false positives.

Another issue in the process is related to the decrease in search speed as the fact-checks database grows. This occurs if the models used do not scale linearly. For this reason, the choice has been made to select faster models at the expense of heavier models, as long as the performance improvement provided by the latter is marginal, i.e., not exceeding 3%. In real-world situations, where the number of pairs of phrases to compare can reach millions, prioritizing scalability over precision is crucial.

In the experiments, it is also observed that a significant challenge is faced regarding the lack of temporal and spatial context. Therefore, in the future, it is necessary to design a strategy that allows the generation and storage of relevant metadata associated with each tweet and phrase in the system. The candidate retrieval process should consider both the semantic relevance of the phrases and the consistency of their spatial and temporal metadata. To address the temporal issue in the current system, the use of temporal windows is suggested to effectively retrieve the topK valid candidates. Although a simple solution, this strategy proves to be effective in retrieving relevant information within the proposed architecture.

The precise identification of entities (entity linking) represents another challenging task for claim matching systems in the political context. Since the model only has access to the information within the given phrase, it is unable to interpret that terms and expressions such as "Feijóo," "the president of the populares," or "the leader of the opposition" refer to the same person, as discussed in Section 4.2. While in some specific cases, the use of dictionary-based approximations may mitigate this problem, the development of more general solutions is required to address this issue.

7. Conclusions

The results of the study indicate that semantic similarity systems are a valuable tool for detecting paraphrases in political discourse and improving the effectiveness of claim matching models that contribute to automated fact-checking. Specifically, the combination of semantic search systems and classifiers for retrieving previous fact-checks can enhance the efficiency and effectiveness of claim matching, enabling fact-checking organizations to detect previously verified claims more quickly and accurately.

The general conclusion that can be drawn from the experiment results is that filtering the model using this type of architecture can be highly beneficial in improving the precision of the retrieved top candidates. The evidence provided by the experiments shows that by using this technique, more appropriate candidates from similar phrases can be retrieved and placed in the top positions, gaining agility without sacrificing precision, making it a particularly effective option for fact-checking. This result may have important practical implications in various domains.

“ Semantic similarity systems are a valuable tool for detecting paraphrases in political discourse and improving the effectiveness of claim matching models that contribute to automated fact-checking ”

The *ClaimCheck* tool employs an approach based on the *sentence-transformers/paraphrase-multilingual-MiniLM-L12-v2* model, which combines semantic search with a classifier to filter the results obtained through the *microsoft/Multilingual-MiniLM-L12-H384* model. The results obtained by *ClaimCheck* are superior to those of other evaluated models, as evidenced by the metrics of *MAP@K*, *Recall@K*, and *Mean Reciprocal Ranking*.

The experience of *ClaimCheck* demonstrates the most successful experimental paths with positive outcomes for the *Newtral* newsroom, while also highlighting some issues that require resolution to improve the effectiveness of the model. These include the lack of context, precise entity recognition, and the need to balance precision with the retrieval of all similar phrases.

The experimental design presented in this article can serve as a starting point for future research in the field of automated fact-checking and the use of semantic similarity systems in identifying false claims in political discourse. This extends beyond extracting previous fact-checks and enabling their reuse more quickly without duplicating the effort but also extends to real-time fact-checking or creating alert systems against misinformation campaigns. Furthermore, this study suggests potential directions for future research to address the challenges present in the proposed model and improve its adoption by specialized fact-checking organizations.

8. References

- Adair, Bill** (2021). “The lessons of Squash, Duke’s automated fact-checking platform”. *Poynter*, 16 June. <https://www.poynter.org/fact-checking/2021/the-lessons-of-squash-the-first-automated-fact-checking-platform>
- Adair, Bill; Li, Chengkai; Yang, Jun; Yu, Cong** (2018). *Automated pop-up fact-checking: challenges & progress*. <https://ranger.uta.edu/~cli/pubs/2019/popupfactcheck-cj19-adair.pdf>
- Agadjanian, Alexander; Bakhru, Nikita; Chi, Victoria; Greenberg, Devyn; Hollander, Byrne; Hurt, Alexander; Kind, Joseph; Lu, Ray; Ma, Annie; Nyhan, Brendan; Pham, Daniel; Qian, Michael; Tan, Mackinley; Wang, Clara; Wasdahl, Alexander; Woodruff, Alexandra** (2019). “Counting the Pinocchios: the effect of summary fact-checking data on perceived accuracy and favorability of politicians”. *Research & politics*, v. 6, n. 3. <https://doi.org/10.1177/2053168019870351>
- Arslan, Fatma** (2021). *Modeling factual claims with semantic frames: definitions, datasets, tools, and fact-checking applications*. Doctoral dissertation. The University of Texas at Arlington. <https://rc.library.uta.edu/uta-ir/bitstream/handle/10106/30765/ARSLAN-DISSERTATION-2021.pdf>
- Babakar, Mevan; Moy, Will** (2016). *The state of automated factchecking. How to make factchecking dramatically more effective with technology we have now*. Full Fact. https://fullfact.org/media/uploads/full_fact-the_state_of_automated_factchecking_aug_2016.pdf
- Baker, Collin F.; Fillmore, Charles J.; Lowe, John B.** (1998). “The Berkeley FrameNet project”. In: *Proceedings of the joint conference of the international conference on computational linguistics and the Association for Computational Linguistics (Coling-ACL)*, pp. 86-90. <https://aclanthology.org/C98-1013.pdf>
- Beltrán, Javier; Míguez, Rubén; Larraz, Irene** (2019). “ClaimHunter: an unattended tool for automated claim detection on Twitter”. *KnOD@WWW. CEUR workshop proceedings*, v. 2877, n. 3. <https://ceur-ws.org/Vol-2877/paper3.pdf>
- Corney, David** (2021a). “How does automated fact checking work?”. *Full Fact*, 5 July. <https://fullfact.org/blog/2021/jul/how-does-automated-fact-checking-work>
- Corney, David** (2021b). “Towards a common definition of claim matching”. *Full Fact*, 5 October. <https://fullfact.org/blog/2021/oct/towards-common-definition-claim-matching>
- Dolan, William B.; Brockett, Chris** (2005). “Automatically constructing a corpus of sentential paraphrases”. In: *Proceedings of the third international workshop on paraphrasing (IWP2005)*, pp. 9-16. <https://aclanthology.org/I05-5002.pdf>
- Floodpage, Sebastien** (2021). “How fact checkers and Google.org are fighting misinformation”. *Google*, 31 March. <https://blog.google/outreach-initiatives/google-org/fullfact-and-google-fight-misinformation>

- Graves, Lucas** (2018). *Understanding the promise and limits of automated fact-checking*. Reuters Institute for the Study of Journalism. Factsheets.
<https://ora.ox.ac.uk/objects/uuid:f321ff43-05f0-4430-b978-f5f517b73b9b>
- Hassan, Aumyo; Barber, Sarah J.** (2021). "The effects of repetition frequency on the illusory truth effect". *Cognitive research: principles and implications*, v. 6, n. 38.
<https://doi.org/10.1186/s41235-021-00301-5>
- 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>
- Hassan, Naeemul; Arslan, Fatma; Li, Chengkai; Tremayne, Mark** (2017). "Toward automated fact-checking: detecting check-worthy factual claims by ClaimBuster". In: *Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining (KDD '17)*. New York: Association for Computing Machinery, pp. 1803-1812.
<https://doi.org/10.1145/3097983.3098131>
- Hövelmeyer, Alica; Boland, Katarina; Dietze, Stefan** (2022). "SimBa at CheckThat! 2022: lexical and semantic similarity based detection of verified claims in an unsupervised and supervised way". In: *CLEF 2022: Conference and labs of the evaluation forum*, 5-8 September, Bolonia, Italia.
<https://ceur-ws.org/Vol-3180/paper-40.pdf>
- Jiang, Ye; Song, Xingyi; Scarton, Carolina; Aker, Ahmet; Bontcheva, Kalina** (2021). "Categorising fine-to-coarse grained misinformation: an empirical study of Covid-19 Infodemic". *Arxiv*.
<https://doi.org/10.48550/arXiv.2106.11702>
- Kazemi, Ashkan; Garimella, Kiran; Gaffney, Devin; Hale, Scott A.** (2021). "Claim matching beyond English to scale global fact-checking". In: *Proceedings of the 59th Annual meeting of the Association for Computational Linguistics and the 11th International joint conference on natural language processing*. Association for Computational Linguistics, pp. 4504-4517.
<https://doi.org/10.18653/v1/2021.acl-long.347>
- Kazemi, Ashkan; Li, Zehua; Pérez-Rosas, Verónica; Hale, Scott A.; Mihalcea, Rada** (2022). "Matching tweets with applicable fact-checks across languages". *Arxiv*.
<https://doi.org/10.48550/arXiv.2202.07094>
- Kessler, Glenn; Fox, Joe** (2021). "The false claims that Trump keeps repeating". *The Washington Post*, 20 January.
<https://www.washingtonpost.com/graphics/politics/fact-checker-most-repeated-disinformation>
- Lan, Zhenzhong; Chen, Mingda; Goodman, Sebastian; Gimpel, Kevin; Sharma, Piyush; Soricut, Radu** (2020). "ALBERT: a lite Bert for self-supervised learning of language representations". In: *Conference paper at International conference on learning representations (ICLR)*. *Arxiv*.
<https://doi.org/10.48550/arXiv.1909.11942>
- Lim, Chloe** (2018). "Checking how fact-checkers check". *Research & politics*, v. 5, n. 3.
<https://doi.org/10.1177/2053168018786848>
- Mansour, Watheq; Elsayed, Tamer; Al-Ali, Abdulaziz** (2022). "Did I see it before? Detecting previously-checked claims over Twitter". *Lecture notes in computer science*, pp. 367-381.
https://doi.org/10.1007/978-3-030-99736-6_25
- Martín, Alejandro; Huertas-Tato, Javier; Huertas-García, Álvaro; Villar-Rodríguez, Guillermo; Camacho, David** (2021). "FacTeR-check: semi-automated fact-checking through semantic similarity and natural language inference". *Arxiv*.
<https://doi.org/10.48550/arXiv.2110.14532>
- Mukherjee, Amit; Sela, Eitan; Al-Saadoon, Laith** (2020). "Building an NLU-powered search application with Amazon SageMaker and the Amazon opensearch service KNN feature". *Amazon SageMaker, artificial intelligence*, 26 October.
<https://aws.amazon.com/es/blogs/machine-learning/building-an-nlu-powered-search-application-with-amazon-sagemaker-and-the-amazon-es-knn-feature>
- Murray, Samuel; Stanley, Matthew; McPhetres, Jon; Pennycook, Gordon; Seli, Paul** (2020). "'I've said it before and I will say it again...': repeating statements made by Donald Trump increases perceived truthfulness for individuals across the political spectrum". *PsyArXiv preprints*, 15 January.
<https://doi.org/10.31234/osf.io/9evzc>
- Nakov, Preslav; Corney, David; Hasanain, Maram; Alam, Firoj; Elsayed, Tamer; Barrón-Cedeño, Alberto; Papotti, Paolo; Shaar, Shaden; Da-San-Martino, Giovanni** (2021). "Automated fact-checking for assisting human fact-checkers". *International joint conference on artificial intelligence*. *Arxiv*.
<https://doi.org/10.48550/arXiv.2103.07769>

- Nakov, Preslav; Da-San-Martino, Giovanni; Alam, Firoj; Shaar, Shaden; Mubarak, Hamdy; Babulkov, Nikolay** (2022). "Overview of the CLEF-2022 CheckThat! Lab task 2 on detecting previously fact-checked claims". In: *CLEF 2022: conference and labs of the evaluation forum*, 5-8 septiembre, Bolonia, Italia.
<https://ceur-ws.org/Vol-3180/paper-29.pdf>
- Nguyen, Vincent; Karimi, Sarvnaz; Xing, Zhenchang** (2021). "Combining shallow and deep representations for text-pair classification". In: *Proceedings of the 19th Annual workshop of the Australasian Language Technology Association*, pp. 68-78.
<https://aclanthology.org/2021.alt-1.7.pdf>
- Phillips, Whitney** (2018). *The oxygen of amplification. Better practices for reporting on extremists, antagonists, and manipulators online*. Data & Society Research Institute.
https://datasociety.net/wp-content/uploads/2018/05/FULLREPORT_Oxygen_of_Amplification_DS.pdf
- Porter, Ethan; Wood, Thomas J.** (2021). "The global effectiveness of fact-checking: Evidence from simultaneous experiments in Argentina, Nigeria, South Africa, and the United Kingdom". *Proceedings of the National Academy of Sciences of the United States of America*, v. 118, n. 37.
<https://doi.org/10.1073/pnas.2104235118>
- Real, Andrea** (2021). "Casado mezcla diferentes estadísticas de empleo para asegurar que hay 4 millones de parados, pero es falso". *Newtral*, 6 octubre.
<https://www.newtral.es/parados-espana-casado-pp-factcheck/20211007>
- Reimers, Nils; Gurevych, Iryna** (2019). "Sentence-bert: sentence embeddings using siamese bert-networks". In: *Proceedings of the 2019 Conference on empirical methods in natural language processing and the 9th International joint conference on natural language processing (EMNLP-IJCNLP)*. Hong Kong, November, pp. 3982-3992.
<https://doi.org/10.18653/v1/D19-1410>
- Shaar, Shaden; Alam, Firoj; Da-San-Martino, Giovanni; Nakov, Preslav** (2021a). "The role of context in detecting previously fact-checked claims". *Arxiv*.
<https://doi.org/10.48550/arXiv.2104.07423>
- Shaar, Shaden; Babulkov, Nikolay; Da-San-Martino, Giovanni; Nakov, Preslav** (2020). "That is a known lie: detecting previously fact-checked claims". In: *Proceedings of the 58th Annual meeting of the Association for Computational Linguistics*, pp. 3607-3618.
<https://doi.org/10.18653/v1/2020.acl-main.332>
- Shaar, Shaden; Haouari, Fatima; Mansour, Watheq; Hasanain, Maram; Babulkov, Nikolay; Alam, Firoj; Da-San-Martino, Giovanni; Elsayed, Tamer; Nakov, Preslav** (2021b). "Overview of the CLEF-2021 CheckThat! Lab task 2 on detecting previously fact-checked claims in tweets and political debates". In: *CLEF 2021: Conference and labs of the evaluation forum*, 21-24 September, Bucharest, Romania.
<https://ceur-ws.org/Vol-2936/paper-29.pdf>
- Sheng, Qiang; Cao, Juan; Zhang, Xueyao; Li, Xirong; Zhong, Lei** (2021). "Article reranking by memory-enhanced key sentence matching for detecting previously fact-checked claims". In: *Proceedings of the 59th Annual meeting of the Association for Computational Linguistics and the 11th International joint conference on natural language processing (volume 1, Long papers)*.
<https://doi.org/10.18653/v1/2021.acl-long.425>
- Sippitt, Amy** (2020). *What is the impact of fact checkers' work on public figures, institutions and the media?*. Africa Check, Chequeado and Full Fact.
<https://fullfact.org/media/uploads/impact-fact-checkers-public-figures-media.pdf>
- Stanford Institute for Human-Centered Artificial Intelligence** (2023). *Artificial intelligence index*. Stanford University.
<https://aiindex.stanford.edu/report>
- The Washington Post** (2018). "Meet the bottomless Pinocchio | Fact Checker". [Video]. *YouTube*, 10 December.
<https://www.youtube.com/watch?v=zoS1sVZRfUU>
- Thorne, James; Vlachos, Andreas** (2018). "Automated fact checking: task formulations, methods and future directions". *Arxiv*.
<https://doi.org/10.48550/arXiv.1806.07687>
- Wardle, Claire** (2018). "Lessons for reporting in an age of disinformation". *Medium*, 28 December.
<https://medium.com/1st-draft/5-lessons-for-reporting-in-an-age-of-disinformation-9d98f0441722>
- Zeng, Xia; Abumansour, Amani S.; Zubiaga, Arkaitz** (2021). "Automated fact-checking: a survey". *Language and linguistics compass*, v. 15, n. 10.
<https://doi.org/10.1111/lnc3.12438>

Constructing counter-hegemony on *Twitter*: Discourse of Ibero-American political women of “change” in the digital environment

Romina Pepe-Oliva; Andreu Casero-Ripollés

Nota: Este artículo se puede leer en español en:

<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87165>

Recommended citation:

Pepe-Oliva, Romina; Casero-Ripollés, Andreu (2023). “Constructing counter-hegemony on *Twitter*: Discourse of Ibero-American political women of ‘change’ in the digital environment”. *Profesional de la información*, v. 32, n. 3, e320304.

<https://doi.org/10.3145/epi.2023.may.04>

Manuscript received on 5th November 2022

Accepted on 2nd January 2023



Romina Pepe-Oliva ✉
<https://orcid.org/0000-0002-1007-2564>

Universitat Jaume I
Facultat de Ciències Humanes i Socials
Av. Vicent Sos Baynat, s/n
12071 Castelló de la Plana, Spain
al367984@uji.es



Andreu Casero-Ripollés
<https://orcid.org/0000-0001-6986-4163>

Universitat Jaume I
Facultat de Ciències Humanes i Socials
Av. Vicent Sos Baynat, s/n
12071 Castelló de la Plana, Spain
casero@uji.es

Abstract

Women are occupying more and more space in the public sphere, not only through pressure on the streets, but increasingly in digital spaces. It is at this intersection that different ways of engaging in politics coalesce with the presence of women who demand their own voice and visibility beyond the mechanisms of traditional politics. Those requiring a transformation of the current political reality are the political women “of change.” One of their main attributes is that they generate and spread counter-hegemonic narratives as a form of empowerment and a way to question the dominant political discourse through digital media. Our objective herein is to analyze, in a comparative way, the use of this communicative strategy to understand its articulation and mechanisms. To do this, we study the discourse of ten Ibero-American political women on *Twitter* who are linked to social change. The methodology is based on the application of the content analysis technique that combines a quantitative dimension with another of a qualitative nature focused on critical discourse analysis. The results show that criticism and denunciation, to give voice to the voiceless and make social problems visible, are the main components of these political actors’ counter-hegemonic discourse on *Twitter*. Likewise, they display a practical and constructive counter-hegemony oriented in applied and positive terms. Finally, the institutional position in the government–opposition axis sharpens or minimizes the use of these types of communication strategies.

Keywords

Political communication; Counter-hegemony; *Twitter*; Political women; Ibero-America; Social media; Discourse; Social change; Public agenda; Activism; Cyberactivism; Political movements; Strategies; Political communication; Feminized politics; Narratives; Political and social change; Gender; Social inequalities.

Funding

This research has been funded by *Ministry of Science and Innovation (Micin)* of Spain, through the *State Research Agency (AEI)*, ref. 10.13039/501100011033, and the *European Regional Development Fund (ERDF)*, program “A way to make Europe”, grant number PID2020-119492GB-I00.



1. Introduction

Despite relevant social and cultural conditioning, women have gained prominence and visibility in public spaces. In recent decades, gender differences in the political arena have decreased notably, especially thanks to the entrance of women to the world of work and the progressive equalization of the educational levels of women and men (**Verge-Mestre; Tormos-Marín**, 2012). This process has led women to increasingly occupy different spaces in governments or parliaments in Western democracies thanks to parity or quota laws implemented in some Ibero-American countries (**Vinuesa-Tejero; Abejón-Mendoza; Sánchez-Calero**, 2011; **Freidenberg; Lajas-García**, 2015).

This has resulted in women playing various roles: working closely with elections, orchestrating campaign rallies, participating in marches and demonstrations in unprecedented numbers, and, of course, increasingly, running for leadership and top positions in the internal structures of political parties (**Rodríguez**, 1999). This phenomenon of integration of women into the democratic procedure, called the feminization of politics (**Rodríguez-García; Navarro-Yáñez**, 2012; **Ochoa**, 2021; **Quevedo-Redondo**, 2021), has been seen as a positive aspect to make it look more human while neutralising the discredit of it. Likewise, citizens expect more sensitive and competent positions from women politicians regarding social policy issues or issues related to feminist demands (**Huddy; Terkildsen**, 1993; **Lawless**, 2004; **Larrondo-Ureta; Orbezo-Terradillos; Peña-Fernández**, 2019). These values that historically have been attributed to women (**Quevedo-Redondo**, 2022) have generated substantial changes in political communication (**Herrnson; Lay; Stokes**, 2003; **Childs**, 2004; **Orbezo-Terradillos; Larrondo-Ureta; Morales-i-Gras**, 2020).

Social and political mobilization has not come alone. It has been the consequence of political and technological transformations that have enabled the emergence of movements that warn of the need for change or transformation (**Castells**, 2009; **Casero-Ripollés**, 2015). Women have participated in most of them, and in many others, women have been the protagonists of marches, demonstrations, mobilizations, or popular movements such as #YoSíTeCreo (**Larrondo-Ureta; Morales-i-Gras; Orbezo-Terradillos**, 2019), #Niunamenos (**Chenou; Cepeda-Másmela**, 2019), #Abortolegal (**Laudano**, 2018), or #Metoo (**Jaffe**, 2018), calling for an end to social discrimination and violence against women.

These feminist movements that emerged on the internet sometimes move to the streets and result from the history of organized women incorporating activism trends (**Rosales**, 2018). Therefore, being a political actor today means using social platforms to gain public visibility more than ever (**Núñez-Puente**, 2011; **Crossley**, 2015; **Kavada**, 2016; **Barker-Plummer; Barker-Plummer**, 2017). In addition, this is where many women politicians could take advantage of the possibilities that digital media provide them, allowing for them to improve their political and social participation (**Friedman**, 2003; **Núñez-Puente; Vázquez-Cupeiro; Fernández-Romero**, 2016) while becoming an “engine” for women’s movements (**Martin; Valenti**, 2013; **Crossley**, 2015).

In addition, digital platforms are relevant in citizens’ political mobilization, proposing themselves as an instrument that citizens can actively use to promote their causes and demands (**Casero-Ripollés; Moret-Soler**, 2022). Thus, the strategic use of digital tools (**Barker-Plummer; Barker-Plummer**, 2017) offers women politicians another way to communicate, expressing their concerns but also introducing their issues on the public agenda (**Aruguete**, 2017; **De-Aguilera; Casero-Ripollés**, 2018) from a different perspective. In this context, the different ways of doing politics emerge as a need for the social transformation of current politics. The political women “of change” demonstrate other possibilities for political intervention beyond the traditional or vertical ones. Instead, they can be associated with those logics of a horizontal nature, which seek to undermine the hegemony of established political forces, betting on socioeconomic stimulation and alternative policies (**Feenstra**, 2016).

Thus, the appropriation of digital media challenges those narratives established by political elites (**Castells**, 2009; **Sampedro**, 2014). This possibility of autonomously and instantly producing and distributing their speeches in digital environments offers, in this case, the women “of change” an active role as producers of their content (**Cammaerts**, 2018).

The traditional political sectors and the mainstream media focus on building and maintaining the dominant discourse (**Van-Dijk**, 2003). From their position of power, they seek to extend their domination not only through economic factors (laws, lobbies, media pressure, etc.), but also through the control of the senses and meanings that circulate socially (**Casero-Ripollés**, 2009). Digital activism and the alternative media openly fight to stop this status quo and promote different visions of a counter-hegemonic nature by disseminating critical voices and alternative manifestations of reality (**Treré**, 2016), offering another way of conceiving and understanding society.

Hegemony is the situation of a class that achieves solid ideological and political unity, allowing it to establish ascendancy over other groups and classes (**Gramsci**, 1972). Its construction is based on discourse and crystallizes in common sense (**Mouffe**, 2008). Hegemony is subjected to consensus and simultaneous processes of resistance (**Elbaum**, 1997). Opposite the ruling class, an alternative class can emerge that aspires to found another “world vision” (**Campione**, 2005), since every hegemonic order can be challenged by counter-hegemonic practices that try to dismantle it to install another form of hegemony. Thus, discourse is a substantial space for analyzing communication strategies to construct counter-hegemony.

“Counter-hegemonic discourses question the traditional ways of doing politics in our society”

Through social media spaces, counter-hegemonic narratives can emerge, thus giving voice to those who have never had one (Fuchs, 2014). Counter-hegemonic narratives are discourses of ruptures against the established order. They emphasize the subaltern character (Moraña, 1998) of the voices that testify to problems of social significance and expose a space of resistance (Colanzi, 2018). This conflict against the hegemonic order involves a double dimension: first, the symbolic part, related to ideas, and second, the material part that relates to the practices and objective positions of activism (Avalos-González, 2019). Thus, these generate new forms of mobilization, leadership, and rhetoric, as well as new forms of approaching electoral campaigns, and therefore, a different place to exercise politics emerges. This “new” way of doing and thinking about politics as a space for building a counterpower (Castells, 2009) enables us to reconceptualize citizenship and democracy entirely.

Counter-hegemony can also be expressed through communicative practices that call for disobedience, to create rhizomatic activism, or to propose an alternative agenda

These communicative dynamics introduce and promote participation actions, placing citizens—in this case, women and their claims—at the centre of the communicative process while endowing them with unprecedented capacities to produce and reproduce messages (Casero-Ripollés, 2015). This represents a step towards public empowerment, as these political actors have an active role as content producers. This use of digital media allows for the generation of a counter-power (Alonso-Muñoz; Casero-Ripollés, 2016), disputing the hegemony held by the traditional media and the political elites, thus generating new meanings and framings activating counter-hegemony. Thus, the public visibility of other ways of engaging in politics and conceiving society increases, challenging hegemonic and unequal discourses (Turley; Fisher, 2018).

2. Methodology

2.1. Objectives

Our general objective is to study and compare how *Twitter* enables the construction of counter-hegemonic narratives, giving voice to other social actors, specifically women politicians, to change and transform social reality while generating new forms of citizen empowerment.

From this, we can formulate two specific objectives:

- O1. Explore how political women “of change” use *Twitter* to build an alternative counter-hegemonic political discourse and narrative.
- O2. Analyse what dimensions and components of the counter-hegemonic discourse are present in the communication strategies of the women “of change” on *Twitter*.

2.2. Research techniques and sample

The methodology relies on the application of the content analysis technique, understood as an interpretive procedure of different communicative products (Piñuel-Raigada, 2002) aimed at exposing, based on objective data, reproducible and valid inferences, based on the characteristics of a set of messages (Krippendorff, 2013). This technique combines a quantitative dimension with a qualitative nature focused on critical discourse analysis (Van-Dijk, 1993; 2006; 2013; Fairclough, 2010). This perspective understands discourse as a form of symbolic power capable of conditioning public opinion as it is a primary tool in building a reality.

This interpretive methodology focuses on delving into the content of public discourse and not on the effects on the public consuming them. The approach focuses on social problems, fundamentally on the role of discourse in the production and reproduction of the abuse of power or domination (Van-Dijk, 2003). Therefore, it is especially suitable for analysing the configuration of counter-hegemonic discourses that seek to dispute the established power.

We understand hegemony as a dynamic based on a dialectic between consensus and conflict of interests (García-Canclini, 1984). It is acquired and has to be constantly maintained through discourse. To this concept must be added counter-hegemony or alternative hegemony (Williams, 1980), which aspires to discredit hegemonic schemes and provide credible alternative understandings that question them (De-Sousa-Santos, 2014). The methodology relies on this concept in order to convert it into a set of operational variables that enable quantitative measurement of its presence in the discourse. We use three large blocks that group 11 variables: the construction of alternative discourses, counter-hegemonic practices and self-proclamation as subjects of change.

To analyse the construction of the counter-hegemonic discourse over a broad period and to better examine its conformation, the period that includes the year 2020 is analysed. We select a constructed sample that alternates by one week each month. In this way, a routine period is chosen, avoiding relevant political events such as electoral periods, to observe how their discursive strategy works. The total number of tweets analysed is 1,559. The sample tweets were obtained using the *Twitonomy Premium* software. The analysis of each of the tweets was performed manually and not automated.

This research takes as its sample a study of the *Twitter* profiles of ten women politicians “for change” in the Ibero-American context:

- Myriam Bregman (Argentina);
- Vilma Ripoll (Argentina);
- Camila Vallejo (Chile);
- Ana Erazo (Colombia);
- Ada Colau (Spain);
- Teresa Rodríguez (Spain);
- Martha Tagle (Mexico);
- Verónica Mendoza (Peru);
- Marisa Matias (Portugal); and
- Verónica Mato (Uruguay).

There are three reasons for this choice. First, they define themselves as critics of the traditional practice of politics and defenders of causes linked to social change. Second, this selection introduces a high geographical diversity by incorporating representatives from eight Ibero-American countries. Finally, it responds to different profiles of presence in the political system (government opposition). These criteria ensure a sample that offers more scope for applying a comparative approach and achieving significant and representative results.

2.3. Model of analysis

The quantitative content analysis is obtained by coding each *Twitter* message in the sample using dichotomous response variables (Yes/No). Table 1 shows the variables enabling the operationalisation of the different dimensions that counter-hegemonic narratives could adopt in the communication strategies of the women politicians studied.

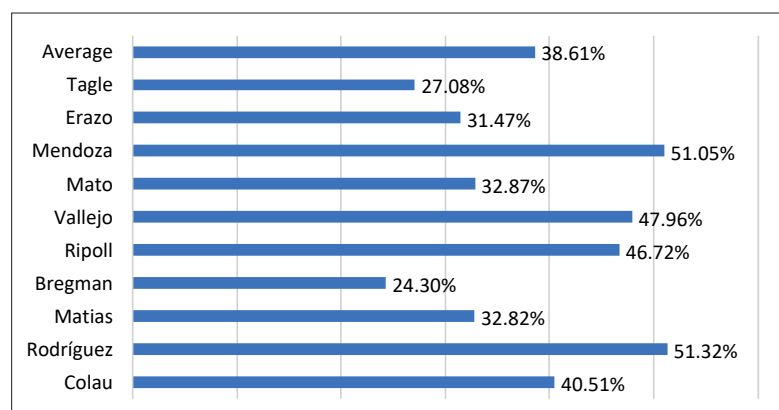
Table 1. Counter-hegemonic discourse analysis model

Nº	Variable
1	It sustains a discursive strategy of differentiation between “us” and “them”: There is a discursive construction that maintains differentiation between who enunciates the tweet and to whom it is transmitted.
2	It criticizes the established power and the economic, political and social elites: it criticizes the power of the political, economic, social or judicial elites.
3	It positions itself as an interlocutor for the voiceless: it positions itself as the spokesperson for those who cannot do so and gives visibility to the problems of those who suffer.
4	It openly denounce inequalities: It make visible and openly denounce inequalities to avoid or lessen the great evils that hunger, violence, racism, and sexism produce.
5	It openly denounces inequalities: It makes visible and openly denounces inequalities to avoid or lessen the great evils that hunger, violence, racism, and sexism produce.
6	It expresses “other” forms and/or values of acting in politics: It exposes new forms of political activity, explaining itself, arguing (this variable links to transparency, accountability, rejecting political privileges or the greater visibility of daily work).
7	Active extra-representative or non-conventional forms for political participation: it invites citizen action and mobilization in which she is also a participant.
8	It vindicates its political condition as a subject “of change”: Her ways of doing politics, and her proposals for transformation are exposed to the public to change the current situation.
9	It proposes an alternative thematic agenda: It is capable of including its themes or defining “other views” in terms of historical memory, city model, environment, industrialization, technology and consumption, housing crisis, decent work, diverse identities, etc.
10	Promotes rhizomatic activism: it proposes, supports, or forms part of self-organization networks, collaboration networks, citizen movements, social pressure systems, etc., linked to network activism.
11	It represents democratic renewal: It is proposed as a representative of a new way of doing things far from political profiles and traditional discourses. They are the antithesis of traditional politics, bringing new values and visions of politics and democracy.

3. Results

Our analysis focuses on three central questions that will enable us to identify counter-hegemonic features in the actors’ discourse. First, we will ask ourselves if they build an alternative discourse. Second, if they communicate counter-hegemonic practices and, finally, if they define themselves as subjects of change or social transformation.

The content analysis applied to the sample shows some relevant aspects in how the women politicians use *Twitter*. Although



Graph 1. Average of the comparison between the categories

all these political actors present themselves as critics of the system, corruption and traditional parties, and they offer alternative proposals in their political discourse, not all the analysed women politicians present aspects qualified or defined as counter-hegemonic discourse. Comparing the results (Graph 1) between the categories in which their discourse clearly enunciates critical traits or breaks with the average (38.61%), only half of the analysed political actors conform to this profile. In this respect, Teresa Rodríguez (51.32%) stands out with a strategy framed in the alternative, criticism, disobedience and questioning established discourses. Along the same lines are Verónica Mendoza (51.05%), Camila Vallejo (47.96%), Vilma Ripoll (46.72%) and Ada Colau (40.51%).

3.1. The construction of an alternative discourse as a communication strategy

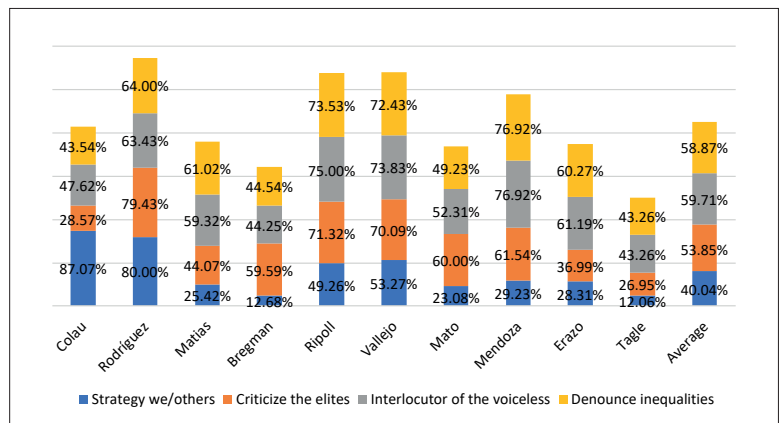
The Ibero-American political women of “change” hold a discursive strategy of differentiation between “us” and “them”. This is a fundamental category in the counter-hegemonic discourse enabling us to observe an understanding of politics, but also the conflict or duality between those who are outside ideology, the ways of seeing and understanding the world and those who do not. Alternative discourses establish distinctions between winners/losers, corrupt/honest, guilty/victims, workers/businessmen, ordinary citizens/established elites or, even more, between traditional and new politics.

Here, Ada Colau (87.07%) and Teresa Rodríguez (80.00%) stand out as the political actors who use this type of practice the most, followed by Camila Vallejo (53.27%) and Vilma Ripoll (49.26%) (Graph 2). All exceed the average of the data analysed in this category (40.04%).

Analysing the data, the fundamental difference between the two Spanish politicians and the two South American ones is their trajectory. Colau and Rodríguez emerged on the political scene from activism or militancy in parties where the differentiation with “others” or the “castes” is constantly present. These are political options born from a commitment to differentiate themselves from those traditional parties of a more institutional nature. Farther away are Ripoll and Vallejo, both belonging to parties with a historical trajectory on the left (*Chilean Communist Party*–Vallejo–, and *Socialist Movement of Workers*–Ripoll–, Argentina) where a historical differentiation is also being built, as long as the oligarchy, the traditional parties in power or “the right” are “the others” for these women politicians.

In the case of Ada Colau (Image 1), the “we” is proposed, in general, as her government group, her work team, and her legislative initiatives, not as belonging to a specific class, but as a member of a team along with the legislative proposals that her group has made together with other similar parliamentary groups. The construction of the “others” is proposed as the opposition, the right, the extreme right, and also those who “scream” in front of “us” who have chosen the “only path”, a space of knowledge and belonging. Her speech refers to the road as the place of “we” that ends the “confrontation”.

In Rodríguez, similar to Colau, there is a profound recurrence of the “we”, but unlike the latter, here “we” is represented by other subjects: “the Andalusians”, “the working classes”, the “ordinary people” and “common people”. Instead, the “others” is used with a clear differentiating intention: “the rich”, “the banks”, “the vulture funds”, “royalty”, “the Crown” and “the central government”. There is a constant appeal to belonging to “her” class, her neighbourhood and her land. In this case, we find the construction of “we” of an ideological and populist nature. In the example analysed (Image 1), there is a tweet accompanied by a photo of a member of the



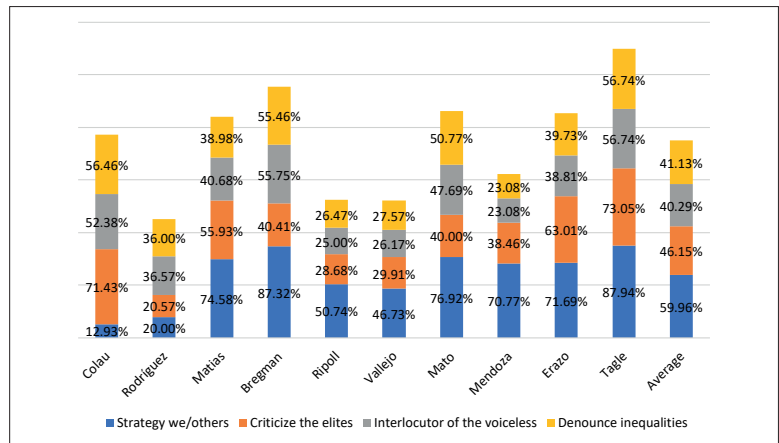
Graph 2. Construction of an alternative discourse as a communication strategy. Positive values



Image 1. We/they discursive construction
<https://twitter.com/AdaColau/status/1213384065224708099>
https://twitter.com/TeresaRodr_/status/1240309276037582848

Royal House, where “they” are the ones who have “taken advantage” to announce “their shame”. In this case, the “we” represented the “united country”, fighting against the “virus” (COVID-19), while they, the Crown, the King, are the “others”, who are talking about money laundering, commissions or fortunes to launder.

The second category of this block enables us to describe whether there is criticism of the established power and the economic, political and social elites. This is another indicator that distinguishes the counter-hegemonic discourse insofar as criticism of the elites. Also, it is significant in political discourses distinguishing a need or search for change. In general terms, this principle is fulfilled and six out of the ten political women analysed criticise the elites and the established power (Graph 2). It should be said that they exceed the average of 53.85%: Teresa Rodríguez (79.43%), Vilma Ripoll (71.32%), Camila Vallejo (70.09%), Verónica Mendoza (61.54%), Verónica Mato (60.00%) and Myriam Bregman (59.59%).



Graph 3. Construction of an alternative discourse as a communication strategy. Negative values

Regarding the communication strategy on *Twitter* to criticise the elites, in general terms, differentiation cannot be made by continents or countries as discourses with traits critical of the elite are common. But there is a specific differentiation between those that do not criticise the elite or do so to a lesser extent (Graph 3), from those that articulate an explicit criticism of the elite, maintaining a detailed questioning profile of the various political and social powers. The latter share the trait of being political actors located outside the government party. However, it is impossible to differentiate the country or continent oriented towards criticism of established power.

Another common characteristic is that the political women of “change” who use this function share common targets of their criticisms. In the profiles with the highest levels of criticism, there is uniformity in the direction of criticism: the opposition, the government in office, the police or the army, businesses, large corporations, health ministers in times of pandemic, international institutions such as the European Union or in Latin America the International Monetary Fund. The following examples reflect these issues.

Thus, the opposition deputy, Verónica Mato, criticises “los legisladores del *Cabildo Abierto*” (the city hall legislators) for preventing judicial action on clarifying the facts during the last Uruguayan Military Dictatorship (Image 2). The criticism relates to certain legislators who implemented “measures” to hinder the task of justice for the actions of certain military and police officers in investigations of crimes against humanity. This type of criticism and denunciation shows how certain powers prevent or block the action of justice to solve the crimes committed by Latin American dictatorships, and is significantly present in the politics of southern Latin America, especially in Argentina, Chile and Uruguay. The call for justice for the crimes committed in dictatorships, and the criticism of the powers that for more than thirty years continue to prevent the facts being uncovered, is present transversally in the counter-hegemonic narrative of these political actors.

Another example of the criticism of elites is the legislator Myriam Bregman from Buenos Aires, an opponent of the Government of the Autonomous City of Buenos Aires. Image 2 shows the criticism of the elites and political powers, in this case, international. The economic interference of the United States on the peoples of Latin America or the Middle East is denounced. The criticism aims at politicians who meet the deadlines to repay *IMF* loans and at creditors who toast this “with champagne”. It uses irony to attack the economic sectors interested in continuing to earn money despite the recession that the country is experiencing and the devastating consequences of economic adjustment that this entails for citizens. On the issue of external debt, this legislator calls for a refusal to comply with payment responsibilities.

The next aspect that enables us to characterise this type of discourse is the categories where the extent to which political



Image 2. Criticism of the elites
https://twitter.com/Veronica_Mato/status/1285400685580881920
<https://twitter.com/myriambregman/status/1228060738100527104>

actors position themselves as interlocutors of the voiceless are analysed and whether they denounce inequalities. Here another characteristic of the counter-hegemonic discourse is evident: it not only recognises the level of social conflict by giving voice to these problems but also positions itself against it. These are specific aspects of the discourses that break with the common sense that operates in society because they tend to expose the most relevant social problems while highlighting the level of conflict and inequality. For this reason, we present both categories together, given their relationship.

The data reveals that the average in the category “interlocutor for the voiceless” is 59.71% (Graph 2), with the following politicians exceeding the average: Verónica Mendoza (76.92%), Vilma Ripoll (75.00%), Camila Vallejo (73.83%), Teresa Rodríguez (63.43%) and Ana Erazo (61.19%). Here a differentiation is detected between the voices of South American politics and European ones. This category is strongly represented by the former because it is closely related to the situation in South American countries of poverty, marginalisation, violence, police, military and human rights abuses.

The category of denouncing inequalities obtains an average of 58.87% (Graph 2). Verónica Mendoza’s percentage (76.92%) is higher. Vilma Ripoll (73.53%), Camila Vallejo (72.43%), Teresa Rodríguez (64.00%), Marisa Matias (61.02%) and Ana Erazo (60.27%) posted a high number of tweets with complaints about social, political or economic inequalities. We cannot make a geopolitical distinction in this category as women politicians from both continents had high percentages. This type of practice – denouncing inequalities – is standard in the profile of more than half of the politicians studied.

Most messages express condemnation of world hunger, wars, immigration, refugees or the violation of human rights. Likewise, they denounce current problems such as unemployment, homelessness, Covid-19 and the lack of access to education. This category, like the previous one, seeks to give visibility or speak out against the problems of the countries regardless of their geographical position. With this vision, the problems highlighted range from the concentration of wealth and the worsening living conditions of families due to the pandemic crisis to the refugee crisis or wars.

In Image 3, we can see how Ana Erazo condemns the proposal for a salary increase for workers made by the Government of Iván Duque in Colombia and contrasts it with the salary rise made to congressmen, contrasting the “misery” of the workers and the “similar” increase for parliamentarians. Vilma Ripoll makes a



Image 3. Report inequalities

https://twitter.com/vilma_ripoll/status/1313520196871753731

<https://twitter.com/AnaErazoR/status/1343975244785135618>



Image 4. Positioning as an interlocutor of the voiceless

https://twitter.com/mmatias_/status/1301124915995856903

https://twitter.com/camila_vallejo/status/1329803847766708224

similar point, “the miserable 7%” increase that the Argentine government proposes for state workers, while economic inflation forecasts reach 40%. This highlights the situation and the standard of living of South American workers. Both leaders use their *Twitter* profiles to complain how low workers’ salaries are and the constant rise in prices they suffer.

The counter-hegemonic discourse of Ibero-American political women of “change” is also positioned as an interlocutor of the voiceless since these political actors put themselves forward as spokespersons for those who cannot do so. The first example (Image 4) is that of Marisa Matias, drawing attention to the immigration crisis and the situation of vulnerability and risk of death, especially in children. Remembering the body “of little Aylan Kurdi on a Turkish beach moved the world” she laments: “The foam of the days erased the image too quickly”. And she regrets that good intentions “also died”. This shows how this political actor denounces and echoes a problem that has not ended. To this day, children continue to arrive alone on the shores of Europe. Continuing the theme of childhood, Camila Vallejo’s tweet also stands out, warning about child crime. This political actor alleges that “no kid is born a thief”, but rather society and the structural and everyday circumstances affecting children push them into crime.

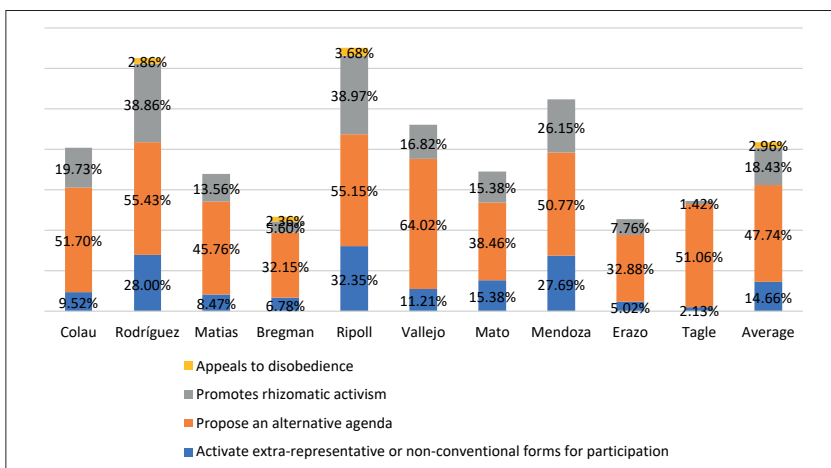
3.2 Counter-hegemonic communication practices for an alternative politics in the digital environment

Counter-hegemony can also be expressed through communicative practices that call for disobedience, to create rhizomatic activism, or to propose an alternative agenda. Therefore, we first analyse whether the tweets of Ibero-American women politicians of “change” contain content that addresses issues or problems characterised as an alternative agenda.

In this category (Graph 4), the politicians who stand out above the average are: Camila Vallejo (64.02%), Teresa Rodríguez (55.43%), Vilma Ripoll (55.15%), Ada Colau (51, 70%), Martha Tagle (51.06%) and Verónica Mendoza (50.77%). They employ *Twitter* to build a discourse that challenges the significant issues raised by media corporations or traditional discourses whether on the agenda of the leading parties or in the conventional media. Therefore, six of the ten women politicians studied appear in this dispute to shape or dismantle current narrative structures.

The posted tweets focus on various problems or issues, which often do not establish differences between the countries or regions of origin of the women politicians since they are common across the world. Among these tweets are condemnations of abuses by employers during the pandemic, the importance of the work of social organisations, the struggles of indigenous peoples, historical memories and human rights, access to digital connectivity among the poor, the ecological crisis or excessive environmental exploitation, the lack of investment in public services, and many other issues, such as the work overload women had to endure during the health crisis or cleaning work in hospitals and medical services.

Concerning these aspects, an example is Image 5, containing the tweet of the Peruvian politician, Verónica Mendoza, demanding a “universal bonus immediately” with the active support and “participation of social organisations”. It calls for the inclusion of the neediest, those who cannot make demands and an inclusive proposal to all social organisations to consider supporting the most vulnerable. Another example within this category is the tweet from Argentinian Vilma Ripoll. In this case, the issue



Graph 4. Counter-hegemonic practices oriented towards an alternative politics in the digital environment. Positive values



Image 5. Alternative agenda proposal
https://twitter.com/Vero_Mendoza_F/status/1258939809852162049
https://twitter.com/vilma_ripoll/status/1329554413770465281



of the environment, ecology and sustainable development is a recurring theme in her messages. Ripoll expresses her support for mobilisation against mega-mining due to the environmental impact in the Chubut region, giving visibility to this issue in her tweets.

Another practice that can be identified as alternative political communication is to know whether the analysed women politicians promote “extra-representative” or unconventional forms of participation in their speeches, such as manifestos, petitions, boycotts, strikes, occupations and street blockades. In this case, the data reveals that only three of the ten politicians propose or activate this type of action in their political communication (Graph 4). Exceeding the average (13.43%) are only Vilma Ripoll (32.35%), Teresa Rodríguez (28%) and Verónica Mendoza (27.69%). In contrast, seven of the ten women do not show signs of encouraging this type of practice in their discourse (Graph 5).

The difference between political actors on *Twitter* promoting or supporting this extra-representative communication from those who do not, connects to their position in public life. Vilma Ripoll (32.35%) and Verónica Mendoza (27.69%) participate in political activity outside the institutional framework. Although both are visible leaders in their party, they do not participate in parliament or official organisations. On the other hand, Teresa Rodríguez (28%) is a regional parliamentarian, but she occupies a minor position far from the government. Another explanatory factor is its link with activism. This is the case of Rodríguez, whose past as an anti-capitalist activist brings her closer to unconventional practices.

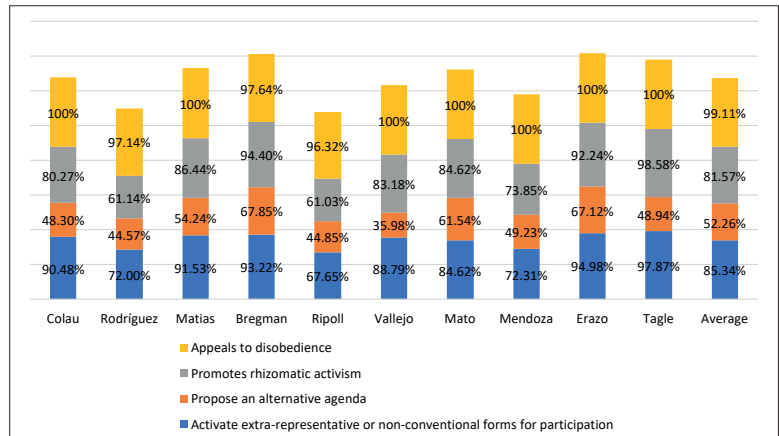
One of the -proponents of this type of extra-representative practice is Vilma Ripoll. She constantly resorts and appeals to promote, participate in and communicate these activities on *Twitter*. In this case, it relates to the resistance to unemployment by the workers of the historic Bauen Hotel on 300 Callao Street in the City of Buenos Aires (Image 6). She claims to feel “proud to have accompanied the struggle in each eviction attempt”.

Those politicians without features in their discourse defined as extra-representative (Graph 5) have institutional positions within political life. Tagle, deputy of the Mexican Parliament; Erazo, Councillor of Cali; Colau, Mayor of Barcelona; Bregman, Buenos Aires legislator; Matias, MEP and Vallejo and Mato, both MEPs in their countries. This factor can cause these political actors to limit their political communication to conventional political practices.

On the other hand, within this scope, we analyse whether they use *Twitter* to promote rhizomatic activism. Four of the ten political actors studied (Graph 4) exceed the average (18,43%). They are Vilma Ripoll (38,97%), Teresa Rodríguez (38,86%), Verónica Mendoza (26,15%) and Ada Colau (19,73%). Comparing this category with the previous one, we find the same political actors promoting rhizomatic activism in their communication practice. Therefore, there is a connection between both types of counter-hegemonic practices in the communication strategies of Ibero-American political women of “change”.



Image 6. Activation of extra-representative or non-conventional forms of participation
https://twitter.com/vilma_ripoll/status/1314245668832186368



Graph 5. Counter-hegemonic practices oriented towards an alternative politics in the digital environment. Negative values



Image 7. Promotion of rhizomatic activism
https://twitter.com/TeresaRodr_/status/1239879284333989890
https://twitter.com/Veronica_Mato/status/1271232558882336771

The political actors that activate this counter-hegemonic practice share information on their *Twitter* accounts and invitations to different activities or events to build ties and help others. They promote, sponsor or invite social and solidarity actions in their messages despite the pandemic crisis: from ordering food or donations to solidarity networks to supporting victims of sexist violence. Thus, Teresa Rodríguez spreads the idea of “support networks” for those who need care, donations or company, generating spaces for interaction and mediation with civil society (Image 7). Likewise, Verónica Mato recalls that “the people need our solidarity”, inviting them “not to give up”. Thus, she stimulates alternative spaces and institutions that help those affected or without resources.

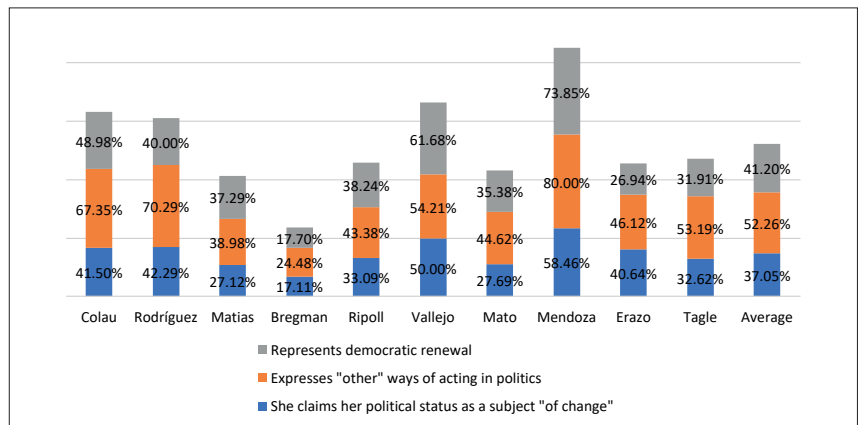
The last aspect in this section analyses the discursive construction of insurgency or disobedience in the tweets of Ibero-American political women of “change”. It is the least used category (Graph 5). Only three of the ten politicians studied have some indications in their political communication on *Twitter* that invites disobedience (Graph 4). Most of the women politicians maintain a discourse following those established within the canons of traditional political discourse. There are no insinuations or statements in favour of civil disobedience, except for marginal percentages.

3.3. Discursive self-proclamation to be agents of “change” or social transformation

This section studies how political actors position themselves or proclaim themselves as agents of “change” or social transformation in their *Twitter* profiles. First, it is analysed whether they express “other” forms and/or values of acting in politics. Five of the ten politicians studied exceed the average (52.26%) (Graph 6). They are Verónica Mendoza (80.00%), Teresa Rodríguez (70.29%), Ada Colau (67.35%), Camila Vallejo (54.21%) and Martha Tagle (53.19%). Half of the political actors build a narrative to express other forms or values in political action. In some cases, they reach 80% as in the case of Verónica Mendoza.

Verónica Mendoza stands out (Image 8) with a speech on changing political proposals. The discourse includes her hometown, peasants, merchants, teachers and workers from all regions, while speaking about ecology, diversity and human rights. She shows an interest in the community and the biosphere of her region and her concern for the people who live there. She explains that the mining project to be developed in the area of Lima “would store toxic waste” and calls on the Judiciary to decide “whether it will protect or put water and life at risk”. Two issues emerge in her speech. First, there is an interest to preserve the environment and its surroundings. Second, there is a direct appeal to justice institutions to show what side they will take. Something similar occurs with Teresa Rodríguez’s tweet about *Bankia* in which she proposes other possible policies in this case. There are several values that the analysed women politicians describe in their *Twitter* accounts which enable them to appear as actors who seek to exercise alternative politics that focus on issues close to the people and their interests. These practices can be understood as an alternative approach to social transformation.

The next category seeks to establish whether these political actors claim their political condition as a subject “of change”. Above the average (37.05%) are Verónica



Graph 6. Discursive self-proclamation to be agents of “change” or social transformation. Positive values

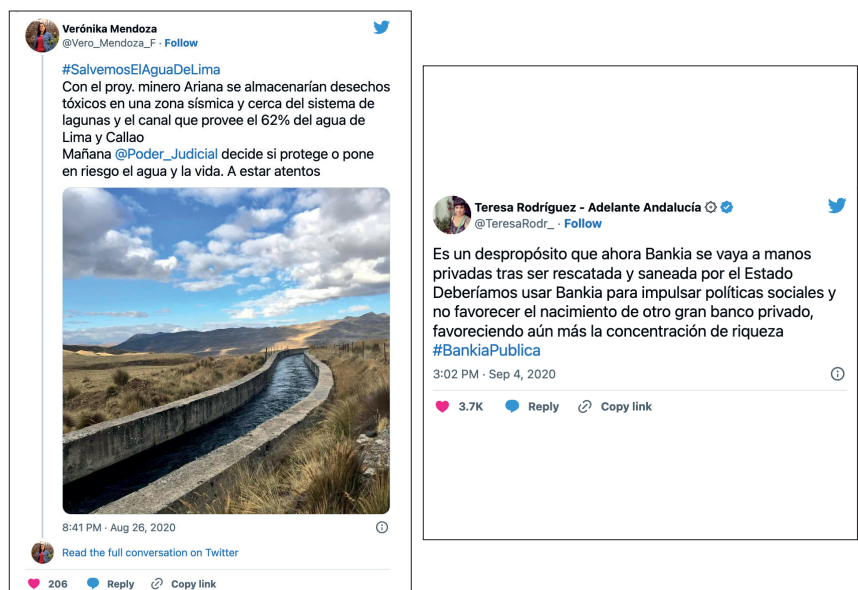


Image 8. Expression of “other” forms and/or values of acting in politics
https://twitter.com/Vero_Mendoza_F/status/1298691953136218118
https://twitter.com/TeresaRodr_/status/1301868352554627072

Mendoza (58.46%), Camila Vallejo (50.00%), Teresa Rodríguez (42.29%), Ada Colau (41.50%) and Ana Erazo (40.64%). Half of the policies studied construct a discourse to proclaim themselves as agents of change and generate messages to give visibility to such an issue (Graph 6).

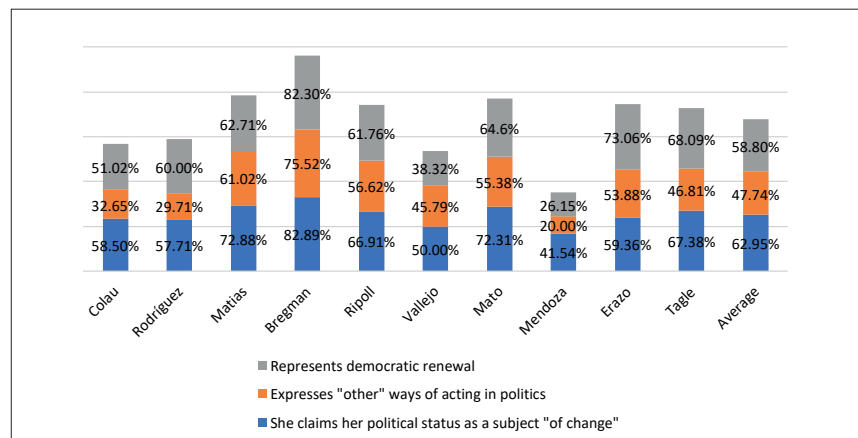
Among these political actors, we can distinguish two groups. On the one hand, Vallejo, Mendoza and Rodríguez generate a narrative from an opposition discourse and pose demands to their countries’ governments. They also highlight actions or legislative proposals to improve citizens’ lives. On the other hand, we find Colau and Erazo, who are representatives of the Government or a part of it and give themselves visibility as subjects of change in their political work. This type of discourse defines itself as change, as a policy of doing things differently from previous governments while being very evident in the political communication on *Twitter* of these political actors. Here the change is understood institutionally, that is to say, as an improvement concerning the previous government and not so much as a process linked to social change, as was the case of the former group of politicians.

The qualitative analysis enables us to observe how Verónica Mendoza expresses the demands of Peruvian civil society that is tired of cases of corruption and the lack of responsibility in traditional politics, asking for “a new democracy and a new State that guarantees citizens’ rights and puts life ahead of profit” (Image 9). She calls on citizens to support a new constitution to foster the country’s regeneration, promoting a profound change in her country. Likewise, Camila Vallejo appeals for a transformation of the budget setting process. She affirms that “because of our Constitution, only the Executive can propose how the money is spent while Congress approves or rejects it”. Also, she criticises that only the Executive can decide how Chilean money is spent. This connects with a significant aspect of democracy: budget transparency.

Finally, the last category of this block aims to show whether the political actors build a story in which they represent democratic renewal. The results show (Graph 6) that only three of the ten politicians outline a discourse in line with the self-proclamation of democratic renewal or transformation. The three politicians exceeding the average (41.20%) are Verónica Mendoza (73.85%), Camila Vallejo (61.68%) and Ada Colau (48.98%).

Of the three leaders who have appealed to this type of narrative, two participate in institutional politics –Camila Vallejo and Ada Colau– while the third –Verónica Mendoza– operates outside parliamentary politics. When analysing their discourses, they deal with issues such as the State model or institutional reorganisation, constitutional reform, human rights, equality policies, autonomy, the independence of the judiciary from other powers or the growing dissemination of disinformation by some media. But they also address other issues, such as repressing vulnerable populations and minorities, original identities or oppressed peoples.

The reforms or new models that they propose make them look more open and committed to citizenship and civil society. Challenging traditional politics, the three leaders favour a discourse of opposing democratic decline, in favour of transparency. However, analysing the general data, we can see that most of the women politicians – seven out of ten – focus on building a critical discourse and proposals related to the insti-



Graph 7. Discursive self-proclamation to be agents of “change” or social transformation. Negative values

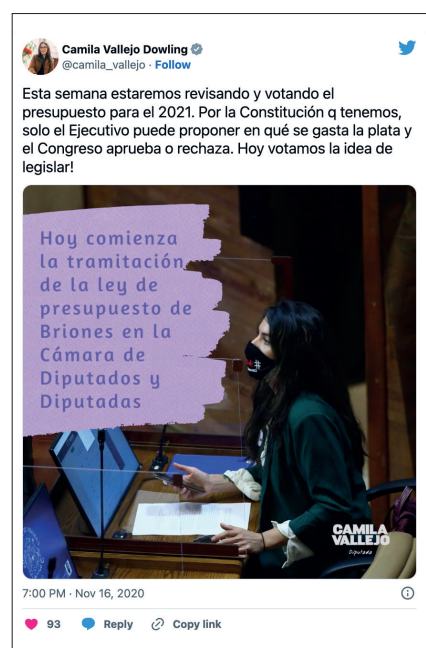


Image 9. Claiming the political condition as a subject “of change”
https://twitter.com/Vero_Mendoza_F/status/1328524266569158656
https://twitter.com/camila_vallejo/status/1328397620813963264

tutions where they can collaborate rather than proclaiming themselves as protagonists of a democratic renewal (Graph 7).

Thus, Ada Colau advocates a debate on a historical question: the monarchy or a republic (Image 10). She brings to light the issue of the state model, alleging that in Barcelona there is “a broad republican majority” that is against honouring the monarchy. She is appealing to another type of discourse: direct criticism of the Crown. She also calls for a “thorough investigation into the malpractice” of the institution, as well as “promoting legal changes to end impunity and the privileges of the Crown”. She claims to promote legal changes so that the monarch is not exempt from responsibility for his or her acts, whether they refer to matters of a public or private nature. By formulating these questions, the narrative of this political actor shows an intention to democratise both the institutions and the people who compose them to avoid abuses such as excess power.

4. Conclusions

This research has sought to make original and relevant contributions to the counter-hegemonic dimension of the discourses of political actors on social networks, especially women politicians “for change” on *Twitter*. Our first contribution is transforming the theoretical concept of counter-hegemony into a set of measurable variables permitting quantitative analysis of its characteristics. The resulting analysis model helps trace attributes of counter-hegemonic or alternative narratives in communicative strategies, which will help the empiric study of this object in future research.

Secondly, the findings permit us to describe the type of counter-hegemonic discourse put into practice by the Ibero-American political women “of change”. The communication strategy of these political actors reveals an intention to differentiate, criticise and denounce since all hegemonic order is susceptible to being questioned by counter-hegemonic practices that try to dismantle it to undermine existing discourses and practices (Mouffe, 2008). Their messages mainly focus on criticising political and economic elites and denouncing inequalities and social conflicts. This last aspect is substantial. Thus, it is evident that this communication strategy recognises social conflict by giving voice to these problems. In addition, this strategy causes the political women “of change” to position themselves against them as these problems cause inequalities. Therefore, they build a discourse to break the status quo or the hegemonic common sense imposed by the dominant elites (Van-Dijk, 2003).

Another relevant finding is that most women politicians –seven out of ten– guide the practice of counter-hegemony towards creating a critical discourse with proposals for social transformation. Moreover, they have a low tendency to proclaim themselves as protagonists of democratic renewal and subjects of change. Their communication is thus more practical than theoretical. In addition, despite promoting counter-hegemonic strategies, Ibero-American political women “of change” restrict their communication within the margins of traditional politics since they hardly make calls for insurgency or insubordination. Most *Twitter* messages of these political actors maintain a discourse complying with the established channels. There are no invitations, insinuations or declarations in favour of civil disobedience, except for marginal percentages. This shows that they employ a counter-hegemony of a positive or constructive nature, rejecting subversive or disruptive modalities.

A fourth relevant contribution is that there are no distinctions in geopolitical terms. All the analysed counter-hegemony variables cross over among regions. Thus, the issues addressed and the criticism recipients are similar regardless of the territorial component. This reveals that Ibero-American women politicians “of change” share a similar agenda concerning counter-hegemonic communication strategies, despite their geographical origin.

Instead, the position of the political actors in the government-opposition axis offers relevant differences in the activation of counter-hegemonic communication strategies on *Twitter*. Thus, those in government opt for more standardised formulas and fewer alternatives, presenting lower levels of variables related to activism. On the other hand, those in opposition or further away from the institutional framework promote unconventional forms of political participation, demands for social change from governments and criticism of elites and established powers. This way, they more actively promote ways of doing and thinking about politics as a space for building a counter-power (Castells, 2009). With this, they finally contribute to the feminisation of politics from social denunciation to promoting democratic regeneration (Quevedo-Redondo, 2022).

The conclusions show that criticism and denunciation to give voice to the voiceless are the main ingredients of the counter-hegemonic discourse on *Twitter* of Ibero-American political women “of change”. This way, they make inequalities and social conflicts visible. Likewise,



Image 10. Representation of democratic renewal
<https://twitter.com/AdaColau/status/1298967593479483392>

“ The politics analyzed employ a counter-hegemonic discourse of a positive or constructive nature ”

these political actors display a practical and constructive counter-hegemony directed towards its application to social problems in positive terms. Finally, the geographical factor does not cause significant differences in these communication strategies, but the position concerning institutional policy does. This last factor sharpens or minimises operating some of the traits of counter-hegemony in its digital communication.

“The women politicians analyzed pursue exercising another way of doing politics on *Twitter*, focusing on issues closer to the people and their interests”

This research helps identify the characteristics of the counter-hegemonic communication strategies of Ibero-American women “of change”. It represents an original contribution to digital political communication because it allows detecting, measuring and defining alternative discourses that question the established power. However, it has limitations since it leaves pending the analysis of these messages on the public. In addition, the influence of this narrative in the digital conversation needs to be studied. Our approach focuses on content analysis, making novel contributions. However, it does not allow us to offer knowledge of other essential aspects to measure the incidence of counter-hegemonic communication on social networks. New research will be necessary to shed light on this relevant topic.

5. Note

1. A rhizome is a descriptive or epistemological model in which the organization of the elements does not follow lines of hierarchical subordination –with a base or root giving rise to multiple branches–, but rather any element can affect or influence any other.

6. References

Alonso-Muñoz, Laura; Casero-Ripollés, Andreu (2016). “La influencia del discurso sobre cambio social en la agenda de los medios. El caso de la Plataforma de Afectados por la Hipoteca”. *Obets: Revista de ciencias sociales*, v. 11, n. 1, pp. 25-51.

<https://revistaobets.ua.es/article/view/2016-v11-n1-la-influencia-del-discurso-sobre-cambio-social-en-la>

Aruguete, Natalia (2017). “The agenda setting hypothesis in the new media environment”. *Comunicación y sociedad*, n. 28, pp. 35-58.

<https://doi.org/10.32870/cys.v0i28.2929>

Avalos-González, Juan-Manuel (2019). “La posibilidad tecnopolítica. Activismos contemporáneos y dispositivos para la acción. Los casos de las redes feministas y Rexiste”. *Comunicación y sociedad*, v. 16, e7299.

<https://doi.org/10.32870/cys.v2019i0.7299>

Barker-Plummer, Bernadette; Barker-Plummer, David (2017). “*Twitter* as a feminist resource: #YesAllWomen, digital platforms, and discursive social change”. In: Earl, Jennifer; Rohlinger, Deana (eds.). *Social movements and media (Studies in media and communications)*. Emerald Publishing Limited, Bingley, v. 14, pp. 91-118. ISBN: 978 1 78743 097 6

<https://doi.org/10.1108/S2050-206020170000014010>

Cammaerts, Bart (2018). *The circulation of anti-austerity protest*. London: Palgrave Macmillan. ISBN: 978 3 3 1970 1226

Campione, Daniel (2005). “Hegemonía y contrahegemonía en la América Latina de hoy. Apuntes hacia una nueva época”. *Sociohistórica*, n. 17-18, pp. 13-36.

<https://www.sociohistorica.fahce.unlp.edu.ar/article/view/SHn17-18a01>

Casero-Ripollés, Andreu (2009). “El control político de la información periodística”. *Revista latina de comunicación social*, n. 64, pp. 354-366.

<https://doi.org/10.4185/10.4185/RLCS-64-2009-828-354-366>

Casero-Ripollés, Andreu (2015). “Estrategias y prácticas comunicativas del activismo político en las redes sociales en España”. *Historia y comunicación social*, v. 20, n. 2, pp. 535-550.

https://doi.org/10.5209/rev_HICS.2015.v20.n2.51399

Casero-Ripollés, Andreu; Moret-Soler, Diana (2022). “La ciudadanía en la comunicación política digital en España: avances y retos pendientes”. En: García-Orosa, Berta (dir.). *25 años de ciberdemocracia en España: estrategias y actores emergentes. Nuevas formas de comunicación, nativos digitales y actores periféricos*. Thomson Reuters Aranzadi, pp. 43-59. ISBN: 978 84 1124 378 0

Castells, Manuel (2009). *Comunicación y poder*. Madrid: Alianza Editores. ISBN: 978 84 206 8499 4

Chenou, Jean-Marie; Cepeda-Másmela, Carolina (2019). “#NiUnaMenos: Data activism from the global South”. *Television & new media*, v. 20, n. 4, pp. 396-411.

<https://doi.org/10.1177/1527476419828995>

Childs, Sarah (2004). “A feminised style of politics? Women MPs in the House of Commons”. *The British journal of politics and international relations*, v. 6, n. 1, pp. 3-19.

<https://doi.org/10.1111/j.1467-856X.2004.00124.x>

- Colanzi, Irma-Cecilia** (2018). "Narrativas contrahegemónicas: enfoque biográfico e investigación cualitativa". *Ficha de cátedra: Seminario de Psicología Experimental*. Facultad de Psicología, UNL.
<http://sedici.unlp.edu.ar/handle/10915/113187>
- Crossley, Alison** (2015). "Facebook feminism: Social media, blogs, and new technologies of contemporary US feminism". *An international quarterly*, v. 20, n. 2, pp. 253-268.
<https://doi.org/10.17813/1086-671X-20-2-253>
- De-Aguilera, Miguel; Casero-Ripollés, Andreu** (2018). "¿Tecnologías para la transformación? Los medios sociales ante el cambio político y social". *Icono14*, v. 16, n. 1.
<https://doi.org/10.7195/ri14.v16i1.1162>
- De-Sousa-Santos, Boaventura** (2014). "La refundación del estado en América Latina". En: Coraggio, José-Luis; Laville, Jean-Louis (orgs.). *Reinventar la izquierda en el siglo XXI: hacia un diálogo Norte-Sur*. Universidad Nacional de General Sarmiento, pp. 281-298. ISBN: 978 987 630 192 3
- Elbaum, Jorge** (1997). "Antonio Gramsci: optimismo de la voluntad y pesimismo de la razón". En: *Cuadernos de comunicación y cultura para teoría y práctica de la comunicación. Elementos para el análisis. Cátedra Aníbal Ford. Cuaderno*, 54. Universidad de Buenos Aires, Argentina.
- Fairclough, Norman** (2010). *Critical discourse analysis. The critical study of language*. New York: Routledge. ISBN: 978 1 4058 5822 9
- Feenstra, Ramón A.; Tormey, Simón; Casero-Ripollés, Andreu; Keane, John** (2016). *La reconfiguración de la democracia: el laboratorio político español*. Granada: Editorial Comares. ISBN: 978 84 9045 441 1
- Freidenberg, Flavia; Lajas-García, Sara** (2015). *De la cuota a la paridad: las reformas para mejorar la representación política de las mujeres en América Latina*. Observatorio de Reformas Políticas en América Latina. Washington, D.C.: Secretaría de Asuntos Políticos de la Organización de los Estados Americanos (SAP/OEA).
<https://reformaspoliticas.org/wp-content/uploads/2019/09/DT-11.pdf>
- Friedman, Elisabeth J.** (2003). "Gendering the agenda: the impact of the transnational women's rights movement at the UN conferences of the 1990s". *Women's studies international forum*, v. 26, n. 4, pp. 313-331.
[https://doi.org/10.1016/S0277-5395\(03\)00077-3](https://doi.org/10.1016/S0277-5395(03)00077-3)
- Fuchs, Christian** (2014). *Social media: a critical introduction*. London: Sage Publications Ltd. ISBN: 978 1 4462 5731 9
<https://doi.org/10.4135/9781446270066>
- García-Canclini, Néstor** (1984). "Gramsci con Bourdieu. Hegemonía, consumo y nuevas formas de organización popular". *Nueva sociedad*, v. 71, pp. 69-78.
https://static.nuso.org/media/articles/downloads/1156_1.pdf
- Gramsci, Antonio** (1972). *Introducción a la filosofía de la praxis*. Ediciones Península. ISBN: 978 8 96875 54 8
- Herrnson, Paul S.; Lay, J. Celeste; Stokes, Atiya-Kai** (2003). "Women running 'as women': Candidate gender, campaign issues, and voter-targeting strategies". *Journal of politics*, v. 65, n. 1, pp. 244-255.
<https://doi.org/10.1111/1468-2508.t01-1-00013>
- Huddy, Leonie; Terkildsen, Nayda** (1993). "Gender stereotypes and the perception of male and female candidates". *American journal of political science*, v. 37, n. 1, pp. 119-147.
<https://doi.org/10.2307/2111526>
- Jaffe, Sarah** (2018). "The collective power of #MeToo". *Dissent*, v. 65, n. 2, pp. 80-87.
<https://doi.org/10.1353/dss.2018.0031>
- Kavada, Anastasia** (2016). "Social movements and political agency in the digital age: A communication approach". *Media and communication*, v. 4, n. 4, pp. 8-12.
<https://doi.org/10.17645/mac.v4i4.691>
- Krippendorff, Klaus** (2013) *Content analysis: an introduction to its methodology*. 3rd ed. California: SAGE Publications. ISBN: 978 1 412983150
- Larrondo-Ureta, Ainara; Morales-i-Gras, Jordi; Orbezo-Terradillos, Julen** (2019). "Feminist hashtag activism in Spain: measuring the degree of politicisation of online discourse on #YoSíTeCreo, #HermanaYoSíTeCreo, #Cuéntalo y #NoEstás-Sola". *Communication and society*, v. 32, n. 4, pp. 207-221.
<https://doi.org/10.15581/003.32.34734>
- Larrondo-Ureta, Ainara; Orbezo-Terradillos, Julen; Peña-Fernández, Simón** (2019). "Palabra de candidata. La limitada influencia del género en el discurso electoral en Twitter". *Comunicación y medios*, v. 28, n. 40, pp. 114-126.
<https://doi.org/10.5354/0719-1529.2019.53046>

- Laudano, Claudia-Nora** (2018). “Aborto y redes: el debate por #AbortoLegal”. *Sociales en debate*, n. 14. http://www.memoria.fahce.unlp.edu.ar/art_revistas/pr.12084/pr.12084.pdf
- Lawless, Jennifer L.** (2004). “Women, war, and winning elections: Gender stereotyping in the post-September 11th era”. *Political research quarterly*, v. 57, n. 3, pp. 479-490. <https://doi.org/10.1177/106591290405700312>
- Martin, Courtney E.; Valentí, Vanessa** (2013). *#FemFuture: Online revolution*. New York: New Feminist Solutions, Barnard Center for Research on Women. <http://bcrw.barnard.edu/wp-content/nfs/reports/NFS8-FemFuture-Online-Revolution-Report-April-15-2013.pdf>
- Moraña, Mabel** (1998). “El boom del subalterno”. En: Castro-Gómez, Santiago; Mendieta, Eduardo (eds.). *Teorías sin disciplina. Latinoamericanismo, postcolonialidad y globalización en debate*. México: Miguel Ángel Porrúa. ISBN: 9789 688428092
- Mouffe, Chantal** (2008). “Critique as counter-hegemonic intervention”. *Transversal multilingual web journal*. <https://transversal.at/transversal/0808/mouffe/en>
- Núñez-Puente, Sonia** (2011). “Activismo y colectivos en red: praxis feminista online y violencia de género”. *Asparkia: investigació feminista*, n. 22, pp. 85-98. <https://www.e-revistas.uji.es/index.php/asparkia/article/view/600>
- Núñez-Puente, Sonia; Vázquez-Cupeiro, Susana; Fernández-Romero, Diana** (2016). “Ciberfeminismo contra la violencia de género: análisis del activismo online-offline y de la representación discursiva de la víctima”. *Estudios sobre el mensaje periodístico*, v. 22, n. 2, pp. 861-877. <https://doi.org/10.5209/ESMP.54240>
- Ochoa, Edurne** (2021). “Feminización de la política”. *Pluralidad y consenso*, v. 11, n. 47, pp. 32-35. <http://revista.ibd.senado.gob.mx/index.php/PluralidadyConsenso/article/view/712>
- Orbegozo-Terradillos, Julen; Larrondo-Ureta, Ainara; Morales-i-Gras, Jordi** (2020). “Influencia del género en los debates electorales en España: análisis de la audiencia social en #ElDebateDecisivo y #L6Neldebate”. *El profesional de la información*, v. 29, n. 2, e290209. <https://doi.org/10.3145/epi.2020.mar.09>
- Piñuel-Raigada, José-Luis** (2002). “Epistemología, metodología y técnicas del análisis de contenido”. *Estudios de socio-lingüística*, v. 3, n. 1. <https://doi.org/10.1558/sols.v3i1.1>
- Quevedo-Redondo, Raquel** (2021). “El estilo retórico femenino en la entrevista política. Una década de aplicación en *Telva*”. *Index.comunicación*, v. 11, n. 1, pp. 271-295. <https://doi.org/10.33732/ixc/11/01Elesti>
- Quevedo-Redondo, Raquel** (2022). “La feminización política y su extrapolación al discurso de líderes y lideresas. Estudio de caso de Yolanda Díaz”. *IC. Revista científica de información y comunicación*, n. 19, pp. 165-188. <https://icjournal-ojs.org/index.php/IC-Journal/article/view/713>
- Rodríguez, Victoria E.** (1999). “Las mujeres y el cambio político en México”. *Revista de estudios de género: La ventana*, v. 1, n. 10, pp. 50-83. <http://www.redalyc.org/articulo.oa?id=88411129004>
- Rodríguez-García, María-Jesús; Navarro-Yáñez, Clemente J.** (2012). “La feminización de la dinámica política municipal. El caso de los municipios españoles”. *Revista internacional de sociología*, v. 70, n. 1, pp. 181-201. <https://doi.org/10.3989/ris.2010.09.11>
- Rosales, María-Belén** (2018). “Ciberactivismo: praxis feminista y visibilidad política en #NiUnaMenos”. *Pléyade (Santiago)*, n. 22, pp. 63-85. <https://doi.org/10.4067/S0719-36962018000200063>
- Sampedro, Víctor** (2014). *El cuarto poder en red: Por un periodismo (de código) libre*. Barcelona: Icaria. ISBN: 978 84 9888 5903
- Treré, Emiliano** (2016). “Distorsiones tecnopolíticas: represión y resistencia algorítmica del activismo ciudadano en la era del big data”. *Trípodos*, n. 39, pp. 35-51. http://www.tripodos.com/index.php/Facultat_Comunicacio_Blanquerna/article/view/379
- Turley, Emma; Fisher, Jenny** (2018). “Tweeting back while shouting back: Social media and feminist activism”. *Feminism & psychology*, v. 28, n. 1, pp. 128-132. <https://doi.org/10.1177/0959353517715875>

Van-Dijk, Teun A. (1993). "Principles of critical discourse analysis". *Discourse & society*, v. 4, n. 2, pp. 249-283.
<https://doi.org/10.1177/0957926593004002006>

Van-Dijk, Teun A. (2003). *Racismo y discurso de las élites*. Barcelona: Gedisa. ISBN: 978 84 74328431

Van-Dijk, Teun A. (2006). "Ideology and discourse analysis". *Journal of political ideologies*, v. 11, n. 2, pp. 115-140.
<https://doi.org/10.1080/13569310600687908>

Van-Dijk, Teun A. (2012). *Discurso y poder*. Barcelona: Gedisa. ISBN: 978 84 97842822

Verge-Mestre, Tània; Tormos-Marín, Raúl (2012). "La persistencia de las diferencias de género en el interés por la política". *Revista española de investigaciones sociológicas*, v. 20 n. 138, pp. 89-108.
<https://doi.org/10.5477/cis/reis.138.89>

Vinuesa-Tejero, María-Lourdes; Abejón-Mendoza, Paloma; Sánchez-Calero, María-Luisa (2011). "Mujeres y política: un binomio con baja representación". *Ámbitos*, n. 20, pp. 127-145.
<http://hdl.handle.net/11441/67584>
<https://doi.org/10.12795/ambitos.2011.i20.07>

Williams, Raymond (1980). *Marxismo y literatura*. Editorial Península. ISBN: 978 987 1501199

Anuario ThinkEPI

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



#EspañalInvadida. Disinformation and hate speech towards refugees on *Twitter*: A challenge for critical thinking

Abdellah Essalhi-Rakrak; Ruth Pinedo-González

Nota: Este artículo se puede leer en español en:
<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87188>

Recommended citation:

Essalhi-Rakrak, Abdellah; Pinedo-González, Ruth (2023). “#EspañalInvadida. Disinformation and hate speech towards refugees on *Twitter*: A challenge for critical thinking”. *Profesional de la información*, v. 32, n. 3, e320310.
<https://doi.org/10.3145/epi.2023.may.10>

Manuscript received on 20th November 2022
Accepted on 23rd January 2023



Abdellah Essalhi-Rakrak ✉
<https://orcid.org/0000-0002-3935-0935>

Universidad de Valladolid
Campus María Zambrano
Pl. de la Universidad, 1
40005 Segovia, Spain
abdellah.essalhi@uva.es



Ruth Pinedo-González
<https://orcid.org/0000-0002-4699-9433>

Universidad de Valladolid
Departamento de Psicología
Campus María Zambrano
Pl. de la Universidad, 1
40005 Segovia, Spain
ruth.pinedo@uva.es

Abstract

Disinformation is not a new phenomenon, but it is widespread in our society since social media have become a media loudspeaker. The outbreak of war in Ukraine has produced a social debate, partly reflected on social networks, about the treatment of Ukrainian refugees compared with other refugees from the South. For this reason, this research proposes a study that follows a qualitative interpretative methodology with an exploratory and descriptive scope that analyses in depth the content of the discourse on *Twitter* about refugees and uses computer-mediated discourse analysis as a technique for obtaining data. Specifically, it compares the content published under the hashtags #NotRefugees and #Refugees, and for this purpose, a data analysis has been carried out using *Atlas.ti*. The results show a differentiated treatment between refugees from the North and the South, as well as a propagation of misinformation through hate speech, in which a favourable and welcoming treatment of Ukrainian refugees is shown, while various types of hate speech towards the rest of the refugees from the South are evident. Hate towards refugees from the South on *Twitter* is justified on the grounds that they threaten national security, economy, and identity. This phenomenon reflects the urgent need to develop critical and ethical digital competence in the face of disinformation in social networks and media, where the development of critical and ethical thinking is essential.

Keywords

Disinformation; Hate speech; *Twitter*; Refugees; Critical thinking; Digital competence; Media literacy; Educommunication; Qualitative methodology; Grounded theory; Content analysis.

Funding

This study has been carried out within the framework of the research project “Truth and ethics in social networks. Perceptions and educational influences in young users of *Twitter*, *Instagram*, and *YouTube* (Internética),” funded by the call for R+D+i projects of the Spanish *Ministry of Science, Innovation, and Universities* 2019 (PID2019-104689RB-I00).

Acknowledgment

Our thanks to Alfonso Gutiérrez Martín and Susana Álvarez Álvarez for their recommendations and comments.



1. Introduction

This study aims to analyse the content disseminated about refugees on the social network *Twitter* to delve deeper into the discursive typologies related to this group. To this end, computer-mediated discourse analysis is used to obtain data. This technique is useful and innovative from a scientific point of view because it analyses the linguistic and pragmatic properties of the content of the interactions on this social network where what prevails is communicative immediacy and the use of colloquialisms (Mancera-Rueda; Pano-Alamán, 2014).

This research is especially important in the Spanish context since, due to its geographical location, there is a constant migratory flow with characteristics that differ from those in other European countries. Furthermore, this phenomenon is a subject of constant public debate, both on social networks and in the conventional media, which has been used on numerous occasions to misinform and manipulate public opinion for political purposes (Lava-Santos, 2021). The following study objectives are proposed:

- to describe the characteristics attributed to refugees on *Twitter*, and
- to analyse the discursive typologies about refugees on *Twitter*.

2. Review of scientific literature

2.1. Communicative misinformation: Anything goes to get your message across and convince

Society is undergoing a major paradigm shift. The so-called knowledge and communication society is giving way to a society characterised by disinformation and 'infocination' –information overload. Hoaxes, information excess, and fake news are the main characteristics that define society's access to information in the media (Amores; Arcila-Calderón; Blanco-Herrero, 2020; Magallón-Rosa, 2021).

The veracity of the media is in crisis, especially at a time when society is immersed in a crisis of values, where instability and social polarisation are the order of the day (Belmonte; Muñoz-Álvarez; Bernárdez-Gómez, 2022). As Pari-Tito, García-Peñalvo and Pérez-Postigo (2022) state, the viral spread of fake news has become one of the great challenges of 21st century communication. The lack of effective mechanisms or strategies to curb its propagation makes it a challenging problem to tackle.

The darkest side of this phenomenon was experienced during the COVID-19 pandemic. Anti-vaccine, conspiratorial, and reactionary discourses posed real risks to health and social welfare because of the virality they achieved (Larrondo-Ureña; Peña-Fernández; Morales-i-Grass, 2021). This expansion of misinformation cannot be explained without considering the potential of its great ally: social networks. As Pari-Tito, García-Peñalvo and Pérez-Postigo (2022) state, even though social networks were created to be spaces for communication between people, they have become one of the main means of exchange of information and news.

Social media possess characteristics that make them perfect allies for misinformation, such as ease of access, immediacy and simplicity of use, and user anonymity (Corbella-Ruiz; De-Juanas-Oliva, 2013; Gamir-Ríos; Tarrullo, 2022). These qualities allow unverified content to go viral as consumers of this information interact with it and share it without effort or analysis.

Alonso-González (2019) states that these characteristics of social media have created a new rhythm of information, i.e., rumours and unverified news that have a minimum of logic begin to spread without any type of control. Likewise, the very design of these tools is so simple and intuitive that they make it easy to acquire a basic knowledge of how to use them, both in their *consumer* and *prosumer* aspects, i.e., users not only consume content but are also content creators and distributors (Martínez-Sala; Segarra-Saavedra; Monserrat-Gauchi, 2018).

The social instability caused by the COVID-19 pandemic was followed by the Russian invasion of Ukraine. The latter has led to a further deterioration in areas such as the economy, the political climate, and social coexistence but has undoubtedly also led to a rise in disinformation (Magallón-Rosa, 2021). The war in Ukraine has made it clear that disinformation is a weapon of war because lies change society's perceptions and have been an instrument of war used since the beginning of the conflict (Arrieta; Montes, 2011). Furthermore, disinformation has been both a defensive and offensive weapon. In short, as Navarrete-Barrutia (2020) asserts in his study on foreigners in Spain, disinformation is an ideal instrument for promoting falsehoods and influencing collective consciousness.

2.2. Hate speech on social media: The scapegoat for our ills

Hiding the truth, manipulating it, or camouflaging it is a phenomenon that has been used in wars and conflicts since the earliest civilisations, but it was during the Second World War that an infrastructure for the organisation and planning of propaganda aimed at manipulating public opinion first appeared (Vergani *et al.*, 2022). Disinformation is a phenomenon that destabilises and polarises society; lies spread through hoaxes and fake news change citizens' perceptions, and society becomes a victim of this manipulation (Badillo-Matos, 2019; Arcila-Calderón *et al.*, 2022).

Disinformation is a phenomenon that destabilises and polarises society; lies spread through hoaxes and fake news change citizens' perceptions

On other occasions, the type of framing offered in a news item has a series of cognitive effects on the reader. **Igartua et al.** (2008) found examples of framing of news about immigration that linked it to the phenomenon of crime (**Igartua et al.**, 2007). Hate speech in social media has been found to promote discrimination against certain groups in society (**Arrieta; Montes**, 2011).

“ Hate speech seeks to harm the dignity of a person or group based on a number of defining characteristics, both inherited and acquired ”

Hate speech can be defined as a form of expression that openly promotes, justifies, and disseminates the disparagement and exclusion of certain social groups based on ethnicity, nationality, religion, sexual orientation, gender, disability, and other characteristics (**Amores; Arcila-Calderón; Blanco-Herrero**, 2020; **Wachs et al.**, 2022). Another definition to consider is that put forward by the European project *Preventing Hate Against Refugees and Migrants (Pharm)*, which aims to prevent hate speech against refugees and migrants. This project defines hate speech as speech that seeks to harm the very dignity of a person or group based on a number of defining characteristics, both inherited and acquired (**Pharm**, 2019).

This project warns of the importance of cyberhate, since, as research has corroborated, there is a clear relationship between hate speech on the internet and the growth of crimes against the groups towards which this hate is directed. Various elements explain the growth of this phenomenon, where the discursive change that has taken place in the public and political sphere is important. In addition, the negative perception of immigration has increased, with people being warned on many occasions that immigrants have links to terrorism (**Pharm**, 2019; **Sánchez-Holgado; Amores; Blanco-Herrero**, 2022).

The *FAD Foundation (FAD)*, (2022) suggests that there are five mechanisms for carrying out the ‘fabrication of hate’:

- creating or exaggerating fear towards a group without justified evidence;
- pointing out supposed culprits who, in fact, have nothing to do with wrongdoing;
- creating unjustified fear and amplifying it until it becomes a great threat;
- predicting the appearance of a great leader appear who will ‘save’ us; and
- acting against the ‘enemy’ without any scruples.

However, it should not be forgotten that all such incitements to hatred, hostility, discrimination, and violence against groups or individuals based on race, ideology, religion, sexual orientation, gender, or disability grounds may constitute crimes in Spain (Article 22.4 of the *Penal Code*) and throughout Europe.

Prejudices, as part of the phenomenon of the fabrication of hate, are of great importance as they form the basis of the discourse and become facilitators of the commission of hate crimes (**Müller; Schwartz**, 2021). Racism and xenophobia cause the most hate crimes in Spain, according to the latest annual study of the *National Office for Combating Hate Crimes* (2022) (**Ondod** hereinafter). Disinformation is a key factor in the promotion of hate speech, especially at times of great socio-economic instability, which encourage the search for scapegoats for the causes of society’s ills. Blame is often placed on migrants through fake news and disinformation (**Narváez-Linares; Pérez-Rufi**, 2022).

As **Álvarez** (2008) observed, manipulation, lies, and power are mechanisms which can be used by the media and which must be addressed through media education of citizens (**Gutiérrez-Martín; Pinedo-González; Gil-Puente**, 2022; **Vuorikari; Kluzer; Punie**, 2022).

2.3. Educommunication: Critical and ethical thinking in the use of the media

Citizens are defenceless if they do not develop the skills, knowledge, and understanding that facilitate a critical and ethical use of media in general and social networks in particular (**Vuorikari; Kluzer; Punie**, 2022). Educommunication becomes part of the basic training or literacy that must be accessible to all citizens so that they can access and use information critically and create media content responsibly and safely. This learning cannot be limited to the simple use of tools and technologies but must aspire to provide the capacity for critical and ethical thinking in the use of information and media (**Gutiérrez-Martín; Pinedo-González; Gil-Puente**, 2022; **Vuorikari; Kluzer; Punie**, 2022).

As **Gutiérrez-Martín** and **Tyner** (2012) argue, this media education should not be reduced to an understanding of the forms of personal or social communication but rather is essential to the development of a critical education in relation to these communicative processes. From the field of information and communication, it is necessary to understand the need for society to develop critical thinking in the use of information and communication technology (ICT) and media, as well as ethical behaviour as responsible and competent citizens. Although this need to develop so-called digital competence in citizens is often mentioned, this training tends to be reduced to the more instrumental and technical dimensions of ICT (**Gutiérrez-Martín; Tyner**, 2012; **Unesco**, 2021).

“ Racism and xenophobia cause the most hate crimes in Spain, according to the latest annual study of the *National Office for Combating Hate Crimes* ”

Consequently, information literacy needs to be developed to access and critically evaluate information and make ethical use of information throughout the process (Wilson *et al.*, 2011). Media literacy is also needed to understand the workings of the media and its economic and ideological interests (Wilson *et al.*, 2011).

“Citizens need critical and ethical training in the use of the internet, especially social networks”

Authors such as Arrieta and Montes (2011) define digital literacy from three perspectives:

- the use of technology;
- critical understanding when using it; and
- the ability to create and communicate digital concepts in a wide variety of formats.

We focus here on the principle of critical understanding, defined as

“the ability to understand, contextualise and critically evaluate the digital media and content with which one interacts” (Montes; Arrieta, 2011, p. 187).

The role of media education should not be limited to instrumental training based on the use of digital media as a medium. An urgent need exists for a critical education of citizens to provide them with the ability and the means to select and contrast the information they access (Zaragoza-Lorca, 2007). The framework of digital competences for citizenship developed by Vuorikari, Kluzer and Punie (2022) within the framework of the *European Commission (DigComp 2.2)* is a necessary and complete approach in its theoretical bases. This update of the DigComp model provides a framework in which critical implementation is crucial in moving away from a purely instrumental perspective on digital competence. The model is made up of a total of 21 competences, grouped into five large blocks:

- search and management of information and data;
- communication and collaboration;
- creation of digital content;
- security; and
- problem solving.

These competences reflect the development of digital competence oriented towards not only technological management but also the importance of internet security and the need for critical thinking in the face of phenomena such as disinformation and fake news.

In short, citizens need critical and ethical training in the use of the internet, especially social networks, to combat phenomena such as disinformation that favour social polarisation and negatively affect social coexistence, especially through the promotion of hate speech (Álvaro-Sánchez, 2018).

3. Method

This study follows an interpretative qualitative methodology with an exploratory and descriptive scope (Hernández-Sampieri; Fernández-Collado; Baptista-Lucio, 2018), which uses computer-mediated discourse analysis as a data collection technique (Pano-Alamán; Moya-Muñoz, 2016; Cantamutto; Vela-Delfa, 2023). This research is structured in four phases:

- a conceptual phase,
- an empirical phase,
- a data analysis phase, and
- an inferential phase (see Figure 1).

The design of the study’s qualitative methodology is a systematic approach based on grounded theory, which is notable for meticulously following an inductive coding process to arrive at a theory for the phenomenon under study (Jorrín-Abellán; Fontana-Abad; Rubia-Avi, 2020; Hernández-Sampieri; Fernández-Collado; Baptista-Lucio, 2018). A descriptive analysis of the main categories in a sample of analysed tweets was also carried out (Gutiérrez-Braojos *et al.*, 2017).

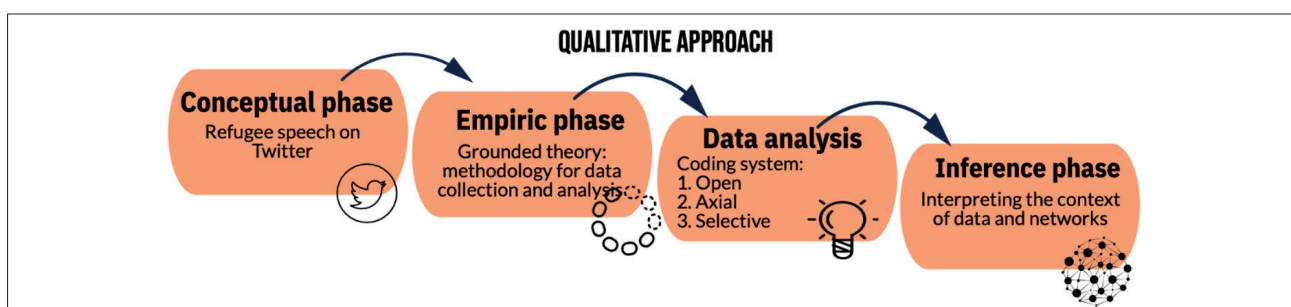


Figure 1. Qualitative research design

3.1. Sampling

This study follows a non-probabilistic convenience sampling approach with two criteria:

- the selection of tweets that fit the research objectives in belonging to a clear discourse on refugees; and
- the establishment of criteria for the interaction of the tweets to select those that have generated at least a minimal level of engagement among *Twitter* users.

These criteria are the number of likes (minimum of 10), comments (minimum of five) and retweets (minimum of three).

Data collection was carried out during February and March 2022. A sample of 344 tweets was obtained using the hashtags #NoSonRefugiados and #SonRefugiados.

3.2. Procedure

First, a systematic literature search on hate speech in social media was carried out. This search allowed us to identify previous studies of interest in the field and identify various research opportunities.

Second, data collection was carried out using the *Tweet Archivist* programme, which allowed the creation of a database of tweets that have been published under the hashtags #NoSonRefugiados and #Refugiados. The social network *Twitter* was selected because its characteristics are congruent with the design of this research:

- it allows any user to access tweets that are posted openly;
- the content of the tweets can be textual, images, or videos, so an in-depth analysis of the messages intended to be conveyed can be conducted;
- hashtags are widely used in this social network (hashtags are texts preceded by the # symbol that help to categorise the subject matter of the tweets, expand the message, and find or detect tweets on a particular topic); and
- *Twitter* reports data that help to quantify the relevance and impact of a tweet, such as the number of likes, comments, and retweets, characteristics that have been taken into account in the selection of the tweets.

These hashtags were selected because of the existence of debates among users about the treatment of refugees, on the one hand, in favour of equal action in the face of the phenomenon (#Refugees) and, on the other hand, the justification of why both groups cannot be treated equally (#NotRefugees).

The qualitative analysis of the data consisted of the following steps (Strauss; Corbin, 2002):

- detailed reading of the data and
- coding of the data based on thematic areas relevant to the study in order to deepen the understanding of the phenomenon under analysis: the discourse towards refugees on *Twitter*.

3.3. Data analysis

The software *Atlas.ti* was used to carry out the data analysis. Figure 2 shows the coding process, which consisted of open, axial, and selective components (Strauss; Corbin, 2002). In this coding process, emerging categories were also detected and defined, as well as memos, comments, and quotes that have allowed a central category to be established. From this, hypotheses and theory development were generated.

As Figure 2 shows, during the open coding process, categories and subcategories are detected from the emerging codes. To facilitate their understanding in the networks presented in the Results section, Table 1, in which the syntax and label of each of the categorical levels are defined, has been elaborated.

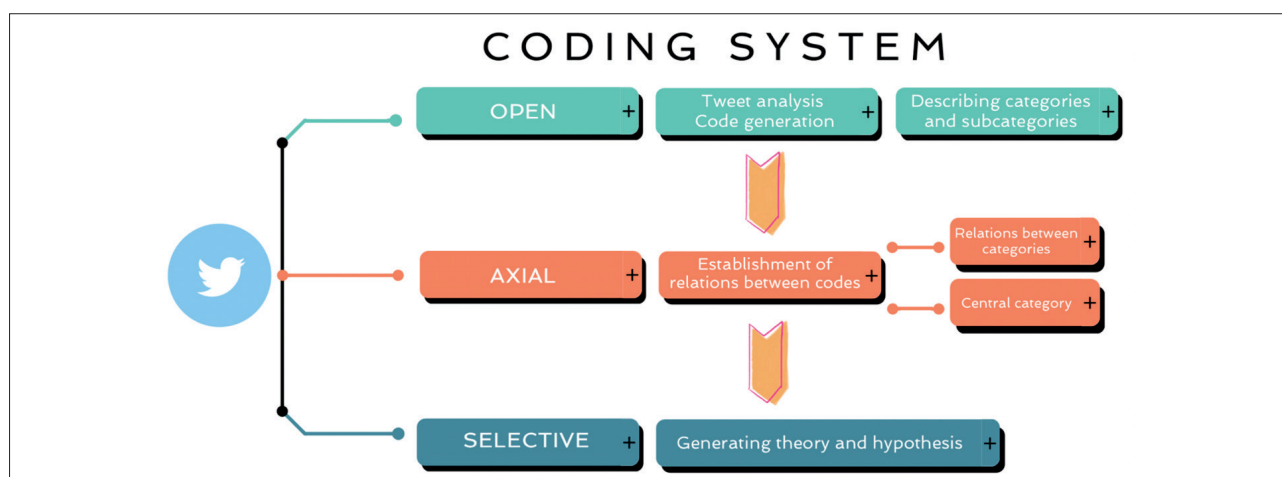
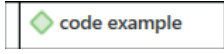

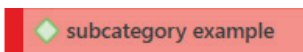


Figure 2. Qualitative data analysis process

Table 1. Categorical levels into which codes, subcategories, and categories are organised

Levels	Syntax for labelling code	Representative example
Code	Minuscule	
Category	Upper case, with colour	
Subcategory	Lower case, same colour as category	

4. Results and discussion

The data analysis follows the coding process (open, axial, and selective) established by grounded theory (Strauss; Corbin; 2002). After elaborating the open coding, two phases were developed in the axial part: a first phase in which the relationships between codes were established considering the research objectives and a second, axial-selective phase, in which the emerging theory of the research carried out was generated (San-Martín-Cantero, 2014).

By means of open coding (an inductive process), we sought to extract the content of the tweets analysed and thus describe the ideas, thoughts, and meanings they reveal (Strauss; Corbin, 2002). This first content analysis process resulted in the extraction of 66 codes. Given the large number of codes, it was decided to develop code groups to facilitate the categorisation process. After comparing the codes with respect to their properties and meanings, a summary of generic concepts and their definitions (see table 2), drawn from the comments elaborated in the analysis process, has been elaborated.

Table 2. Codes, categories, and definitions derived from the categorisation process

Codes	Categories	Definitions of the categories
Aid, Ease of access to aid, Family, Humanitarian aid, Humanity, Order, Refugees solidarity.	Positive attitudes and values.	Attitudes and values.
Disorder, Warning, Conquest of Spain, Criminals, Hostile, Invasion, Invasion; Irony, Armed struggle, Profit, Lies, They are not welcome, They are not refugees, Danger, Looters, Supremacism, Terrorism.	Negative attitudes and values.	Set of beliefs, convictions and conceptions that have been detected in the codes of the analysed tweets.
Tenderness, Joy, Love, Help, Sorrow.	Positive feelings and emotions.	Feelings and emotions: Set of feelings and emotions detected in the content of the tweets issued by users in response to the refugee phenomenon.
Hostile, Helplessness, Rage, Fear, Sorrow, Anger.	Negative feelings and emotions.	
Blondes, Baby, White, Christian, Family, Female, Women and girls, Refugees.	Ukrainian refugees.	Refugee characteristics: A set of traits or characteristics associated with refugees.
Criminals, Butchers, Men.	Other refugees.	
Weapons, Trojan Horse, Butchers, Conquest of Spain, Criminals, Disorder, Islamic State, War, Illegal immigration, Invasion, ISIS, Islamists, Mass arrival, Armed struggle, Mafia, Mob, Crowd, Danger, Refugees vs Invasion, Border security, Terrorism, Violence.	National security reasons.	Reasons why hatred is expressed: Hate speech towards refugees from the South as a global concept has generated different types, classified into subcategories. Each shows a motive or reason for justifying hatred towards refugees.
Criminals, Living life without working, Profit, Looters.	Problem for the national economy.	
Crusade against Islam, Invasion, Nation.	Nationalist exaltation.	
Apes, Whites vs Blacks, Blacks.	Racist.	The subcategory of "racism" has a peculiarity in that it is not a justification as such, but simply hate speech based on racial adjectives.
Left politicians, Government of Spain.	Spanish organisations supporting immigration.	
ISIS, Islamists, Muslims, Crusade against Islam, Islamic State, Christian.	Hatred against religion.	

This process provides a starting point for the next phase: axial coding. At this point, the aim is to identify the relationships between the categories and subcategories obtained from the analysis of the tweets in the open coding (San-Martín-Cantero, 2014). According to Strauss and Corbin (2002), it is important to bear in mind that categories and subcategories represent a phenomenon. Therefore, through the emerging categories, the phenomenon of the discourse on *Twitter* about refugees is developed in parallel.

Axial coding was carried out in two phases: (1) establishment of the type of relationship between the codes and their subcategories and categories (this process is based on the table obtained from the open coding) and (2) development of the relationship between the categories and subcategories. Comments and memos elaborated throughout the qualitative analysis process are considered. At this point, the elaboration of networks is carried out. Their presentation is organised according to the objectives of this research.

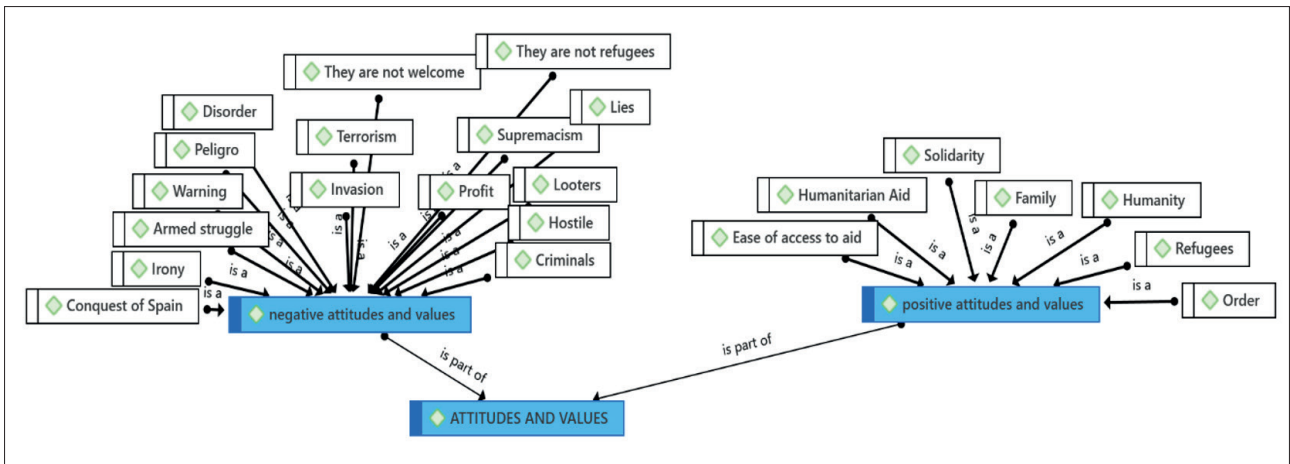


Figure 3. Attitudes and values network

4.1. Attitudes and values towards refugees

First, and following the inductive process of grounded theory outlined by **San-Martín-Cantero** (2014), the first network is obtained (see Figure 3). It shows how codes are connected to positive or negative attitudes and values (subcategories) through an ‘is a’ relationship. It can then be seen that both subcategories have an ‘is a part of’ relationship. This shows that both types are part of attitudes and values as a general concept. This dual structuring shows a division in terms of positive and negative values and attitudes. It is posed as a way of polarising and dividing society into good and bad, with the latter being identified as a threat (i.e., invasion, disorder, danger). This content is related to what was previously stated by **FAD** (2022) concerning negative values and attitudes helping to create a great threat, pointing out culprits that have nothing to do with any wrongdoing, and creating an unjustified fear that becomes a threat or enemy against which we must direct ourselves. The network codes (see Figure 3) are evidence of this (**FAD**, 2022).

4.2. Feelings and emotions towards refugees

The second network (see Figure 4) maintains the same structure as the previous one. All codes distinguish between negative and positive emotions and feelings, although, for example, ‘grief’ can be linked to both since it is expressed in both senses. The codes have a relationship of possessions to their subcategories. These are linked as part of the first-level category, ‘emotions and feelings’.

This content clearly reflects the importance of emotional and sentimental content on social networks. In this regard, **Martínez-Rodrigo, Segura-García and Sánchez-Martín** (2011) observe that there are three reasons why emotions are so predominant in social networks. The first is immediacy as a dynamic in the functioning of the networks, where emotional content permeates the consumer quickly and without reflection. The second is participation, which arises from the rapid interaction between users, which invites rapid communication. The third is personalisation, as people understand social media as a space for free expression. This last conception of networks as an area where freedom of expression has no limits is related to the conception of unreality in their use (**Bustos-Martínez et al.**, 2019). All of this reflects a perception of impunity, translated into anonymous profiles that make the propagation of hate speech even more feasible (**Blanco-Alfonso; Rodríguez-Fernández; Arce-García**, 2022).

Consequently, the use of emotional content reinforces the impact of tweets on users. This effect is also detected in the use of visual content as a way of reinforcing and making the message visible, an element that has also been detected in this study, as shown in Figures 5 and 6. In Figure 5, we can read a tweet and see the image that accompanies it; it can be seen that it conveys negative emotional content, e.g., fear, hostility, or anger. In contrast to this tweet, we find

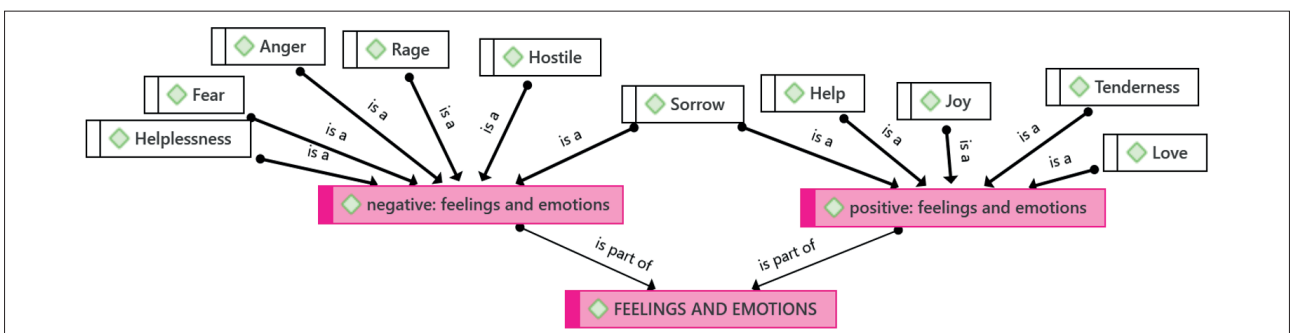


Figure 4. Network of emotions and feelings

Figure 6, where both the written text and the image aim to convey opposing emotions. On the one hand are positive emotions towards the women and children from Ukraine, provoking tenderness, help, etc., and are the opposite emotions, such as fear and hostility.

This use of visual content as a means of impact corroborates the hypothesis put forward by researchers such as **Gamir-Ríos, Tarullo and Ibáñez-Cuquerella (2021)**, who argue that audiovisual content is a great mechanism for promoting disinformation by increasing the possibility of penetrating deeper into the social subconscious. **Larrosa (2007)** defends the need to provide citizens with tools to understand the visual world, given its great potential to promote disinformation and hate speech.

This constant appeal to emotions in the content awakens a collective perception that leads to misinformation and social polarisation (**Castillo-de-Mesa et al., 2021**). Research by **Molina-Cañabate and Magallón-Rosa (2019)** suggests that information based on the promotion of hate speech on *Twitter* can arouse a series of irrational and emotional impulses that contribute to increasing the virality and visibility of such speeches.

These highly emotional discourses make it impossible for information verification agencies, such as *Maldita.es*, to intervene. This is because they are a type of message that is not subject to verification, connected to emotion and opinion, a channel through which disinformation and hate speech can be transmitted.

4.3. Refugee characteristics

In the third network (see Figure 7), the following process of relating categories emerges. The codes are associated with existing types of refugees. This shows a clear segmentation according to refugee type. Consequently, both refugee typologies have an oppositional connection, as they represent two different phenomena in terms of characteristics. Nevertheless, both subcategories are part of ('is part of') the category 'refugee characteristics'. Therefore, this network reflects how, in *Twitter* discourses, distinct qualities are assigned to each type of refugee.

This network shows that the discourse analysed on *Twitter* shows a differentiated treatment between two typologies of refugees: on the one hand, those coming from Ukraine and, on the other hand, those coming from other places, especially from the south. The latter are criminalised and seen as 'butchers' and 'criminals', while refugees from Ukraine are viewed more positively, with characteristics such as 'family' and 'women and girls'. **Valdez-Apolo, Arcila-Calderón and Amores (2019)** conclude that immigrant status is used in the media to discredit or disqualify certain refugees.



Figure 5. Tweet using an image to reinforce the message. Source: *Twitter* <https://twitter.com/Sigriddmrd/status/1432818999759306757>

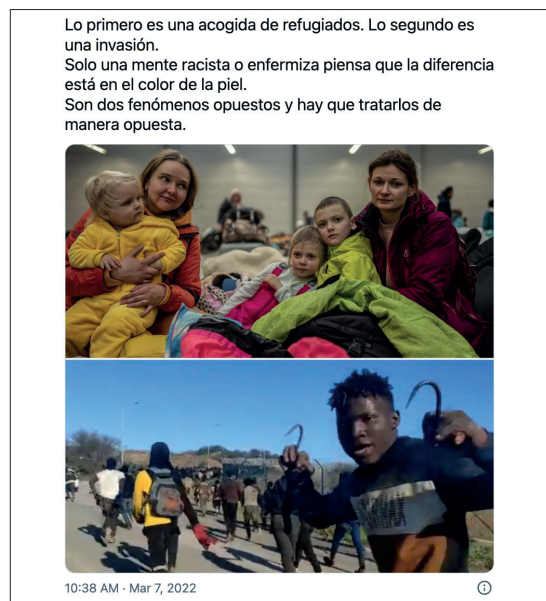


Figure 6. Tweet using images to reinforce a message. Source: *Twitter* <https://onx.la/d197d>

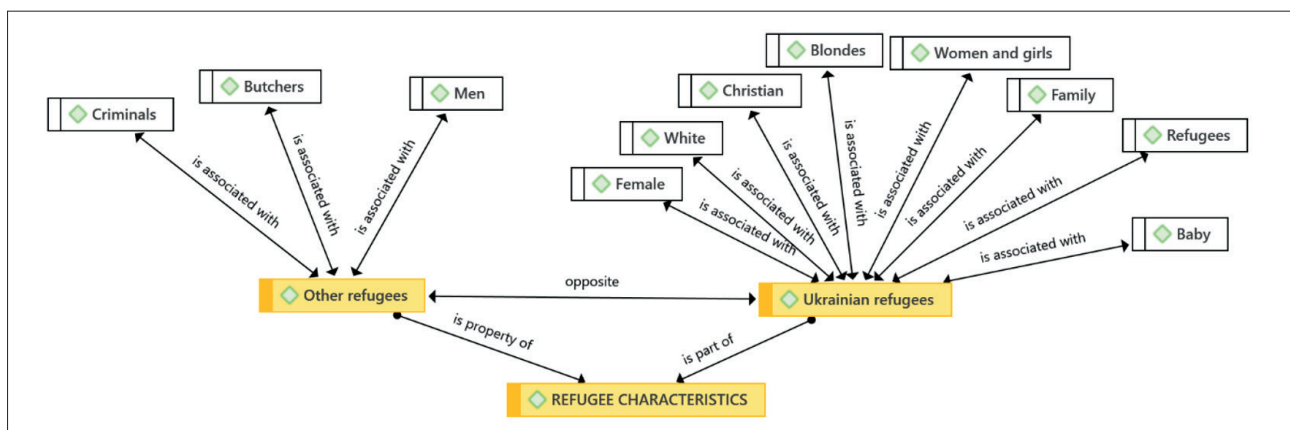


Figure 7. Network of refugee characteristics

People from Ukraine are treated more favourably in terms of the characteristics associated with them. The discourse against other refugees has a clear negative focus in the same dimensions as the previous one. These ideas are discussed by **Rebollo-Díaz** (2021), who identifies social networks as a medium in which a negative discourse about refugees spreads hatred and prejudice towards them. However, the same is not true of the treatment of Ukrainian refugees. As **Delfino** (2022) argues, the response to these refugees has had a clear focus on solidarity and welcome, which is clearly reflected in public opinion, as well as in social media. However, given that this is such a recent phenomenon, there are no precise studies on how they have been treated in the media. This pronounced dualism can also be seen in the proliferation of hate speech against refugees from the south.

4.4. Motives for expressing hatred

The fourth network has a very high complexity, given the large number of codes generated in the open coding, despite having carried out a code fusion process, as in the rest of the networks. For this reason, the exposition of the network was divided into three different figures (Figures 8, 9 and 10) to make it easier to understand them.

In Figure 8, two subcategories emerge in the third network. All codes have an associative relationship with their subcategories, as they are linked to both racist and hateful labels against the Islamic religion. The latter are linked to the first level category, “reasons for expressing hatred”, as a way of justifying such discrimination.

The second part of this fourth network (see Figure 9) shows how three subcategories emerge from the codes. In the first, refugees are shown as a ‘problem for the national economy’; in the second, there is an appeal to ‘nationalist exaltation’; and finally, in the third, hatred is spread towards ‘Spanish entities that support emigration’. This means that their codes express economic problems (i.e., looters) or discourse of nationalist exaltation (i.e., nation). Subsequently, there is the subcategory ‘Spanish entities that support immigration’, where their codes have a possession connection, i.e., both codes are organisations that defend immigration. This is related as a way or justification by which hatred is expressed, linked to the main category that has emerged in the network.

Finally, Figure 10 shows the different codes that give meaning to and are associated with the second-level category ‘reasons for the country’s security’. This entire volume shows the different emerging reasons that create a homogeneous discourse. This is reflected as a subcategory, forming a reason why hatred is expressed and justified on the social network *Twitter*.

In short, the three networks observed show the six subcategories defined as the reasons why hate is expressed: (1) racism, (2) hatred against religion, (3) perceived threats to the economy, (4) nationalist exaltation, (5) rejection of Spanish entities that support this immigration, and (6) perceived threats to the country’s security. These subcategories that emerge from the analysis of the discourse disseminated on *Twitter* are closely related to the classification proposed by **Santos-Alvarado** (2021). The networks obtained in this research show the clear importance of discourses linked to refu-

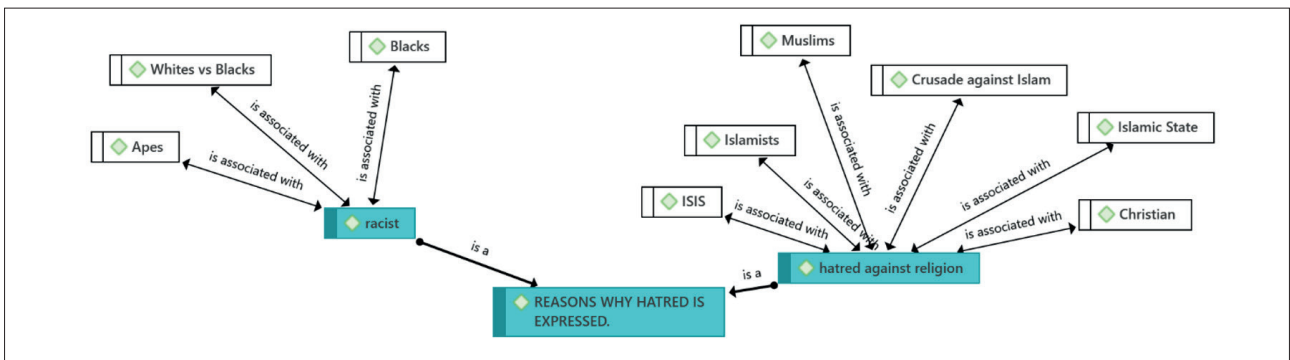


Figure 8. First network of reasons why hate speech is expressed

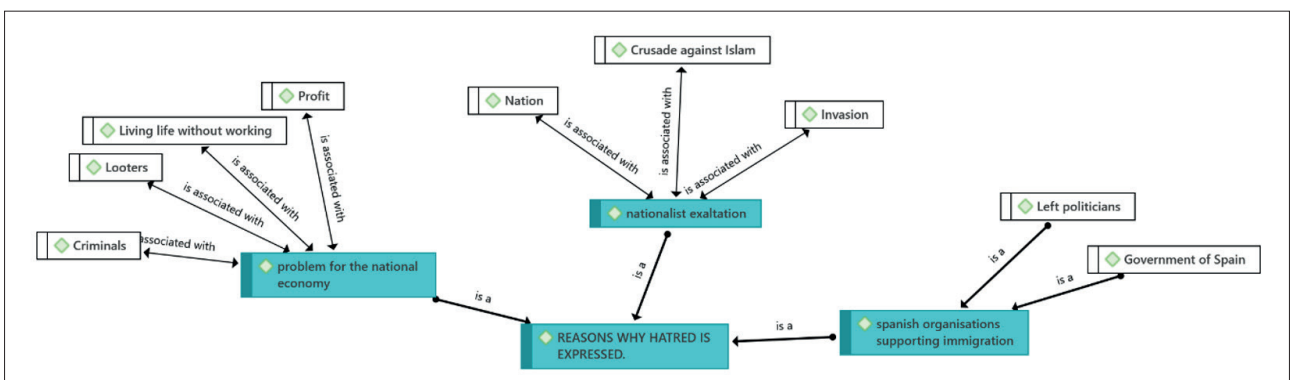


Figure 9. Second network of reasons why hate speech is expressed

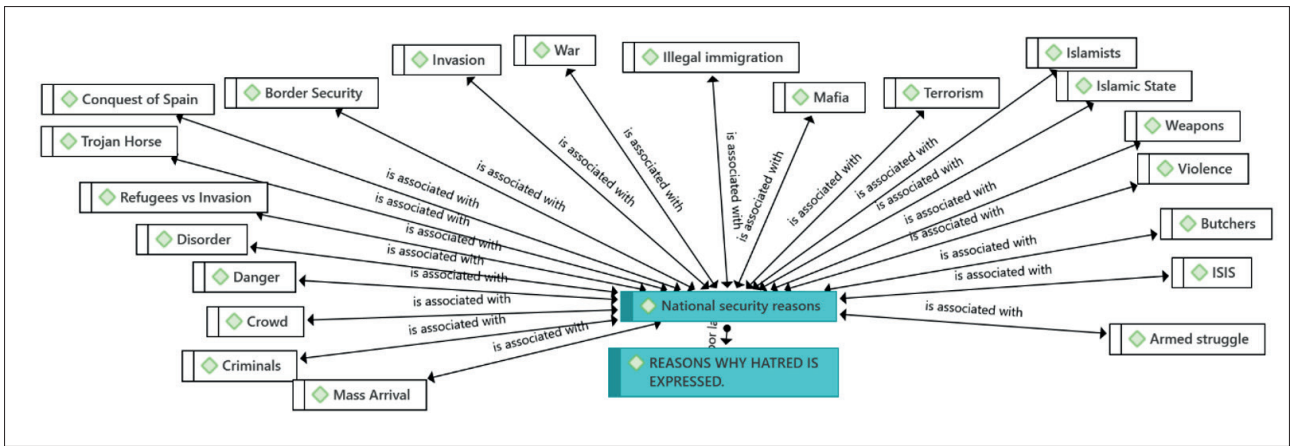


Figure 10. Third network of reasons for hate speech.

gees as a threat to security. In Figure 10, the subcategory that identifies these refugees from the south as a problem for security is the one with the highest density (number of related codes). The threat to the economy as a justification for hatred against refugees from the south is another emerging category, which is the second highest in qualitative density. Finally, the threat to national identity, as a variable of analysis and as defined by Santos-Alvarado (2021), encompasses two subcategories: hatred against religion and national exaltation. Both views seek to identify refugees from the south as a problem for cultural values, as they are labelled as Muslim and radical people who threaten national Christian values (see Figure 11). This approach is one way in which Ukrainian refugees are justified in receiving more favourable treatment since they have close identity ties to nationals in terms of religion, physical appearance, and other characteristics.

All this development of refugees as a threat is linked to the findings of Valdez-Apolo, Arcila-Calderón and Amores (2019), whose study found that, in general terms, a representation of these displaced people as a threat clearly predominates. These authors establish the threat in the same areas (economic, security, and identity) that have been detected in this study. The danger in the economic sphere is that refugees are seen as a burden on national finances, given their social class. This is a clear sign of aporophobia, the hatred of poor or low-income people.

The threat to cultural identity has a strong symbolic value, strongly linked to the Islamic religion in our study. However, the idea of refugees as a security problem has a more realistic character, i.e., that the threat poses a risk to the integrity of the country and people, as they are portrayed as aggressive or criminal. As it affects something so direct and close to people’s reality, it is probably the discourse with the highest density, as established by the research results. Valdez-Apolo, Arcila-Calderón and Amores (2019) establish that the security of a country is the issue that generates the most hatred since it affects what is most sensitive: the supposed physical integrity and the very life of the citizens of the country hosting the refugees.

The blaming of nationals has a lower, but still remarkable, density (number of codes linked to the subcategory). As Santos-Alvarado (2021) argues, this discursive typology is essential to understand social polarisation as a symptom of the rise of hate speech. This last detail is evident in Spanish society according to Ondod (2021) data, where hate crimes have increased by 43.6% since 2013.

Lastly, we find racist hate speech, which has a low density. This low frequency is directly related to one of the issues raised by Valdez-Apolo, Arcila-Calderón and Amores (2019), where explicit racism is politically incorrect, such as that analysed in the study (for example, apes, whites versus blacks), and is largely rejected in advanced societies. Therefore, this racist discourse often tends to be diluted or disguised with other discursive typologies, such as relating the refugee as an economic or security threat. On other occasions, the tweets themselves claim not to be racist and that they only seek to defend European refugees against those they do not consider refugees, but only illegal immigrants without any right to enter the country (see Figure 12). This is clear evidence of the finding of Valdez-Apolo, Ar-



Figure 11. Threat to cultural values tweet. Source: Twitter <https://onx.la/4bd31>

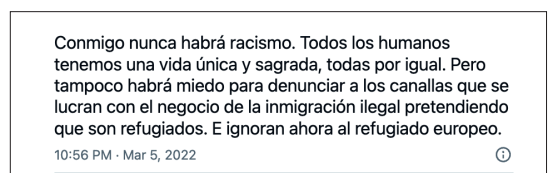


Figure 12. Tweet denying racism towards refugees from the South. Source: Twitter

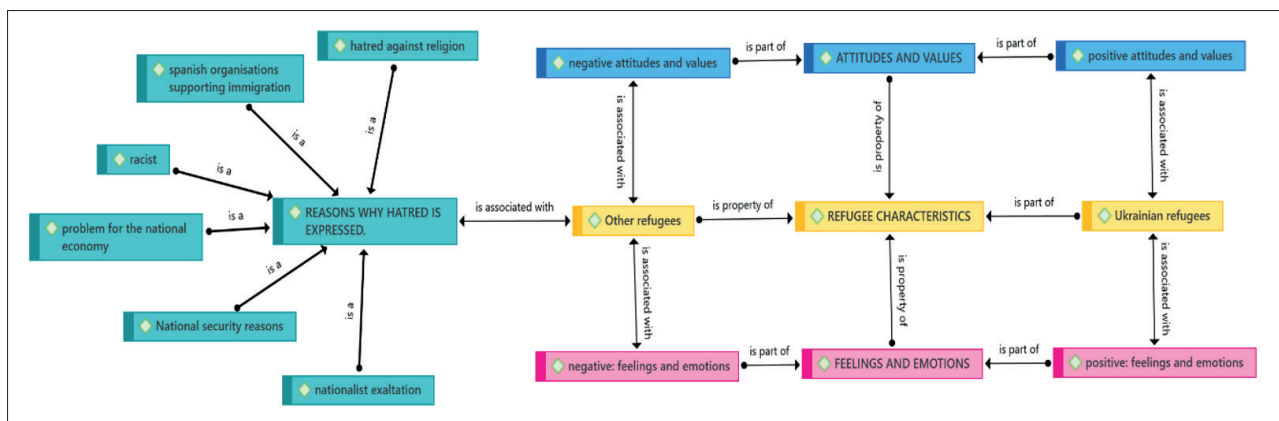


Figure 13. Network of subcategories and emerging categories

cila-Calderón and Amores (2019) that immigrant status is used in the media as a way of discrediting or disqualifying certain refugees, given the negative stereotypes created about this group (Fernández-Fernández; Revilla-Guijarro; Andaluz-Antón, 2020).

4.5. Summary and relationship between categories and subcategories

Once the results obtained in accordance with the objectives set out have been presented, we proceed to develop the relationship between the categories and subcategories. To this end, the comments, quotations, and memos elaborated throughout the analysis have been considered. As a result, the networks that emerge from the categories and subcategories and the type of relationship between them are shown in Figure 13. It is clear how the motives for hate speech are associated with the rest of the refugees. In addition, they are linked to negative attitudes/values and emotions/feelings. Conversely, Ukrainian refugees are characteristically associated with positive attitudes/values and emotions/feelings.

Despite being considered an axial coding process, the last phase of coding, called selective coding, is already being established. This last stage is not independent of the previous two but rather is an extension of them with a higher level of abstraction (San-Martín-Cantero, 2014). This allows us to develop a theory about refugee discourse on *Twitter* based on the last network obtained after the axial process. Following Strauss and Corbin (2002), our final network, composed of the relationship between the subcategories and categories, allows us to establish the central category that will define the theory of our study of the discourses.

The resulting theory on selective coding emphasises that there is a differentiated treatment of refugees on *Twitter*. Ukrainians are associated with certain characteristics (i.e., family, women, and children), as well as a clearly positive focus on attitudes and values (i.e., humanity, order) and emotions and feelings (i.e., tenderness, love). Other refugees receive an opposite discourse, linked to negative values and attitudes (i.e., danger, disorder) and negative emotions and feelings (i.e., rage, anger). The refugees coming from the south receive a series of hate speeches motivated by various justifications: (1) because of their religion, (2) pointing the finger at Spanish entities to blame for this immigration, (3) racism, (4) problem for the security of the country, (5) problem for the economy and (6) nationalist exaltation.

This typology of discourses shapes the various ways in which *Twitter* users find justifications for hate speech against refugees from the south. This reflects the urgent need to carry out media literacy training aimed at all citizens from an emancipatory and transformative approach (Barbas-Coslado, 2012). Consequently, it is necessary to rethink the models of digital skills training since, as Gutiérrez-Martín and Tyner (2012) argue, they are often reduced to the technological dimension, where the focus is on the instrumental control of the programmes.

“The victims of this hatred must be the focus of the debate, as they are the ones who suffer the violation of the most fundamental human rights, such as the right to life, security, and non-discrimination”

5. Conclusions

This study has demonstrated the dualism in the treatment of refugees and hate speech towards groups from the South on the social network *Twitter*. This work is a methodological and scientific novelty, as computer-mediated discourse analysis has made it possible to capture and analyse informal and colloquial content that users of this social network share in their posts in an immediate and probably not very reflective way. The normality with which hate is poured out on social networks is not an innocuous phenomenon, given that hate crimes in Spain have grown exponentially, making it an enormously dangerous phenomenon for social coexistence.

In this way, it should be recalled that the victims of this hatred must be the focus of the debate, as they are the ones who suffer the violation of the most fundamental human rights, such as the right to life, security, and non-discrimination.

In this regard, Spain, from an institutional and legal point of view, is showing significant concern for this phenomenon. The creation of the *National Office for Combating Hate Crimes Ondad*, which reports to the Spanish *Ministry of the*

Interior, was an important step forward in this regard in 2013. Regarding this same body, it is necessary to highlight the annual reports that show the increase in figures, as well as the recently approved *II Action plan to combat hate crimes (2022-2024)*, where the treatment of the victim of hate crimes becomes the centre of the strategy of action. Added to this is the new *Organic Law 15/2022*, comprehensive for equal treatment and non-discrimination, which establishes a special emphasis in articles 22 and 53 on aggressions or actions that incite hatred through social networks.

Failure to confront hate and the manipulation of information and communication media means normalising discrimination against many individuals and groups

However, the regulatory and judicial structure of the state is not enough if it does not consider the training of citizens in what is known as media education or media literacy, especially regarding the critical and ethical use of the media, which will make it possible to combat phenomena such as disinformation.

With respect to the spread of disinformation on networks and the urgent need for media literacy, this study has detected the need to provide specific training on the analysis of audiovisual content in order to help citizens understand the use of the audiovisual world in the media. This concept is becoming more important, especially in social media, which use and sometimes abuse visual content as a format for information and on other occasions for disinformation. If this approach is not considered, citizens are condemned to become easily manipulated subjects.

It should be noted that this research aims to detect these needs in the digital training models of our society, as well as the fight against misinformation, especially that which involves hate speech. Failure to confront hate and the manipulation of information and communication media means normalising discrimination against many individuals and groups. This is consistent with the so-called Sustainable Development Goals (SDGs) and, in particular, SDG 16, 'Peace, justice and strong institutions' (UN, 2015). The aims are to promote the development of both peaceful and inclusive societies and create justice accessible to all, with an emphasis on the need for effective institutions to meet these challenges.

6. Limitations and future lines of research

The non-probabilistic sampling of the research cannot guarantee control over how the sample is constructed. Furthermore, the number of tweets that make up the sample may not be representative of hate speech about refugees on *Twitter* and therefore may be inadequate to generalise. For these reasons, this study should be considered an exploratory approach that can be a starting point in the study of this phenomenon. There are some limitations associated with the characteristics and functioning of the *Twitter* platform. The anonymity of the accounts does not allow us to know which individuals, groups, or bots are behind these types of publications. The social network *Twitter* usually uses tags that categorise posts, but given that only the content of two tags (#Refugees and #NotRefugees) has been analysed in this study, it is likely that not all tweets conveying hate speech towards refugees have been accessed.

With respect to possible future lines of research, continuity with the theme of this work is suggested. The main idea is to replicate the study but bear in mind the need to overcome the limitations indicated. Enlargement of the sample size will be fundamental.

It is worth pointing out the possibility of continuing to study hate speech in the media towards other social groups, such as Unaccompanied Foreign Minors, the LGTBQI+ group, and members of particular religions, ideologies, and social classes. Another possible future line of research is the analysis of why citizens become involved in broadcasting this type of discourse, as this would be very useful information for designing prevention proposals aimed at minimising this phenomenon.

In short, this article suggests great possibilities for research to try to detect the needs of citizens in the face of the challenges posed by the information and communication media, especially with respect to disinformation.

7. References

- Alonso-González, Marián** (2019). "Fake news: desinformación en la era de la sociedad de la información". *Ámbitos: revista internacional de comunicación*, n. 45, pp. 29-52.
<https://doi.org/10.12795/Ambitos.2019.i45.03>
- Álvarez, Freddy** (2008). "La ética y los medios de comunicación". *Alteridad: revista de educación*, v. 3, n. 1, pp. 6-19.
<https://www.redalyc.org/articulo.oa?id=467746250002>
- Álvaro-Sánchez, Sandra** (2018). "The public sphere in the epoch of the algorithmic hyper-mediation: Fake news, disinformation, misinformation and the commodification of behaviour". *Hipertext*, n. 17, pp. 74-82.
<https://doi.org/10.31009/hipertext.net.2018.i17.07>
- Amores, Javier J.; Arcila-Calderón, Carlos; Blanco-Herrero, David** (2020). "Evolution of negative visual frames of immigrants and refugees in the main media of Southern Europe". *Profesional de la información*, v. 29, n. 6.
<https://doi.org/10.3145/epi.2020.nov.24>

- Arcila-Calderón, Carlos; Sánchez-Holgado, Patricia; Quintana-Moreno, Cristina; Amores, Javier J.; Blanco-Herrero, David** (2022). "Hate speech and social acceptance of migrants in Europe: Analysis of tweets with geolocation. [Discurso de odio y aceptación social hacia migrantes en Europa: Análisis de tuits con geolocalización]". *Comunicar*, v. 30, n. 71, pp. 21-35. <https://doi.org/10.3916/C71-2022-02>
- Arrieta, C. Adolfo; Montes, V. Donicer** (2011). "Alfabetización digital: uso de las TIC's más allá de una formación instrumental y una buena infraestructura". *Revista colombiana de ciencia animal*, v. 3, n. 1, pp. 180-197. <https://doi.org/10.24188/recia.v3.n1.2011.360>
- Badillo-Matos, Ángel** (2019). *La sociedad de la desinformación: propaganda, fake news y la nueva geopolítica de la información*. Madrid: Real Instituto Elcano, 14 mayo. <https://www.realinstitutoelcano.org/documento-de-trabajo/la-sociedad-de-la-desinformacion-propaganda-fake-news-y-la-nueva-geopolitica-de-la-informacion>
- Barbas-Coslado, Ángel** (2012). "Educomunicación: desarrollo, enfoques y desafíos en un mundo interconectado". *Foro de educación*, v. 10, n. 14, pp. 157-175. <https://www.redalyc.org/articulo.oa?id=447544618012>
- Belmonte, María-Luisa; Muñoz-Álvarez, José-Santiago; Bernárdez-Gómez, Abraham** (2021). "Educando en valores: el respeto, la tolerancia, la igualdad y la interculturalidad como pilares básicos". *Brazilian journal of development*, v. 7, n. 1, pp. 9516-9529. <https://doi.org/10.34117/bjdv7n1-644>
- Blanco-Alfonso, Ignacio; Rodríguez-Fernández, Leticia; Arce-García, Sergio** (2022). "Polarización y discurso de odio con sesgo de género asociado a la política: análisis de las interacciones en *Twitter*". *Revista de comunicación*, v. 21, n. 2, pp. 33-50. <https://doi.org/10.26441/rc21.2-2022-a2>
- Bustos-Martínez, Laura; De-Santiago-Ortega, Pedro-Pablo; Martínez-Miró, Miguel-Ángel; Rengifo-Hidalgo, María-Sofía** (2019). "Discursos de odio: una epidemia que se propaga en la Red. Estado de la cuestión sobre el racismo y la xenofobia en las redes sociales". *Mediaciones sociales*, v. 18, pp. 25-42. <https://doi.org/10.5209/meso.64527>
- Cantamutto, Lucía; Vela-Delfa, Cristina** (2023). "Corpus de interacciones digitales: sistematización de técnicas para recoger datos en *WhatsApp*". *Cuadernos.info*, n. 54, pp. 117-139. <https://doi.org/10.7764/cdi.54.53165>
- Castillo-de-Mesa, Joaquín; Méndez-Domínguez, Paula; Carbonero-Muñoz, Domingo; Gómez-Jacinto, Luis** (2021). "Homofilia, polarización afectiva y desinformación en *Twitter*. Caso de estudio sobre la crisis migratoria #Openarms". *Redes*, v. 32, n. 2, pp. 153-172. <https://doi.org/10.5565/rev/redes.913>
- Corbella-Ruiz, Marta; De-Juanas-Oliva, Ángel** (2013). "Redes sociales, identidad y adolescencia: nuevos retos educativos para la familia". *Estudios sobre educación*, n. 25, pp. 95-113. <https://dialnet.unirioja.es/servlet/articulo?codigo=4532318>
- Delfino, Micaela** (2022). "Respuesta europea al éxodo ucraniano: ¿Distinción entre refugiados?". *Boletín informativo del grupo de jóvenes investigadores*, n. 15, pp. 9-12. <http://sedici.unlp.edu.ar/handle/10915/137264>
- España** (2022). "Ley 15/2022, de 12 de julio, integral para la igualdad de trato y la no discriminación". *Boletín oficial del estado*, n. 167, 13 julio. https://www.boe.es/diario_boe/txt.php?id=BOE-A-2022-11589
- Fernández-Fernández, Ángel; Revilla-Guijarro, Almudena; Andaluz-Antón, Lucía** (2020). "Análisis de la caracterización discursiva de los relatos migratorios en *Twitter*. El caso Aquarius". *Revista latina de comunicación social*, n. 77. <https://www.doi.org/10.4185/RLCS-2020-1446>
- Fundación FAD Juventud** (2022). *Discurso de odio en redes*. <https://fad.es>
- 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: revista de estrategias, tendencias e innovación en comunicación*, n. 23, pp. 97-118. <https://doi.org/10.6035/adcomunica.6299>
- Gamir-Ríos, José; Tarullo, Raquel; Ibáñez-Cuquerella, Miguel** (2021). "Multimodal disinformation about the otherness on the Internet: The spread of racist, xenophobic and islamophobic fake news in 2020". *Anàlisi: quaderns de comunicació i cultura*, n. 64, pp. 49-64. <https://doi.org/10.5565/rev/analisi.3398>

- Gutiérrez-Braojos, Calitxo; Martín-Romera, Ana; Salmerón-Pérez, Honorio; Casasepère-Satorres, Antoni; Fernández-Cano, Antonio** (2017). "Análisis temático de la investigación educativa soportada por *Grounded Theory*". *Bordon*, v. 69, n. 1, pp. 83-102.
<https://doi.org/10.13042/Bordon.2016.41035>
- Gutiérrez-Martín, Alfonso; Pinedo-González, Ruth; Gil-Puente, Cristina** (2022). "ICT and media competencies of teachers. Convergence towards an integrated MIL-ICT model. [Competencias TIC y mediáticas del profesorado. Convergencia hacia un modelo integrado AMI-TIC]". *Comunicar*, n. 70, pp. 21-33.
<https://doi.org/10.3916/C70-2022-02>
- Gutiérrez-Martín, Alfonso; Tyner, Kathleen** (2012). "Media education, media literacy and digital competence. [Educación para los medios, alfabetización mediática y competencia digital]". *Comunicar*, v. 38, n. 19, pp. 31-39.
<https://doi.org/10.3916/C38-2012-02-03>
- Hernández-Sampieri, Roberto; Fernández-Collado, Carlos; Baptista-Lucio, Pilar** (2018). *Metodología de la investigación*. México: McGraw Hill. ISBN: 978 1 4562 2396 0
- Igartua, Juan-José; Cheng, Lifen; Moral, Félix; Fernández, Iciar; Frutos, Francisco J.; Gómez-Isla, José; Otero, José** (2008). "Enmarcar la inmigración en las noticias y sus efectos socio-cognitivos". *Palabra clave*, v. 1, n. 1, pp. 87-107.
<https://palabraclave.unisabana.edu.co/index.php/palabraclave/article/view/1418>
- Igartua, Juan-José; Muñiz-Muriel, Carlos; Otero-Parra, José-Antonio; De-la-Fuente-Juan, Montse** (2007). "El tratamiento informativo de la inmigración en los medios de comunicación españoles. Un análisis de contenido desde la Teoría del Framing". *Estudios sobre el mensaje periodístico*, v. 13, pp. 91-110.
<https://revistas.ucm.es/index.php/ESMP/article/view/ESMP0707110091A>
- Jorrín-Abellán, Iván-Manuel; Fontana-Abad, Mónica; Rubia-Avi, Bartolomé** (coords.). (2020). *Investigar en educación*. Madrid: Síntesis. ISBN: 978 84 13576329
- Larrondo-Ureta, Ainara; Peña-Fernández, Simón; Morales-i-Gras, Jordi** (2021). "Desinformación, vacunas y Covid-19. Análisis de la infodemia y la conversación digital en Twitter". *Revista latina de comunicación social*, n. 79.
<https://doi.org/10.4185/RLCS-2021-1504>
- Larrosa, Jorge** (2007). "Las imágenes de la vida y la vida de las imágenes: tres notas sobre el cine y la educación de la mirada". *Educación y realidad*, v. 32, n. 2, pp. 7-22.
<https://www.redalyc.org/pdf/3172/317227046002.pdf>
- Lava-Santos, David** (2021). "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>
- Magallón-Rosa, Raúl** (2021). "La agenda que queda fuera de los medios. Desinformación y 'fact-checking' sobre (in)migración". *Migraciones*, n. 52, pp. 59-87.
<https://doi.org/10.14422/mig.i52.y2021.003>
- Mancera-Rueda, Ana; Pano-Alamán, Ana** (2014). "Las redes sociales como corpus de estudio para el análisis del discurso mediado por ordenador". *Janus. Estudios sobre el Siglo de Oro*, anexo 1, pp. 305-315.
<https://www.janusdigital.es/anexos/contribucion.htm?id=28>
- Martínez-Rodrigo, Estrella; Segura-García, Rosario; Sánchez-Martín, Lourdes** (2011). "El complejo mundo de la interactividad: emociones y redes sociales". *Revista mediterránea de comunicación*, v. 2, n. 1, pp. 171-190.
<https://doi.org/10.14198/MEDCOM2011.2.10>
- Martínez-Sala, Alba-María; Segarra-Saavedra, Jesús; Monserrat-Gauchi, Juan** (2018). "Los millennials como prosumers y adprosumers en las redes sociales corporativas". *Cuadernos.info*, n. 43, pp. 137-159.
<https://doi.org/10.7764/cdi.43.1335>
- Molina-Cañabate, Juan-Pedro; Magallón-Rosa, Raúl** (2019). "Procedimientos para verificar y desmontar informaciones falsas basadas en el discurso del odio. El caso de *Maldita Migración*". *Revista de la Asociación Española de Investigación de la Comunicación*, v. 6, n. 12, pp. 95-122.
<https://doi.org/10.24137/raeic.6.12.5>
- Müller, Karsten; Schwarz, Carlo** (2021). "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>
- Narváez-Llinares, Ángel; Pérez-Rufí, José-Patricio** (2022). "Fake news y desinformación sobre migración en España: prácticas del discurso xenófobo en redes sociales y medios online según la plataforma *Maldita Migración*". *Estudios sobre el mensaje periodístico*, v. 28, n. 4, pp. 841-854.
<https://doi.org/10.5209/esmp.82845>

- Navarrete-Barrutia, Mercedes J.** (2020). "Fake news y extranjería: la desinformación como respuesta al estigma social". *Revista inclusiones*, v. 7, pp. 286-305.
<https://revistainclusiones.org/index.php/inclu/article/view/1710>
- Ondod** (2022). *II Plan de acción de lucha contra los delitos de odio 2022-2024*. Madrid: Ministerio del Interior. Secretaría de Estado de Seguridad. Oficina Nacional de Lucha contra los Delitos de Odio.
<https://www.interior.gob.es/opencms/pdf/servicios-al-ciudadano/Delitos-de-odio/descargas/II-PLAN-DE-ACCION-DE-LUCHA-CONTRA-LOS-DELITOS-DE-ODIO.pdf>
- ONU** (2015). *Objetivos de desarrollo sostenible*.
<https://www.un.org/sustainabledevelopment/es/objetivos-de-desarrollo-sostenible>
- Pano-Alamán, Ana; Moya-Muñoz, Patricio** (2016). "Una aproximación a los estudios sobre el discurso mediado por ordenador en lengua española". *Tonos digital*, n. 30, pp. 277-298.
<https://dialnet.unirioja.es/servlet/articulo?codigo=5327425>
- Pari-Tito, Fernando; García-Peñalvo, Francisco-José; Pérez-Postigo, Gerber** (2022). "Análisis bibliométrico de la desinformación mediática y fake news en las redes sociales". *Revista universidad y sociedad*, v. 14, n. S2, pp. 37-45.
<https://rus.ucf.edu.cu/index.php/rus/article/view/2755>
- Pharm** (2019). *Preventing hate against refugees and migrants*.
<https://pharmproject.usal.es>
- Rebollo-Díaz, Carolina** (2021). *Tuiteando sobre refugiados. Una comparación internacional de discursos, imaginarios y representaciones sociales*. Tesis doctoral, Universidad de Huelva.
<https://dialnet.unirioja.es/servlet/tesis?codigo=291417>
- Sánchez-Holgado, Patricia; Amores, Javier J.; Blanco-Herrero, David** (2022). "Online hate speech and immigration acceptance: A study of Spanish provinces". *Social sciences*, v. 11, n. 11, 515.
<https://doi.org/10.3390/socsci11110515>
- San-Martín-Cantero, Daniel** (2014). "Teoría fundamentada y *Atlas.ti*: recursos metodológicos para la investigación educativa". *Revista electrónica de investigación educativa*, v. 16, n. 1, pp. 104-122.
http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1607-40412014000100008&lng=es&tlng=es
- Santos-Alvarado, Orlando-Nikolais** (2021) "Xenofobia y racismo hacia (y por) inmigrantes venezolanos residentes en Perú a través de *Twitter*". *Global media journal México*, v. 18, n. 34, pp. 160-184.
<https://doi.org/10.29105/gmjmx18.34-8>
- Strauss, Anselm; Corbin, Juliet** (2002). *Bases de la investigación cualitativa: técnicas y procedimientos para desarrollar la teoría fundamentada*. Universidad de Antioquia. Contus. ISBN: 958 655 623 9
<https://diversidadlocal.files.wordpress.com/2012/09/bases-investigacion-cualitativa.pdf>
- Unesco** (2021). *Media and information literate citizens: Think critically, click wisely!* ISBN: 978 92 3 100448 3
<https://unesdoc.unesco.org/ark:/48223/pf0000377068>
- Valdez-Apolo, María-Belén; Arcila-Calderón, Carlos; Amores, Javier J.** (2019). "El discurso del odio hacia migrantes y refugiados a través del tono y los marcos de los mensajes en *Twitter*". *Revista de la Asociación Española de Investigación de la Comunicación*, v. 6, n. 12, pp. 361-384.
<https://doi.org/10.24137/raic.6.12.2>
- Vergani, Matteo; Perry, Barbara; Freilich, Joshua; Chermak, Steven; Scrivens, Ryan; Link, Rouven** (2022). "Protocol: mapping the scientific knowledge and approaches to defining and measuring hate crime, hate speech, and hate incidents". *Campbell systematic reviews*, v. 18, n. 2.
<https://doi.org/10.1002/cl2.1228>
- Vuorikari, Riina; Kluzer, Stefano; Punie, Yves** (2022). *DigComp 2.2. The digital competence framework for citizens. With new examples of knowledge, skills, and attitudes*. Luxembourg: European Commission.
<https://doi.org/10.2760/115376>
- Wachs, Sebastian; Wejetstein, Alexander; Bilz, Ludwig; Gámez-Guadix, Manuel** (2022). "Adolescents' motivations to perpetrate hate speech and links with social norms. [Motivos del discurso de odio en la adolescencia y su relación con las normas sociales]". *Comunicar*, v. 30, n. 71.
<https://doi.org/10.3916/C71-2022-01>
- Wilson, Carolyn; Grizzle, Alton; Tuazon, Ramon; Akyempong, Kwame; Cheung, Chi-Kim** (2011). *Alfabetización mediática e informacional: currículum para profesores*. Francia: Unesco. ISBN: 978 959 18 0787 8
<https://unesdoc.unesco.org/ark:/48223/pf0000216099>

Blood donors wanted: narrative innovation on *TikTok* to enable mobilization

Raquel Martínez-Sanz; Amaia Arribas-Urrutia

Nota: Este artículo se puede leer en español en:
<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87171>

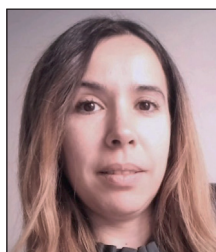
Recommended citation:

Martínez-Sanz, Raquel; Arribas-Urrutia, Amaia (2023). "Donors wanted: narrative innovation on *TikTok* to enable mobilization". *Profesional de la información*, v. 32, n. 3, e320305.

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

Manuscript received on 8th November 2022

Accepted on 28th February 2023



Raquel Martínez-Sanz

<https://orcid.org/0000-0002-4753-0282>

Universidad de Valladolid
Facultad de Filosofía y Letras
Plaza del Campus, s/n
47011 Valladolid, Spain
raquel.martinez.sanz@uva.es



Amaia Arribas-Urrutia ✉

<https://orcid.org/0000-0001-9452-8364>

Universidad de Valladolid
Facultad de Filosofía y Letras
Plaza del Campus, s/n
47011 Valladolid, Spain
amaya.arribas@uva.es

Abstract

The explosive success of *TikTok* is one of a long list of phenomena that are transforming audiovisual relationships, creation, and consumption (Gómez-García; Vicent-Ibáñez, 2022). At the same time, social networks are demonstrating a capacity for outreach, awareness, and activism through innovative narratives that impact the audience. Therefore, we ask if and how, in addition to entertaining and being a place to find the latest trends, *TikTok* can help promote health activism linked to blood donation—an act of social responsibility—the collection of which has been in sharp decline for years (Carter *et al.*, 2011; Huis-in 't-Veld *et al.*, 2019). Through multimodal discourse analysis (Kress, 2012), we determine the characteristics, significance, and communicative resources of the *tiktoks* that the platform grouped under the hashtag #donasangre [#DonateBlood], ultimately comparing them with the perception and relationship that a large group of young university students expressed in focus groups. Among the results, we highlight how content on *TikTok* is reinterpreted and appropriated, as well as the false myths perpetuated among young people that keep them from donating and reinforce the importance of health institutions taking an active role in the online conversation and integrating narrative innovation into their content creation dynamics. Qualitative analysis of comments (6,215) revealed that there is an audience that “comes together” (Juris, 2012) and whose members reaffirm their status as donors and whose experience tempers the idea of donating being stressful and scary.

Keywords

Health communication; Social media; Social networks; Blood donation; Recipients; *TikTok*; Social transformation; Activism; Misinformation; Digital citizenship; Innovation; Storydoing; Narratives.

Funding

This article is a result of:

1. The R+D+i project “Innovation ecosystems in the communication industries: actors, technologies and configurations for the generation of innovation in content and communication. Innovacom” [“Ecosistemas de innovación en las industrias de la comunicación: actores, tecnologías y configuraciones para la generación de innovación en contenido y comunicación. Innovacom”], funded by the Spanish State Research Agency [Agencia Estatal de Investigación] (PID2020-114007RB-I00), MCIN AEI/10.13039/501100011033.
2. The author Raquel Martínez-Sanz is a beneficiary of a Requalification Grant (2022–2024) funded by the Next-GenerationEU European Commission.



1. Introduction

Social networks are a powerful channel for the dissemination of important content (**Gutiérrez-Lozano; Cuartero, 2020; Mannell; Ahmad; Ahmad, 2018; Syrdal; Briggs, 2018**); their dynamics have been able to turn each user, either individually or corporately, into a valuable driver of thoughts, lifestyles, or information, thereby challenging the dominance of traditional media (**Dubovi; Tabak, 2021; Vandenbosch; Eggermont, 2016; Campos-Freire et al., 2016**). In addition to inevitably losing control of the message through its exposure to the social audience's comments, reinterpretation, and criticism (**Zulli; Zulli, 2020**), the current information ecosystem is faced with the phenomenon of disinformation (**Seminar-Córdova, 2021; Gisondi et al., 2022**), which encompasses both the spread of malicious information or infoxication –an instigator of hate speech– (**Azzimonti; Fernandes, 2023; Porroche-Escudero, 2017**) and a lack of cited sources for a topic.

Altruistic blood donation has a limited media presence (**Bomfim-de-Souza; Santoro-Domingo, 2020**), especially considering the vital role it plays in any public health system (**WHO, 2022**). The fact that blood cannot be manufactured and that it expires and the variety of applications for it in an increasingly longer-living population make hemoderivatives a valuable asset, dependent upon the sense of responsibility and civic-mindedness of citizens (**Eser et al., 2010**). Healthcare organizations have addressed this concern mainly by launching massive communication campaigns (**Weidmann et al., 2022**). However, the emergence of platforms such as *TikTok*, the sixth most important social network globally (**Hootsuite; We are Social, 2022**), which is popular with very young age groups (**IAB, 2022**), has forced institutions to adapt and innovate content with the aim of capturing the attention of this age group (**Waheed et al., 2020; Yuan et al., 2016**), which is especially valuable because it is a potential asset when it comes to both donating and communicating the altruistic values and responsibility that define this act (**Padilla-Garrido et al., 2021; Rael et al., 2021**).

Health is increasingly socially valuable. It is understood as not just the absence of disease but a state of complete physical, mental, and social well-being (**Francis, 2021; Rojas-Rajs; Jarillo-Soto, 2013**). Health communication concerns how the recipient can take action, either by developing a new healthy habit or by eliminating an established harmful habit (**Martínez-Beleño; Sosa-Gómez, 2016**). And despite the features that digital media offers for carrying out health activism, these tools are underutilized by healthcare organizations (**Sobowale et al., 2020; Strauck-Franco; Guillén-Arruda, 2017**).

With this in mind, the main objective of this research is to analyze the narratives about altruistic blood donation constructed on *TikTok*, a social network on which entertainment, ephemeral attention (**Navarro-Robles; Vázquez-Barrio, 2020**), and trivial topics (**Olivares-García; Méndez-Majuelos, 2020**) reign, but which has shown a great capacity for raising awareness (**Hautea et al., 2021**) and disseminating information (**Martínez-Sanz; Buitrago; Martín-García, 2023**). Secondly, we hope to:

- O1. Identify the narrative innovation aimed at mobilization that has been developed on *TikTok*.
- O2. Analyze the conversation generated on *TikTok* around the topic and verify the audiovisuals' mobilizing effect.
- O3. Explore young people's stance on altruistic blood donation and compare it with the stories uploaded to *TikTok*.

In addition, several research questions are posed to help further explore the topic. They are as follows:

- Q1: What is the purpose of posting videos about blood donation on *TikTok*?
- Q2: What communicative resources do content creators use on *tiktoks* about blood donation to mobilize the audience; do they combat the myths associated with this practice?
- Q3: Does the audience show a commitment to action through their comments?

Ultimately, we hope that the knowledge derived from this study will improve institutional health communication focused on the importance of being a blood donor, pointing out digital narrative innovations that further the mobilization and perceptions of the target audience. With all this in mind, we hope to demonstrate *TikTok's* capacity to raise awareness of and promote healthy and humanitarian lifestyles.

2. Theoretical framework

2.1. Social media health activism: giving blood to give life

The concept of health has evolved extensively in recent decades, taking on the idea of not only the absence of disease but also the development of a state of physical, mental, and social well-being (**WHO, 1986**). In this sense, health communication focuses its efforts on promoting the adoption of healthy lifestyles by increasing public knowledge, helping people learn health skills, spreading the word about the benefits of behavioral change, and refuting myths and misconceptions (**Francis, 2021; Rojas-Rajs; Jarillo-Soto, 2013**).

The misinformation provided in the health field is the result of either the unintentional propagation of erroneous content or the premeditated dissemination of intentionally misleading messages (**Salaverría, 2021**). Social networks have created an information ecosystem full of data mixed with opinions and often unverified claims. Thus, public health today is faced with the challenge of curbing and disproving false, inaccurate, or incomplete news (**Seminar-Córdova, 2021**;

Porroche-Escudero, 2017). This situation is exacerbated by health agencies' segregation of information, posing a challenge for their institutional communication, given that they must redouble their efforts to inform, prevent, and motivate effectively in the interest of collective health.

“ The various stakeholders who produce content can generate the same level of social impact on social media as health professionals ”

It is possible to approach the experience of health communication using an informational model—already outdated—that is aimed at the adoption of content that inspires action through messages that come mainly from authorized sources, physicians, and health institutions, and the extensive use of the mass media, or using a relational model (**Veron**, 1996) that involves recognizing that messages are reinterpreted, whereby discourses and meanings have a greater capacity to be heard, read, or perceived. Social networks, which play an important role in the social production of meaning, are technological mediations that construct social relationships (**Arribas-Urrutia; Islas-Carmona; Gutiérrez-Cortés**, 2019).

In the Internet, social movements have found an excellent channel for their exposure and collaborative strategies, turning cyberspace into an interactive, collaborative, and participatory space (**Sola-Morales; Zurbano-Berenguer**, 2020). The objective of digital activism, which offers the user a new way to live out digital citizenship, is communicating, promoting, and defending social, political, or cultural causes, among others, in the digital environment (**Candón-Mena; Benítez-Eyzaguirre**, 2016). It is worth remembering that health professionals are not the only ones who can play a key role in the process of social change in health—the various stakeholders who produce meaning, capable of having a great impact on society, can, as well (**Mannell; Ahmad; Ahmad**, 2018).

Social networks have been particularly active during the recent Covid-19 health crisis, creating such dissimilar effects as providing misinformation (**Gisondi et al.**, 2022), raising awareness (**Hautea et al.**, 2021), or promoting activities beneficial to health (**Martínez-Sanz; Buitrago; Martín-García**, 2023). Consequently, health activism aims at a social transformation that leads citizens to develop a collective conscience about a certain health issue—in our case, altruistic blood donation.

Blood's health applications are numerous, including transfusions to alleviate blood loss during surgery or trauma, the production of drugs and vaccines, and the treatment of blood diseases (**WHO**, 2022). In Europe, blood is collected through a voluntary and unremunerated system that has experienced a sharp decline in recent years (**Carter et al.**, 2011). In addition, during the Covid health crisis, the fear of contagion, social isolation measures, and the increased demand for plasma, used to improve the recovery of patients admitted for Covid, caused stocks to fall sharply, greatly increasing mass calls for assistance (**Weidmann et al.**, 2022; **Waheed et al.**, 2020).

In addition to having a better state of health, which makes them the most suitable subjects for donation in terms of quality, young people are a population segment with a higher rate of social involvement (**Padilla-Garrido et al.**, 2021; **Rael et al.**, 2021). However, as research by **Greffin et al.** (2021) and **Hupfer, Taylor, and Letwin** (2005) has shown, there are significant barriers to motivating them to action, the most repeated of which are the fear of pain, needles, and having a bad experience. Likewise, the lack of information and role models negatively condition the intention to donate. **Kalargirou et al.** (2014) suggest targeting recruitment campaigns to explain the real asset needs. A survey of 28 European Union member states revealed that, on average, 30% of non-donors would come forward to donate to alleviate stock shortages (**Huis-in't-Veld; De-Kort; Merz**, 2019).

The factors that influence behavior and whose approaches drive our study has been examined from the point of view of social psychology and communication. Thus, **Grunig and Hunt's** (2003) situational theory of audiences focuses on the environment as the main element in transforming a passive individual into an active one, whereas social cognitive theory (**Bandura**, 2004) points out the importance of modeling and self-efficacy, defined as the self-confidence to overcome fears and take action. Thibaut and Kelley warn, using the theory of social exchange (**Garza-Guzmán**, 2009), that all human relationships are shaped by cost-benefit analysis and the comparison of alternatives, predicting people's behavior when it comes to perceived rewards and costs. And Katz, Blumler, and Gurevitch's uses and gratifications theory (**Sheldon; Newman**, 2019) explains media consumption, including that of social networks, based on the satisfaction of the specific wants and needs of audiences, who take an active role.

TikTok is the social network that has generated the most interaction in 2022, with an increase of 107%, with 250% more views than in the previous year (**IAB**, 2022). The power that this network of short videos has to enhance youth expression, identity development, and community membership makes *TikTok* the ideal platform for spreading activism. Traditional methods for mobilization have been adapted to the digital environment, developing new ways of communicating and producing content that transform collective action and emphasize the subject's power to become a player (**Arribas-Urrutia; Islas-Carmona; Gutiérrez-Cortés**, 2019).

2.2. Innovative narratives: from storytelling to storydoing on *TikTok*

Creativity and innovation in content promote its consumption (**Gómez-García; Vicent-Ibáñez**, 2022). Young people produce and search for viral, fresh, ephemeral, and fun content on *TikTok*. Likewise, they value online audiovisual content for its spatiotemporal immediacy, its thematic variety, and their identification with the creators (**Navarro-Robles; Vázquez-Barrío**, 2020), which is associated with a creative demand. The stories that are produced and disseminated

on social networks are designed to engage an audience that prioritizes rapid consumption and viralization, turning them into echo chambers (Rodríguez-Cano, 2018). Consequently, as Gutiérrez-Lozano and Cuartero (2020) point out, a transformation is taking place in the way in which young people relate to audiovisual products.

With *storydoing*, the *tiktoker* can take stronger ownership of the cause they are promoting, making it more credible while also developing their creative qualities

The creative principle of empathy leads the receiver to identify with the story and its protagonists. Empathy humanizes the relationship and develops a deep understanding of people's motivations and needs. Storytelling has the ability to instill empathy (empathetic storytelling) in the audience, and through it, to move them. The work of Burgess, Rogers, and Jeffries (2022) demonstrates that storytelling is a useful resource for the topic of health because it awakens solidarity in people. Storytelling is ideally suited to institutional communication by showing the more human face of the company, its corporate values, its actions, and its social programs to the community, thus engendering greater trust (Zabala-Cia; Lorenzo-Sola; González-Pacanowski, 2022).

However, storytelling has a greater impact if, in addition to narrating, it provides boots-on-the-ground proof of what is being done or advocated for. "Storydoing" consists of showing videos of the actions, testimonials, or rationales that spur action. Rodríguez-Ríos and Lázaro-Pernias (2022) define it as a new narrative discourse that involves people in a cause and encourages them to participate by co-creating and, in the corporate case, spreading the brand story through their social networks.

With *storydoing*, the *tiktoker* has the advantage of taking stronger ownership of the cause they are promoting, making it more credible. The user develops their creative qualities and turns the application into a socializing agent of content (Moreno-Barreneche, 2022). Thus, *TikTok* users take ownership of the content they post, edit it with user-friendly programs, and give it a meaning with the aim of influencing and advocating for a certain cause, with a tinge of humor and entertainment in most cases (Olivares-García; Méndez-Majuelos, 2020).

3. Methodology

Given that *TikTok* is a social network in the process of consolidation and its academic study is in its early stages, a mixed methodology with an exploratory sequential design (Creswell; Plano-Clark, 2011) in two phases is proposed to respond to the objectives and research questions posed herein. The first phase aims to study university students' perceptions, beliefs, and attitudes about altruistic blood donation through the use of several discussion groups. The second is a multimodal speech analysis (Kress, 2012) of the audiovisuals posted with the hashtag #donasangre. The hashtag, characteristic of social networks, is defined as

"a conduit for individuals to locate, organize and collectively contribute information flows [...] resulting in thematic and affective audiences that converge around a topic or event" (Zulli; Zulli, 2020, p. 3).

Seven focus groups involving 44 young university students younger than 25 years of age were organized in the facilities of the Universidad de Valladolid. The access-controlled convenience sample sought gender parity (45.5% men and 54.5% women) and a proportional representation of donors and non-donors in accordance with the most widespread figures: 3 out of 10 university students in Spain have donated at some point (Padilla-Garrido *et al.*, 2021). The sessions had an average duration of 50 minutes, were recorded, and subsequently were transcribed. Participants signed an informed consent form that explained the purpose of the study and how the data collected would be processed. In addition, to understand their connection to blood donation and availability and, above all, to encourage reflection on this topic, students were invited to complete a brief survey two weeks before the focus groups were held.

Then, we collected the first 100 videos with the hashtag #donasangre suggested by *TikTok* between September 22 and 29, 2022; this was carried out without logging in to avoid having the researcher's previous searches or interests condition the results. To identify the intentionality, message framing, innovative narrative resources, and impact, a worksheet was prepared for this purpose based on the work of Hautea *et al.* (2020) and Rael *et al.* (2021). Regarding the study of creators, behavioral engagement was calculated (Dubovi; Tabak, 2021) and the type (institutional or personal) and thematic nature of each profile was identified. Finally, through qualitative observation, we determined whether the comments made on each post prompted the audience to commit to mobilizing, as evidenced by an intention to donate and/or to pass on the message (Durántez-Stolle; Martínez-Sanz; Rodríguez-de-Dios, 2022).

The result was a sample of 100 audiovisuals from 88 different profiles, which generated a total of 6,215 manually monitored comments. After evaluating the descriptive characteristics of the dataset, the authors of this paper, together and then in isolation, recorded the emergent themes; the semiotic interaction between images, text, and sound (Moreno-Barreneche, 2022); and the narrative innovation (Trauth-Taylor, 2021) of the first 12 *tiktoks*. Subsequently, the responses were discussed, the analysis sheet was reviewed, and a taxonomy of the intention of the inferred message, type of innovation, and thematic nature of the author's profile was created.

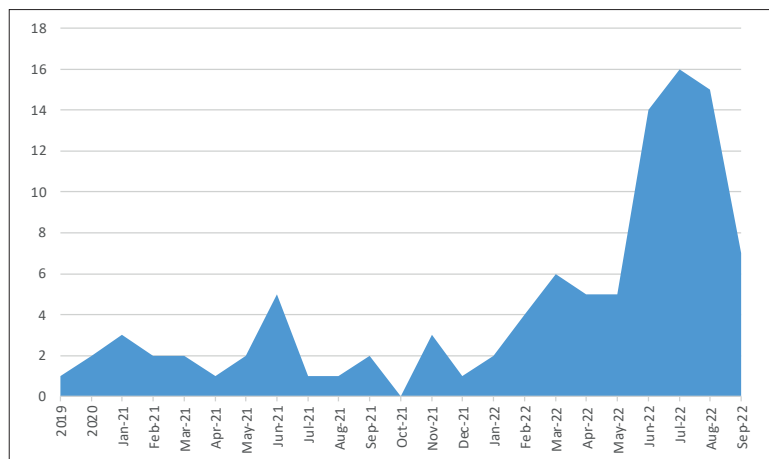
4. Results

4.1. The role of *TikTok* in normalizing blood donation

The videos about blood donation given priority by *TikTok*, i.e., those that were placed in the first 100 positions and, consequently, were easier to access, broke down as follows: 19 were from institutional profiles—donation centers and compulsory education schools, mainly— whereas the remaining 81 were from personal accounts, of which 30.9% were linked to profiles of healthcare professionals, mostly from medicine and nursing. The rest did not express any affiliation, beyond wanting to communicate an experience, need, or knowledge of the subject.

Of the profiles, 10.2% (9 out of 88) dedicated more than one post to the topic of blood donation; this occurred both with healthcare (@bancodesangrehcsb) and educational (@purezarealejos) institutions and on the profiles of healthcare professionals (@cris_carrillo) and non-healthcare professionals (@tanit_tb).

The videos suggested by *TikTok* were mostly from the previous six months (68%). This breakdown (Graph 1) allowed us to confirm how the platform's algorithms work, tending to prioritize the latest content and the accounts with the most followers, since the only three posts older than two years were made by the creators (@jose_eduardo_derbez, @rudyrumanoficial, and @alexis_moralejo) who had a large community of users; the peaks of activity that occurred in the month of June had a thematic explanation as they were related to the celebration of World Blood Donor Day on June 14.



Graph 1. Breakdown of posts about blood donation suggested by *TikTok*

The in-depth study of the messages' intention based on the multimodal analysis permitted us to establish three categories of content:

- Donor's testimony. In first person, the protagonist talked about the experience of donating along with the steps followed, making the viewer a participant in the whole process: from entering the facility (hospital, bus, or donation center) and registering informed consent to the actual extraction on the donation chairs and subsequent snack. The cheerful and casual attitude of the *tiktokers*, mostly young and sometimes accompanied by friends, projects an image of normality and contrasts with the sense of fear and anxiety traditionally associated with this practice (Image 1). Although it is true that a positive bias stands out in these stories, there were quite a few who expressed being upset because they had been rejected for some reason (low hemoglobin, recent travel, or sexual partners in quick succession). These situations, even if they were a one-off, also send a message: that of strict control and the priority of health safety over the individual's desire to donate. The rest of the criticisms were about the wait times or the unfriendliness of the health personnel.
- Authorized disclosure. The explanation of the main issues related to blood donation, such as the individual requirements that must be met by the potential donor (weight and age), the most common reasons for rejection (hemoglobin, medicine intake, or travel), or blood compatibility, among others, were mostly provided by personal profiles linked to health, medicine, and nursing, primarily. They also frequently mentioned myths in an attempt to make users aware of the existence of erroneous beliefs that are widespread among the population. With their posts, these profiles developed an informative function that is not just limited to the video in question but goes on by answering the many questions asked on their walls.
- Empathic awareness. A third category related to the intention of the post analyzed was found in the clear call to donate. This request was strongly emotionally charged, highlighting the personal well-being achieved with this act and/or what it would mean for the recipients in terms of health and life. In the latter situation, the advocacy fell to sick people, sometimes children, who related some specific aspect of their experience (the treatment followed, the progress that had been made in their disease, their hospital stays, etc.) always in a positive and hopeful mood. This subgroup of profiles received comments mainly of encouragement and admiration for their struggle, in addition to achieving the highest rates of commitment to mobilization (Table 1). The emotional nature of the videos is enhanced by expressions such as "you can save three lives" or "tomorrow it could be you, it could be me, or even someone in your family", which appeal directly to the receiver with the goal of not leaving them indifferent.

“ Young *tiktokers* project a happy and relaxed attitude when it comes to blood donation, in contrast to the sensation of fear and anxiety traditionally associated with this practice ”



Image 1. Representation of message categories according to their intention: testimony, disclosure, and awareness
Source: The profiles of @elenadmiguel, @donasangre.donavida, and @silviacarreras

Translations:

1. Form with personal data
2. Can I donate if I drink alcohol or smoke marijuana? You can donate provided that more than 12 hours have passed since you last drank or smoked.
3. Give blood, give life!

In conjunction with the above, we found that the topics most frequently addressed (in the following order) were focused:

(a) on the protocol followed for donation (40%), in which phlebotomy itself was very much front and center. The *tiktokers* gravitate toward streaming how the medical staff inserted the IV and how the extracted blood was stored in constantly moving bags. Although this content was not planned with an informative purpose in mind per se, they did produce the sense of closeness and understanding of a subject and procedures that the general population is not familiar with, as evidenced by the numerous questions that were commented on these posts.

(b) The second thematic block corresponded to the call for donation (31%), which was focused on attracting the attention of potential donors by presenting a specific circumstance: a shortage, a sick person, a visit of a mobile unit, or a rare blood group, among other examples.

<https://acortar.link/Bxq8sv>

(c) In third place, we found the topic focused on the intrinsic qualities of donation (29%), where questions such as requirements, previous advice, benefits, or myths associated with altruistic blood donation were addressed. In these videos, there was a clear intention to transfer knowledge, for which simple approaches, lists, and visual aids that summarized the principle key ideas were used.

Entertainment is a constant. Although it is true that it was not an end in itself for creators, it was the main method used to capture and retain the user's attention. Resources such as dances, group choreographies, visual effects, or sketches helped to bring a smile to the viewer's face and keep the audience waiting to see how the audiovisual developed. They were widely used, on both personal and institutional profiles. A good example of this is the works made by schoolchildren and posted by their respective schools.

<https://acortar.link/ySalot>

Within the blood donation content, we found that outreach was mainly supported with demonstrations and examples, establishing comparisons, and the use of common references: "that thing that looks like a lemon-lime *Calipo* is the plasma concentrate", said @perdidue (Image 2). Explanations were provided using simple vocabulary, and the main concerns or most widespread erroneous beliefs were explained. Similarly, we highlight the authority that *tiktokers* who worked as health workers used by appearing with some recognizable element of their profession –a gown, a stethoscope, the interior of an ambulance, etc.– that identified them as leaders.

Health professionals on *TikTok* go beyond an informative function to also respond to questions asked on their walls

The innovation applied to the creation of content aimed at health mobilization on *TikTok* had three features:

- Re-editing. This was especially evident when it came to sound, incorporating music and familiar and/or loud voices from third parties as another element of the creation. The ability to make duets and interacting with prerecorded content is also part of the re-editing phenomenon, although this specific situation did not occur in our sample.

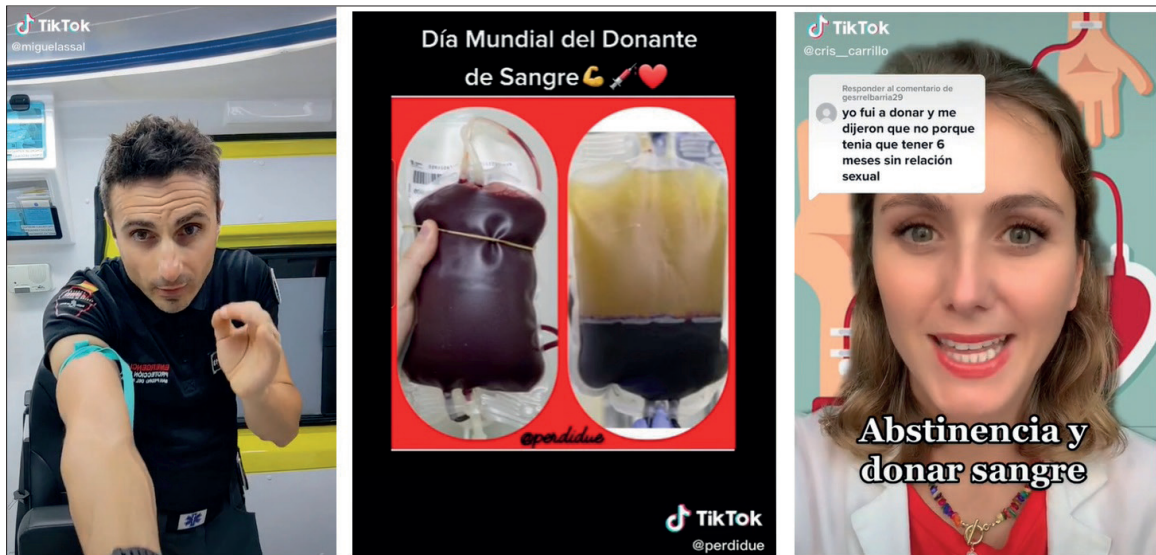


Image 2. Resources on which outreach was based
 Source: @miguelassal, @perdidue, and @cris__carrillo profiles

Translations:

2. World Blood Donor Day

3. I went to donate and they told me I couldn't because I had to have not had sexual relations for 6 months. Abstinence and donating blood

- Editing. The creators had outstanding audiovisual editing skills and presentation skills, such as incorporating on-screen advertisements, applying different visual effects, interspersing still images, or creating careful transitions between scenes.
- Reinterpretation. The most viral challenges, trends, and suggestions on the networks were applied and transformed to suit users' needs. Thus, we found dances and choreographies, such as a video (Image 1) in which several young people positioned themselves to the right and left, under a "yes" or "no", depending on the phrase being said; parodies of everyday situations; unboxing sessions, in which a young person unpacked something and commented on the package received after donating; and jokes, etc., about blood donation.

4.2. Effects on the conversation and purpose-driven action in the audience

After analyzing the narrative and the innovative nature of the primary *tiktoks* about donation, we identified their effect by studying the comments. Specifically, we focused on the type of reaction the post caused by differentiating between direct action (mobilization) and affirmation (model). The former is the serious and immediate commitment of a user to take action in support of blood donation. This is what occurred in 62 comments on 19 different videos.

They attested to the commitment of the receiver and the impact of the content consumed, using expressions such as "I donated blood just this week, and they treat you super amazingly!" or "I'm going on Tuesday" (Durántez-Stolle; Martínez-Sanz; Rodríguez-de-Dios, 2022). It should be noted that the audiovisuals that proportionally presented a greater intention to donate were those that made a call to action (Table 1), specifically those that emphasized the need for assets (low reserves) and humanized the recipients by presenting specific cases of people who were sick.

Table 1. Mobilization and affirmation rate by message theme

	Total comments	Comments about taking action	Videos with comments about taking action	Comments about affirmation	Videos with comments about affirmation
Qualities of the donation	2,421	10 (0.41%)	5 out of 29	48 (1.98%)	12 out of 29
Call to action	1,204	41 (3.40%)	8 out of 31	47 (3.91%)	9 out of 31
Donation process	2,590	11 (0.42%)	6 out of 40	60 (2.32%)	8 out of 40
Totals	6,215	62	19	155	29

With regard to affirmation, or in other words, the expression of pride in one's status as a donor¹, there are comments that are positive references. The emergence of role models—even more so when traits are shared—had a great ability to reduce fear (Igartua; Guerrero-Martín, 2022) and reinforce self-efficacy (Bandura, 2004), defined as the belief in one's capabilities to make a change. Comments of this type occurred on 155 occasions, coinciding with 29 different videos where the characteristic that most frequently recurred was the first-person account. Statements such as

"I am a donor and the truth is that it makes you feel really good I encourage you all to do it" or "every 3 months, I go. Very important"

show the audience's desire to support an act embraced by people from very different walks of life.

The amplification of the “go donate” message from the selected hashtag #donasangre also occurs in the spontaneous remarks of *TikTok* users, which, on some occasions, were equally as valuable or more valuable than the video that sparked the conversation. One example is this discussion on the profile of @miguelassal (Image 3).

Additionally, *TikTok*'s robust popularity among teenagers (IAB, 2022) is confirmed by the high number of comments from users who indicated that they were looking forward to turning 18 years old so that they could donate:

“Wait another year or two!!! ever since I was 12, I've wanted to donate blood 🥺 when I turn 18, I'm going to go donate”.

4.3. Evidence of commitment in the virtual and in-person environment

Comparing the meanings inferred from the most relevant *tiktoks* about blood donation with the perception expressed by young university students when asked in focus groups about donating blood revealed interesting points of convergence.

Students self-identified as caring and felt that they had the “ability and interest to help others” (man, group 3). *TikTok* videos starring young people donating blood showed ordinary men (33.3%) and women (66.7%) eager to contribute and happy to do so

<https://acortar.link/3Oj6GL>

This attitude was especially significant in women, the group that, in the focus groups, most frequently expressed their reluctance to donate for fear of possible pain or dizziness at the sight of blood (“I'd pass out”, woman, group 1), as was echoed in numerous comments on social networks. In contrast, lack of time and commitment was the response repeated most frequently among male university students in regard to why they do not donate.

The naturalness of the videos' protagonists helped to normalize the act of blood collection by showing it as an everyday occurrence that does not require superhuman effort or a particularly significant investment of time. Furthermore, these testimonies helped eradicate some false beliefs that were repeatedly expressed in the focus groups with university students, such as having to fast first or the risk of contagion, the mention of which undermines the high respect for and trust placed in the Spanish public health system.

We observed a general rejection of receiving compensation for blood donation because of the loss of meaning that it would entail.

“Absolutely not, that would not get me to donate. I think it would defeat the purpose” (woman, group 6).

On *TikTok*, this possibility is not even mentioned in any of the top 100 videos, although the comments do ask, in isolation, about the possibility of receiving financial compensation. Such comments sparked little debate and were met with responses such as “Saving a life is priceless”, which emphasizes the selflessness inherent in the word “donate.”

Finally, we point out that tragedies increase immediate cooperation, as many of the young people were able to remember events that had appeared in the mass media, on television, and on social networks, primarily, that awakened their conscience. On *TikTok*, 9% of the videos showed people who were living with a critical situation, generally related to cancer, reminding and explicitly asking the audience to donate. It was these direct and emotional appeals in particular that generated the highest percentage of comments (Table 1) expressing a strong commitment to donate.



Image 3. Answers that the audience provided to a user's question.

Source: The profile of @miguelassal

Translations:

1. Follow

2. does donating blood have any risks?

Respond

3. no, provided that you follow the proper protocol and do it in a sanitary place with a sanitized needle and syringe

Respond

4. Of course there is a risk, improving your well-being and your karma. You run the risk of being happy, or even being proud of yourself.

Respond

5. There's no risk at all

Respond

See more

5. Discussion and conclusions

The phenomenon of digital health activism in the context of blood donation on *TikTok* follows what **Juris** (2012) describes as “aggregation logic”, in which a common position, encapsulated in a hashtag and integrated by the spontaneous remarks of various users, creates a shared identity capable of reaching beyond the screen and which, in our specific case, manifests itself in the adoption of a healthy and charitable habit (**Eser et al.**, 2010; **Kalargirou et al.**, 2014). Despite all of this, healthcare institutions’ lack of prominence as content creators is striking –one might wonder why they are not present or prominent on *TikTok*. This means that the online visibility of this act depends on the goodwill of individual profiles that, by leveraging their audiovisual skills and interest in combating misinformation (**Salaverría**, 2021), use their experience as donors, recipients, or health professionals to highlight the benefits of donation.

The narrative innovation of the audiovisuals analyzed was based on the phenomena of appropriation, editing, and reinterpretation, which enhanced the replication of content. Likewise, the approach that some videos used showed a transition from storytelling to storydoing, and with it, the demand of a receiver who is not satisfied with just being caught up in the story but rather seeks facts and co-participation (**Rodríguez-Ríos; Lázaro-Pernias**, 2022).

Consulting young university students confirmed the persistence of false myths related to blood donation, and ascertained the level of involvement in and limitations of access to blood donation. Fear and lack of time were the main barriers, in line with the results of **Padilla-Garrido et al.** (2021) in Spain, **Greffin et al.** (2021) in the German context, and **Duboz and Cunéo** (2010) in the French context. The presence of blood and/or needles is uncomfortable, so **Lemmens et al.** (2005) recommend avoiding showing them when developing recruitment campaigns. However, the analyzed *tiktoks* that had the greatest visibility showed how the protagonists experienced the different steps of the process, including blood collection, in great detail. The comments that this inspired, far from reflecting rejection, provided a way for the audience to affirm “I am also a donor” and, ultimately, to be a role model.

The *tiktoks* relied primarily on the emotional, in line with the main motivating feature of viralization (**Arjona-Martín; Méndiz-Noguero; Victoria-Mas**, 2020). Several studies have shown that positive experiences, especially those related to healthcare in donation centers, lead to repetition and sharing by word of mouth among friends and family (**Weidmann et al.**, 2022). In addition to expanding their audience, the dissemination of these experiences in a forum such as social networks exponentially increase the leaders who are visible to young people. Investigating the influence that these digital profiles have opens up a range of possibilities in future research considering that the theory of planned behavior (TPB) assumes that subjective norms, understood as perceived social pressure, determine the individual’s important decisions (**France et al.**, 2014).

Healthcare professionals featured on *TikTok* took on the role of authoritative voices. With a relaxed tone and an explanatory and direct narrative, they performed a primarily informative task, supported by their audience (**Martínez-Sanz; Buitrago; Martín-García**, 2023). People undergoing treatment were another of the groups of protagonists. Their highly empathetic stories stood out owing to the support they received in the comments, generating the highest mobilization rates. As shown by **Durántez-Stolle, Martínez-Sanz, and Rodríguez-de-Dios** (2022), perceived similarity moderates the effects of emotional appeal when it comes to a narrative of pride and satisfaction (donor protagonist) and solidarity and empathy (recipient protagonist), as well as producing a greater intention to spread the message and become a donor.

Zulli and Zulli (2020) point out that, on *TikTok*, socialization is driven by content, thereby reducing interpersonal connections on the platform. In this sense, we found a number of disparate comments, removed from the debate or from the proposal of ideas. On the contrary, asking questions and talking about personal experiences whose value lies in reinforcing the feeling of community (**Mannell; Ahmad; Ahmad**, 2018) and empowering its members (**Francis**, 2021) were prevalent. Meanwhile, the continued raising of doubts, some associated with the false myths acknowledged in the focus groups, highlighted the need for healthcare institutions to increase their efforts to have a presence wherever the potential donors are.

“The explicit call to donate is strongly emotionally charged, especially for those creators who are or have been ill, achieving the highest levels of commitment to act. They advocate donation through their positive and hope-filled talent, and in turn receive comments full of encouragement and admiration”

6. Note

1. For the sake of accuracy and validity, we did not count the responses of those donors who said that they were donors in the past but, for some reason, were no longer donors at that time or who had tried and had been rejected. For example:

“I donated once, but because of anemia I can’t anymore 🙄🙄🙄🙄”

“for 10 years I did not miss a donation appointment. Now I’m undergoing treatment that prevents me from doing it”

7. References

- Arjona-Martín, José-Borja; Méndiz-Noguero, Alfonso; Victoria-Mas, Juan-Salvador** (2020). "Virality as a paradigm of digital communication. Review of the concept and update of the theoretical framework". *Profesional de la información*, v. 29, n. 6. <https://doi.org/10.3145/epi.2020.nov.07>
- Arribas-Urrutia, Amaia; Islas-Carmona, Octavio; Gutiérrez-Cortés, Fernando** (2019). "De prosumidores a observadores: una tendencia emergente en Internet y en los jóvenes ecuatorianos. Resultados del estudio del *World Internet Project*, Ecuador". *Revista latina de comunicación social*, n. 74, pp. 969-996. <https://doi.org/10.4185/rlcs-2019-1367>
- Azzimonti, Marina; Fernandes, Marcos** (2023). "Social media networks, fake news, and polarization". *European journal of political economy*, v. 76, 102256. <https://doi.org/10.1016/j.ejpoleco.2022.102256>
- Bandura, Albert** (2004). "Health promotion by social cognitive means". *Health education & behavior*, v. 31, n. 2, pp. 143-164. <https://doi.org/10.1177/1090198104263660>
- Bomfim-de-Souza, Mariluce-Karla; Santoro-Domingo, Pablo** (2020). "Donación de sangre y medicina transfusional en la prensa española". *Revista española de comunicación en salud*, v. 11, n. 1, pp. 9-19. <https://doi.org/10.20318/recs.2020.4717>
- Burgess, Simon; Rogers, Marg; Jefferies, Diana** (2022). "Empathy and narrative: a discussion of contemporary stories from childcare and healthcare". *International electronic journal of elementary education*, v. 14, n. 5, pp. 631-646. <https://doi.org/10.26822/iejee.2022.268>
- 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>
- Candón-Mena, José; Benítez-Eyzaguirre, Lucía** (2016). *Activismo digital y nuevos modos de ciudadanía: una mirada global*. Institut de la Comunicació (InCom-UAB). ISBN: 978 84 944171 8 4 https://incom.uab.cat/publicacions/20160000_0000_ebook12.cat.html
- Carter, Moira C.; Wilson, Jeniffer; Redpath, Gordon S.; Hayes, Paul; Mitchell, Carol** (2011). "Donor recruitment in the 21st century: challenges and lessons learned in the first decade". *Transfusion and apheresis science*, v. 45, pp. 31-43. <https://doi.org/10.1016/j.transci.2011.06.001>
- Creswell, John W.; Plano-Clark, Vicki L.** (2011). *Designing and conducting mixed methods research* (2nd. ed.). Thousand Oaks, CA: Sage. ISBN: 978 1 4129279 1 8
- Dubovi, Ilana; Tabak, Iris** (2021). "Interactions between emotional and cognitive engagement with science on *YouTube*". *Public understanding of science*, v. 30, n. 6, pp. 759-776. <https://doi.org/10.1177/0963662521990848>
- Duboz, Priscilla; Cunéo, Bernard** (2010). "How barriers to blood donation differ between lapsed donors and non-donors in France". *Transfusion medicine*, v. 20, pp. 227-236. <https://doi.org/10.1111/j.1365-3148.2010.00998.x>
- Durántez-Stolle, Patricia; Martínez-Sanz, Raquel; Rodríguez-de-Dios, Isabel** (2022). "Effects of narrative persuasive frames on *Twitter* regarding blood donation: pride versus empathy and people versus numbers". *Profesional de la información*, v. 31, n. 4. <https://doi.org/10.3145/epi.2022.jul.15>
- Eser, Bülent; Kurnaz, Fatih; Kaynar, Leilagül; Yay, Mehmet; Sivgn, Serdar; Ünal, Ali; Çetin, Mustafa** (2010). "Are university students a favorable target group for blood donation campaigns?". *Turkish journal of hematology*, v. 27, n. 4, pp. 275-281. <https://doi.org/10.5152/tjh.2010.42>
- France, Janis L.; Kowalsky, Jennifer M.; France, Christopher R.; McGlone, Sarah T.; Himawan, Lina K.; Kessler, Debra A.; Shaz, Beth H.** (2014). "Development of common metrics for donation attitude, subjective norm, perceived behavioral control, and intention for the blood donation context". *Transfusion*, v. 54, pp. 839-847. <https://doi.org/10.1111/trf.12471>
- Francis, Diane B.** (2021). "Twitter is really therapeutic at times: examination of black men's *Twitter* conversations following hip-hop artist Kid Cudi's depression disclosure". *Health communication*, v. 36, n. 4, pp. 448-456. <https://doi.org/10.1080/10410236.2019.1700436>
- Garza-Guzmán, María-Rosalía** (2009). "Fuentes genealógicas y teorías esenciales de la comunicación interpersonal". *Razón y palabra*, v. 67. <https://www.redalyc.org/pdf/1995/199520725003.pdf>

- Gisondi, Michael A.; Barber, Robert; Faust, Jeremy-Samuel; Raja, Ali; Strehlow, Matthew; Westafer, Lauren; Gottlieb, Matthias** (2022). "A deadly infodemic: social media and the power of Covid-19 misinformation". *Journal of medical internet research*, v. 24, n. 2.
<https://doi.org/10.2196/preprints.35552>
- Gómez-García, Salvador; Vicent-Ibáñez, Mireya** (2022). "¿A qué juega Netflix? La convergencia de contenidos a través de los videojuegos en su estrategia multietapa". *Revista de comunicación*, v. 21, n. 2, pp. 115-134.
<https://doi.org/10.26441/rc21.2-2022-a6>
- Greffin, Klara; Schmidt, Silke; Schönborn, Linda; Muehlan, Holger** (2021). "'Blood for blood'? Personal motives and deterrents for blood donation in the German population". *International journal of environmental research and public health*, v. 18, n. 8, 4238.
<https://doi.org/10.3390/ijerph18084238>
- Grunig, James E.; Hunt, Tood** (2003). *Dirección de relaciones públicas*. Gestión 2000. ISBN: 84 8088 395 2
- Gutiérrez-Lozano, Juan-Francisco; Cuartero, Antonio** (2020). "El auge de Twitch: nuevas ofertas audiovisuales y cambios del consumo televisivo entre la audiencia juvenil". *Ámbitos. Revista internacional de comunicación*, v. 50, pp. 159-175.
<https://doi.org/10.12795/ambitos.2020.i50.11>
- Hautea, Samantha; Parks, Perry; Takahashi, Bruno; Zeng, Jim** (2021). "Showing they care (or don't): affective publics and ambivalent climate activism on TikTok". *Social media + society*, v. 7, n. 2.
<https://doi.org/10.1177/205630512111012344>
- Hootsuite; We are Social* (2022). *Digital report 2022: el informe sobre las tendencias digitales, redes sociales y mobile*.
<https://bit.ly/3zN5VQK>
- Huis-in't-Veld, Elisabeth M. J.; De-Kort, Wim L. A. M.; Merz, Eva-María** (2019). "Determinants of blood donation willingness in the European Union: a cross-country perspective on perceived transfusion safety, concerns, and incentives". *Transfusion*, v. 59, n. 4, pp. 1273-1282.
<https://doi.org/10.1111/trf.15209>
- Hupfer, Maureen E.; Taylor, Daniel W.; Letwin, Jim A.** (2005). "Understanding Canadian student motivations and beliefs about giving blood". *Transfusion*, v. 45, n. 2, pp. 149-161.
<https://doi.org/10.1111/j.1537-2995.2004.03374.x>
- IAB* (2022). *Estudio de redes sociales 2022 de IAB Spain*.
<https://bit.ly/3T37MIO>
- Igartua, Juan-José; Guerrero-Martín, Iñigo** (2022). "Personal migrant stories as persuasive devices: effects of audience-character similarity and narrative voice". *Journal of social and political psychology*, v. 10, n. 1, pp. 21-34.
<https://doi.org/10.5964/jspp.8237>
- Juris, Jeffrey S.** (2012). "Reflections on #occupy everywhere: social media, public space, and emerging logics of aggregation". *American ethnologist*, v. 39, n. 2, pp. 259-279.
<https://doi.org/10.1111/j.1548-1425.2012.01362.x>
- Kalargirou, Aikaterini A.; Beloukas, Apostolos I.; Kosma, Alexandra G.; Nanou, Cristina I.; Saridi, Maria I.; Kriebardis, Anastasios G.** (2014). "Attitudes and behaviours of Greeks concerning blood donation: recruitment and retention campaigns should be focused on need rather than altruism". *Blood transfusion*, n. 12, pp. 320-329.
<https://doi.org/10.2450/2014.0203-13>
- Kress, Gunther** (2012). "Multimodal discourse analysis". In: Gee, James-Paul; Handford, Michael (eds.). *The Routledge handbook of discourse analysis*, pp. 35-50. Routledge. ISBN: 978 0 3674738 3 9
- Lemmens, Karin P. H.; Abraham, Charles; Hoekstra, Trynke; Ruiters, Robert A. C.; De-Kort, Wim L. A. M.; Brug, Johannes; Schaalma, Herman P.** (2005). "Why don't young people volunteer to give blood? An investigation of the correlates of donation intentions among young nondonors". *Transfusion*, v. 45, n. 6, pp. 945-955.
<https://doi.org/10.1111/j.1537-2995.2005.04379.x>
- Mannell, Jenevieve; Ahmad, Lida; Ahmad, Ayesha** (2018). "Narrative storytelling as mental health support for women experiencing gender-based violence in Afghanistan". *Social science & medicine*, v. 214, pp. 91-98.
<https://doi.org/10.1016/j.socscimed.2018.08.011>
- Martínez-Beleño, Carlos-Andrés; Sosa-Gómez, María-Silvina** (2016). "Aportaciones y diferencias entre comunicación en salud, comunicación para el desarrollo y para el cambio social". *Revista de comunicación y salud*, v. 6, pp. 71-82.
[https://doi.org/10.35669/revistadecomunicacionysalud.2016.6\(1\).71-82](https://doi.org/10.35669/revistadecomunicacionysalud.2016.6(1).71-82)

- Martínez-Sanz, Raquel; Buitrago, Álex; Martín-García, Alberto** (2023). "Comunicación para la salud a través de *TikTok*. Estudio de influencers de temática farmacéutica y conexión con su audiencia". *Revista mediterránea de comunicación*, v. 14, n. 1.
<https://www.doi.org/10.14198/MEDCOM.23435>
- Moreno-Barreneche, Sebastián** (2022). "Identidades estetizadas, interacciones mediatizadas: un abordaje de *Instagram* desde la sociosemiótica". *Comunicación y medios*, v. 31, n. 45, pp. 129-141.
<https://doi.org/10.5354/0719-1529.2022.64515>
- Navarro-Robles, María; Vázquez-Barrio, Tamara** (2020). "El consumo audiovisual de la Generación Z. El predominio del vídeo online sobre la televisión tradicional". *Ámbitos. Revista internacional de comunicación*, v. 51, pp. 10-30.
<https://doi.org/10.12795/ambitos.2020.i50.02>
- Olivares-García, Francisco-José; Méndez-Majuelos, Inés** (2020). "Análisis de las principales tendencias aparecidas en *TikTok* durante el periodo de cuarentena por la Covid-19". *Revista española de comunicación en salud*, supl. 1, pp. 243-252.
<https://doi.org/10.20318/recs.2020.5422>
- OMS (1986). *Carta de Ottawa. Primera conferencia mundial de promoción para la salud*.
<https://www.paho.org/hq/dmdocuments/2013/Carta-de-ottawa-para-la-apromocion-de-la-salud-1986-SP.pdf>
- OMS (2022). *Disponibilidad y seguridad de la sangre*. Organización Mundial de la Salud.
<https://www.who.int/es/news-room/fact-sheets/detail/blood-safety-and-availability>
- Padilla-Garrido, Nuria; Fernández-Herrera, María-Dolores; Aguado-Correa, Francisco; Rabadán-Martín, Inmaculada** (2021). "Motivators, barriers and communication channels for blood donation in relation to students at a university in Spain". *Transfusion and apheresis science*, v. 60, n. 6.
<https://doi.org/10.1016/j.transci.2021.103270>
- Porroche-Escudero, Ana** (2017). "Problematizando la desinformación en las campañas de concienciación sobre el cáncer de mama". *Gaceta sanitaria*, v. 31, n. 3, pp. 250-252.
<https://doi.org/10.1016/j.gaceta.2016.11.003>
- Rael, Christine T.; Pierre, Dominique; Frye, Victoria; Kessler, Debra; Duffy, Louisa; Malos, Nick; Van-Tieu, Hong** (2021). "Evaluating blood donor experiences and barriers/facilitators to blood donation in the United States using *YouTube* video content". *Transfusion*, v. 61, n. 9, pp. 2650-2657.
<https://doi.org/10.1111/trf.16568>
- Rodríguez-Cano, César-Augusto** (2018). "Los usuarios en su laberinto: burbujas de filtros, cámaras de eco y mediación algorítmica en la opinión pública en línea". *Virtualis: revista de cultura digital*, v. 8, n. 16, pp. 57-76.
<https://www.revistavirtualis.mx/index.php/virtualis/article/view/250/218>
- Rodríguez-Ríos, Antonio; Lázaro-Pernias, Patricia** (2022). "Nuevas narrativas en publicidad. El storydoing como modelo de comunicación que favorece la cocreación del relato de marca entre los usuarios de *Instagram*". *AdComunica*, n. 24, pp. 57-80.
<https://doi.org/10.6035/adcomunica.6554>
- Rojas-Rajs, Soledad; Jarillo-Soto, Edgar** (2013). "Comunicación para la salud y estilos de vida saludables: aportes para la reflexión desde la salud colectiva". *Interface. Comunicação, saúde, educação*, v. 17, n. 46, pp. 587-599.
<https://doi.org/10.1590/s1414-32832013000300008>
- Salaverría, Ramón** (2021). *Entender y combatir la desinformación sobre ciencia y salud*. Ministerio de Ciencia e Innovación.
<https://dadun.unav.edu/handle/10171/60223>
- Seminario-Córdova, Renzo** (2021). "Noticias falsas y su impacto en la sociedad en épocas de pandemia". *Social innova sciences*, v. 2, n. 2, pp. 6-17.
<https://socialinnovasciences.org/ojs/index.php/sis/article/view/49>
- Sheldon, Pavica; Newman, Megan** (2019). "*Instagram* and American teens: understanding motives for its use and relationship to excessive reassurance-seeking and interpersonal rejection". *The journal of social media in society*, v. 8, n. 1.
<https://www.thejsms.org/index.php/JSMS/article/view/423/209>
- Sobowale, Kunmi; Hilliard, Heather; Ignaszewski, Martha J.; Chokroverty, Linda** (2020). "Real-time communication: creating a path to covid-19 public health activism in adolescents using social media". *Journal of medical internet research*, v. 22, n. 12.
<https://doi.org/10.2196/21886>
- Sola-Morales, Salomé; Zurbano-Berenguer, Belén** (2020). "Activismo digital y feminismo: un análisis comparado de ciber campañas contra el acoso callejero en España, Marruecos y Chile". *Comunicación. Revista internacional de comunicación audiovisual, publicidad y estudios culturales*, v. 1, n. 18.
<https://doi.org/10.12795/comunicacion.2020.i18.01>

- Strauck-Franco, María-Antonia; Guillén-Arruda, Claudia-Alexandra** (2017). "Twitter como herramienta para la comunicación en salud en el contexto del fomento de la donación de órganos en el Perú". *ZER: Revista de estudios de comunicación*, v. 22, n. 42, pp. 189-209.
<https://doi.org/10.1387/zer.17841>
- Syrdal, Holly A.; Briggs, Elten** (2018). "Engagement with social media content: a qualitative exploration". *Journal of marketing theory and practice*, v. 26, n. 1-2, pp. 4-22.
<https://doi.org/10.1080/10696679.2017.1389243>
- Trauth-Taylor, Kathryn** (2021). "Innovators must be powerful storytellers". *Research-technology management*, v. 64, n. 3, pp. 43-47.
<https://doi.org/10.1080/08956308.2021.1893057>
- Vandenbosch, Laura; Eggermont, Steven** (2016). "The interrelated roles of mass media and social media in adolescents' development of an objectified self-concept: a longitudinal study". *Communication research*, v. 43, n. 8, pp. 1116-1140.
<https://doi.org/10.1177/0093650215600488>
- Verón, Eliseo** (1996). *La semiósis social. Fragmentos de una teoría de la discursividad*. Gedisa. ISBN: 978 84 743250 2 7
- Waheed, Usman; Wazeer, Akhlaaq; Saba, Noore; Qasim, Zahida** (2020). "Effectiveness of WhatsApp for blood donor mobilization campaigns during Covid-19 pandemic". *ISBT science series*, v. 15, n 4, pp. 378-380.
<https://doi.org/10.1111/voxs.12572>
- Weidmann, Christian; Derstroff, Marie; Klüter, Harald; Oesterer, Martin; Müller-Steinhardt, Michael** (2022). "Motivation, blood donor satisfaction and intention to return during the Covid-19 pandemic". *Vox sanguinis*, v. 117, n. 4, pp. 488-494.
<https://doi.org/10.1111/vox.13212>
- Yuan, Shan; Chang, Shelley; Uyeno, Kasie; Almqvist, Gay; Wang, Shirong** (2016). "Blood donation mobile applications: are donors ready?". *Transfusion*, n. 56, v. 3, pp. 614-621.
<https://doi.org/10.1111/trf.13387>
- Zabala-Cia, Oroitz; Lorenzo-Sola, Francisco; González-Pacanowski, Toni** (2022). "Interactividad en redes sociales para crear relaciones de confianza: ayuntamientos de Navarra en tránsito". *Revista de ciencias de la comunicación e información*, v. 27, pp. 25-44.
<https://doi.org/10.35742/rcci.2022.27.e246>
- Zulli, Diana; Zulli, David-James** (2020). "Extending the Internet meme: conceptualizing technological mimesis and imitation publics on the TikTok platform". *New media & society*, v. 24, n. 8, pp. 1872-1890.
<https://doi.org/10.1177/1461444820983603>

Dialnet | métricas
 Nueva edición 2020

Fundación Dialnet

dialnet.unirioja.es/metricas

Political conversations and regret: A qualitative evaluation on the aftermath of political discussions on social media

Manuel Goyanes; Porismita Borah; Homero Gil de Zúñiga

Nota: Este artículo se puede leer en español en:
<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87210>

Recommended citation:

Goyanes, Manuel; Borah, Porismita; Gil de Zúñiga, Homero (2023). "Political conversations and regret: A qualitative evaluation on the aftermath of political discussions on social media". *Profesional de la información*, v. 32, n. 3, e320302.

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

Manuscript received on 13th December 2022

Accepted on 24th March 2023



Manuel Goyanes ✉
<https://orcid.org/0000-0001-8329-0610>

Universidad Carlos III de Madrid
Departamento de Comunicación
Madrid, 133
28903 Getafe (Madrid), Spain
mgoyanes@hum.uc3m.es



Porismita Borah
<https://orcid.org/0000-0002-1140-4233>

Washington State University
Murrow College of Communication
Pullman, USA
porismita@gmail.com



Homero Gil de Zúñiga
<https://orcid.org/0000-0002-4187-3604>

Universidad de Salamanca (DRU), España
hgz@usal.es
Pennsylvania State University, USA
hgzn@psu.edu
Universidad Diego Portales, Chile
homero.gildezuniga@mail.udp.cl

Abstract

The fundamental role of political discussions in democracy has been frequently highlighted by quantitative and qualitative literature at the intersection of political communication and media effects. Most research has revolved around whether, and if so under what conditions, social media platforms constitute public spaces where democracy can be nurtured and promoted. Building on this literature and underscoring the importance of individuals' self-effects theories, this qualitative study, based on 42 in-depth interviews, clarifies how social media users navigate political discussions and their ulterior affective and cognitive processes, introducing the notion of *political discussion regret*. Specifically, this concept fundamentally emphasizes the sterility of partaking in political discussions as the main motivation for users' cognitive lamentation, which indirectly cancels the presumed muscle of social media as the sphere of public and private political discussion and deliberations. Implications of the study's findings and main theoretical consequences for the political discussion literature are also provided.

Keywords

Political discussion; Regret; Political discussion regret; Social media; Political persuasion; Social networks; Political communication; *WhatsApp*; *Twitter*; *Facebook*.

Funding

This work has benefited from the support of the *Spanish National Research Agency's* Program for the "Generation of Knowledge and the Scientific and Technological Strengthening Research + Development" Grant PID2020-115562GB-I00. The third author is funded by the *Beatriz Galindo Program* from the *Spanish Ministry of Science, Innovation & Universities*, and the *Junta de Castilla y León*.



Note

Responsibility for the information and views set out in this study lie entirely with the authors.

1. Introduction

In recent years the gateways for news use have mushroomed on social media, substantially altering citizens' consumption patterns and challenging users' objective and subjective political knowledge (Lee; Diehl; Valenzuela, 2022). In the case of Spain, the share of people who used *Facebook* for news neared 44% in 2020, a slight decline from the previous year –47%– (Statista, 2021). Beyond these practices of news distribution and consumption (Guallar *et al.*, 2016), research has also suggested that social media enable users to partake in political discussions when exposed to news posted, linked, or shared across these platforms (Valenzuela; Kim; Gil de Zúñiga, 2012); regardless of whether it takes place in public spaces –as the case of *Twitter* and *Facebook*–, or in more private, backstage environments –such as *WhatsApp* (Treré, 2020).

However, not all users, who consume news, discuss politics actively, as some users may choose either to ignore the news, contribute to the discussions, or attend discussions as mere spectators (Jordá; Goyanes, 2022). Within these practices, media effects literature has documented that cognitive elaboration is an important indirect mechanism to be politically informed when exposed to news or engage in political discussions (Eveland, 2001) and, arguably, one of the most examined self-effects (Shah, 2016; Valkenburg, 2017). In short, people use social media for an array of motivations (Rojas, 2015; Gil de Zúñiga; Valenzuela; Weeks, 2016), triggering behavioral and cognitive consequences, such as political discussion, cognitive elaboration, and learning outcomes.

Quantitative research examining the antecedents and consequences of political discussion have been prolific in the last decade (Conover; Searing; Crewe, 2002; Eveland, 2004; Valenzuela; Kim; Gil de Zúñiga, 2012; Rojas, 2015). Despite this vast knowledge, what remains unclear is to what extent and under what conditions social media political discussions lay fertile paths for engendering negative self-effects, such as *regret*. As Valkenburg, points out,

“few communication theories have conceptualized how creating or sending messages for the purpose of communicating to others may affect oneself” (Valkenburg, 2017, p. 478).

This study takes on this call and focuses on how users navigate political discussions and how the dynamic interplay of practices of persuasion, tolerance, and incivility, articulates an interesting message release effect (Pingree, 2007): social media political discussion regret.

Ultimately, feelings of regret provide an indirect mechanism to calibrate citizens' perceptions on the intrinsic value of holding political discussions on social media, enabling us to better theorize about users' cognitive and emotional sensemaking in the aftermath of such interactions, whether they were civil or uncivil. Based on a set of 42 of qualitative interviews, our findings first illustrate the main rationales behind the initiation of political exchanges on social media and how the political thematic patterns set in the agenda are the most common triggers, typically with strong ties in *WhatsApp*, and to a lesser extent with weak ties on *Twitter*.

Our study also shows that when individuals participate in political discussions, some of them activate affective and cognitive mechanisms to reflect upon such interactions. Due to this reflection on political discussions, sometimes individuals may realize that they could have handled certain situations differently, which we broadly define as social media political discussion regret. With this concept we add nuance to prior findings by showing that most users find political discussions on social media sterile, typically lamenting the participation in the *proper political discussion itself* rather than about its terms, origin and nature. This study contributes to the budding literature on self-effects (Pingree, 2007; Shah, 2016; Valkenburg, 2017), proposing a new concept to examine the potential muscle political exchanges have in activating users' feelings of regret.

2. Literature review

2.1. The dynamics of social media discussions

Defined as the

“episodes of political conversation and discussion that take place between the non-elite members of political community” (Schmitt-Beck, 2008, p. 341),

political discussion is considered as a major component of the political process (Valenzuela; Kim; Gil de Zúñiga, 2012). Scholars have consistently argued that in addition to enabling knowledge exchanges, political discussions also involve interpretive frameworks helping to process novel information (Eveland, 2004; Valenzuela; Kim; Gil de Zúñiga, 2012; Masip; Ruiz-Caballero; Suau, 2019).

Due to the growth of users on social media platforms, an increasing number of people discuss politics in these ecologies, which allow for

“creation and exchange of user-generated content” (Hampton; Shin; Lu, 2017, p. 1092).

Scholars believed that social media platforms have constituted online public spheres in which a deliberative democracy can be nurtured and promoted (**Hampton; Shin; Lu**, 2017). A defining characteristic of a deliberative democracy speaks to the free flow of information, which provides accountability and justification for the political order (**Chambers**, 2003). Therefore, free and public exchange of political information, ideas, and ideals (i.e., political discussions) on social media may provide an arena for the sustainability of democratic societies (**Halpern; Gibbs**, 2013).

Not all users who consume news discuss politics actively, as some users may choose either to ignore the news, contribute to the discussions, or attend discussions as mere spectators

A plethora of studies has probed the motivations, antecedents, and consequences of political discussions on social media platforms. Many argue that motivations and antecedents of political discussions include information seeking, opinion formation, self-presentation, persuading others, as well as entertainment and relaxation (e.g. **Conover; Searing; Crewe**, 2002; **Rojas**, 2015). Social media political discussions also have various consequences, both at the societal and at the individual levels, which involve

- political participation (e.g. **Brundidge**, 2010; **Valenzuela; Kim; Gil de Zúñiga**, 2012);
- civic engagement (e.g. **Gil de Zúñiga; Valenzuela; Weeks**, 2016);
- political knowledge acquisition (e.g. **Eveland**, 2004); and
- increased political efficacy (e.g. **Zhou; Pinkleton**, 2012).

Prior research argues that frequent political discussions on social media provide individuals exposure to political viewpoint diversity, which is beneficial for the development of discussion network heterogeneity (**Kim; Chen**, 2015). A heterogeneous discussion network, in turn, can enhance digital citizens' political tolerance and "awareness of rationales for their own political opinions," providing possibility of "reasoned public opinion and deliberative democracy" (**Kim; Chen**, 2015, p. 2347). Moreover, during these processes of information exchanges, individuals' political knowledge can also increase (**Park**, 2017). Political discussion online can enhance coordinated political efforts, such as civic engagement and political participation (**Su; Xiao**, 2022). This is due to not only the functioning of the social media affordances in channeling social pressure and reinforcement, which effectively persuade participants to engage in activism (**Halpern; Valenzuela; Katz**, 2017; **Gil de Zúñiga; Ardèvol-Abreu; Casero-Ripollés**, 2021), but also the fact that frequent and consistent political discussion can help construct shared identity among digital citizens (**Halpern; Valenzuela; Katz**, 2017).

However, political discussion can also lead to negative outcomes. For instance, as political discussions are not always heterogeneous, many people may tend to discuss with those who share similar backgrounds, potentially reducing their cognitive dissonance (**Boulianne; Koc-Michalska; Bimber**, 2020). These homogeneous discussions would facilitate the formation of echo chambers online, strengthening ideological polarization (**Taylor; Mantzaris; Garibay**, 2018), escalating populism (**Boulianne; Koc-Michalska; Bimber**, 2020), and intensifying the viral dissemination of misinformation (**Wang; Song**, 2020).

Moreover, even when discussions may fall within the more positive context of heterogeneous networks, not all discussions on social media platforms are polite. As many social media platforms afford anonymity, their users may ease in expressing themselves online in terms of politeness and interpersonal interaction norms and moral code (**Hwang; Kim; Kim**, 2018). These uncivil online discussions may trigger moral judgment of others, strong moral condemnation, interpersonal hostility, and moral indignation (**Ng; Detenber**, 2005), all of which further erode justice and legitimacy of discussions, hinder the expression of disagreements, and jeopardize a more deliberative democracy (**Hwang; Kim; Kim**, 2018).

In sum, what we know thus far about the effects of social media political discussions and the main dynamics of discussing politics in such ecologies revolves around the main, general antecedents and effects, addressing both positive and negative outcomes, while little is known as to the self-effects of social media political discussions, such as political regret.

2.2. Regret as self-effect

In general lines, regret has been observed as a corollary aspect within the choice-making literature, especially as part of emotional experiences while making decisions (**Connolly; Zeelenberg**, 2002). Although it is commonly believed that people usually make behavioral decisions to maximize expected utility such as economic profit, emotional pleasure, or political benefit, many asserted that emotions such as impulse and whim can also lead individuals to some decisions, which further elicits regret (**Landman**, 1987). Among several of these emotional dimensions in decision making (i.e., relief, anxiety) the one that arguably has attracted most research attention is regret (**Connolly; Zeelenberg**, 2002).

Regret has been studied from various perspectives. For instance, marketing science has framed regret as post-choice evaluation when considering forgone alternatives in the comparison between products (**Inman; Dyer; Jia**, 1997). Psychologists defined regret as a cognitively based emotion that helps people realize that the negative situation is due to their behavior and that they would have been better off if they chose a different action (**Zeelenberg**, 1999). These conceptualizations collectively entail a general defining character of regret: unlike other emotions such as anger or fear, regret requires more cognitive elaboration (**Connolly; Zeelenberg**, 2002). In other words, regret might be viewed as a normal,

inevitable, and direct consequence of rationality, which more likely occurs when an individual realizes that a situation is not ideal or that his/her behaviors were inappropriate (Connolly; Zeelenberg, 2002)). Therefore, regret is considered a matter of “intertwined reason and emotion” (Wang *et al.*, 2020, p. 5).

What remains unclear is to what extent and under what conditions social media political discussions lay fertile paths for engendering negative self-effects, such as regret

Regret has a series of consequences, including

- facilitating people’s ability to rapidly learn from negative outcomes (O’Connor; McCormack; Feeney, 2014);
- increasing prosocial behaviors (Martínez; Zeelenberg; Rijsman, 2011);
- making people anticipate future possible regrets and take different decisions to avoid feeling regretful again (Zeelenberg; Inman; Pieters, 2001); and
- facilitating individuals’ goal setting (Lecci; Okun; Karoly, 1994).

In a nutshell, people who experienced regret would have a higher chance to alter their future behaviors to avoid similar experiences of regret.

Although research on political regret on social media is scanty, the concept is relevant for an established area of research on self-effects (Gil de Zúñiga; Molyneux; Zheng, 2014; Valkenburg, 2017). Specifically, self-effects are

“the effects of messages on the cognitions (knowledge or beliefs), emotions, attitudes, and behavior of the message creators/senders themselves” (Valkenburg, 2017, p. 487).

Within this literature several studies have shown how expressing online can indeed affect oneself. For example, self-expression via blogging can increase the perception of support (Baker; Moore, 2008), or a higher sense of self-esteem (Schmitt-Beck, 2008). Although the extant research primarily supports the positive impact of self-effects, online political expression can have certain negative consequences such as a feeling of regret. Political regret can be thought of as release effects (Pingree, 2007), which refers to the “cognitive, attitudinal, emotional and behavioral side effects” (Valkenburg, 2017, pg. 482) that can occur after someone has sent a message or posted on social media. Combining both the literature on regret and self-effects in communication, this study aims to inductively understand the cognitive aftermath of political discussion. Accordingly, this study poses the following research questions:

RQ1. Under what circumstances do social media users regret their political discussions?

RQ2. Beyond regret, which other self-effects are salient in the aftermath of holding social media political discussions?

3. Method

Data for this study comes from a corpus of in-depth, semi-structured interviews with Spanish social media users (n = 42). Spain is one of the liberal democracies with the greatest affective polarization in the world (Miller, 2020), a fact that may have a significant impact on the political discussion emerged in social media. In addition, more than 40 million citizens use social media (Statista, 2023), turning Spain in a fertile soil for partaking in political discussion that many may end regretting. As the main objective of this study is to navigate the complexity of social media political discussions from citizens’ perspectives, we rely on rich data from interviews rather than quantitative observations from survey data, as most prior scholarship has done (Kim; Chen, 2015; Rojas, 2015; Gil de Zúñiga; Valenzuela; Weeks, 2016). Accordingly, our methodology enables us to capture and distill users’ emotional reactions stemming from political discussions on social media, “across a whole range of situations and contexts” (Couldry, 2004, p. 110).

Drawing on the work of Boczkowski and colleagues, for data collection we relied on two different, but complementary techniques: snowball sampling, and random interviews with potential respondents in public places, including libraries, bookstores, or coffee shops (Boczkowski; Mitchelstein; Matassi, 2018). Regarding snowball sampling, this method was implemented due to the lockdown restrictions implemented in Spain during the COVID-19 outbreak (April, 2020). Although snowball sampling is not often used to achieve this type of representation, a growing number of qualitative studies have implemented it to gather qualitative evidence online (Goyanes; Demeter, 2022) due to its effectiveness, cost, and external validity. Face-to-face interviews in public spaces were carried out in June 2020, and participants were randomly selected, to maximize their representation and diversity in accordance with the Spanish population.

All interviews were carried out and transcribed by a team of research assistants and eventually coded and analyzed by the first author. Interviews were carried out until saturation of ideas was achieved, following three main ethical principles: all participants were granted anonymity, were fully informed about the intended nature of the study, and permission was requested to use verbatims. Accordingly, names reported in the results are pseudonyms. Interviewees from the snowball sampling were conducted in two different digital platforms: *Google Meets* and *Skype*. The length of interviews was similar in the snowball sampling, public interviews, and across the two different digital platforms, and lasted between 25-40 minutes each. By complementing face-to-face and online interviews a balanced diversity of participants was satisfactorily achieved.

As previously stated, interviews followed a semi-structured protocol in which respondents were asked about their expe-

periences with political discussions on social media, their cognitive effects (regret), and the general emotional consequences they have generated (see appendix for the interview guide in Spanish). All interviews were carried out in Spanish and all verbatims were team traduced from this language. For the selection of potential participants, we rely on Patton's, (2002) recommendations to design a diverse sample. The diversity of participants, based on key demographic variables, such as gender, age, education, region, income, and work positions, enabled us to create a rich and diverse sample, ensuring quality data to inform our research questions (see Table 1).

After data transcription, a thematic analysis (Braun; Clarke, 2006) was implemented. Thematic analysis typically encompasses six phases, which were followed to identify, analyze, and report patterns (i.e., themes) within data. These six phases were developed in subsequent steps to first familiarize oneself with the data for generating the initial codes. Then, after inductively generating themes, we interactively reviewed their congruence, defined and named the main ones. After and before finishing the data report, we discussed the main findings with two independent researchers to clarify potential inconsistencies and clarify conceptual issues. Throughout the interviews, different thematic patterns emerged: the motivations to consume news on social media, the motivations to regret political discussions on social media, and other complementary emotional effects rooting from the political discussions. We structured the results accordingly.

Table 1. Sociodemographic of the sample vs Spanish Census

Socio-demographics	Sample (100%)	Spanish census	Regiones incluidas
Age (range = 18–65)		Range (0–86.1)	Andalucía Aragón Canarias Cantabria Castilla la Mancha Castilla y León Cataluña Comunidad Valenciana Extremadura Galicia Murcia
18–29	38.1% (n = 16)	10.3%	
0–49	50.0% (n = 21)	37.9%	
50–65	11.9% (n = 5)	37.9%	
Female	50.0% (n = 21)	50.9%	
Education		Range (0–100%)	
High school or less	40.5% (n = 17)	39.9%	
Some college	33.3% (n = 14)	22.9%	
College degree or more	26.2% (n = 11)	37.2%	
Employment status		Range (0–100%)	
Working	83.3% (n = 35)	86.22%	
Unemployed	16.7% (n = 6)	13.78%	

4. Results

4.1. Motivations and platforms for partaking in political discussions

After inquiring about respondents' political discussions on social media, our interview guide first focused on clarifying the main patterns of social media use. Not surprisingly, all our respondents declared using different social media platforms hinging on the intrinsic gratifications they cover. *Twitter*, *Facebook*, *WhatsApp*, *Instagram* or *YouTube* were, above all, the most referenced, but other more recent ones such as *TikTok* or *Twitch* were also mentioned by many of our participants. Beyond the common patterns of social media use addressed by our respondents—and systematically portrayed by extant research—, such as for entertainment or for being in touch with “friend and relatives,” most of our respondents used *Twitter* and *Facebook* as the most salient platforms for being informed about public affairs and politics.

However, there are some important nuances among our interviewees when it comes to assessing the appropriate context of social media for political discussion purposes. According to our evidence, some respondents raised critical concerns regarding the misleading nature of the news circulating massively on social media which, according to them, might spark an uninformed public opinion. Sebastián, an unemployed mechanic from Valencia, acknowledged that he “does not trust the news shared on social media,” and consequently, he claimed not to follow current news on social media, but he typically watched the news on television instead. Ana, a university sophomore from the Canary Islands, said that social media was not a suitable place for getting news and said that when she wanted to be informed about a particular (political) issue, she typically sought more actively that content on the Internet.

When respondents were asked about the frequency and nature of the political discussion they had on social media, our evidence shows an incredible diversity of opinions. Indeed, most of our respondents acknowledged their involvement in political discussions, providing three main rationales:

- because they enjoy political deliberations and think they can deliver informed opinions on certain topics,
- to challenge deceptive arguments or persuasive discourses that may mislead public opinion (i.e., misinformation) and
- because they are open to learning from others' perspectives (i.e., being persuaded), to reinforce their ideas, and to attempt to persuade others.

Susana, a schoolteacher from Santiago de Compostela, perfectly portraits her main trigger, when she blatantly explains that many people share misinformation about political leaders because they dislike them. Francisco, a carpenter from

Zamora, said that the COVID-19 health crisis had activated an array of perspectives on how to manage the pandemic and, interestingly, he thought that such diversity of perspectives “provides him a more ample and complex vision about the current situation”. Accordingly, he valued others’ thoughts, and, on many occasions, he directly acknowledged changing his opinion about different political issues, such as the restrictions of the Spanish government during the lockdown.

The most typical platform for discussing politics is, by far, *WhatsApp*. Most respondents, who initially declared having political discussions, brought up this platform, arguing that their affordances permeated the close discussion with “people I care,” as Alba, a senior architect from Segovia said. The connection with strong ties in this platform furnishes and an environment “of trust and security in which you can show your political ideals with any fear of retaliation” as Fernando, a car dealer from Zaragoza, mentioned. Despite the secure context in which those discussions emerge, Paula, a bartender from Madrid, believed that such interactions were “exhausting and pointless” as she considered that “it is very difficult to learn something”. Besides using *WhatsApp*, some respondents also used *Twitter* and *Facebook*. Two participants, who mentioned *Twitter* as the most appropriate platform, considered that the constant flow of news and the availability of information threads motivated their willingness to share their political opinion.

Another participant, Manuela, a tourist guide from Toledo, said that *Facebook* was also a potential platform for political expression, explaining how she elaborated a text to persuade people to cast their vote:

“When I discussed politics, I used *Facebook*, since it allows you to write long texts. For example, when we were in a situation of political blockade and people said they would not go to vote, I wrote a text talking about the rights and duties of citizens and social responsibility to encourage people to vote.”

4.2. Political discussion, outcomes, and regret

When we inquired about the main topics of respondents’ discussions, most of them related the thematic patterns to the political issues set in the agenda. Current issues about Spanish democratic institutions, such as laws discussed in Parliament, government performance, political party leaders’ opinions, and policy agendas were commonly brought up by many of our interviewees. In addition, many other respondents also referred to news events related to immigration, economics, gender, racism, education, or social rights. While there is an ample spectrum of thematic patterns, those above mentioned were arguably the most common topics that our participants discuss, and also represent the most salient topics our interviews eventually end up regretting discussing.

David, an industrial worker from A Coruña, pointed out that the topics he discusses the most were typically those revolving around education, health (mostly Covid-19) immigration and the related “agendas of *Vox*¹ and *Podemos*”² Fernando shared a similar thought, saying that many people and party politics prioritized the “extermination” of the universal health care system, especially, “for those most disadvantaged who do not hold a work permit [i.e., immigrants].” Efrén, a waiter from Barcelona, referred, as an example, to the coalition agreement between the *Socialist Party* and *Podemos* and the great number of ministries that the new government has. While for some respondents, political discussions on those topics were sporadic, for other participants—mostly depending on the political diversity of their network—were recurrent, even repetitive. All in all, our findings illustrate that the agenda set by the media, and the political differences that such content activate, are typically the main factors that facilitate our subjects’ political discussions.

Our evidence seems to suggest that those who frequently discussed politics typically did so with people “they care,” such as family or friends. “I have discussed with both direct family members and lifelong friends,” Sonia, a nurse from Albacete, said. Manuela also indicated that she had discussed political issues with close friends who had polarized opinions “and always want to be right,” while Paula revealed she typically discussed such issues with her father “because he has a different ideology.” These discussions among strong ties, are fundamentally taken place on instant message applications, mainly, *WhatsApp*, and the main rationale referred by our participants is because this platform enables privacy.

However, other participants said that they did not only discuss political issues with friends or family, but also with strangers, mostly on *Twitter*. “I mostly discuss about politics with unknown users,” said José, an unemployed TV producer from Cáceres. Similarly, Sergio, an industry worker from Murcia, said that he usually engages in political discussions with unknown users, because the people he knows on *Twitter* are politically aligned. Finally, there is a notable portion of respondents who clearly acknowledged that they had never participated in political discussion on social media, arguing that they perceived such interactions as “pointless” and “uncivil.”

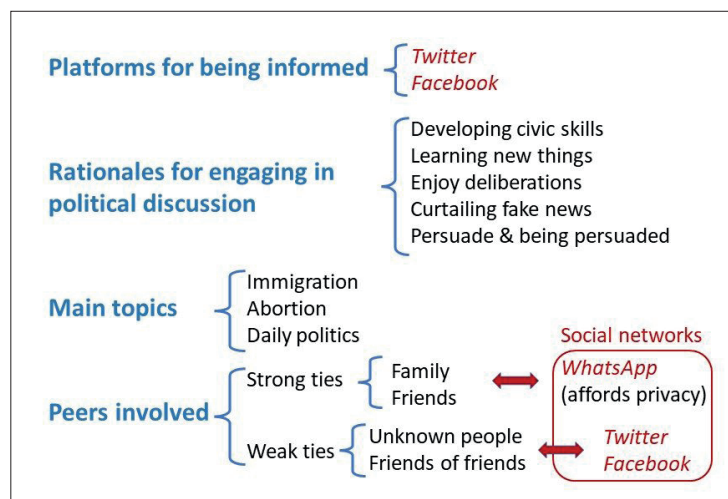


Figura 1. Discusión política en las redes sociales

Different people make different evaluations of social media political discussions they get engaged in and they end up regretting. To evaluate the usefulness of these discussions for the participants themselves we need to consider that different participants might hold different ideas about the purpose of this type of discussion. First, some respondents found political discussions on social media platforms useless for *changing others' minds* (i.e., persuade). However, some interviewees acknowledged that such discussions had moderated their viewpoints about a topic (for instance, abortion). Fernando said:

“Indeed, they have been very useful to me [the discussions], since I have learned other points of view and have been enriched by the opinions of those who think differently. Somehow, I have moderated my perspective.”

Other participants, Fátima, a senior nurse from Barcelona, believed that even if none of the participants had changed their mind, “discussions are good for better understanding others' viewpoints.”

Learning new information is also a crucial outcome of engaging in political discussions, although a significant number of participants blatantly state the contrary. Sebastián said:

“In my opinion, discussions on social media are typically not very useful. They serve to test the waters, but they do not generally generate value in establishing rational debates and sharing different and moderate points of view”.

The new information acquired may be about the political facts or news participants discuss, other people's views, or different arguments that help participants to reinforce their previous position. Sofía, a taxi driver from Castellón, said:

“As I have little political knowledge, the few times that I have discussed politics, I have obtained new information that I previously lacked, for instance about the measures implemented during the pandemic”.

Many participants felt that *sharing information with others* was a common effect of engaging on social media political discussions, while for many others *developing civic skills* was the ultimate outcome.

Interestingly, civic skills such as tolerance, participation in reasonable debates and civil exchange of opinions, are common elements brought by many of our interviewees. For instance, Mariña, an unemployed schoolteacher from Vigo, believed that she had learned from engaging in political discussions about how to better communicate, especially “by being calm and more resilient.” While some interviewees reported that they had taught or showed others how to be more tolerant, some others said that it was impossible to have civic discussions with some individuals because of the dynamics social media ecologies foster (i.e., misunderstandings, incentives for vehement and passionate arguments, not a good space for rational debate), as the following quote illustrates:

“What I have learned is that you have to know with whom and where to discuss politics. Ultimately, I think I have learned that on social media people tend to be more passionate and it is difficult to establish a rational dialogue.”

When individuals participate in political discussions on social media, some of them activate effective and cognitive mechanisms to reflect upon such interactions and is what we broadly define as *social media political discussion regret*. Specifically, and firstly, social media political discussion regret entails a common response to political discussions that aim to facilitate users' sensemaking of the perceivable role of social media in democratic processes. While many internet users may regret a specific behavior, for example, being more moderate during a discussion, social media political regret fundamentally encapsulates challenging the fact itself: it entails the cognitive and effective elaboration about the paucity of the value of discussing on social media. There are many reasons for regretting a political discussion on social media, but beyond the assumed importance of uncivil interactions (i.e., behavior), most of our respondents clarified that what they regretted the most was not the terms and nature of the conversation, but *the conversation itself*. This illustrative finding suggests that social media are no longer, if they have ever been, the deliberative spaces to foster pro-democratic values, as reflected in the following responses:

“I do not regret the terms of the discussion, but the discussions themselves because generally, they do not have any value” [Sebastián, unemployed mechanic]

“Yes, sometimes I regret having discussions about politics because generally it just creates tension and they don't bring me much new, let's say they don't evolve” [Susana, schoolteacher]

“I regret it because the political discussion has wasted my time. They don't have value” [Alba, senior architect]

Being this the most salient rationale for activating social media political discussion regret, many other testimonies also mentioned some other motivations. Alberto, a lawyer from Sevilla, regretted the vocabulary he used when discussing politics with a friend. Specifically, Alberto has labeled his friend “fascist”, something Alberto now believes should not have done, introducing an example of social costs.

What we know thus far about the effects of social media political discussions and the main dynamics of discussing politics in such ecologies revolves around the main, general antecedents and effects, addressing both positive and negative outcomes, while little is known as to the self-effects of social media political discussions, such as political regret

“The discussion was about the Parental PIN Law³. My friend mentioned that on certain issues parents should have a say over their children. Although this person did not agree with the ideals behind that Law, I labeled him fascist.”

Similarly, Alejandra regretted a conversation because it had “transpired to real-life” and affected her friendship, as she preferred to refrain from discussing politics with her friend anymore. Some other participants reflected upon other political discussions they ended up regretting. Sergio, for instance, regretted discussing political issues shared by his colleagues on *WhatsApp* about *Vox* or *Podemos*.

“It is common for some of my contacts to share opinions or proposals of these parties and use them as a throwing weapon, as if one or the other were ‘my party,’ ‘my team,’ ‘my life’.”

Finally, some interviewees pointed out how social media political discussions were different from offline discussions. On social media, there are more instances of misunderstandings, which might lead to one of the discussants being particularly offended. As Manuela said,

“We are confined and if I discuss with someone in my family, I feel bad, I cannot see him in person and speak and reason in person, it would be all very different. By *WhatsApp*, the words are misunderstood many times.”

Other participants emphasized the fact that on social media, people were more likely to misbehave or get more “passionate” or be more “intransigent.” For most “regretters,” political discussion with strong ties starts after some friend or relative shares some news links or audio-message, usually on *WhatsApp*, and people getting involved in weak tie discussion starts on *Twitter* because someone shares a news link or there is already a discussion going on there.

Table 2. Motivations for activating social media political discussion regret

Rationale	Explanation	Example quote
The discussion itself	Lack of value of the political discussions	“I do not regret the terms of the discussion, but the discussions themselves because generally they do not have any value”
	Waste of time	“I think I’ve wasted my time during the discussion”
	Nothing new to learn	“I did not learn nothing new...”
The uncivil nature of the discussion	Incivility of the conversation	“The discussions I had where not very civil, I mean, people on social media do not behave as they should, in my opinion”
	Insulting others	“...suddenly, he started to insult me”
The fact that the discussion happened on social media	Misunderstandings	“We are confined and if I discuss with someone in my family, I feel bad, I cannot see him in person and speak and reason in person, it would be all very different. By <i>WhatsApp</i> , the words are misunderstood many times”
	Social media exacerbate passion among participants	“Some people when they engage in discussions on social media become more passionate about their perspectives and sometimes this ends badly”
Effects	Explanation	Example quote
In-person	Social media discussion transpires into real life interactions	“We keep discussing [in-person] and get angry to each other during the day”
Online	Break of a dyad relation in-person and online (i.e., unfriending, blocking, unfollowing)	“The great thing about social media is that you can block people, and that is actually what I did”

Altogether, users participate in political discussions on social media for many reasons, both with strong and weak ties. Some respondents received positive outcomes from such interactions, while the vast majority fully negated their value. However, when it comes to political discussion regret, all our evidence points to a clear direction: what respondents regretted the most is to have engaged in these discussions themselves, rather than the nature and terms of the conversations, especially with strong ties or people they care.

Despite the complexity of univocally addressing common features between participants who have acknowledged regretting about social media discussions, we found certain commonalities that are interesting to highlight. First the most obvious is that they have all engaged in political discussions on social media in the past and have somewhat strong political convictions, regardless of their political ideology inclination (right or left leaning). In addition, our evidence also suggests that most of these “regretters” are young adults who report heavy social media use, and who typically adopt a passive attitude towards social media discussions, but who are particularly inclined to partake in such interactions when certain levels of intolerance or lies are significantly breached.

4.3. Complemental self-effects of social media political discussion

The cognitive elaboration of social media “regretters” also fosters other emotional self-effects that are important to understand the aftermath of political discussions on social media. Despite some of our respondents’ statements that such

discussions do not typically generate any kind of emotional response in their real lives, and thus do not affect them to a great extent, others acknowledge that other affective emotions are saliently connected to the relationship established between the peers involved (strong vs. weak ties) and the tension generated in the discussion. Indeed, those respondents who strongly regret past political discussions are more prone to neglect future political interactions, fundamentally due to their lack of perceived value, while those less apprehensive do not mention their future intentions or simply say that they will discuss politics if they think they should.

Most self-effects that our respondents agreed upon revolve around “emotional ones,” as Sebastián mentioned, and fundamentally entail annoyance, anger, tension, anxiety, and sadness. Accordingly, if the discussion materializes with strong ties, the reported self-effects are typically annoyance, anger, and tension, while if the discussion emerges with weak ties, the reported effects revolve around anxiety and sadness. Here, the interviewees directly connected those effects with being attacked or insulted on social media.

Regarding strong ties emotional self-effects, when José discussed with a friend about the political measures during the lockdown on *Facebook* he felt “momentarily” angry (i.e., only that day), and was anxious about the radical position of his friend. Likewise, Susana, who supports liberal politics, acknowledged that she discussed political issues with her father, who strongly supports Vox, on social media, and when they got together, they “keep discussing and get angry at each other during the day.”

Finally, when it comes to weak ties, sadness and anxiety are the most common and complementary responses to social media political discussions. For instance, Sergio said that during “heated discussion, even though they were unknown people,” he sometimes felt anxious for being blatantly exposed and attacked. He also provided an example about how usually the conversation evolves on *Twitter*:

“On *Twitter*, other people can join the thread and talk. Sometimes it happens that other users enter the conversation and start to discuss [a topic] with you, let’s say from the side of the person with whom you are discussing, and they come to attack instead of debate. Such situations have caused me a bit of anxiety.”

Similarly, Raquel, an unemployed waitress from Alicante, got sad when she realized that there were

“bad people, who spend most of the times on social media insulting people they do not know”.

She went on to introduce one of the most common tactics to curtail uncivil behavior, above all, with weak ties: blocking and unfriending, saying that

“The great thing about social media is that you can block people, and that is actually what I did.”

5. Discussion

Political discussions on social media are fairly widespread and yet the examination of individuals’ cognitive self-effects is relatively scarce. Although previous studies have examined the main antecedents and political outcomes of engaging in political interactions on social media (**Conover; Searing; Crewe**, 2002; **Rojas**, 2015), in this study we put the spotlight on a presumably important issue, but relatively neglected in the literature as an emotional and cognitive factor: *regret* (**Wang et al.**, 2020). Specifically, drawing upon qualitative data from in-depth interviews with adults from diverse backgrounds and geographical areas of Spain, this study aims to better understand under what circumstances social media users hold political discussions that they end up regretting, providing some insightful empirical findings and one tentative construct explication.

At the empirical level, our findings first capture and describe the main rationale behind the initiation of political discussions on social media. Developing civic skills, learning, enjoying public or private deliberations, curtailing misinformation, and persuading and being persuaded, are the most salient motivations. Prior scholarship has indeed suggested the relevance of these motivations in triggering users’ discussions (**Taylor; Mantzaris; Garibay**, 2018; **Wang; Song**, 2020) but our study clearly theorizes about their constructive nature: all of them are positive for a healthy democracy. Beyond the display of uncivil behaviour, also very common in these virtual spaces (**Goyanes; Borah; Gil de Zúñiga**, 2021), attitudes towards political interactions are fairly positive, as most citizens who engage in discussions believe they can enrich civic life, whether by means of providing their point of view or by implementing corrective actions to reduce the circulation of fabricated content. Our findings thus provide a fresh and optimistic breath to the typically apocalyptic discourse around social media, highlighting users’ proactive behavior in their attempt to contribute to the deliberative process of democracy.

Similarly, as also suggested by extant research (**Goyanes; Demeter**, 2020), political discussions on social media emerge from an amalgam of news and political issues trending in a particular moment. These thematic patterns tend to speak to the diverse political exchanges in these ecologies and their connection with news consumption (**Hampton; Shin; Lu**, 2017). In addition, our findings also suggest that such political discussions are

“Regret might be viewed as a normal, inevitable, and direct consequence of rationality, which more likely occurs when an individual realizes that a situation is not ideal or that his/her behaviors were inappropriate”

linked to the established connection between the peers involved, which directly speak to the backstage (in *WhatsApp* with strong ties) and frontstage (mostly in *Twitter* with weak ties) of political discussions on social media (Pont-Sorribes; Besalú; Codina, 2020; Treré, 2020).

Our findings illustrate that the affordances of social media enable diverse discussion dynamics which trigger different consequences. For instance, users are more prone to discuss political issues in private environments, especially with people they know or care about, and less inclined to commence political discussions in public forums with strangers, in which civic terms are not fully warranted. Despite this general inclination, the social costs associated to this generalized pattern is paradoxical: while people are more inclined to discuss politics with strong ties in private environments such as *WhatsApp* (because participants know each other and feel safe), the personal costs in terms of unfriending and flaming are undoubtedly higher than engaging in conversations with the general public, for instance on *Twitter*.

At a theoretical level, our study fleshes out the concept of social media political discussion regret to examine the aftermath of political exchanges. Specifically, this conceptual contribution aims to examine the nature, nuances, and characteristics of political conversations that users end up regretting. While beyond cognitive elaboration (e.g., Eveland, 2001) extant research has limitedly examined self-effects (Pingree, 2007; Shah, 2016; Valkenburg, 2017), our study seeks to capture under what circumstances users that held political discussions ultimately reflect upon them, causing the feeling of regret. As reported in our findings, users found most political discussions sterile, despite the pro-democratic rationales that motivated their engagement. Accordingly, there was a surprising gap between users' intended contribution to the public (*Facebook* and *Twitter*) or private (*WhatsApp*) sphere and the outcome of such political exchanges: people's prior motivations are promising, yet aftermath outcomes are fairly disappointing.

The concept of 'social media political discussion regret' also hones a strand of public and academic discourse revolving its potential effects. At times, this discourse essentially takes for granted that social media are fully responsible for diverse deleterious effects in the democratic process, as if they were fully and irresponsibly managed by artificial intelligence and biased algorithms. Contrary to these assumptions, which are indeed critical, social media political discussion regret primarily emphasizes the human nature of social media conversations, and the fact that beyond the presumed uncivil nature of many discussions (Goyanes; Borah; Gil de Zúñiga; 2021), what citizens end-up lamenting is not the terms and conditions of the political exchanges, but what is even more worrying, the conversation itself. By focusing on users' cognitions and emotions, our study clearly portrays a widespread rejection of the democratic value of political exchanges and deliberation opportunities on social media ecologies. For many, political discussions on social media become essentially pointless, and that resonates as one of the main motivations as for why users end up regretting.

However, as suggested in our results, many respondents acknowledged that they have never participated in political discussions (N = 8), and thus the theorization on regret is drawn from the rest of participants (N = 34). Accordingly, future research on social media discussion and regret should focus specifically on participants that manifest some level of regret to capture further nuances and motivations of such behavior. In addition, since the interview guide addressed questions that may resonate in participants' minds as normative, their responses may include, in some instances, desirability bias. However, as extant research has largely shown, people have different motivations to discuss politics, and those motivations, in fact, have different effects (Gil de Zúñiga; Valenzuela; Weeks, 2016). Accordingly, our findings should be read in terms of expectations: while many participants seem naïve in their motivations to discuss politics on social media, once they engage in such discussions the outcome is rather disappointing.

Our findings may also point to a double mechanism of political discussions regret, which directly speaks to the irrelevance some discussions trigger. On the one hand, some respondents may not even reflect upon the political discussions they participated in, should they not be directly inquired about their nature. Once such interrogation is raised, participants may connect the absurdity and lack of value of their past discussions in the present time, reporting feelings of regret. This cognitive mechanism could be associated with low levels of regret, thus triggering limited emotional and cognitive effects, and presumably not preventing future political exchanges. A large portion of respondents may fit well in this category, and future studies may focus on how different levels of discussion engagement may intervene to activate different levels of regret.

On the other hand, other respondents may display stronger political regret, which may be directly associated with the cognitive burden political discussions have triggered and the impact they had in participants' everyday lives. Some of these consequences may include, for instance, online and offline unfriending and strong discussions translated to the real world, especially with strong ties. Such a strong cognitive burden may have a subsequent impact on users' behavior, limiting or fully inhibiting future political discussions on social media. Future quantitative studies may be better equipped to dig into these nuances and provide empirical evidence regarding the potential mechanisms.

“ There was a surprising gap between users' intended contribution to the public (*Facebook* and *Twitter*) or private (*WhatsApp*) sphere and the outcome of such political exchanges: people's prior motivations are promising, yet aftermath outcomes are fairly disappointing ”

In conclusion, the findings presented in this study provide a much-needed nuance over some of the main effects of engaging in political discussions on social media, as well as important motivations that trigger people's future participation in this type of political discussion. In doing so, *social media political discussion regret* emerges as a vital variable that future research should consider when casting a more nuanced light on the effects of social media deliberative and discussion processes in democracy.

6. Notes

1. Vox is a right-wing populist party.
2. Podemos is a left-wing populist party.
3. The parental pin is a written request in which parents ask directors of educational centers to inform them in advance, through express authorization, about any subject, talk, workshop, or activity about gender identity, feminism or LGBTBI diversity, in such a way that parents can give their consent for their child to attend or not.

7. References

- Baker, James R.; Moore, Susan M.** (2008). "Blogging as a social tool: A psychosocial examination of the effects of blogging". *Cyberpsychology & behavior*, v. 11, n. 6, pp. 747-749.
<https://doi.org/10.1089/cpb.2008.0053>
- Boczkowski, Pablo J.; Mitchelstein, Eugenia; Matassi, Mora** (2018). "'News comes across when I'm in a moment of leisure': Understanding the practices of incidental news consumption on social media". *New media & society*, v. 20, n. 10, pp. 3523-3539.
<https://doi.org/10.1177/1461444817750396>
- Boulianne, Shelley; Koc-Michalska, Karolina; Bimber, Bruce** (2020). "Right-wing populism, social media and echo chambers in Western democracies". *New media & society*, v. 22, n. 4, pp. 683-699.
<https://doi.org/10.1177/1461444819893983>
- Braun, Virginia; Clarke, Victoria** (2006). "Using thematic analysis in psychology". *Qualitative research in psychology*, v. 3, n. 2, pp. 77-101.
<https://doi.org/10.1191/1478088706qp063oa>
- Brundidge, Jennifer** (2010). "Encountering 'difference' in the contemporary public sphere: The contribution of the internet to the heterogeneity of political discussion networks". *Journal of communication*, v. 60, n. 4, pp. 680-700.
<https://doi.org/10.1111/j.1460-2466.2010.01509.x>
- Chambers, Simone** (2003). "Deliberative democratic theory". *Annual review of political science*, v. 6, n. 1, pp. 307-326.
<https://www.annualreviews.org/doi/10.1146/annurev.polisci.6.121901.085538>
- Connolly, Terry; Zeelenberg, Marcel** (2002). "Regret in decision making". *Current directions in psychological science*, v. 11, n. 6, pp. 212-216.
<https://doi.org/10.1111/1467-8721.00203>
- Conover, Pamela-Johnston; Searing, Donald D.; Crewe, Ivor M.** (2002). "The deliberative potential of political discussion". *British journal of political science*, v. 32, n. 1, pp. 21-62.
<https://doi.org/10.1017/S0007123402000029>
- Couldry, Nick** (2004). "Theorising media as practice". *Social semiotics*, v. 14, n. 2, pp. 115-132.
<https://doi.org/10.1080/1035033042000238295>
- Eveland, William P.** (2001). "The cognitive mediation model of learning from the news: Evidence from nonelection, off-year election, and presidential election contexts". *Communication research*, v. 28, n. 5, pp. 571-601.
<https://doi.org/10.1177/009365001028005001>
- Eveland, William P.** (2004). "The effect of political discussion in producing informed citizens: The roles of information, motivation, and elaboration". *Political communication*, v. 21, n. 2, pp. 177-193.
<https://doi.org/10.1080/10584600490443877>
- Gil de Zúñiga, Homero; Ardèvol-Abreu, Alberto; Casero-Ripollés, Andreu** (2021). "WhatsApp political discussion, conventional participation and activism: exploring direct, indirect and generational effects". *Information, communication & society*, v. 24, n. 2, pp. 201-218.
<https://doi.org/10.1080/1369118X.2019.1642933>
- Gil de Zúñiga, Homero; Molyneux, Logan; Zheng, Pei** (2014). "Social media, political expression, and political participation: panel analysis of lagged and concurrent relationships". *Journal of communication*, v. 64, n. 4, pp. 612-634.
<https://doi.org/10.1111/jcom.12103>

- Gil de Zúñiga, Homero; Valenzuela, Sebastián; Weeks, Brian E.** (2016). "Motivations for political discussion: Antecedents and consequences on civic engagement". *Human communication research*, v. 42, n. 4, pp. 533-552.
<https://doi.org/10.1111/hcre.12086>
- Goyanes, Manuel; Borah, Porismita; Gil de Zúñiga, Homero** (2021). "Social media filtering and democracy: Effects of social media news use and uncivil political discussions on social media unfriending". *Computers in human behavior*, v. 120, e106759.
<https://doi.org/10.1016/j.chb.2021.106759>
- Goyanes, Manuel; Demeter, Marton** (2022). "Beyond positive or negative: Understanding the phenomenology, typologies and impact of incidental news exposure on citizens daily lives". *New media & society*, v. 24, n. 3, pp. 760-777.
<https://doi.org/10.1177/1461444820967679>
- Guallar, Javier; Suau, Jaume; Ruiz-Caballero, Carlos; Sáez, Albert; Masip, Pere** (2016). "Re-dissemination of news and public debate on social networks". *El profesional de la información*, v. 25, n. 3, pp. 358-366.
<https://doi.org/10.3145/epi.2016.may.05>
- Halpern, Daniel; Gibbs, Jennifer** (2013). "Social media as a catalyst for online deliberation? Exploring the affordances of Facebook and YouTube for political expression". *Computers in human behavior*, v. 29, n. 3, pp. 1159-1168.
<https://doi.org/10.1016/j.chb.2012.10.008>
- Halpern, Daniel; Valenzuela, Sebastián; Katz, James E.** (2017). "We face, I tweet: How different social media influence political participation through collective and internal efficacy". *Journal of computer-mediated communication*, v. 22, n. 6, pp. 320-336.
<https://doi.org/10.1111/jcc4.12198>
- Hampton, Keith N.; Shin, Inyoung; Lu, Weixu** (2017). "Social media and political discussion: when online presence silences offline conversation". *Information, communication & society*, v. 20, n. 7, pp. 1090-1107.
<https://doi.org/10.1080/1369118X.2016.1218526>
- Hwang, Hyunseo; Kim, Youngju; Kim, Yeojin** (2018). "Influence of discussion incivility on deliberation: An examination of the mediating role of moral indignation". *Communication research*, v. 45, n. 2, pp. 213-240.
<https://doi.org/10.1177/0093650215616861>
- Inman, Jeffrey J.; Dyer, James S.; Jia, Jianmin** (1997). "A generalized utility model of disappointment and regret effects on post-choice valuation". *Marketing science*, v. 16, n. 2, pp. 97-111.
<https://doi.org/10.1287/mksc.16.2.97>
- Jordá, Beatriz; Goyanes, Manuel** (2022). "The rear window effect: How users respond to political discussions and persuasive discourses in social media". *International journal of communication*, v. 16, pp. 4361-4379.
<https://ijoc.org/index.php/ijoc/article/view/19027>
- Kim, Yonghwan; Chen, Hsuan-Ting** (2015). "Discussion network heterogeneity matters: Examining a moderated mediation model of social media use and civic engagement". *International journal of communication*, v. 9, pp. 2344-2365.
<https://ijoc.org/index.php/ijoc/article/view/3254>
- Landman, Janet** (1987). "Regret: a theoretical and conceptual analysis". *Journal for the theory of social behaviour*, v. 17, n. 2, pp. 135-160.
<https://doi.org/10.1111/j.1468-5914.1987.tb00092.x>
- Lecci, Len; Okun, Morris A.; Karoly, Paul** (1994). "Life regrets and current goals as predictors of psychological adjustment". *Journal of personality and social psychology*, v. 66, n. 4, pp. 731-741.
<https://doi.org/10.1037/0022-3514.66.4.731>
- Lee, Sangwon; Diehl, Trevor; Valenzuela, Sebastián** (2022). "Rethinking the virtuous circle hypothesis on social media: Subjective versus objective knowledge and political participation". *Human communication research*, v. 48, n. 1, pp. 57-87.
<https://doi.org/10.1093/hcr/hqab014>
- Martínez, Luis M. F.; Zeelenberg, Marcel; Rijsman, John B.** (2011). "Behavioural consequences of regret and disappointment in social bargaining games". *Cognition and emotion*, v. 25, n. 2, pp. 351-359.
<https://doi.org/10.1080/02699931.2010.485889>
- Masip, Pere; Ruiz-Caballero, Carlos; Suau, Jaume** (2019). "Active audiences and social discussion on the digital public sphere. Review article". *El profesional de la información*, v. 28, n. 2.
<https://doi.org/10.3145/epi.2019.mar.04>
- Miller, Luis** (2020). *Polarización en España: más divididos por ideología e identidad que por políticas públicas*. EsadeEcPol – Center for Economic Policy and Political Economy.
<https://dobetter.esade.edu/es/polarizacion-espana>

- Ng, Elaine W. J.; Detenber, Benjamin H.** (2005). "The impact of synchronicity and civility in online political discussions on perceptions and intentions to participate". *Journal of computer-mediated communication*, v. 10, n. 3. <https://doi.org/10.1111/j.1083-6101.2005.tb00252.x>
- O'Connor, Eimear; McCormack, Teresa; Feeney, Aidan** (2014). "Do children who experience regret make better decisions? A developmental study of the behavioral consequences of regret". *Child development*, v. 85, n. 5, pp. 1995-2010. <https://doi.org/10.1111/cdev.12253>
- Park, Chang-Sup** (2017). "Do social media facilitate political learning? Social media use for news, reasoning and political knowledge". *The journal of social media in society*, v. 6, n. 2, pp. 206-238. <https://thejsms.org/index.php/JSMS/article/view/292>
- Patton, Michael-Quinn** (2002). "Two decades of developments in qualitative inquiry: A personal, experiential perspective". *Qualitative social work*, v. 1, n. 3, pp. 261-283. <https://doi.org/10.1177/1473325002001003636>
- Pingree, Raymond J.** (2007). "How messages affect their senders: A more general model of message effects and implications for deliberation". *Communication theory*, v. 17, n. 4, pp. 439-461. <https://doi.org/10.1111/j.1468-2885.2007.00306.x>
- Pont-Sorribes, Carles; Besalú, Reinald; Codina, Lluís** (2020). "WhatsApp como canal de información política en España: credibilidad, perfil de usuarios y compartición de contenidos". *Profesional de la información*, v. 29, n. 6. <https://doi.org/10.3145/epi.2020.nov.19>
- Rojas, Hernando** (2015). "Egocentric publics and perceptions of the world around us". In: Gil de Zúñiga, Homero (ed.). *New agendas in communication: New technologies & civic engagement*. New York, NY: Routledge, pp. 93-102. ISBN: 978 1 315750927 <https://doi.org/10.4324/9781315750927>
- Schmitt-Beck, Rudiger** (2008). "Interpersonal communication". In: Kaid, Lynda-Lee; Holtz-Bacha, Christina (eds.). *Encyclopedia of political communication*, Los Angeles: Sage, pp. 341-350.
- Shah, Dhavan V.** (2016). "Conversation is the soul of democracy: Expression effects, communication mediation, and digital media". *Communication and the public*, v. 1, n. 1, pp. 12-18. <https://doi.org/10.1177/2057047316628310>
- Statista** (2023). "Share of news consumers who used Facebook for news in selected European countries from 2016 to 2022". Statista. <https://www.statista.com/statistics/298038/platforms-used-to-share-online-news-uk>
- Statista** (2023). "Number of social media users in Spain from 2019 to 2028". Statista. <https://www.statista.com/statistics/569077/spain-number-social-media-users>
- Su, Yan; Xiao, Xizhu** (2022). "Interacting effects of political social media use, political discussion and political trust on civic engagement: Extending the differential gains model". *International communication gazette*, v. 84, n. 3, pp. 206-226. <https://doi.org/10.1177/1748048521993118>
- Taylor, Cameron E.; Mantzaris, Alexander V.; Garibay, Ivan** (2018). "Exploring how homophily and accessibility can facilitate polarization in social networks". *Information*, v. 9, n. 12. <https://doi.org/10.3390/info9120325>
- Treré, Emiliano** (2020). "The banality of WhatsApp: On the everyday politics of backstage activism in Mexico and Spain". *First Monday*, v. 25, n. 1. <https://doi.org/10.5210/fm.v25i12.10404>
- Valenzuela, Sebastián; Kim, Yonghwan; Gil de Zúñiga, Homero** (2012). "Social networks that matter: Exploring the role of political discussion for online political participation". *International journal of public opinion research*, v. 24, n. 2, pp. 163-184. <https://doi.org/10.1093/ijpor/edr037>
- Valkenburg, Patti M.** (2017). "Understanding self-effects in social media". *Human communication research*, v. 43, n. 4, pp. 477-490. <https://doi.org/10.1111/hcre.12113>
- Wang, Junkai; Zheng, Bowen; Liu, Hefu; Yu, Lingling** (2020). "A two-factor theoretical model of social media discontinuance: role of regret, inertia, and their antecedents". *Information technology & people*, v. 34, n. 1. <https://doi.org/10.1108/ITP-10-2018-0483>

Wang, Xiaohui; Song, Yunya (2020). "Viral misinformation and echo chambers: the diffusion of rumors about genetically modified organisms on social media". *Internet research*, v. 30, n. 5, pp. 1547-1564.

<https://doi.org/10.1108/INTR-11-2019-0491>

Zeelenberg, Marcel (1999). "Anticipated regret, expected feedback and behavioral decision making". *Journal of behavioral decision making*, v. 12, n. 2, pp. 93-106.

[https://doi.org/10.1002/\(SICI\)1099-0771\(199906\)12:2<93::AID-BDM311>3.0.CO;2-S](https://doi.org/10.1002/(SICI)1099-0771(199906)12:2<93::AID-BDM311>3.0.CO;2-S)

Zeelenberg, Marcel; Inman, J. Jeffrey; Pieters, Rik G. M. (2001). "What do we do when decisions go awry: Behavioral consequences of experienced regret". In: Weber, Elke U.; Baron, Jonathan; Loomes, Graham (eds.). *Conflict and tradeoffs in decision making*, Cambridge: Cambridge University Press. pp. 136-155. ISBN 0 521 77238 9

Zhou, Yushu; Pinkleton, Bruce E. (2012). "Modeling the effects of political information source use and online expression on young adults' political efficacy". *Mass communication and society*, v. 15, n. 6, pp. 813-830.

<https://doi.org/10.1080/15205436.2011.622064>

8. Annex. Interview guide (in Spanish)

Preguntas generales

Edad, género, situación laboral

Uso de redes sociales y uso de redes sociales para consumir noticias

¿Usa redes sociales? ¿Cuáles? ¿Por qué? ¿Para qué usa las redes sociales? ¿Alguna de ellas las usa en especial para mantenerse informado del día a día? ¿Cuáles? ¿Por qué esa/esas y no otra/otras? ¿Cuál cree que es la mejor red social para mantenerse informado? ¿Por qué?

Discusión política en redes sociales: motivaciones y temas

¿Alguna vez ha discutido sobre política en redes sociales? ¿Por qué ha discutido? ¿Alguna razón en particular? ¿Me podría poner un ejemplo? ¿Discute sobre política en redes sociales de manera habitual? ¿Por qué? ¿Me podría poner un ejemplo? ¿Cuál es, si alguna, la red social que más utiliza para discutir sobre política? ¿Por qué esa y no otra? ¿Me podría poner un ejemplo? ¿Me podría decir sobre qué temas ha discutido sobre política en redes sociales? ¿Alguna temática en particular sobre la que discuta en especial? ¿Por qué? ¿Me podría poner un ejemplo?

Discusión política en redes sociales: personas y utilidad

Cuando ha discutido sobre política, ¿con quién lo ha hecho? ¿Qué tipo de persona se trataba? ¿Ha discutido sobre política de manera reiterada con alguna persona? ¿Por qué? ¿Bajo su punto de vista, las discusiones sobre política que usted ha tenido en redes sociales le han sido útiles? ¿Por qué? ¿Considera que a través de discusiones políticas en redes sociales usted ha aprendido algo distinto o digno de destacar? ¿Por qué? ¿Y considera que ha enseñado algo a otras personas (su punto de vista, tolerancia etc.)? ¿Por qué?

Discusión política en redes sociales y arrepentimiento

¿Se arrepiente de mantener discusiones sobre política en redes sociales? ¿Por qué? ¿Me puede poner un ejemplo de discusión sobre la que se sienta arrepentido? ¿De qué tema en particular se trataba? ¿Y a través de que red social? ¿Con qué persona o personas ha mantenido esa discusión que se arrepiente? ¿Cómo le ha afectado esa discusión política en su día a día? ¿Y en tu estado emocional?, ¿Me puede poner un ejemplo?



<http://www.profesionaldelainformacion.com>

Bienvenido a EPI
Revista científica internacional

e-ISSN: 1699-2407
<https://doi.org/10.3145/EPI>

Revista internacional de
Información y Comunicación
indexada por WoS Social Sciences Citation Index (Q2),
Scopus (Q1) y otras bases de datos

Factor de impacto JCR:
JIF 2021=3,596

Scopus/SCImago Journal Rank:
SJR 2021=0,831

Using algorithms to identify social activism and climate skepticism in user-generated content on *Twitter*

Nuria Villagra; Ana Reyes-Menéndez; Jorge Clemente-Mediavilla; Dimitrina J. Semova

Nota: Este artículo se puede leer en español en:

<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87155>

Recommended citation:

Villagra, Nuria; Reyes-Menéndez, Ana; Clemente-Mediavilla, Jorge; Semova, Dimitrina J. (2023). "Using algorithms to identify social activism and climate skepticism in user-generated content on *Twitter*". *Profesional de la información*, v. 32, n. 3, e320315.

<https://doi.org/10.3145/epi.2023.may.15>

Manuscript received on 4th November 2022

Accepted on 27th April 2023



Nuria Villagra

<https://orcid.org/0000-0002-2613-1530>

Universidad Complutense de Madrid
Facultad de Ciencias de la Información
Avda. Complutense, 3
28040 Madrid, Spain
nuriavillagra@ucm.es



Ana Reyes-Menéndez

<https://orcid.org/0000-0003-0636-9573>

Universidad Rey Juan Carlos
Facultad de Ciencias Jurídicas y Sociales
Campus de Vicálvaro
P^o Artilleros, s/n
28032 Madrid, Spain
ana.reyes@urjc.es



Jorge Clemente-Mediavilla ✉

<https://orcid.org/0000-0002-9819-1129>

Universidad Complutense de Madrid
Facultad de Ciencias de la Información
Avda. Complutense, 3
28040 Madrid, Spain
jclemen@ucm.es



Dimitrina J. Semova

<https://orcid.org/0000-0001-8355-5121>

Universidad Complutense de Madrid
Facultad de Ciencias de la Información
Avda. Complutense, 3
28040 Madrid, Spain
djivkova@ucm.es

Abstract

Climate change has become an issue of great relevance in society in recent years, and the data provided by the scientific community recommend acting as soon as possible and forcefully. Scientists, politicians, the media, and thanks to the new media, citizens and other social agents participate in the debate on this issue. Despite the data and general consensus in the scientific community, the climate change debate is highly polarized, with skeptical voices denying or questioning climate change and using social media to amplify the reach of their message. This can encourage misinformation and polarization. This study tries to identify the key indicators of social skepticism around climate change through the analysis of users' social activism and behavioral patterns on *Twitter*. We analyze keywords, frequency, topics, and categories from a sample of 78,168 tweets. The results show, first, that there is an overlap of topics, with 24 of the 28 topics grouped in the intertopic distance map; second, that the size of the topics is relatively small and linked to specific events; and, third, that there is a significant political presence, especially from the United States. This work therefore contributes to the analysis of communication on *Twitter* about opinions against climate change.

Keywords

Climate change; Climate skeptics; Skepticism; Climate communication; Linguistic corpus; Algorithms; Social networks; Activisms; Indicators; Social media; Behavior patterns; Opinions; Politics; Polarization; *Twitter*.

Funding

This work has been carried out within the project 2021-1-ES02-KA220-YOU-000029048 *Green Artivism*. This project has been funded with support from the *European Commission (EC)*. This publication reflects the views only of the author, and the *EC* cannot be held responsible for any use which may be made of the information contained herein.



1. Introduction

At present, social platform users actively share information about their activities, ideas, and personal experiences through their smartphones (Dubrofsky; Wood, 2014), which leads to the generation of massive amounts of data. In recent years, such user-generated content has been used extensively for research in the social sciences (Schmid, 2016) and has been analyzed in terms of both language aspects and contexts. The results of such analysis provide a comprehensive understanding of users' behavior (Schmid, 2016; Terkourafi; Haugh, 2019).

Social identity, which is an essential component of self-concept, stems from an individual's perception of their membership to a social group(s), as well as from the significance that this individual ascribes to that membership (Tajfel, 1974). Therefore, social identity explains how identification works from the individual, interactional, and institutional perspectives (Jenkins, 2014). Accordingly, social identity has been defined as the individual's concept of the self with respect to specific aspects of social behavior (Tajfel, 1981; Kastanakis; Balabanis, 2012; Singh *et al.*, 2021).

Today, social activism is increasingly created through users' interactions with others via social media platforms, such as *Twitter* (HerdaĜdelen *et al.*, 2013; Saura; Rodríguez-Herráez; Reyes-Menéndez, 2019).

Among the major characteristics of social media platforms is users' organization into networks, that is, communities that share common interests. This makes social platforms a valuable source of data for social scientists investigating different cultural and social issues (Ntontis *et al.*, 2018).

Another important characteristic of social platforms is that their users collectively create and interact with content. This content, referred to as user-generated content (UGC), includes any content created by social platform users that is publicly shared with other users (Reyes-Menéndez *et al.*, 2020). This makes social platforms such as *Twitter* structured communities where UGC offers an enriched source of users' activities (Fujita; Harrigan; Soutar, 2018).

Recent decades have witnessed the emergence of social activism (Reyes-Menéndez; Saura; Álvarez-Alonso, 2018; Pearce *et al.*, 2019; Moernaut *et al.*, 2022). The rapid development of this new social activism has been supported by the emergence and rapid spread of new means of communication, such as email, websites, or social platforms. By facilitating the rapid spread of content around the globe, these new channels have proven to provide a valuable opportunity for people to shape their collective support (Van-de-Donk *et al.*, 2004).

Among other social platforms, *Twitter* is a very popular platform on which users can share information about their ideas, activities, opinions (Aswani *et al.*, 2018; Monfort; Villagra; López-Vázquez, 2019; Reyes-Menéndez; Saura; Álvarez-Alonso, 2018), and location (HerdaĜdelen *et al.*, 2013). Through *Twitter*, users have the opportunity to share information and their thoughts. Accordingly, there is a growing need to investigate these user interactions on *Twitter* from the perspective of social science (Stieglitz *et al.*, 2018).

The information shared by *Twitter* users can include tweets (original content created by a user) or retweets (content that users share with others); see Table 1 for a summary of such types of interaction.

Table 1. Major types of interaction on *Twitter* by grade of engagement

Interaction	Description	Presence on Twitter	UGC type
Profile mention	User A makes a mention of User B's profile in their tweet	@[Original Profile]	Profile
Tweet	A <i>Twitter</i> user writes an online post (tweet)	[Original Tweet]	Text
Retweet	User A shares User B's tweet in their profile, thereby expanding the audience	RT: @[Original Profile] Original Tweet]	Text
Like	User A presses the "like" button on User B's tweet	Like [Original Tweet]	Action
Hashtag	User A includes a tweet with a hashtag (#). Clicking the hashtag gives access to all comments published with this hashtag	#[Hashtag]	Text

Source: Reyes-Menéndez *et al.* (2020).

The analysis of the network structure on social platforms enables an analysis of not only the individual activities of specific users but also their social activism. Starting social activism on a social platform is facilitated by the joint performance of the following three effects: network structure, collaboration, and the interaction of users (Ntontis *et al.*, 2018; Saura; Reyes-Menéndez; Álvarez-Alonso, 2019). Social activism can emerge around a social profile, for instance, @GretaThunberg (Olesen, 2022), or a hashtag (#), such as #WorldEnvironmentDay (Reyes-Menéndez; Saura; Álvarez-Alonso, 2018).

The issue of climate change has grown in importance in recent decades, causing greater social concern about the effects it may have in the future.

Public interest in climate change is growing, and the European Union, for example, has earmarked significant funding within the *Horizon 2020* program for the study of this issue. The *European Green Deal Research and Innovation Program* funded a study aimed at collecting data on climate change and human opinions via *Twitter*, spanning 13 years and including more than 15 million tweets spatially distributed around the world. The variables analyzed were geo-location, user gender, climate change stance and sentiment, aggressiveness, deviations from historical temperature,

topic modeling, and information about environmental disaster events. The data provided by scientists, dramatic extreme weather events, and the reports published periodically by the *Intergovernmental Panel on Climate Change (IPCC)*, have highlighted the need to act as soon as possible and forcefully. As pointed out by **Eide and Kunelius** (2021), the year 2018 represented a turning point toward a general discourse on the subject, further promoting activism and conveying a message of urgency. Just like the *European Parliament*, instead of talking about climate change, they talk about a “climate emergency.” This idea of urgency and risk if action is not taken quickly is also enhanced by movements such as *Fridays for Future*, which have had significant social repercussions and have demonstrated the importance of citizen mobilization, in this case led by young people. All this favors an activist stance in relation to the issue that no longer involves traditional agents such as nongovernmental organizations (NGOs) but rather citizens.

Despite the data and general consensus in the scientific community, the climate change debate is highly polarized

Despite the data and general consensus in the scientific community, the climate change debate is highly polarized (**Dunlap; McCright**, 2011; **Elgesem; Steskal; Diakopoulos**, 2015; **Hoggan**, 2009; **Washington; Cook**, 2011; **Moernaut et al.**, 2022; **Pearce et al.**, 2019), and along with the voices that promote social awareness, there is also a current of thought that denies or questions climate change and downplays its effects or the role that people themselves have on it (anthropogenic climate change theory).

Social platforms have become an environmental protest space where users express their opinions and concerns about this topic. For example, the findings of research on the #WorldEnvironmentDay tag include conclusions that, among all the Sustainable Development Goals (SDGs), those of most concern to users are related to the environment and public health, such as climate change, global warming, extreme weather, water pollution, deforestation, climate risks, acid rain, and massive industrialization.

In this sense, relying on the aforementioned research, it becomes clear that the *Twitter* platform offers an opportunity to analyze UGC related to environmental issues such as climate change (**Pearce et al.**, 2019; **Moernaut et al.**, 2022) from a user perspective by enabling an analysis of both types of interactions: those organized around a profile and those organized around a hashtag.

In this context, we seek in this work to understand the social skepticism around climate change through an analysis of users’ social activism and behavioral patterns. The analysis was performed on a UGC corpus of a total of 78,168 tweets using textual analysis techniques. The first of these techniques was latent Dirichlet allocation (LDA), a machine-based technique, applied in combination with a corpus linguistic approach. We also performed discourse analysis using the log-likelihood and mutual information (MI) statistical measures.

The research question (RQ1) addressed in the present study is: What are the key indicators of social skepticism around climate change according to the analysis of users’ social activism and behavioral patterns on *Twitter*?

In what follows, we explain the theoretical framework of the present study (Section 2). This is followed by a description of the data collection process and the methodology (Section 3). The results are reported in Section 4. Section 5 presents the discussion. Conclusions, limitations of the present study, and directions for further research are presented in Section 6.

2. Theoretical framework

Previous research has focused on understanding social activism around climate change (**Reyes-Menéndez; Saura; Álvarez-Alonso**, 2018; **Pearce et al.**, 2019; **Moernaut et al.**, 2022). However, the conversation on climate change is highly polarized (**Elgesem; Steskal; Diakopoulos**, 2015; **Pearce et al.**, 2019; **Moernaut et al.**, 2022). In general, people adopting these two positions are referred to as accepters/believers and skeptics in literature. Authors such as **Washington** and **Cook** (2011) question the use of the term “skeptics” and propose that it would be more correct to call those who oppose the theory of anthropogenic climate change “deniers”. However, we use the term “skeptics” herein to refer to both those who deny as well as those who question or minimize the scientific data or theories that indicate that climate change is taking place, because this term is most commonly used in previous studies (**Capstick; Pidgeon**, 2014; **Kaiser; Rhomberg**, 2016; **Moernaut et al.**, 2022; **Van-Eck; Feindt**, 2022).

To better understand which factors can influence a person to adopt one position or another, various studies have been carried out. On the one hand, the influence of political ideology on climate change opinion has been studied (**Anderson; Huntington**, 2017; **Van-Eck; Feindt**, 2022; **Whitmarsh; Corner**, 2017). In general, the literature that relates positions on climate change with ideology distinguishes between left and right or liberal and conservative (**Elgesem; Steskal; Diakopoulos**, 2015; **Matthews**, 2015). The results indicate a tendency for those who defend more conservative positions to show less concern about climate change than those who defend a leftist position. However, more work is needed since one study carried out in Germany by **Engels et al.** (2013) found a negative correlation between political participation and skepticism. Another factor that has been studied is the influence of geographical region. **Whitmarsh** and **Capstick** (2018), for example, state that there is more climate skepticism in Western countries. A study carried out by **Hagen, Mittel** and **Pijawka**, (2016) in different countries of the European Union (Spain, the Netherlands, the United Kingdom, and

Germany) and studies carried out in the United States (**Smith; Leiserowitz, 2012**) and Great Britain (**Corner; Markowitz; Pidgeon, 2014; Capstick et al., 2015**) also highlight skepticism in public opinion and further suggest that it has become especially marked over the last two decades.

Some of the factors that are argued to be possible reasons for the greater skepticism in public opinion in recent years are

- news in the media and skeptical positions defended by politics (**Corner; Markowitz; Pidgeon, 2014**) or the scientific community (**Lahsen, 2013**);
- a lack of commitments, which were postponed to subsequent summits, at the Copenhagen *UNFCCC* in 2009 (**Van-Eck; Feindt, 2022**); and
- the climategate case (**Grundmann, 2013; Matthews, 2015; Van-Eck; Feindt, 2022**).

It is relevant that the level of education and scientific knowledge are not important factors to explain this position (**Kahan et al., 2012; Whitmarsh, 2011**), and contrary voices can even be heard within the scientific community itself (**Lahsen, 2013**), something that also has been able to contribute to increasing the level of skepticism. Additionally, opinions that deny climate change have had greater acceptance.

The documents analyzed by **McCright and Dunlap (2003)**, produced by 14 different conservative think-tanks between 1990 and 1997, conclude that climate skeptics challenge the science of global warming by:

- treating supporting evidence as weak or nonexistent;
- highlighting the potential net benefits that might result if climate change should occur; and
- clarifying that policies designed to address climate change would be economically harmful and ineffective.

Given this increase in climate skepticism, various studies have tried to establish a categorization or typology for it, although a consensus has yet to be achieved owing to the different viewpoints and attitudes associated with climate skepticism (**Matthews, 2015**). **Capstick and Pidgeon (2014)** distinguish two categories:

- epistemic skepticism: those who question science; and
- response skepticism: those who question the value of acting to prevent climate change.

Lahsen (2013) analyzes the positions defended by scientists and distinguishes two types:

- mainstream scientists, who show moderate levels of skepticism; and
- contrarian scientists, who show a high level of skepticism.

On the other hand, **Matthews (2015)** analyzes the communication from climate skeptics in blogs and distinguishes three degrees of skepticism:

- lukewarmers: who believe that pollution is affecting the planet and will continue to do so but that its impact is less than what the experts predicted; therefore, these scientists do not deny climate change but understand that the generated concern is exaggerated;
- moderate skeptics: who do not consider global warming to be a problem, believe that it has been exaggerated, and distrust the scientific theories that defend it; they understand that climate change has occurred throughout history but depends more on natural processes than on human action; and
- strong skeptics: who do not believe in the opinions of climate scientists or activists and think they are dishonest and fraudulent.

Social networks have brought about a change in traditional communication structures, making it possible for messages to be spread by citizens so that they coexist alongside the messages of traditional gatekeepers (legacy news media, companies, political parties, or the scientific community). Social media promote a more interaction-oriented and open horizontal communication than legacy media (**Dahlberg, 2001**). Especially over the last decade, it has been observed that people consult information on social networks to search for information and understand and discuss different scientific topics (**Anderson; Huntington, 2017; Su et al., 2015**). This represents a great opportunity because it enables social debate on relevant issues such as climate change, but at the same time it can contribute to misinformation and polarization. **Williams et al. (2015)** propose that the online debate on climate change is polarized with each group of believers/skeptics considering the position of their opponents to be illegitimate or unnatural. Social media platforms make it easier for anti-climate-change activists to spread their ideas than it would be in legacy news media (**Moernaut et al., 2022**).

In their work, **Bolsen and Shapiro (2017)** review the climate change topic in the US news media and the emergence of related frames in the public discourse, focusing on divisions and highlighting the role that events, journalistic practices, technological changes, and individual-level factors such as ideology and identity have played in fostering polarization. They identify the core challenges facing communicators who seek to build consensus for action on climate change and highlight the most viable solutions for generating efficient messages.

“ We have named the categories of topics to understand social skepticism around climate change through the analysis of users’ social activism and behavioral patterns ”

In “The US news media, polarization on climate change, and pathways to effective communication,” **Bolsen and Shapiro** (2017) review the results obtained from various studies over the years regarding the debate taking place about climate change on online platforms and social networks. Regarding *YouTube* uses in the United States, they identified that post-video discussions among members of the *YouTube*-viewing public tend to debate the science of climate change regardless of its relevance to the content of the videos to which they are attached (**Bolsen; Shapiro**, 2017). In other words, the public is using *YouTube*—and likely other social media discussion platforms— not to deliberate but rather to campaign for increased activism or skepticism about climate change.

One of the recommendations they make is that communicating the existence of a scientific consensus about human-caused climate change shifts the public’s belief toward the scientific consensus.

There has been extensive research on social activism on social platforms (**Hardaker; McGlashan**, 2016; **Drakett et al.**, 2018; **Fujita; Harrigan; Soutar**, 2018). As discussed above, an investigation of shared or social activism requires an analysis of the language used in UGC, including publications, posts, and interactions (**Hardaker; McGlashan**, 2016; **Kapoor et al.**, 2018).

There are many studies and much evidence showing that *Twitter* is the platform preferred by activists or social movements, acting as a real collaborative activist arena:

- **Li et al.** (2021) and **Xiong** (2019) analyze its use as a tool for feminist social movements;
- **Skill, Passero and Francisco** (2021) and **Crew** (2014) emphasize *Twitter*’s use as the platform for materializing environmental activism;
- **Zoller and Casteel** (2021) investigate a social media campaign for health activism in *Twitter*; and
- **Sinpeng** (2021) describes young political activism.

Twitter is an emerging space with an important role in the climate change debate. It allows its users to share opinions and information about climate change. Several studies on this topic have been published in recent years (**Kirilenko; Stepchenkova**, 2014; **Pearce et al.**, 2014; **Williams et al.**, 2015; **Anderson; Huntington**, 2017; **Moernaut et al.**, 2022), but given the importance and popularity of this platform for consulting and exchanging information on climate change, more work focusing on *Twitter* is needed (**Veltri; Atanasova**, 2017).

Although different elements of interaction can serve as objects of such analysis, the means most frequently used to identify relevant content are keywords, whether they incorporate a hashtag or not (**Zappavigna**, 2015; **Palos-Sánchez et al.**, 2018; **Reyes-Menéndez; Saura; Álvarez-Alonso**, 2018; **Saura; Reyes-Menéndez; Álvarez-Alonso**, 2018; **Wu et al.**, 2021). That said, it is also possible to use other elements of interaction (Table 2).

Table 2. Previous research according to the type of social interaction element studied

Authors	Interaction type	Social network	Category	Year
Hardaker; McGlashan (2016)	Profile (@CCriadoPerez)	<i>Twitter</i>	Threats	2016
Karami et al. (2020)	Hashtag (#scflood)	<i>Twitter</i>	National disasters	2020
Muralidharan et al. (2011)	Comments	<i>Twitter, Facebook</i>	Natural disasters	2011
Reyes-Menéndez; Saura; Álvarez-Alonso (2018)	Hashtag (#WorldEnvironmentDay)	<i>Twitter</i>	Environment	2018
Singh; Shula; Mishra (2018)	Hashtag (#BeefBin, #FoodSafety)	<i>Twitter</i>	Food waste	2017
Wu et al. (2021)	Comments	<i>Sina Weibo</i>	Urban waste	2021

Of all the identified movements about climate change, special attention has been paid to those regarding anti-climate views or climate skepticism. The main platform from which the analyzed content has been obtained is *Twitter*. For this, user profiles (**Hardaker; McGlashan**, 2016) or hashtags (**Singh et al.**, 2018) have been used.

Despite its advantages, *Twitter* can also contribute to disinformation and polarization. **Williams et al.** carried out an interesting study analyzing user opinions on *Twitter* and concluded that active users (either skeptics or believers) show strong attitudes in their discussions about climate change,

“characterized by strong attitude-based homophily and widespread segregation of users” (**Williams et al.**, 2015, p. 135).

Anderson and Huntington (2017) also analyze the sentiment of comments on *Twitter*, finding a persistent presence of incivility and sarcasm. The authors find that these characteristics are more frequent among skeptics and those who mention right-leaning politics in their profiles.

It is common to analyze comments in important moments such as weather events (**Anderson; Huntington**, 2017; **Capstick; Pidgeon**, 2014; **Reyes-Menéndez; Saura; Álvarez-Alonso**, 2018; **Moernaut et al.**, 2022), *Conference of the Parties (COP)* summits (**Kaiser; Rhomborg**, 2016; **Wozniak; Wessler; Lück**, 2017), the publication of *IPCC* reports (**O’Neill et al.**, 2015; **Newman**, 2017), or *climategate* (**Porter; Hellsten**, 2014).

Following previous research that framed their data extraction around specific important climate dates (**Reyes-Menéndez; Saura; Álvarez-Alonso**, 2018; **Moernaut et al.**, 2022), we expected to find an active debate as we collected data during and from *World Environment Day* in 2022.

2.1. Hypothesis development

As argued by Lakoff (2004), a shared reality is created through words and their specific uses in a discourse. Accordingly, an analysis of language opens up a way to understand the shared reality, as well as the underlying shared identity, of its participants. As mentioned above, social identity is shaped by individuals' perceptions of their belonging to a social group or groups and by the significance they attach to this (Tajfel, 1974). In this respect, Grover *et al.* (2019) argues that users' exposure to certain *Twitter* content can reinforce their previous opinions, thus causing a polarization of such views. A parallel process that can also occur is acculturation, which is defined as adaptations of an individual's views and opinions under the influence of individuals or groups from other cultural backgrounds. This suggests that such interactions should be carefully investigated (Stieglitz *et al.*, 2018). Interestingly, a study that used the information system success model showed that the influence of UGC can occur on not only the user but also organizational and social levels (Alalwan *et al.*, 2017). Therefore, online social movements can be investigated through the analysis of UGC on social platforms.

Contrary to the aforementioned studies, another paradigm that can be very useful in terms of providing a holistic perspective is that of information management (Dwivedi; Kapoor; Chen, 2015; Pace; Buzzanca; Fratocchi, 2016). From this perspective, what matters the most is not the management of information but rather the ways in which information must be provided to initiate changes in individuals' behavior. Through a review of the literature on climate skeptics, we identified an important research gap in previous research, specifically regarding social activism and climate skepticism in UGC. To address this research gap, we investigate herein the association among topics that determine the social skepticism around climate change. Our aim is to identify relevant users' social activism and behavioral patterns.

To this end, in our application of the holistic approach, we focus on both differences (Wu; Su, 1993; Arora *et al.*, 2019; Grover *et al.*, 2019) and correlations (Bouma, 2009; Iyengar; Sood; Lelkes, 2012).

With regard to the latter, the hypothesis tested in the present study is as follows:

(H1): There will be correlations between the UGC topics that identify the social skepticism around climate change through the analysis of users' social activism and behavioral patterns.

The data were collected from *Twitter*, with a focus on keywords related to social skepticism around climate change.

To collect keywords linked to anti-climate activism, we carried out an initial screening in which we obtained the 20 most mentioned and the 20 most relevant anti-climate-change hashtags (Vanhala *et al.*, 2020; Blasco-Arcas *et al.*, 2022). After analyzing these, we obtained the final hashtags that were used to extract the data: #ClimateHoax, #ClimateFraud, #ClimateBrawl, #Klimaathysterie, #ClimatePanic, #ClimateAlarm, and #ClimateFraud.

The collected tweets were then analyzed using corpus linguistics tools. In doing so, we adopted the approach previously proposed by Fujita, Harrigan and Soutar (2018). We also drew on the analysis of the texts about feminism using computational techniques carried out by Al-Nakeeb and Mufleh (2018) and the corpus linguistics and discourse analysis carried out by Hardaker and McGlashan (2016) regarding the social identity of users in the #MeToo movement, based on the comments published by @CCriadoPerez. To validate the existence of a social identity related to social skepticism around climate change through this analysis of users' social activism and behavioral patterns on *Twitter*, we complemented our analysis with log-likelihood statistical measures (Iyengar; Sood; Lelkes, 2012) and mutual information (Wu; Su, 1993; Bouma, 2009).

3. Methodology

3.1. Data collection

Following the work of Reyes-Menéndez, Saura and Álvarez-Alonso (2020), we extracted a sample of tweets to collect the data for subsequent linguistic analysis with keywords related to the climate skeptic movement between *World Environment Day* (June 5) and October 2, 2022. The optimal sample size was determined using previous studies (Saura; Rodríguez-Herráez; Reyes-Menéndez, 2019; Hardaker; McGlashan, 2016). The criteria used to extract the initial tweets collected in the present study are presented in Table 3, resulting in 78,168 tweets. To collect the database of tweets, we used *Python* 3.7.0.

Table 3. Sampling criteria for each type of interaction

Interaction type	Function
Tweet	Every tweet published with climate skeptic keywords
Retweet (RT) with no #	Every retweet mentioning the keywords
Retweet (RT) with #	Any retweet mentioning the keywords with hashtags

Source: Based on Hardaker y McGlashan (2016) and Reyes-Menéndez, Saura y Álvarez-Alonso (2020).

Next, since our aim was not to analyze multimedia content, a series of quality filters were applied to clean the data, and we eliminated images and videos (Saura; Reyes-Menéndez; Álvarez-Alonso, 2018). To increase the quality of the data, we also removed URLs from the tweets. The *Python* and *Pandas* software libraries were used for data cleaning. Specif-

ically, the commands to select or replace columns and indices to reshape lost or empty values and to debug repeated or unnecessary data were run. Finally, since retweets represent users' opinions and individual behaviors, they were analyzed separately. Table 4 presents examples of the tweets included in the final sample.

Table 4. Sample tweets included in the final sample

User	Date/time	Tweet
Matrix_backup	2022-07-20 00:21	Fecking hell!!!! It's all about FAKE man-made climate change this morning 😞 Be afraid, Be very afraid! 😞😞 😞😞 #ClimateCult #climatefraud
zoetnet	2022-06-02 02:19	Ice in #Antarctica has been increasing since 1979. #climatefraud #globalwarmingfraud

3.2. Corpus linguistics and the latent Dirichlet allocation model

Corpus linguistics (CL), a subfield of linguistics that combines both quantitative (Jia, 2018; Saura; Reyes-Menéndez; Álvarez-Alonso, 2018) and qualitative (Baker *et al.*, 2008) research methods, focuses on the analysis of large amounts of linguistic data referred to as corpora (McEnery; Hardie, 2013; Reyes-Menéndez; Saura; Filipe, 2019).

The latent Dirichlet allocation (LDA) model is a widely used quantitative technique in corpus linguistics. Initially developed by Pritchard, Stephens, and Donnelly, (2000) as a machine-based technique, LDA was subsequently improved and expanded upon by Blei, Ng, and Jordan (2003). With the support of artificial intelligence (AI), this model enables the identification of both keywords and topics linked to them. The key assumption of LDA is that the topics in a database are not observable a priori and should be analyzed through a probabilistic model (Reyes-Menéndez *et al.*, 2020). Specifically, the model aims to determine the number of times a given word is repeated in a corpus or an individual document. The mathematical model developed in *Python* establishes the variables in the sample; these variables are known as latent variables. These variables are used to determine the number of topics identified by the algorithm, considering the importance of those variables (Jia, 2018). In this way, once the algorithm determines the total number of words and that of repeated words, as well as the number of each of the most frequent words that occur before and after the identified words, each topic is assigned a name. The quality of the data is important for the quality of our model, so we preprocess the data by removing symbols with regular expressions and performing tokenization and delete punctuation and create *N*-grams (bigrams and trigrams), applying lemmatization and removing stop-words. Using a standardized process of the LDA model based on grounded theory studies, each topic's name is derived from the words within each of the clusters identified.

Thus, the LDA model consists of the following two steps: First, all keywords present in the corpus are obtained. Second, the topics linked to these keywords are identified (Reyes-Menéndez *et al.*, 2020). To identify topics in a maximally objective way, the mathematical distribution shown in Equation (1) is employed.

$$p(\beta_{1:K}, \theta_{1:D}, z_{1:D}, w_{1:D}) = \prod_{i=1}^K p(\beta_i) \cdot \prod_{d=1}^D p(\theta_d) \cdot \sum_{n=1}^N p(z_{d,n} | \theta_d) p(w_{d,n} | \beta_{1:K}, z_{d,n}) \quad (1)$$

β_i is the distribution of word in topic i among a total of K topics

θ_d is the proportion of topics in document d among a total of D documents

z_d is the topic assignment in document d

$z_{d,n}$ is the topic assignment for the n th word in document d among a total of N words

w_d is the observed words for document d

$w_{d,n}$ is the n th word for document d

In the next step, to identify the topics that make up the dataset, we used Gibbs sampling [Equation (2); Jia, 2018] using the Mac version of the Python software LDA 1.0.5.

$$p(\beta_{1:K}, \theta_{1:D}, z_{1:D} | w_{1:D}) = \frac{p(\beta_{1:K}, \theta_{1:D}, z_{1:D}, w_{1:D})}{p(w_{1:D})} \quad (2)$$

4. Results

This section reports the results we obtained on the keywords and frequency related to social skepticism around climate change through the analysis of users' social activism and behavioral patterns identified in our corpus (Section 4.1), the topics (Section 4.2), and the corresponding categories, social activism, and behavioral patterns (Section 4.3)

4.1. Keywords and frequency

We carried out an analysis of the keywords in the corpus, considering the importance of the fact that keywords express user behavior and the linguistic importance of the terms (Reyes-Menéndez *et al.*, 2020).

In this same line, the frequency of a term’s occurrence in a text is a key measure in corpus linguistics. Frequency is assumed to highlight users’ social identity (McEnery; Hardie, 2013). Here, frequency is defined as the number of times a word appears in a given text (Baker et al., 2008). Table 5 lists the frequencies of 10 main words identified in our data.

Table 5. Frequency of 10 main words

Rank	Word	Similar words	Frequency
1	ClimateBrawl	Brawl, ClimateBrawlers	4,246
2	ClimateHoax	ClimateHoaxWe, UnrealClimateHoax	3,461
3	ClimateCrisis	NoClimateCrisis, Crisis, KlimaatCrisis	3,525
4	ClimateEmergency	NoClimateEmergency, Emergency	3,050
5	ClimateScam	Scam, Scammers, Scams	1,567
6	WEF	WorldEconomicForum, WefCrimeAgainstHumanity, FTheWef	1,205
7	Resist	TheGreatResist, Resistance	1,006
8	Science	JunkScience, Sciences	801
9	Farmers	Farmer, FarmerProtests, NoFarmersNoFood	766
10	Support	Supporting, Supports, Supporters	735

As seen in Table 5, the most frequent term in our data is the keyword “ClimateBrawl” (4,246 times) that was previously identified as a hashtag to be extracted. This was also the case with the hashtag “ClimateHoax” (3,461 times).

Other keywords that were not extracted as hashtags were “ClimateCrisis” (3,525 times), “ClimateEmergency” (3,050 times), and “ClimateScam” (1,567 times).

The fact that “WEF,” which corresponds to “World Economic Forum,” is present 1,205 times is interesting, as this is a global forum for economic development. Also, “science” is present 801 times, in terms such as “JunkScience.” Additionally, “farmers” is mentioned 766 times, showing the interest that food production has for users.

4.2. Topics

Topics in a corpus are clusters of words linked to each other. Accordingly, topics are intrinsically related to their keywords (Reyes-Menéndez et al., 2020). To find topics in our database, the LDA model and its corresponding Equation (1) were used (Section 3.2).

Next, to evaluate our LDA model, we used the metric referred to as topic coherence, which measures the relative distance between words within a topic (Syed; Spruit, 2017; Rama-Maneiro; Vidal; Lama, 2020). It is rare to see a coherence of 1 or +0.9 unless the words being measured are either identical words or bigrams. The overall coherence score of a topic is the average of the distances between words. We attain a value of 0.34 in our LDAs, since there is no strong topic correlation; in other words, the distance between words within topics is not very close.

To determine whether the identified topics are relevant key indicators of social skepticism around climate change using the analysis of users’ social activism and

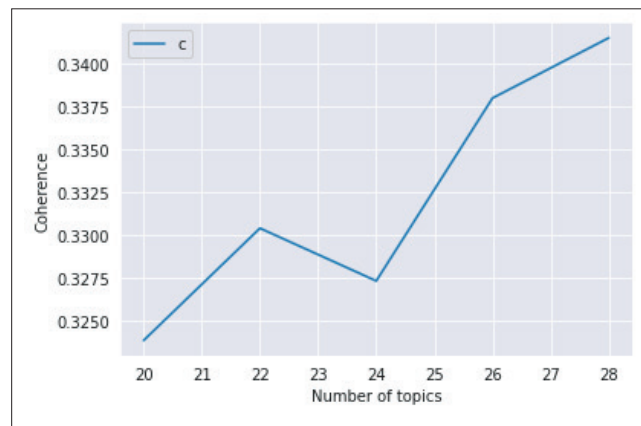


Figure 1. Coherence of topics

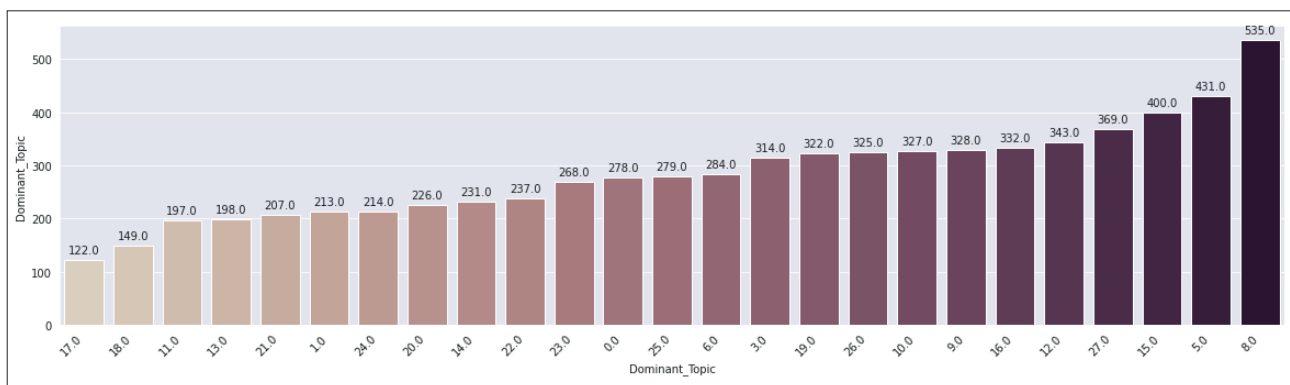


Figure 2. Bar graph of topics

behavioral patterns on *Twitter*, we relied on the measure of coherence. This function, built in *Python*, searches for an optimal number of topics in the dataset. The graph (Figure 1) shows 28 topics as optimal, with a coherence score of ~0.34 listing the ideal number of topics that will compose the social skepticism around climate change using the analysis of users' social activism and behavioral patterns.

As seen in Figure 2, the 28 identified topics have different contributions to the overall research. The topic with the greatest contribution is topic 8.0.

Below, we present the contribution of the 28 topics identified (Table 6) in the tweets database. We also highlight the main keywords that make up each topic, and each topic has been assigned a name with a randomized controlled process (Jia, 2018).

Table 6. Topic contribution and keywords of the topics

Topic name	Topic num.	Keywords	Topic contribution
Trump	0.0	Climatescam, climate, change, instill, fly, fear, trump, covid	0.654
Temperatura	1.0	climatefraud, colour, potus, sea_level, temperature	0.637
Story	2.0	climatehoax, climatescam, climate, change, globalwarme, week, happen, story, enjoy, destroy	0.652
Coal	3.0	climatescam, climatehoax, climate, change, eat, summer, coal, support, fart, trillion	0.649
Diesel	4.0	climatehoax, climatescam, climate, change, liberal, wef, year, people, diesel, investment	0.628
Media	5.0	climatescam, climatecrisis, climatehoax, climateaction climate, change, support, bbcnew, lie, travel	0.692
Brainwash	6.0	climatehoax, climatescam, brainwash, hear, climate, emergency, global_warme, crap, power, try	0.648
Fakenews	7.0	climatehoax, climatecrisis, climate, change, fakenew, people, life, today, save, love	0.658
Lie	8.0	climatehoax, climatescam, agenda, fear, end, support, weather, lie, climate, control	0.600
War	9.0	climatehoax, climatescam, support, joke, wake, year, war, emergency, climate, resist	0.665
Private jet	10.0	climatescam, climatecult, climatehoax, climate, change, emergency, private_jet, brain, starve, years_ago	0.607
Wef	11.0	climatehoax, climate, change, wefpuppet, ton, ocean, globalwarming, game, year, freedom	0.597
Politics	12.0	climatehoax, climatescam, climatecrisis, fuck, change, climate, retweet, politician, potus, human	0.635
Covid	13.0	climatescam, climatecrisis, climate, change, covid, basisscholen, agenda, laten, thegreatreset, charge	0.613
Dream	14.0	climatecrisis, climatehoax, climateaction, climategrifter, al_gore, climate, year, dream, support, straight	0.659
Bullshit	15.0	climatehoax, climatescam, climatecrisis, change, bullshit, listen, support, believe, climate, scam	0.630
Cows	16.0	climate, change, charge, fall, cow, carbon, alarmist, hypocrisy, weekend, work	0.659
Education	17.0	climatehoax, climatescam, climate, crisis, basisscholen, laten, thegreatreset, surprise, spot	0.611
Green new deal	18.0	climatehoax, runderen, die, geoengineering, still_legal, change, climate, support, greennewdeal, geoengineere	0.591
Biden farmers	19.0	climatescam, climatehoax, climatecrisis, climate, change, agenda, farmer, prediction, money, biden	0.625
Tax	20.0	Climatescam, tax, sky, record, gas_price, accord, warm, climate, vaccine, temperature	0.601
Narrative	21.0	climatehoax, climatecrisis, climatechange, climatescam, climate, change, narrative, speech, pay, state	0.627
Negative	22.0	climatescam, climatehoax, eat, basisscholen, laten, thegreatreset, agenda, insect, support	0.613
Support	23.0	climatehoax, climatecrisis, climatescam, climate, change, auspol, stop, die, wrong, support	0.644
Red painting rule	24.0	climatehoax, climatescam, climate, change, red, painting, rule, truth, list, night	0.638
Globalwarming	25.0	climatehoax, climatescam, climatecrisis, climate, change, window, sun, globalwarme, support, truth	0.618
Administration	26.0	climatecrisis, climatehoax, climate, change, hoax, buy, administration, excuse, open, support	0.631
Obama	27.0	climatescam, climatehoax, climatecult, climate, change, support, obama, control, rest, science	0.640

Figure 4 shows the intertopic distance map. This visualization represents the different topics and the distance between them. Similar topics appear closer and dissimilar topics farther away. The relative size of a topic's circle in the plot corresponds to the relative frequency of that topic in the corpus. In our case, topics 14.0 and 18.0 appear closer and not far from topic 23.0, while topic 19.0 is far away. The overlap among the remaining topics (1.0, 2.0, 3.0, 4.0, 5.0, 6.0, 7.0, 8.0, 9.0, 10.0, 11.0, 12.0, 13.0, 15.0, 16.0, 17.0, 20.0, 21.0, 22.0, 24.0, 25.0, 26.0, 27.0, and 28.0) is also noteworthy. Specifically, the rule of thumb was as follows: the shorter the distance between the central node and the topics, the greater the link between them (Al-Nakeeb; Mufleh, 2018).

This visualization reveals that topic 19.0 "biden farmers" in Table 7 is isolated and has a greater distance from the other topics, while topics 14 "dream," 18 "green new deal," and 23 "support" lie in the same quadrant and can form a category of topics.

The remaining topics (“trump,” “temperature,” “story,” “coal,” “diesel,” “media,” “brainwash,” “fakenews,” “lie,” “war,” “private jet,” “wef,” “politics,” “covid,” “dream,” “bullshit,” “cows,” “education,” “tax,” “narrative,” “negative,” “Red painting rule,” “globalwarming,” “administration,” and “Obama”) overlap and, therefore, can form a third category of topics.

4.3. Categories of topics, social skepticism, and behavioral patterns around climate change

We have named the categories of topics to understand social skepticism around climate change using the analysis of users’ social activism and behavioral patterns. Furthermore, to identify the different categories in which the topics fall, a name has been assigned through a randomized controlled process (Jia, 2018).

The groupings of topics explained above serve as the basis for the development of the categories of the social identity and behavioral patterns. In this way, we obtained the following three different categories:

- Biden;
- *Green New Deal*;
- Hoax.

Which topics correspond to which categories is explained in Table 7.

Table 7. Categories of topics

Categories	Topic num.	Topic tag	Description
Biden	19.0	Biden	This category presents the proposal developed by Biden to fund farmers’ losses among other proposals. Opinions on Biden are mostly critical.
<i>Green New Deal</i>	14.0 18.0 23.0	Dream <i>Green New Deal</i> Support	This category shows the importance of the <i>Green New Deal</i> and the positive aspect of this proposal for those who consider it necessary. While the detractors use critical comments to express their perception of the proposal and its unsustainable nature for companies and individuals. This is especially notable in fuels, economy, supplies.
Hoax	0.0 1.0 2.0 3.0 4.0 5.0 6.0 7.0 8.0 9.0 10.0 11.0 12.0 13.0 15.0 16.0 17.0 20.0 21.0 22.0 24.0 25.0 26.0 27.0	Trump Temperature Story Coal Diesel Media Brainwash Fakenews Lie War Private jet WEF Politics Covid Bullshit Cows Education Tax Narrative Negative Red painting rule Global warming Administration Obama	This category includes all the other topics that are also linked by the hashtags (#) included in the search and other hashtags such as #ClimateCrisis or #ClimateScam. The fact that these hashtags are strongly present in all the topics means that they are presented as a consolidated category determined by the strong use of hashtags linked to anti-climate-change activism.

As can be seen in Table 7, there are a series of topics that are well determined and cohesive and that are of considerable size. These topics are “dream,” “green new deal,” and “support,” and they belong to the category *Green New Deal*.

Table 8 below shows some examples of tweets that belong to the different categories identified as Biden, *Green New Deal* and Hoax.

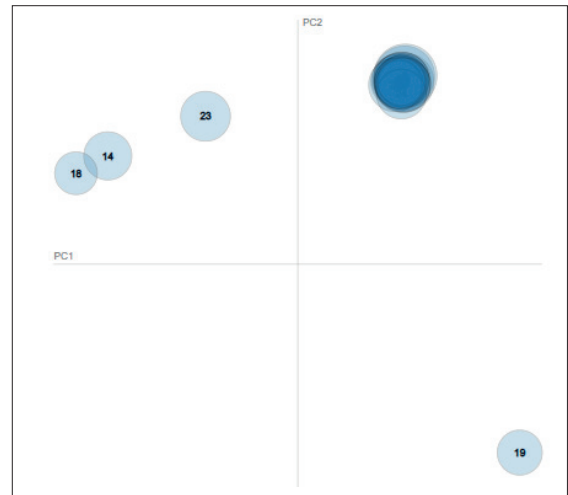


Figure 3. Intertopic distance map

Table 8. Sample of tweets per category

Cat.	User	Date/time	Tweet
Biden	Jon Tveten	2022-07-11 09:44	Green tyranny has finally provoked mass reactions, and the first government has fallen after imposing insane policies that wrecked the food supply for its people. @Thomas_Lifson #ClimateFraud
	Genuinedavid	2022-06-21 22:44	Biden goes begging countries for (dirtier than ours) oil while simultaneously exporting our own clean domestic oil, thus pumping much more c02 via redundant tanker ships. #ClimateFraud Why Is The United States Still Exporting Fuel?
	Jorj X McKie	2022-06-10 18:59	More #ClimateFraud from the #BidenAdministration RT @ClimateDepot: Climate Depot's Morano: "After years of claiming you can't challenge 'the science' they now claim you can't challenge their 'solutions!'"
Green New Deal	Kevin Killough	2022-07-07 02:57	I want to set up a charity to help European victims of green energy policies. Even if I could crowdsource or something, I'd have no idea how to get it to those who are suffering under these policies. #ClimateBrawl #GreenEnergyKills
	Steve Tatum	2022-10-04 12:47	They're shutting down our pipelines, canceling our drilling projects, blowing up our gas prices, and destroying our economy just for their <i>Green New Deal</i> Climate Cult Hoax. #LetsGoBrandon #GasPrices #ClimateHoax #VoteGOP2022
Hoax	Mel	2022-09-29 17:17	Just a friendly reminder to NEVER trust the mainstream media... #FakeNews #Climate-Hoax
	<i>Blinkered Britain</i>	2022-09-21 02:09	Be afraid very afraid. 'My Carbon': An approach for inclusive and sustainable cities 'COVID-19 was the test of social responsibility' #WEF #TalkRadio #TalkTV #GBNews #ClimateHoax #ClimateChange #GreenAgenda

It should be highlighted that the topic "biden farmers," owing to its position in the intertopic distance map and its size, is a singleton category, while the rest of the topics overlap, composing the last category referred to as Hoax. This category includes all of the other topics that are also linked by the hashtags (#) included in the search as well as other hashtags such as #ClimateCrisis or #ClimateScam. The fact that these hashtags are strongly present in all the topics means that they are presented as a consolidated category determined by the strong use of hashtags linked to anti-climate-change activism.

5. Discussion

In the present study, we used a systematic literature review to identify, evaluate, and synthesize social skepticism around climate change indicators through an analysis of users' social activism and behavioral patterns on *Twitter*. Our study answers **Veltri** and **Atasanova** (2017) call for research efforts to better understand UGC content on Twitter.

Some previous research (**Capstick; Pidgeon**, 2014; **Anderson; Huntington**, 2017; **Wozniak; Wessler; Lück**, 2017; **Reyes-Menéndez; Saura; Álvarez-Alonso**, 2018; **Moernaut et al.**, 2022) has been developed around events related to climate change. In the same line, this research has obtained satisfactory results with data extracted during *World Environment Day* in 2022.

Numerous previous investigations have linked climate skepticism on *Twitter* with issues of political ideology (**Whitmarsh; Corner**, 2017; **Van-Eck; Feindt**, 2022), and even with determining the discourse on the basis of the political position of the users and the parties (**Elgesem; Steskal; Diakopoulos**, 2015; **Matthews**, 2015). This is in line with the results obtained herein because the political presence is evident in the topic analysis (Section 4.2) with topics such as topic 19.0 "biden farmers," "trump," "politics," and "obama."

With respect to the categories of topics identified (Section 4.3), one of them is the "*Green New Deal*." This category identified in our results has not been analyzed in previous research dealing with social skepticism and behavioral patterns around climate change, so this opens new lines of research on skepticism about climate change, the communication that is carried out in this sense about climate change, climate on *Twitter*, and the *Green New Deal*.

“ Social media platforms make it easier for anti-climate change activists to spread their ideas than it would be in legacy news media ”

Some of the practical implications of this work are the application of the results for the development of public and private policies by institutions, governments, or companies that are concerned about climate change. Social media platforms are a space where awareness of the climate change can be promoted, but also where the opposite effect can be achieved by amplifying anti-climate change views. As **Moernaut et al.** (2022) point out, in contrast to traditional media, social networks allow anti-climate change activists to easily disseminate their ideas.

Another of the practical implications would be related to the education of those who think that climate change is not a worrying issue but rather a lie. In this sense, it is also necessary to implement communication strategies based on expert opinions that reduce the polarization highlighted by several studies on this topic (**Dunlap; McCright**, 2011; **Moernaut et al.**, 2022; **Pearce et al.**, 2019) and fight fake news. However, two main difficulties must be taken into account:

- that *Twitter* can promote polarization and misinformation (**Williams et al.**, 2015; **Anderson; Huntington**, 2017) as users tend to search for opinions similar to their own in order to reinforce them (**Grover et al.**, 2019); and
- that previous studies highlight that the degree of education and scientific knowledge is not a decisive factor in explaining whether people take a position for or against.

There will be correlations between the UGC topics that determine the social skepticism around climate change through the analysis of users' social activism and behavioral patterns

Moreover, within the scientific community there are opinions that deny or question the importance of climate change (**Lahsen**, 2013) and that, therefore, may encourage those opinions to have greater credibility.

Among the theoretical implications would be the development of new research based on the results obtained, (e.g., the relationship between political events and the polarization of opinion on climate change) and the fact that, by using data from *Twitter*, it is possible to analyze the discourse of climate change skeptics.

Among the limitations of this work are the number of data extracted, the hashtags used, the language of the extraction, the date selected, and the analysis carried out, which does not identify whether the comments are positive or negative, thus we cannot know whether they are for or against the arguments presented in the topics.

Future lines of research could include the modeling of the different topics identified and a model that integrates opinions based on political preferences, as well as longitudinal analysis using data extracted from the different editions of WED on *Twitter* to determine how the conversation and behavioral patterns evolved.

6. Conclusions

In this study, we used machine learning and artificial intelligence techniques to review 78,168 tweets to identify the keys of social skepticism around climate change indicators through an analysis of users' social activism and behavioral patterns on *Twitter*. These results were analyzed in depth to address the aim of this research.

Based on our results, we were able to answer hypothesis H1, that there will be correlations between the UGC topics that identify the social skepticism around climate change through the analysis of users' social activism and behavioral patterns. Specifically, we identified 28 topics that, in turn, could be grouped into 3 categories that identify the social skepticism around climate change through the analysis of users' social activism and behavioral patterns (Table 7).

The investigation has produced a series of results that confirm the proposed hypothesis.

In addition, some relevant conclusions have been obtained. The first is that 24 of the 28 topics are overlapping on the intertopic distance map. The second is that the size of the topics is relatively small and linked to specific events. The third is that there is a significant political presence, especially from the United States.

There is a group of topics, 24 of the 28, that appear superimposed such that, although they use different words and therefore form different clusters, they have a close relationship and, therefore, appear not only close but also superimposed. This opens up the possibility for new research focused only on these topics to better understand the reason for this overlap, although it does not permit their combination into a single larger topic.

The size of the topics is relatively small. There is no big theme, which means that attention is divided among the 28 themes developed by the skeptics opposed to climate change. This may be due to the fact that each of the user groups defends a viewpoint on a specific climate change topic, without any of them having gone viral, or because of their temporary nature. Themes arise, but none stay around for a long time. For example, in the political arena, one can mention the different political leaders in their corresponding topics ("trump", "biden", and "obama").

In relation to this point, note that the small size of the topics may be related to the relationship between the communication actions against climate change and the specific facts related to it, for example, Biden's proposal to support farmers in the topic "biden farmers", the "covid" health crisis, or the "wef" meeting. Regarding this point, the important presence of the United States, with its different presidents, also stands out, while there is no mention of other leaders of other countries.

We identified 28 topics that, in turn, could be grouped into 3 categories that identify the social skepticism around climate change through the analysis of users' social activism and behavioral patterns

7. References

Al-Nakeeb, Ohood A. M. S.; Mufleh, Basher A. H. (2018). "Collocations and collocational networks of characters: A corpus-based feminist stylistic analysis". *Language in India*, v. 18, n. 9. <https://onx.la/05e12>

- Alalwan, Ali A.; Rana, Nripendra P.; Dwivedi, Yogesh K.; Algharabat, Raed** (2017). "Social media in marketing: A review and analysis of the existing literature". *Telematics and informatics*, v. 34, n. 7, pp. 1177-1190.
<https://doi.org/10.1016/j.tele.2017.05.008>
- Anderson, Ashley A.; Huntington, Heidi E.** (2017). "Social media, science, and attack discourse: How *Twitter* discussions of climate change use sarcasm and incivility". *Science communication*, v. 39, n. 5, pp. 598-620.
<https://doi.org/10.1177/1075547017735113>
- Arora, Anuja; Bansal, Shivam; Kandpal, Chandrashekhar; Aswani, Reema; Dwivedi, Yogesh K.** (2019). "Measuring social media influencer index-insights from *Facebook*, *Twitter* and *Instagram*". *Journal of retailing and consumer services*, v. 49, pp. 86-101.
<https://doi.org/10.1016/j.jretconser.2019.03.012>
- Aswani, Reema; Kar, Arpan K.; Ilavarasan, P. Vigneswara; Dwivedi, Yogesh** (2018). "Search engine marketing is not all gold: Insights from *Twitter* and *SEOClerks*". *International journal of information management*, v. 38, n. 1, pp. 107-116.
<https://doi.org/10.1016/j.ijinfomgt.2017.07.005>
- Baker, Paul; Gabrielatos, Costas; Khosravini, Majid; Krzyżanowski, Michał; McEnery, Tony; Wodak, Ruth** (2008). "A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press". *Discourse and society*, v. 19, n. 3, pp. 273-306.
<https://doi.org/10.1177/0957926508088962>
- Blasco-Arcas, Lorena; Lee, Hsin-Hsuan M.; Kastanakis, Minas N.; Alcañiz, Mariano; Reyes-Menéndez, Ana** (2022). "The role of consumer data in marketing: A research agenda". *Journal of business research*, v. 146, pp. 436-452.
<https://doi.org/10.1016/j.jbusres.2022.03.054>
- Blei, David M.; Ng, Andrew Y.; Jordan, Michael I.** (2003). "Latent Dirichlet allocation". *Journal of machine learning research*, n. 3, pp. 993-1022.
<https://dl.acm.org/doi/10.5555/944919.944937>
- Bolsen, Toby; Shapiro, Mathew A.** (2018). "The US news media, polarization on climate change, and pathways to effective communication". *Environmental communication: A journal of nature and culture*, v. 12, n. 2, pp. 149-163.
<https://doi.org/10.1080/17524032.2017.1397039>
- Bouma, Gerlof** (2009). "Normalized (pointwise) mutual information in collocation extraction". *Proceedings of GSCL*, pp. 31-40.
- Capstick, Stuart B.; Pidgeon, Nicholas F.** (2014). "What is climate change scepticism? Examination of the concept using a mixed methods study of the UK public". *Global environmental change*, v. 24, pp. 389-401.
<https://doi.org/10.1016/j.gloenvcha.2013.08.012>
- Capstick, Stuart B.; Whitmarsh, Lorraine; Poortinga, Wouter; Pidgeon, Nicholas F.; Upham, Paul** (2015). "International trends in public perceptions of climate change over the past quarter century". *Wiley interdisciplinary reviews: Climate change*, v. 6, n. 1, pp. 35-61.
<https://doi.org/10.1002/wcc.321>
- Carew, Joanne** (2014). "Online environmental activism in South Africa: A case study of the #iam4rhinos *Twitter* campaign". *Global media journal: African edition*, v. 8, n. 2, pp. 207-230.
<https://doi.org/10.5789/8-2-169>
- Corner, Adam; Markowitz, Ezra; Pidgeon, Nicholas F.** (2014). "Public engagement with climate change: the role of human values". *Wiley interdisciplinary reviews: climate change*, v. 5, n. 3, pp. 411-422.
<https://doi.org/10.1002/wcc.269>
- Dahlberg, Lincoln** (2001). "The internet and democratic discourse: exploring the prospects of online deliberative forums extending the public sphere". *Information, communication & society*, v. 4, n. 4, pp. 615-633.
<https://doi.org/10.1080/13691180110097030>
- Drakett, Jessica; Rickett, Bridgette; Day, Katy; Milnes, Kate** (2018). "Old jokes, new media - Online sexism and constructions of gender in Internet memes". *Feminism and psychology*, v. 28, n. 1, pp. 109-127.
<https://doi.org/10.1177/0959353517727560>
- Dubrofsky, Rachel E.; Wood, Megan M.** (2014). "Posting racism and sexism: authenticity, agency and self-reflexivity in social media". *Communication and critical/cultural studies*, v. 11, n. 3, pp. 282-287.
<https://doi.org/10.1080/14791420.2014.926247>
- Dunlap, Riley E.; McCright, Aaron M.** (2011). "Climate change denial". In: Lever-Tracy, Constance (ed.). *Routledge handbook of climate change and society*. Routledge, pp. 144-160. ISBN: 978 0 203876213
- Dwivedi, Yogesh K.; Kapoor, Kawaljeet K.; Chen, Hsin** (2015). "Social media marketing and advertising". *The marketing review*, v. 15, n. 3, pp. 289-309.
<https://doi.org/10.1362/146934715X14441363377999>

- Eide, Elisabeth; Kunelius, Risto** (2021). "Voices of a generation the communicative power of youth activism". *Climatic change*, v. 169, n. 6.
<https://doi.org/10.1007/s10584-021-03211-z>
- Elgesem, Dag; Steskal, Lubos; Diakopoulos, Nicholas** (2015). "Structure and content of the discourse on climate change in the blogosphere: the big picture". *Environmental communication*, v. 9, n. 2 pp. 169-188.
<https://doi.org/10.1080/17524032.2014.983536>
- Engels, Anita; Hüther, Otto; Schäfer, Mike; Held, Hermann** (2013). "Public climate-change skepticism, energy preferences and political participation". *Global environmental change*, v. 23, n. 5, pp. 1018-1027.
<https://doi.org/10.1016/j.gloenvcha.2013.05.008>
- Fujita, Momoko; Harrigan, Paul; Soutar, Geoffrey N.** (2018). "Capturing and co-creating student experiences in social media: a social identity theory perspective". *Journal of marketing theory and practice*, v. 26, n. 1-2, pp. 55-71.
<https://doi.org/10.1080/10696679.2017.1389245>
- Grover, Purva; Kar, Arpan K.; Dwivedi, Yogesh K.; Janssen, Marijn** (2019). "Polarization and acculturation in US election 2016 outcomes. Can Twitter analytics predict changes in voting preferences?". *Technological forecasting and social change*, v. 145, pp. 438-460.
<https://doi.org/10.1016/j.techfore.2018.09.009>
- Grundmann, Reiner** (2013). "'Climategate' and the scientific ethos". *Science, technology & human values*, v. 38, n. 1, pp. 67-93.
<https://doi.org/10.1177/0162243911432318>
- Hagen, Bjoern; Middel, Ariane; Pijawka, David** (2016). "European climate change perceptions: public support for mitigation and adaptation policies". *Environmental policy and governance*, v. 26, n. 3, pp. 170-183.
<https://doi.org/10.1002/eet.1701>
- Hardaker, Claire; McGlashan, Mark** (2016). "Real men don't hate women: Twitter rape threats and group identity". *Journal of pragmatics*, n. 91, pp. 80-93.
<https://doi.org/10.1016/j.pragma.2015.11.005>
- Herdağdelen, Amaç; Zuo, Wenyun; Gard-Murray, Alexander; Bar-Yam, Yaneer** (2013). "An exploration of social identity: the geography and politics of news-sharing communities in Twitter". *Complexity*, v. 19, n. 2, pp. 10-20.
<https://doi.org/10.1002/cplx.21457>
- Hoggan, James; Littlemore, Richard** (2009). *Climate cover-up: the crusade to deny global warming*. Greystone Books Ltd. ISBN: 1553654854
- Iyengar, Shanto; Sood, Gaurav; Lelkes, Yphtach** (2012). "Affect, not ideology: a social identity perspective on polarization". *Public opinion quarterly*, v. 76, n. 3, pp. 405-431.
<https://doi.org/10.1093/poq/nfs038>
- Jenkins, Richard** (2014). *Social identity*. New York: Routledge. ISBN: 978 0 415448499
- Jia, Susan S.** (2018). "Leisure motivation and satisfaction: a text mining of yoga centres, yoga consumers, and their interactions". *Sustainability*, v. 10, n. 12, 4458.
<https://doi.org/10.3390/su10124458>
- Kahan, Dan M.; Peters, Ellen; Wittlin, Maggie; Slovic, Paul; Ouellette, Lisa-Larrimore; Braman, Donald; Mandel, Gregory** (2012). "The polarizing impact of science literacy and numeracy on perceived climate change risks". *Nature climate change*, v. 2, n. 10, pp. 732-735.
<https://doi.org/10.1038/nclimate1547>
- Kaiser, Jonas; Rhomberg, Markus** (2016). "Questioning the doubt: Climate skepticism in German newspaper reporting on COP17". *Environmental communication*, v. 10, n. 5, pp. 556-574.
<https://doi.org/10.1080/17524032.2015.1050435>
- Kapoor, Kawaljeet-Kaur; Tamilmani, Kuttimani; Rana, Nripendra P.; Patil, Pushp; Dwivedi, Yogesh; Nerur, Sridahr** (2018). "Advances in social media research: past, present and future". *Information systems frontiers*, v. 20, n. 3, pp. 531-558.
<https://doi.org/10.1007/s10796-017-9810-y>
- Karami, Amir; Shah, Vishal; Vaezi, Reza; Bansal, Amit** (2020). "Twitter speaks: a case of national disaster situational awareness". *Journal of information science*, v. 46, n. 3, pp. 313-324.
<https://doi.org/10.1177/0165551519828620>
- Kastanakis, Minas N.; Balabanis, George** (2012). "Between the mass and the class: Antecedents of the 'bandwagon' luxury consumption behavior". *Journal of business research*, v. 65, n. 10, pp. 1399-1407.
<https://doi.org/10.1016/j.jbusres.2011.10.005>

- Kirilenko, Andrei P.; Stepchenkova, Svetlana O.** (2014). "Public microblogging on climate change: One year of *Twitter* worldwide". *Global environmental change*, v. 26, pp. 171-182.
<https://doi.org/10.1016/j.gloenvcha.2014.02.008>
- Lahsen, Myanna** (2013). "Anatomy of dissent: a cultural analysis of climate skepticism". *American behavioral scientist*, v. 57, n. 6, pp. 732-753.
<https://doi.org/10.1177/0002764212469799>
- Lakoff, Robin T.** (2004). *Language and woman's place: text and commentaries*. Oxford University Press, USA. ISBN: 0195167570
- Li, Manyu; Turki, Nadia; Izaguirre, Casandra R.; DeMahy, Chloe; Thibodeaux, Brooklyn L.; Gage, Taylor** (2021). "Twitter as a tool for social movement: an analysis of feminist activism on social media communities". *Journal of community psychology*, v. 49, n. 3, pp. 854-868.
<https://doi.org/10.1002/jcop.22324>
- Matthews, Paul** (2015). "Why are people skeptical about climate change? Some insights from blog comments". *Environmental communication*, v. 9, n. 2, pp. 153-168.
<https://doi.org/10.1080/17524032.2014.999694>
- McCright, Aaron M.; Dunlap, Riley E.** (2003). "Defeating Kyoto: the conservative movement's impact on U.S. climate change policy". *Social problems*, v. 50, n. 3, pp. 348-373.
<https://doi.org/10.1525/sp.2003.50.3.348>
- McEnery, Tony; Hardie, Andrew** (2013). "The history of corpus linguistics". In: Allan, Keith (ed.) *The Oxford handbook of the history of linguistics*. Oxford: Oxford University Press, pp. 727-745. ISBN: 978 0 191750434
<https://doi.org/10.1093/oxfordhb/9780199585847.013.0034>
- Moernaut, Renée; Mast, Jelle; Temmerman, Martina; Broersma, Marcel** (2022). "Hot weather, hot topic. Polarization and skeptical framing in the climate debate on *Twitter*". *Information, communication & society*, v. 25, n. 8, pp. 1047-1066.
<https://doi.org/10.1080/1369118X.2020.1834600>
- Monfort, Abel; Villagra, Nuria; López-Vázquez, Belén** (2019). "Exploring stakeholders' dialogue and corporate social responsibility (CSR) on *Twitter*". *Profesional de la información*, v. 28, n. 5.
<https://doi.org/10.3145/epi.2019.sep.13>
- Muralidharan, Sidharth; Rasmussen, Leslie; Patterson, Daniel; Shin, Jae-Hwa** (2011). "Hope for Haiti: an analysis of *Facebook* and *Twitter* usage during the earthquake relief efforts". *Public relations review*, v. 37, n. 2, pp. 175-177.
<https://doi.org/10.1016/j.pubrev.2011.01.010>
- Newman, Tood P.** (2017). "Tracking the release of IPCC AR5 on *Twitter*: users, comments, and sources following the release of the working group. Summary for policymakers". *Public understanding of science*, v. 26, n. 7, pp. 815-825.
<https://doi.org/10.1177/0963662516628477>
- Ntontis, Evangelos; Drury, John; Amlôt, Richard; Rubin, G. James; Williams, Richard** (2018). "Emergent social identities in a flood: implications for community psychosocial resilience". *Journal of community and applied social psychology*, v. 28, n. 1, pp. 3-14.
<https://doi.org/10.1002/casp.2329>
- Olesen, Thomas** (2022). "Greta Thunberg's iconicity: performance and co-performance in the social media ecology". *New media & society*, v. 24, n. 6, pp. 1325-1342.
<https://doi.org/10.1177/1461444820975416>
- O'Neill, Safron; Williams, Hywel T. P.; Kurz, Tim; Wiersma, Bouke; Boykoff, Maxwell** (2015). "Dominant frames in legacy and social media coverage of the IPCC fifth assessment report". *Nature climate change*, v. 5, n. 4, pp. 380-385.
<https://doi.org/10.1038/nclimate2535>
- Pace, Stefano; Buzzanca, Stefano; Fratocchi, Luciano** (2016). "The structure of conversations on social networks: between dialogic and dialectic threads". *International journal of information management*, v. 36, n. 6, pp. 1144-1151.
<https://doi.org/10.1016/j.ijinfomgt.2016.04.001>
- Palos-Sánchez, Pedro; Saura, José-Ramón; Reyes-Menéndez, Ana; Vázquez-Esquivel, Ivonne V.** (2018). "Users acceptance of location-based marketing apps in tourism sector: an exploratory analysis". *Journal of spatial and organizational dynamics*, v. 6, n. 3, pp. 258-270.
<https://www.jsod-cieo.net/journal/index.php/jsod/article/view/144>
- Pearce, Warren; Holmberg, Kim; Hellsten, Iina; Nerlich, Brigitte** (2014). "Climate change on *Twitter*: topics, communities and conversations about the 2013 IPCC working group 1 report". *PLoS one*, v. 9, n. 4, e94785.
<https://doi.org/10.1371/journal.pone.0094785>

- Pearce, Warren; Niederer, Sabine; Özkula, Suay M.; Sánchez-Querubín, Natalia** (2019). "The social media life of climate change: platforms, publics, and future imaginaries". *Wiley interdisciplinary reviews: climate change*, v. 10, n. 2, e569. <https://doi.org/10.1002/wcc.569>
- Porter, Amanda J.; Hellsten, Iina** (2014). "Investigating participatory dynamics through social media using a multideterminant 'frame' approach: the case of Climategate on YouTube". *Journal of computer-mediated communication*, v. 19, n. 4, pp. 1024-1041. <https://doi.org/10.1111/jcc4.12065>
- Pritchard, Jonathan K.; Stephens, Matthew; Donnelly, Peter** (2000). "Inference of population structure using multilocus genotype data". *Genetics*, v. 155, n. 2, pp. 945-959. <https://doi.org/10.1093/genetics/155.2.945>
- Rama-Maneiro, Efrén; Vidal, Juan C.; Lama, Manuel** (2020). "Collective disambiguation in entity linking based on topic coherence in semantic graphs". *Knowledge-based systems*, v. 199, 105967. <https://doi.org/10.1016/j.knsys.2020.105967>
- Reyes-Menéndez, Ana; Correia, Marisol B.; Matos, Nelson; Adap, Charlene** (2020). "Understanding online consumer behavior and eWOM strategies for sustainable business management in the tourism industry". *Sustainability*, v. 12, n. 21, 8972. <https://doi.org/10.3390/su12218972>
- Reyes-Menéndez, Ana; Saura, José-Ramón; Álvarez-Alonso, César** (2018). "Understanding #WorldEnvironmentDay user opinions in Twitter: a topic-based sentiment analysis approach". *International journal of environmental research and public health*, v. 15, n. 11. <https://doi.org/10.3390/ijerph15112537>
- Reyes-Menéndez, Ana; Saura, José-Ramón; Filipe, Ferrao** (2019). "The importance of behavioral data to identify online fake reviews for tourism businesses: a systematic review". *PeerJ computer science*, v. 5, e219. <https://doi.org/10.7717/peerj-cs.219>
- Saura, José-Ramón; Reyes-Menéndez, Ana; Álvarez-Alonso, César** (2018). "Do online comments affect environmental management? Identifying factors related to environmental management and sustainability of hotels". *Sustainability*, v. 10, n. 9, 3016. <https://doi.org/10.3390/su10093016>
- Saura, José-Ramón; Rodríguez-Herráez, Beatriz; Reyes-Menéndez, Ana** (2019). "Comparing a traditional approach for financial brand communication analysis with a big data analytics technique". *IEEE access*, v. 7, n. 1. <https://doi.org/10.1109/ACCESS.2019.2905301>
- Schmid, Hans-Jörg** (2016). "Why cognitive linguistics must embrace the social and pragmatic dimensions of language and how it could do so more seriously". *Cognitive linguistics*, v. 27, n. 4, pp. 543-557. <https://doi.org/10.1515/cog-2016-0048>
- Singh, Akshit; Shukla, Nagesh; Mishra, Nishikant** (2018). "Social media data analytics to improve supply chain management in food industries". *Transportation research part E: logistics and transportation review*, v. 114, pp. 398-415. <https://doi.org/10.1016/j.tre.2017.05.008>
- Singh, Devinder-Pal; Kastanakis, Minas N.; Paul, Justin; Felix, Reto** (2021). "Non-deceptive counterfeit purchase behavior of luxury fashion products". *Journal of consumer behaviour*, v. 20, n. 5, pp. 1078-1091. <https://doi.org/10.1002/cb.1917>
- Sinpeng, Aim** (2021). "Hashtag activism: social media and the #FreeYouth protests in Thailand". *Critical Asian studies*, v. 53, n. 2, pp. 192-205. <https://doi.org/10.1080/14672715.2021.1882866>
- Skill, Karin; Passero, Sergio; Francisco, Marie** (2021). "Assembling Amazon fires through English hashtags. Materializing environmental activism within Twitter networks". *Computer supported cooperative work*, v. 30, n. 5-6, pp. 715-732. <https://doi.org/10.1007/s10606-021-09403-6>
- Smith, Nicholas; Leiserowitz, Anthony** (2012). "The rise of global warming skepticism: exploring affective image associations in the United States over time". *Risk analysis*, v. 32, n. 6, pp. 1021-1032. <https://doi.org/10.1111/j.1539-6924.2012.01801.x>
- Stieglitz, Stefan; Mirbabaie, Milad; Ross, Björn; Neuberger, Christoph** (2018). "Social media analytics-challenges in topic discovery, data collection, and data preparation". *International journal of information management*, v. 39, pp. 156-168. <https://doi.org/10.1016/j.ijinfomgt.2017.12.002>
- Su, Leona-Yi-Fan; Akin, Heather; Borssard, Dominique; Scheufele, Dietram A.; Xenos, Michael A.** (2015). "Science news consumption patterns and their implications for public understanding of science". *Journalism & mass communication quarterly*, v. 92, n. 3, pp. 597-616. <https://doi.org/10.1177/1077699015586415>

- Syed, Shaheen; Spruit, Marco** (2017). "Full-text or abstract? Examining topic coherence scores using latent Dirichlet allocation". In: *IEEE International conference on data and advanced analytics*.
<https://doi.org/10.1109/DSAA.2017.61>
- Tajfel, Henry** (1974). "Social identity and intergroup behaviour". *Social science information*, v. 13, n. 2, pp. 65-93.
<https://doi.org/10.1177/053901847401300204>
- Tajfel, Henry** (1981). *Human groups and social categories. Studies in social psychology*. Cambridge: Cambridge University Press. ISBN: 0521280737
- Terkourafi, Marina; Haugh, Michael** (2019). "Quo vadis, pragmatics?". *Journal of pragmatics*, v. 145.
<https://doi.org/10.1016/j.pragma.2019.04.005>
- Van-de-Donk, Wim; Loader, Brian D.; Nixon, Paul G.; Rucht, Dieter** (2004). *Cyberprotest: new media, citizens and social movements*. Routledge. ISBN: 978 0 415297851
- Van-Eck, Christel W.; Feindt, Peter H.** (2022). "Parallel routes from Copenhagen to Paris: climate discourse in climate sceptic and climate activist blogs". *Journal of environmental policy & planning*, v. 24, n. 2, pp. 194-209.
<https://doi.org/10.1080/1523908X.2021.2000376>
- Vanhala, Mika; Lu, Chien; Peltonen, Jaakko; Sundqvist, Sanna; Nummenmaa, Jyrki; Järvelin, Kalervo** (2020). "The usage of large data sets in online consumer behaviour: a bibliometric and computational text-mining-driven analysis of previous research". *Journal of business research*, v. 106, pp. 46-59.
<https://doi.org/10.1016/j.jbusres.2019.09.009>
- Veltri, Giuseppe A.; Atanasova, Dimitrinka** (2017). "Climate change on *Twitter*: content, media ecology and information sharing behaviour". *Public understanding of science*, v. 26, n. 6, pp. 721-737.
<https://doi.org/10.1177/0963662515613702>
- Washington, Haydn; Cook, John** (2011). *Climate change denial*. Earthscan. ISBN: 978 1 849713368
- Whitmarsh, Lorraine** (2011). "Scepticism and uncertainty about climate change: dimensions, determinants and change over time". *Global environmental change*, v. 21, n. 2, pp. 690-700.
<https://doi.org/10.1016/j.gloenvcha.2011.01.016>
- Whitmarsh, Lorraine; Capstick, Stuart** (2018). "Perceptions of climate change". In: Clayton, Susan; Manning, Christie (eds.). *Psychology and climate change. Human perceptions, impacts and responses*. Amsterdam: Elsevier, pp. 13-33. ISBN: 978 0 128131305
<https://doi.org/10.1016/B978-0-12-813130-5.00002-3>
- Whitmarsh, Lorraine; Corner, Adam** (2017). "Tools for a new climate conversation: a mixed-methods study of language for public engagement across the political spectrum". *Global environmental change*, v. 42, pp. 122-135.
<https://doi.org/10.1016/j.gloenvcha.2016.12.008>
- Williams, Hywel T. P.; McMurray, James R.; Kurz, Tim; Lambert, F. Hugo** (2015). "Network analysis reveals open forums and echo chambers in social media discussions of climate change". *Global environmental change*, v. 32, pp. 126-138.
<https://doi.org/10.1016/j.gloenvcha.2015.03.006>
- Wozniak, Antal; Wessler, Harmut; Lück, Julia** (2017). "Who prevails in the visual framing contest about the United Nations climate change conferences?". *Journalism studies*, v. 18, n. 11, pp. 1433-1452.
<https://doi.org/10.1080/1461670X.2015.1131129>
- Wu, Ming-Wen; Su, Keh-Yih** (1993). "Corpus-based automatic compound extraction with mutual information and relative frequency count". In: *Proceedings of rocling VI Computational linguistics conference*, pp. 207-216.
<https://aclanthology.org/O93-1009>
- Wu, Zezhou; Zhang, Yan; Chen, Qiaohui; Wang, Hao** (2021). "Attitude of Chinese public towards municipal solid waste sorting policy: a text mining study". *Science of the total environment*, v. 756, 142674.
<https://doi.org/10.1016/j.scitotenv.2020.142674>
- Xiong, Ying; Cho, Moonhee; Boatwright, Brandon** (2019). "Hashtag activism and message frames among social movement organizations: semantic network analysis and thematic analysis of *Twitter* during the #MeToo movement". *Public relations review*, v. 45, n. 1, pp. 10-23.
<https://doi.org/10.1016/j.pubrev.2018.10.014>
- Zappavigna, Michele** (2015). "Searchable talk: the linguistic functions of hashtags". *Social semiotics*, v. 25, n. 3, pp. 274-291.
<https://doi.org/10.1080/10350330.2014.996948>
- Zoller, Heather M.; Casteel, Diana** (2021). "#March for our lives: health activism, diagnostic framing, gun control, and the gun industry". *Health communication*, v. 37, n. 4.
<https://doi.org/10.1080/10410236.2020.1871167>

Femitags in the networks and in the streets: 50 hashtags for feminist activism in Latin America

Guiomar Rovira-Sancho; Jordi Morales-i-Gras

Note: Este artículo se puede leer en español en:
<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87179>

Recommended citation:

Rovira-Sancho, Guiomar; Morales-i-Gras, Jordi (2023). "Femitags in the networks and in the streets: 50 hashtags for feminist activism in Latin America". *Profesional de la información*, v. 32, n. 3, e320319.

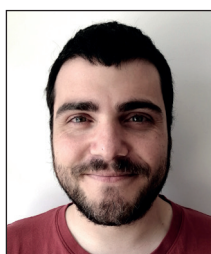
<https://doi.org/10.3145/epi.2023.may.19>

Article received on November 10th 2022
Approved on May 16th 2023



Guiomar Rovira-Sancho ✉
<https://orcid.org/0000-0002-4046-3531>

Universitat de Girona
Facultat de Dret
Campus Montilivi
17003 Girona, Spain
guiomar.rovira@udg.edu



Jordi Morales-i-Gras
<https://orcid.org/0000-0003-4173-3609>

Cámara Bilbao University Business School
Licenciado Poza, 17, pl. 1
48011 Bilbao, Spain
morales.jordi@gmail.com

Abstract

The aim of this article is to characterize the main hashtags of Mexican feminist activism as part of the Latin American connected crowds, on the basis of a review of the most popular trends between 2016 and 2021 on *Twitter*. A total of 50 hashtags have been selected that, due to their repeated use, are characterized as *femitags*, that is, performative meta-discursive identifiers that mainly do three things: disseminate and tune in to frames of protest; extend women's voices and slogans; and mobilize synchronic, diachronic, and transnational repertoires of collective action. These *femitags* appear in situated contexts, usually around events such as new cases of violence against women and calls to take to the streets, but with their extension, they become meta-communicative labels that articulate different activist practices. This study reveals not only their functions, but also their role as articulators of networked mobilizations, demonstrating the profound online and real-life intertwining of the fourth wave of Mexican, Latin American, and Spanish-speaking feminism.

Keywords

Hashtags; Feminism; Online activism; Network activism; Cyberactivism; Social media; Women rights; Female demands; Social networks; *Femitags*; *Twitter*; Connected crowds; Femicides; #MeToo; #NiUnaMenos; #SiMeMatan; Mexico; Latin America.

Funding

This project has received funding from the European Union's framework programme for research and innovation *Horizon 2020* (2014-2020) under the *Marie Skłodowska-Curie Grant Agreement* n. 893348, "Online feminist global actions. Potentials and pitfalls of the #MeToo campaign", *Femonmetoo*.

1. Introduction

One of the most relevant transnational social movements in Latin America in recent years is the wave of feminist protests (García-González, 2021; Chávez-Rodríguez, 2017; Revilla-Blanco, 2019). This is a phenomenon of contentious collective action led by a very broad spectrum of women against patriarchal violences that is characterised by an overlapping and continuous occupation of online and in-person spaces. These "feminist connected crowds" (Rovira-Sancho, 2018) use hashtags not only to give meaning to their protests, but also as tools for collective action, which today is "connective action" (Bennett; Segerberg, 2012). In particular, the most used hashtags in Mexico, which in many cases correspond to the most used in Latin America, point to a common agenda that has at its core the fight against femicide with a capacity for action and communication without strong leadership, in networks free of scale. Despite being owned by profit-ori-



ented tech corporations, feminist crowds are appropriating the digital public sphere (Laudano, 2019; Mendes; Ringrose; Keller, 2019; Portillo *et al.*, 2022) and using hashtags to promote their demands and sustain their networked coordination. Understanding what these meta-discursive hashtags are and what functions they fulfil in the period of maximum effervescence of the fourth wave of feminism in Mexico is the aim of this study.

As has happened with other digital communication tools, hashtags may not last or may be replaced by other forms (Chen *et al.*, 2018), which is why it is essential for academics to document their appearance, implementation and extension in the peak years of feminist protests in Mexico, between April 2016, with the first march of the cycle under the hashtag #VivasNosQueremos (“We Want Ourselves Alive”), and the present in which this study was conducted in 2022.

2. Theoretical framework

2.1. Feminisms in networks

In the so-called fourth wave of feminism (Munro, 2013; Cochrane, 2014; Varela, 2019; Zimmerman, 2017) organisations no longer structure communication in the feminist movement; on the contrary, as Clark (2016) points out, communication itself, from blogs to mailing lists to hashtags, has become an organisational infrastructure: networks convene and are the convening. Baer (2016, p. 19) notes that women’s online activism re-establishes the basis for collective feminist politics.

Despite the gaps in access to technology, in digital networks women bypass the gatekeepers of speech: media, politicians, experts and any authorisation to speak. In the spaces of networked communication, voices and claims appear that raise

“the violent encounter of the equality of the logos” (Rancière, 1996, p. 54),

because what is said and denounced, for example, by victims of sexual violence, can no longer be interpreted as laments or babbling when a huge amount of testimonies reiterate and reveal a structural condition.

In no movement has the politicisation gained online been more important than in the wave of outraged women shaking the world in the second decade of the 21st century (Kim, 2017; Thrift, 2014; Barker-Plummer; Barker-Plummer, 2015; Baer, 2016). The distribution of voices and peer-to-peer connection politicises women and makes them responsive to a broad, diverse, proliferating and translocal feminist agenda.

Digital activism is often analysed from a strictly network-centric perspective or as a minor resource of street action. This article assumes that activism today is simultaneously online and offline (Bonila; Rosa, 2015; Van-Laer; Van-Aelst, 2010). The web is not a “time out” for activism or “real action”, but the place where campaigns and mobilisations take root, give feedback and spread, as co-presence is no longer necessary to activate a protest (Earl; Kimport, 2011).

In the age of digital networks, there is a trend towards “the personalisation of politics” (Bennett, 2012), where organisations lose their power of representation and common people demand to speak in their own voice. People speak for themselves and “the personal becomes political”, to use the famous slogan of the second wave of feminism. Women communicate, link, tell and protest in the first person and from their embodied experience. The prefiguration of another possible world, where all voices count and there are no hard leaderships, is part of the new logic of networked action of the connected multitudes.

In digital spaces, women expose their grievances and come together to form “intimate publics” (Khoja-Moolji, 2015). Activism through hashtags generates affective communities, appealing to common emotions and shared meanings (Pacharissi, 2016). Through these networked practices, the “feminisation of collective action” (Rovira-Sancho, 2018) is fostered, i.e., a growing sensitivity against overlapping forms of power and violence.

Hashtag activism (Peroni; Rodak, 2020; De-Kosnik; Feldman, 2019) is often dismissed as “slacktivism” (Knibbs, 2013; Mulla, 2018), i.e., easy and inconsequential couch or click activism, such as giving a “like” or signing an online letter.

Against this view, Mendes, Ringrose, and Keller (2019, p. 74) show the “emotionally draining” labour and time effort of those who sustain feminist hashtag campaigns. Undoubtedly, the online dimension of social protest has expanded forms of participation and levels of involvement in political activism, some more intense and some less so. Activists combine in hashtags denunciations and testimonies, all kinds of images, from art to photographs or drawings, infographics, posters to call for participation, portraits of disappeared or murdered women to give them names and faces, sound and videos in songs, slogans, protests, in a multimodal repository of messages that constructs new “rules of sentiment” (Hochschild, 1979; Morales-Sánchez, 2022).

“The most used hashtags in Mexico, which in many cases correspond to the most used in Latin America, point to a common agenda that has at its core the fight against feminicides, with a capacity for action and communication without strong leadership, in networks free of scale.”

The aggregation of emancipatory connected crowds is often spontaneous and organic (as opposed to that promoted by the social influence industry and marketing), characterised by the articulation of weak ties (**Granovetter**, 1984). This implies difficulties in achieving continuity or generating stable organisation. Hashtags are precisely the discursive elements that for specific periods of time sustain this weak structuring that enables and makes protest possible.

It is women who put feminist hashtags into circulation and turn them into an enabler, a piece of collective articulation that functions as a hinge for the online/offline

2.2. Activism and hashtags

Digital networks offer communication and message exchange spaces between users, delimited by the protocols of their technical architecture and by algorithms, currently oriented towards the data economy.

However, communicative practices influence the shaping of technical devices, which are adapted to the demands of the market and its users. As **Lomborg** (2011) explains, *Twitter* started out as a space for expressing and reading opinions in “tweets”. It was the users who started to spread tweets of others, which led to the directionality function of the @ sign and the RT (retweet) function being incorporated into the design. The hashtag or # tag was implemented as a means to help categorise information, but with its widespread use it became a “user-created metadiscursive convention” (**Brock**, 2012, p. 534).

Using hashtags, anyone can access content that has not been produced by the people they follow on *Twitter* and aggregate messages for a wider community than just their followers. Searching for a hashtag provides access to tweets that have used it. And at the same time, you can add something to that intertextual string. The meaning of a hashtag is not fixed, it mutates and diversifies in its iteration, it can expand as an “empty signifier” (**Laclau**, 1996) or lose its initial function and disappear, or even turn against its creators. In this way, the growing visibility of feminism has also put into circulation an online misogyny that is quickly propagated and normalized (**Banet-Weiser**; **Miltner**, 2016; **Gill**, 2016). The backlash against feminism often appropriates hashtags to change their meaning (**Ganzer**, 2014). Phenomena such as cyber harassment, trolling, dissemination of data and images of women

“are some of the challenges that grow proportionally to the opportunities generated by technologies,” explain **Silva-Reis** and **Natansohn** (2019, p. 394).

Hashtags, as elements of “semantic condensation” (**Pfleguer**, 2021), compress their meaning into a few words and are quickly recognisable. Some hashtags create around them a performative constellation or a “virtual settlement” (**Jones**, 1997), a conversation anchored in this discursive element that proliferates and diversifies. When shared by a broad community, hashtags are not only search terms but also frames of meaning, as they facilitate an “archival and semiotic” process (**Bonila**; **Rosa**, 2015, p. 5).

The tactical use of hashtags by women’s struggles has allowed for framing, linking protests and extending agendas transnationally. Hashtags drive and shape connectivity (**Sundén**; **Paasonen**, 2019). As a strategy of connection, they enable “a second level of amplification” of feminist discourses to break into the public sphere (**Barker-Plumer**; **Barker-Plumer**, 2018).

Already since the cycle of protests that began in late 2010 with the Arab Spring (**Castells**, 2012),

“the hashtag has become a tool for those trying to promote social or political change” (**Chen et al.**, 2018, p. 199).

Papacharissi and **Oliveira** (2012) analyse the hashtag #Egypt that accompanied the revolution before and after the resignation of President Hosni Mubarak. During the protests in Turkey in defence of Gezi Park, the use of hashtags allowed the outrage to spread (**Oz**, 2016). **Candón-Mena** (2019) analyses the gestation of a global digital culture linked to the Internet, driven by creativity and technopolitics.

The strength of this cycle of feminist protests and their hashtags is rooted in the voice and agency of women, not in technological devices, which are precisely *hacked*¹ in their most predictable functions (consumer and entertainment oriented) to extend their technopolitical power. It is women who put feminist hashtags into circulation and turn them into an enabler, a piece of collective articulation that functions as a hinge for the online/offline, i.e. *on-life* (**Briones**, 2022) continuity of action and protest.

2.3. Femitags: the hashtags of feminism

From a pragmatic perspective, the most stable and repeated hashtags in the Mexican feminist community can be analysed in terms of their functions, as a toolbox for connective action.

As a versatile and multifunctional indicator, a hashtag can serve as a call for a march, documentation and reflection on what happened, a didactic instance for political concepts, a place to share experiences and show affection, evidence of violence and a case file, a complaint to the state and denunciation, a list of aggressors, motivation and enthusiasm to take to the streets.

The arsenal of hashtags most used by the feminist connected crowds (shared on feminist *Instagram*, feminist *Facebook*, as well as on blogs and other platforms) are what we call *femitags*. We do not use the word *femtags*, because they are not hashtags marked by the feminine condition of their enunciators, but *femitags*, because of their feminist character, as a political instrument available to anyone, making a parallel with the difference pointed out by **Lagarde** (2005) between femicide (the murder of a woman) and feminicide, the latter term being the legal marker of the ultimate violence against a woman for the fact of being perceived and treated as a “woman”.²

“The arsenal of hashtags most used by the feminist connected crowds (shared on feminist *Instagram*, feminist *Facebook*, as well as on blogs and other platforms) are what we call *femitags*”

It is because of their reiteration and extension that *femitags* function, no longer as simple indexical elements, but as “metacommunicative tags” (**Dear; Hoffman; Goodman**, 2014). By becoming stable, *femitags* can be analysed as “rhetorical genres”, discursive devices that help

“communicating parties to reach an intersubjective understanding of the situation”, according to **Lomborg** (2011).

A rhetorical genre is only realised to the extent that communicative practices, conventions, and expectations are socially shared and recognised among a group of users and are recurrently expressed. *Femitags* and their sub-genres are therefore a collaborative achievement, even if there are localised groups or individuals who initiated them. They are inappropriate and inappropriable, they do not respect author or authority, and their success or continuity depends on the crowd that sustains them. As a network act, they are a performance, not a programme: they occur and gain value the moment they are collectively activated.

Our analysis seeks to trace the small success story of some *femitags* as discursive markers of a protest campaign, beyond their appearance on digital social networks, as hinges or gears that sustain online/in-person action. The goal of this paper is to show the pragmatic versatility of the Mexican feminist crowds’ arsenal of hashtags over a period of time, as a meta-communicative toolbox. We are interested in highlighting how they emerge, negotiate their meanings and establish their connective action-oriented functions.

Although the focus of this research is on agency and pragmatic use, it is imperative to reiterate that these hashtags circulate within the framework of applications such as *Twitter*, *Facebook*, or *Instagram*, which are applications whose algorithms are designed to profit from data (**Zuboff**, 2015).

2.4. Some hashtags of global feminism

Hashtags for collective action enable “memetic disruption” (**Thrift**, 2014, p. 1091), breaking the silence and calling for the incorporation of new voices and protests. This is not to say that everyone has access to technology, nor that those who do have access to technology have successful conversations; marginalisation and silencing often stifle their power. As **Crenshaw** (1991, p. 1246) points out:

“Where systems of domination of race, gender and class converge, as they do in the experiences of abused women of colour, intervention strategies based solely on the experiences of women who do not share the same class or race will be of limited help to women who face different obstacles by race and class.”

One hashtag that points to the various overlapping forms of oppression is #WhyIStayed, where women tell why they endure domestic violence. It started in September 2014, with the video of an NFL player beating his girlfriend in a lift. At the time Beverly Gooden, an African American blogger, tweeted her own experience with her ex-husband and added #WhyIStayed. Testimonies began to be shared of how lack of financial means makes it impossible for many women to leave abusive partners:

“I had to plan my exit for months before I had a place to go and money for the bus to get there. #WhyIStayed” (in **Conley**, 2017, p. 31).

In this hashtag, **Clark** (2016) finds a form of first-person storytelling that allows for resonance and empathy. The same is analysed by **Keller, Mendes, and Ringrose** (2018) in the case of the hashtag #BeenRapedNeverReported, where participants felt empowered personally and collectively.

In the case of African American feminism, black *femitags* connect, multiply and deepen what it is to end racism as a woman and how the struggle against oppression is also a class struggle. **Conley** (2017) documents how digital feminist spaces have served to radicalise and mobilise with the hashtags #SolidarityIsForWhiteWomen, #BlackPowerIsForBlackMen, #WhyIStayed and #YouOkSis. As Conley explains, the hashtags of black feminism are:

“...are thresholds between the lived and the liveable dehumanisation; they are sites of struggle for the politics of representation. They function as a way of renewing histories and interventions across time and space. They express desires to break social norms of violence and marginalisation, and of belonging” (2017, p. 29).

Sonia **Reverter** and María **Medina-Vicent** (2020) have analysed 35 hashtags of transnational feminism, markers of these connected crowds that began to spread their campaigns in global digital networks, many of them in English. Among

the most relevant are #BringBackOurGirls, created in Nigeria after 276 girls were abducted by Boko Haram in April 2014. Others are #SlutWalk, #FreePussyRiot, #IAM-Malala, #YesAllWoman, #NiUnaMenos (“not a single one less”)... Each one responds to a singular moment and a viral campaign.

A lot of literature analyses hashtags in English. This is why we consider it essential to document the variety and functions of Mexican and Latin American feminist hashtags, in many cases shared with Spain, in the broad Spanish-speaking region. We start from Mexico, our place of ethnographic observation, to enter this tangled web of performative markers.

3. Methodology

For this research, we started with a set of 40 hashtags that we first identified using digital ethnography techniques and observant participation in feminist protests in Mexico City throughout the sample years, from April 2016, the date of the first major feminist self-convocation called on networks: the National Mobilisation against Machista Violence #VivasNos-Queremos (“We Want Ourselves Alive”), and the end of 2021, after the mobilisation for the *International day against violence against women*. We followed the actions live and in the digital sphere during all these years, and for this study we reviewed the feminist hashtags that have been trending. We also analysed programmatically which were the ones in concurrence that were most present and implemented a “snowballing” strategy.

Using *Twitter’s* academic API, we collected 2.5 million tweets issued between January 2016 and October 2021 in which at least one of the hashtags from our initial package was used. The tweets obtained contained no less than 275,271 hashtags used on 5,451,566 occasions in total. The most used hashtags over this 5-year period are included in Table 1.

To analyse all this, we built a data model and dashboard with *PowerBi* (<https://powerbi.microsoft.com>) to interactively explore the geographic reach of each hashtag, its volume and temporality, its contents and main authors and its linkage with other hashtags. A sample of the data obtained is shown in Figure 1.

As criteria for the final selection of hashtags, we have eliminated those that are single words (#Women, #Feminism, etc.) and those that make up backlash communities, such as #AsiNo and #EllasNoMeRepresentan (the latter two emerged against direct action protests in Mexico) or non-feminist ones such as #NiUnoMenos.

The top 5 users of this period correspond to 5 personal accounts, indicating that the leaderships in these connected crowds are contingent and distributed, a phenomenon we have analysed in a previous paper (Rovira-Sancho; Morales-i-Gras, 2023).

“The interest of this research is not to analyse the tweets that have accompanied each hashtag, but rather the *femitags* as performative tools that articulate different repertoires for action and communication, that “do things with words” (Austin, 1982), in the networks and beyond them”

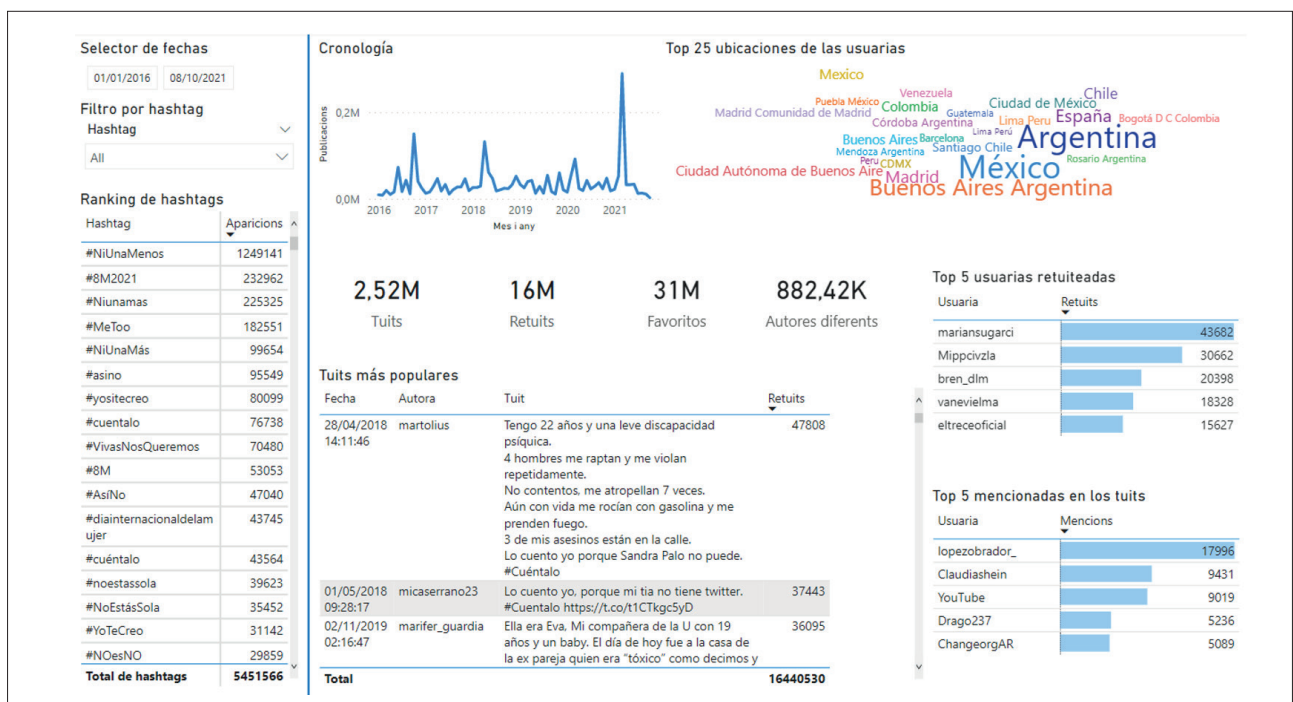


Figure 1. Screenshot of the *PowerBI* dashboard

It should be noted that these hashtags can have different syntax and variations, as well as accents and capital letters: #NosQueremosVivas or #VivasNosQueremos. #YoTeCreo (“I believe you”), #AmigaYoTeCreo (“my friend, I believe you”), #YoSiTeCreo (“I do believe you”), #HermanaYoTeCreo (“sister, I believe you”), or even marked by country: #MeTooMx or #YoTeCreoVenezuela (“I believe you Venezuela”). They can also add dates such as #8M2021, and most of the time they go together, in strings.

To contrast our results, we also reviewed the research of **Esquivel-Domínguez** (2019), who compiled the main feminist hashtags in Mexico up to 2019, based on the following premises: that they were related to protests against violence against women and girls, that they had been trending in Mexico and in some cases internationally, that they had been picked up by the digital media, that they had been used by the media, and that they had been used by the media in the country’s media.

In the end, we elaborated a battery of hashtags specific to the Mexican connected crowds that were revealed in many Latin American cases. These *femitags* are not only indexing elements but “metacommunicative tags”, since they are characterised by: having been used repeatedly in one or more of the three main feminist campaigns: against feminicide, for the legalisation of abortion, against sexual violence. They are at the same time transnational hashtags, with different origins, mostly in Spanish-speaking regions.

The interest of this research is not to analyse the tweets that have accompanied each hashtag, but rather the *femitags* as performative tools that articulate different repertoires for action and communication, that “do things with words” (**Austin**, 1982), in the networks and beyond them.

We will begin with a brief genealogical description of some particularly relevant *femitags*. According to **Foucault**:

“Genealogy is a particular investigation of those elements that we tend to feel are without history. It is a work of demythologisation that sets out to dislodge the hidden meanings and false semblances under which a historical reality is concealed” (1997, p. 15).

We will trace their appearance in specific contexts of activism, the way in which they are the fruit of intertexts with density in time and space, using secondary sources and data from our own ethnography.

After tracing a brief genealogy of some *femitags*, we will go on to elaborate a more general characterisation of the 50 in our list (see Table 2) by cross-referencing their main attributes.

4. Classification and genealogy of some *femitags*

In this section we trace the intertextual origin and transnational extension of the most emblematic *femitags*. Emphasising the online/in-person imbrication of feminist activism, we will start with a first classification:

1. *Femitags* to take the streets, i.e. that constitute calls for collective action beyond the networks, and
2. Narrative *femitags* “to tell and to count us” in the networks.

4.1. *Femitags* to take the streets

#VivasNosQueremos began as an activist campaign launched by the collective *Mujeres Grabando Resistencias* (“Women Engraving Resistances”) in Mexico City in July 2014. Its aim was, according to their words on their *Facebook* page,³ to print posters with the engraving technique in linoleum

“with clear and understandable messages against violence against women, feminicides and for our right to self-defence, in order to intervene in the streets of Mexico and other countries of Abya Yala”.

As **Rovetto** (2015, p. 17) explains,

“they received seventeen posters that they circulated on the social network FB and distributed through feminist collectives in Mexico, Europe and other Latin American countries to print and paste on the streets or make political interventions in public spaces”.

As a direct antecedent to this campaign, Rovetto points to the wallpapering of posters in Ciudad Juárez in April 2011, with the portraits of disappeared and murdered women by the civil association *Nuestras Hijas de Regreso a Casa* (“Our Daughters Back Home”) and the hashtag #HastaEncontrarles (“Until They are Found”). These images were circulated on social networks and printed on stencils, posters, and silkscreens. In 2014, the enforced disappearance of the 43 students from Ayotzinapa in Mexico circulated #VivosLosQueremos (“We Want Them Alive”, masculine). The feminine, first-person plural version of #VivasNosQueremos trended on *Twitter* to call for the first big Mexican Violet Spring march on 26 April 2016 (**García-González**, 2021; **Pfleger**, 2021).

“#NiUnaMenos opened the space for a broad agenda, not only against feminicides but also for the legalisation of abortion, against women’s precariousness and indebtedness, generating new campaigns such as #AbortoLegal (“legal abortion”), or the strike on 8 March”

Inspired by *Mujeres Grabando Resistencias* (“Women Engraving Resistances”), in Argentina they also held a feminist linoleum engraving workshop in 2015 with the same slogan. And after the success of the Mexican march in 2016, they incorporated #VivasNosQueremos to their call for #NiUnaMenos (“Not a Single One Less”) to take to the streets on 3 June of that year.

#NiUnaMenos clearly stands out as the most relevant *femitag* of the whole wave of Latin American feminist connected crowds. Its genealogy is also an example of the online/offline overlapping and transnational dimension of feminist protests, with historical memory. The phrase “Not a Single One More!” was written at the beginning of this century by the poet Susana Chávez, mother of a young victim of feminicide in Ciudad Juárez, Mexico. This mother was murdered in 2011 while demanding justice for her daughter. The *United Nations Economic Commission for Latin America and the Caribbean (Eclac)* chose the phrase from her poem to publish the 2007 report on violence against women, “Not a Single One More! The right to live a life free of violence in Latin America and the Caribbean.”

Years later, in 2015, the variant “Not a Single One Less” (which emphasises life: “Not a single one in life”), positioned itself in Argentina above Ni Una Más (which alludes to death: “Not one more dead”). Furthermore, #NiUnaMenos in Argentina bridged the struggle of the *Mothers of the Plaza de Mayo* and their demand for the “appearance alive” of those who disappeared during the military dictatorship of 1976-1983 (Fuentes, 2019).

Ni Una Menos was also the title of a reading marathon in Buenos Aires against the rising numbers of feminicides in April 2015. In May of that year, after the murder of 14-year-old Chiara Páez at the hands of her boyfriend, journalist Marcela Ojeda (@MarcelitaOjeda) posted on *Twitter*: “Actresses, politicians, artists, businesswomen, social leaders... women, all of them, bah... aren’t we going to raise our voices? THEY ARE KILLING US”.

A group of journalists with many followers on social networks decided to call for a mobilisation against feminicides on 3 June 2015. According to Natalucci and Rey (2018, p. 16),

“in order to unify the slogans, the organising committee centralised the different calls and published an updated list of meeting points on the official *Twitter* (@niunamenos) and *Facebook* (Ni Una Menos) accounts”.

680 politicians posted a tweet with a poster, 834 social organisations and 2,137 celebrities from Argentina and 280 from Latin America tweeted #NiUnaMenos and made it a global trend. Some 300,000 people marched across the country on 3 June 2015.

According to Laudano (2019), it was then that a feminist hashtag became a global trend for the first time:

“the cyber-action that became a successful trending topic with repercussions within national political spheres and international resonance will immediately become the initiating tweet to be emulated in different causes” (p. 360).

Thereafter, many women across the continent joined the cry of #NiUnaMenos and organised their own mobilisations in their countries. At the local level, as Garibotti and Hopp (2019, p. 186) show, #NiUnaMenos opened the space for a broad agenda, not only against feminicide but also for the legalisation of abortion, against precariousness and women’s indebtedness, generating new campaigns such as #AbortoLegal (“legal abortion”) or the strike on 8 March. What began as a call for a march spread through networks has become a broad social movement, with many local organisations and at the same time a global campaign. #NiUnaMenos has called for mobilisations in the streets of Bolivia, Brazil, Chile, El Salvador, Guatemala, Mexico, Paraguay, Peru, Puerto Rico, and Uruguay, among others (Rovira-Sancho; Morales-i-Gras, 2022). And it has jumped the ocean, reaching Spain, Turkey, Germany, and Italy. According to Marcela Fuentes (2019), #NiUnaMenos

“configures a multisited, multiplatform, (a)synchronic performance constellation that aims to hack patriarchy” (p. 176).

At the height of the #NiUnaMenos mobilisation in Argentina, a work stoppage was called on 19 October 2015, following the murder of 16-year-old Lucía Pérez, with the hashtag #NosotrasParamos (“We Stop”) and #MiercolesNegro (“Black Wednesday”). This replicated the #BlackMonday strike by women in Poland that same month against the criminalisation of abortion. After the success of the strike, both Polish and Argentinean women and women of other nationalities in a *Facebook* group decided to call for an International Women’s Strike for the following 8 March 2017. According to Annunziata (2020), the call for the strike in networks

“constituted a milestone that connected misogynist violence with the economic and social violence of capitalism, exploitation and labour precariousness” (2020, p. 160).

The strike on *Women’s Day* in 2017 and 2018 spread with great success in Argentina and Spain, but also in Mexico, Colombia, Venezuela, Ecuador, and the United States (Garrido-Ortolá, 2022). On 9 March 2020, in Mexico, the *Twitter* profile of a small women’s collective, Las Brujas del Mar (“The Witches of the Sea”), called for #UnDiaSinNosotras (“A Day Without Us”), with great success.

“The *femitags* we call “for stories and counts” acquire metonymic power: the part is the whole, one testimony grows and re-presents (presents again) all the others, that is to say, it condenses a discourse that generates a redundancy cascade with new voices”

#NiUnaMás, the original phrase from the poem by Susana Chávez, the struggling mother from Ciudad Juárez, has not been as widely used as Ni Una Menos in recent years at a global level, but it has had a constant presence in Mexico's social networks and acquired maximum power between 2019 and 2020, with a clear function of indexing new cases of femicide victims. As a confrontation with the state, #NiUnaMás has served as a marker of impunity. In turn, this *femitag* bridged two distinct communities: networked, urban and young feminism, and the struggle of the families of victims of femicide and disappearance, the latter being a more traditional, non-feminist, and territorially anchored movement (Rovira-Sancho; Morales-i-Gras, 2023).

“ The present tense of the *femitags* opens a *Jetztzeit*, in Walter Benjamin's. It shows the prefigurative will of connected action, here and now, opening up the moment and making the present the place of struggle and listening, without waiting any longer or projecting ”

There are a variety of other *femitags* calling for action in the streets. #NoMeCuidanMeViolan (“they don't take care of me, they rape me”) had a one-off, but high intensity impact in Mexico in August 2019, following three rape allegations against members of the police forces. This hashtag appeared at the same time as the self-convocation of the #Brillanteada (“glittered”) march, after an activist sprayed the chief of police with glitter. The mobilisation was also accompanied by #MeCuidanMisAmigasNoLaPolicia (“my friends look after me, not the police”). These hashtags circulated intensely and in strings, while in the streets of Mexico City groups of women threw incendiary cocktails at public buildings, urban transport and graffitied monuments.

Hashtags with key dates for taking to the streets are the most abundant in this wave of feminist crowds: for example, #8M (and its variants, which can include the year and even the place) for the annual call to march on *International women's day*; #25N as the *International day against violence against women*; #28S for the *Global day of action for the decriminalisation of abortion*. Sometimes the date is moved forward or postponed by a day, depending on the year and the call to march. Date hashtags allow the simultaneous and delocalised articulation of the mobilisations of a translocal and global feminist agenda with a common code.

Date hashtags also indicate specific local calls, such as #3J, corresponding to the huge demonstration on 3 June 2015 in Argentina, or #24A for 24 April 2016 in Mexico. The dates do not stand alone, they are usually accompanied by other *femitags* that drive them and that define the appropriate action and repertoire to be used, usually a march, but it can also be other repertoires, such as strike, as indicated by the successful call for #HuelgaFeminista8M (“Feminist Strike 8M”), for *International women's day*, #HuelgadeMujeres (“Women's Strike”), #ParoInternacionalDeMujeres (“International Women's Stoppage”), #NosotrasParamos (“We Stop”), #UnDiaSinNosotras (“A Day without Us”), extended between the years 2017 to 2020.

4.2. Narrative *femitags* for “stories and counts”⁴

There is another sub-genre of *femitags* that do not call for taking to the streets, but rather stand as digital *story-telling* campaigns (Polletta, 2006) or “ethical testimony” (Núñez-Puente; Fernández-Romero, 2017). It is a networked discursive action that goes viral and adopts the first person singular, mainly to break the silence about gender-based, sexual and feminicidal violence.

These *femitags* implement the tactic⁵ of telling stories and counting the number of stories, which unveils and builds a collaborative archive in the first person of something whose evidence was so scattered and silenced that it could not be assembled as such. Each account tells its own story and at the same time adds to the overall story. The *account is kept* in the additive sense, the testimonial account is *taken into account* and *told*, especially about sexual harassment and violence. In this sense, against any generalisation, the *femitag* campaigns are based on the singular case and the denunciation by iteration. The feminist politicisation of their activists occurs when the aggregative power of individual stories reveals a structural condition.

The *femitags* we call “stories and counts” acquire metonymic power: the part is the whole, one testimony grows and re-presents (presents again) all the others, that is to say, it condenses a discourse that generates a redundancy cascade with new voices. These *femitags* also trace frameworks of meaning that highlight stereotypes and the complicity of the state in gender violence and its impunity. They show the continuity between jealousy in dating, harassment and femicide, in what has come to be called the “violentometer”,⁶ with its varying degrees of the same problem.

In our archive we have found two sub-genres of *femitags* of stories and counts, those that serve to break the silence in the face of sexual violence and those that stage the seriousness of feminicidal violence by unveiling the cynicism of its authorised normalisation. We will call the latter *femitags* of paradoxical reflexivity.

4.3. *Femitags* of paradoxical reflexivity

Femitags that confront the ineffectiveness, omission and complicity of the state, the media, and institutions in the face of feminicidal violence, showing the vexatious conditions that stigmatise women, belong to this category. For example, #ViajoSola (“I Travel Alone”), #SiMeMatan (“If I am Killed”) or #MisSeñasParticulares (“My Personal Details”).

#ViajoSola emerged after the disappearance and murder of two young Argentinian women, Marina Menegazzo and María José Coni, in Ecuador on 22 February 2016 (Piñeiro-Otero; Martínez-Rolán, 2016). Outrage at the media's treatment of the event led a networked community to highlight the fact that two women do not "travel alone" together. The Mexican collective *Plumas Atómicas* (2016) ("Atomic Feathers") made a viral video that accompanied the Facebook hashtag: "#ViajoSola. Violence against women."

It was in this same context that, on 1 March 2016, Paraguayan student Guadalupe Acosta wrote a letter on Facebook entitled "Ayer me mataron" ("Yesterday I was killed"), a post that Facebook deleted a month later (Piñeiro-Otero; Martínez-Rolán, 2016, p. 23). The dissemination of this letter, both on digital networks and in various media, made the femitag #SiMeMatan a virtual settlement around which an outraged crowd of women gathered again, ready to show how many ways a misogynist crime is usually justified.

A year later, this femitag resurfaced and trended again in May 2017 following the femicide of National Autonomous University of Mexico student Lesvy Berlin Osorio, when authorities exposed elements of her private life that appeared to blame her for her death. #SiMeMatan accompanied the rabid student demonstrations that led to the boyfriend being tried and imprisoned for femicide.

In these chilling tweets, each woman gives the potential reasons she could be blamed if killed, such as being bisexual and indigenous, or being an investigative journalist and not choosing motherhood, or having been diagnosed with a mental disorder and being divorced. In researching the content of the tweets of this hashtag, Kadic (2019) found that

"one of the first things that stands out is the overwhelming use of verbs that we use on a daily basis, such as 'estar', 'viajar', 'andar', 'beber', 'tener', 'vivir', 'hablar' (to be, 'to travel', 'to walk', 'to drink', 'to have', 'to live', 'to speak', 'to have', 'to live', 'to speak')."

In September 2017, one of the young women who had participated in the #SiMeMatan tweet campaign, Mara Castilla, a student in Puebla, was raped and killed by a Cabify taxi driver. She had written on Twitter: "#SiMeMatan is because I liked to go out at night and drink a lot of beer...".

Another femitag of paradoxical reflexivity that anticipates the horror and denounce the hypothetical possibility of disappearance and femicide is #MisSeñasParticulares. It refers to a 2018 report where it is acknowledged that, out of more than 9 thousand registered cases of disappearance of women since 1968 in Mexico, more than 61% of the files do not specify any physical traits that could facilitate their search. Two years later, in February 2020, at the height of the protests against the femicide of young Ingrid Escamilla, this femitag was recovered and became a trend. Here is an example of one of the tweets:

"#MisSeñasParticulares ("My Personal Details") in case you have to look for me: 1.57 Brown eyes Short, straight, painted hair. Tattoo of 1 heart with 2 hands on my right thigh. A mole just above my belly button and a mole under my right eye. I wear brackets #11assassinadasaldía ("11 Murdered a Day") #Niunamás."

4.4. Femitags to break the silence on sexual violence

In October 2015, the femitag #PrimeiroAsseido appeared in Brazil, created by journalist Juliana de Faria, from the feminist collective *Think Olga*, to accompany and defend a 12-year-old girl who suffered a barrage of obscene insults on Twitter after her participation in the *MasterChef* youth competition. The supportive journalist explained that she had also been a victim of harassment when she was a minor. Within days, more than 100,000 women had shared their own experiences of sexual violence. According to what was published, the average age of the first harassment in Brazil was 9.7 years old. And this happened not only in the street, but also in the family, where girls are supposed to feel safe (Ruiz-Navarro, 2016).

A few months later, accompanying the call to take to the streets on 24 April 2016 in Mexico, the Spanish version of this femitag went viral when Colombian journalist Catalina Ruiz-Navarro wrote on Twitter: "When and how was your first harassment? Today from 2pmMX using the hashtag #MiPrimerAcoso ('My First Harassment'). We all have a story, raise your voice!". While a self-convened crowd of women marched in 40 cities across the country, the testimonies followed one after another on the networks. With #MyFirstHarassment, the experience of "realising" managed to transform the damage or shame into political power in the networks and simultaneously in the streets.

But the most successful femitag of this kind, with the longest duration and global reach, came at the end of 2017: #MeToo was able to jump context, iterate and mutate over the last 5 years. Because of its brevity and simplicity, #MeToo is not only read, it is seen: it becomes an iconic unit, it becomes a meme. Launched on Twitter from the United States on 15 October 2017 by Alyssa Milano, it crossed borders and languages. Its ability to produce and connect individual stories was articulated in the Mexican case with other hashtags: #YoSiTeCreo and its variants, or #NoEstásSola, femitags more specifically dedicated to generating affective community, which came from the Spanish context, with the #Cuéntalo campaign in 2018, unleashed after the so-called La Manada⁷ sentence.

Unlike #MyFirstHarassment, the Mexican #MeToo not only recounts a situation of harassment or violence, but also names the perpetrator. In this sense, it took the repertoire of protest known as *escrache*⁸ to digital networks. The different hashtags that emerged from #MeToo in Mexico⁹ had an unexpected intensity between 24 March and 10 April 2019

and showed the ineffectiveness (social, cultural, legal) of access to justice for women victims of sexual harassment and rape. The Mexican #MeToo, with all its variants by workplace and educational centre, had an enormous impact on public opinion and the media (Rovira-Sancho, 2023). It also marks the beginning of the misogynist wave in digital networks directed against women activists and journalists: discrediting them, threatening them and making them the target of new aggressions (Pedraza, 2019).

Table 1. Axes of relevance of the collected *femitag*s

	Against femicide	For abortion	Against sexual violence	In 1st and 2nd person singular	In 1st person plural	Past tense	Present tense	Future tense	Paradox and denunciation	Affective community	Add cases (Archive)	Collective identity
#NiUnaMenos	X											
#NiUnaMás	X										X	
#MeToo			X	X						X	X	
#YoSiTeCreo, #YoTeCreo			X	X			X			X		
#AmigaYoTeCreo												
#Cuéntalo			X	X			X			X	X	
#VivasNosQueremos	X				X		X			X		
#8M, #8M2020, #8M2021, #DiaInternacionalDeLaMujer	X	X	X									X
#HuelgaFeminista8M, #HuelgadeMujeres, #ParoInternacionalDeMujeres	X	X	X						X			X
#NosotrasParamos, #UnDiaSinNosotras	X	X	X		X				X			X
#NoEstasSola			X	X			X			X		
#NOesNO			X				X		X			
#MarchaFeminista												X
#3J, #24A	X		X									
#ViolenciaMachista	X		X								X	
#NoMeCuidanMeViolan, #MeCuidanMisAmigas			X	X			X		X	X		
#UnVioladorNoSeraGobernador			X					X				
#ViolenciaDeGénero			X									
#BastaYa	X		X				X				X	
#IngridEscamilla	X										X	
#DiaContraLaViolenciaDeGenero, #25N	X	X	X									
#JusticiaPatriacal	X		X								X	
#SeVaACaer	X	X	X					X		X		
#AbortoLegalYa, #AbortoLegal, #28S		X					X					
#MiercolesNegro	X											
#MachismoMata	X						X				X	
#MiraComoNosPonemos			X		X					X	X	
#Metoo, #MeTooMx			X	X						X	X	
#MexicoFemicida	X						X				X	
#VivasLasQueremos	X				X					X		
#EstaEsNuestraManada			X		X					X		X
#FuimosTodas					X	X				X		X
#ViajoSola	X			X			X		X	X		
#SiMeMatan	X			X				X	X	X		
#MisSeñasParticulares	X			X					X	X		X
#SeráLey		X						X	X			
#MareaVerde		X										X
	Feminist agenda			Verbal Person and Tense				Specificity/Emphasis				

5. Femitags of the crowds connected from Mexico

In Table 1 we have marked the axes of relevance of the *femitags* collected. Regarding the feminist agenda, for example, we have only indicated whether the *femitag* initiates and has a special impact on a specific campaign of the three main ones found:

- against femicide,
- for the legalisation of abortion, and
- against sexual violence.

If this is not the case, it is understood that it is generalist and includes all of them, such as #8M and other dates of large mobilisations or #ViolenciaMachista (“Misogynist Violence”), #ViolenciaDeGenero (“Gender Violence”), #JusticiaPatriarcal (“Patriarchal Justice”). It is worth highlighting the relevant axis of #HuelgaFeminista (“Feminist Strike”), #ParoInternacionaldeMujeres (“International Women’s Stoppage”), #NosotrasParamos (“We Stop”), as it marks a milestone of global mobilisation and a new issue on the feminist agenda that will become fundamental: that of undervalued reproductive work and wage inequality, tracing a new framework: the relationship between patriarchy and capitalism.

With Latin America being the second most lethal region for women in the world, it is not surprising to see the wide deployment of *femitags* to denounce femicides that take on a transnational dimension: #NiUnaMenos or #MiércolesNegro originating in Argentina, #ViajoSola, which emerged in Ecuador, #NiUnaMas in Mexico.

It is also common to create hashtags with the names of victims, such as #IngridEscamilla in Mexico, as well as many others with names and surnames that we represent in this hashtag.

Against femicide we also find #NiUnaMas, #VivasNosQueremos, #MachismoMata, #MexicoFemicida (“Mexico Femicidal”), #VivasLasQueremos, #SiMeMatan, #MisSeñasParticulares.

Femitags specifically designed to denounce sexual violence are: #MeToo (originating in the United States), #MeTooMx and variants (which in Mexico were up to 40 in 2019), #YoSiTeCreo, #AmigaYoTeCreo, #YoTeCreo, #Cuéntalo (“Tell It”) (originating in Spain but viralised throughout Latin America), #NoEstasSola (“You Are Not Alone”), #NoEsNo (“No Is No”), #NoMeCuidanMeViolan, #MeCuidanMisAmigas, #MiraComoNosPonemos (“See How We Get”) (typical of Argentinean women activists), #EstaEsNuestraManada (“This Is Our Pack”), #UnVioladorNoSeráGobernador (“A Rapist Will Not Be Governor”) (the latter against the electoral nomination of a Mexican politician accused of rape (Portillo *et al.*, 2022).

For its part, the campaign for the legalisation of abortion has had its arsenal of transnational *femitags*, most of them driven by the Argentinean crowds, calling to take to the streets and put pressure on the legal apparatus of each country: #AbortoLegalYa (“Legal Abortion Now”), #AbortoLegal (“Legal Abortion”), #SeraLey (“It Will Be Law”), #MareaVerde (“Green Tide”), #PañuelazoInternacional (“International Scarf-Raising”), #28S, #26S.

5.1. The personalisation of politics or the personal is the political

We can also observe the recurrent use of the first and second person in *femitags*, especially in the sub-genre for stories and counts. The “I” and the “you” are present, even combined, in hashtags such as #MeToo, #YoTeCreo or #Cuéntalo. Undoubtedly, this type of rhetorical resource coincides with what has been considered typical of digital networks: the personalisation of politics, which leaves aside mediating structures and expresses itself from the first person, from individual experience, making the personal the political. The plural, the “we” (feminine) appears in the *femitags* where the desire for unity of the collective is shown and the desire that none of them should be missing as a slogan of struggle: #NiUnaMenos (from us), #NiUnaMás (from us), #VivasNosQueremos, #FuimosTodas (“We Were All”), #EstaEsNuestraManada, #NosotrasParamos.

5.2. Building affective community and feminist awareness

Silence-breaking hashtags often appear in concurrence with those whose main function is to emphasise the value of women’s voices and confirm the existence of an affective community based on empathy: #MeToo, #YoSiTeCreo, #YoTeCreo, #AmigaYoTeCreo, #Cuéntalo, #NoEstasSola, although in some way all *femitags* fulfil the purpose of connecting emotions and motivating to action, as do the following: #MeCuidanMisAmigas, #EstaEsNuestraManada, #FuimosTodas, #MarchaFeminista (“Feminist March”). At the same time, some *femitags* warn and inform by creating a community that teaches itself and raises awareness among peers, in a clear distributed pedagogical function, reiterating examples in #NoesNo, #BastaYa (“Enough is Enough”), #JusticiaPatriarcal, #MachismoMata, #ViolenciaDeGenero.

5.3. Filing of reports as a marker of impunity

Several hashtags have the function of indexing new reports of cases of violence in a growing digital archive: #MeToo and its variants, #MiPrimerAcoso, or #MiraComonosponemos do so for sexual violence. A *femitag* whose priority function is the indexing of new femicides is #NiUnaMás. It is usually accompanied by hashtags with the names of the victims. These seek not

Some *femitags*, because of their intertextual quality, show the historical density of protests, invoke previous struggles, and weave a global dimension between local mobilisations and sometimes distant causes

only to denounce and call for justice, but also to recover the memory and humanise the murdered women. For example, in the case of #IngridEscamilla, the hashtag managed to change the denigration of the leaked images of her tortured corpse for other images that associate her memory with aspects of her life, landscapes of peace and beauty (*Signa_Lab*, 2020).

5.4. Foreshadowing and *jetztzeit* (= the time of now)

Most of the *femitags* with conjugated verbs are in the present tense: #NoesNo, #YoSiTeCreo, #Cuéntalo, #NosotrasParamos, #NoEstasSola, #NoMeCuidanMeViolan, #MeCuidanMisAmigas, #ViajoSola, #VivasNosQueremos, #Estaesnuestramanada, #Miracomonosponemos, #Vivaslasqueremos, #MachismoMata.

Only in one case does a verb appear in the past tense: #FuimosTodas. This hashtag was punctual and specific: it appeared in mid-August 2019 in Mexico in the face of the media and institutional rejection of the direct action of women in the streets, breaking and burning. #FuimosTodas allowed a networked community to collectively take on the destruction and graffiti to alleviate the criminalisation of protest, at a time of enormous backlash.

The present tense of the *femitags* opens a *Jetztzeit*, in Walter Benjamin's terms (Benjamin, 2008, p. 51). It shows the prefigurative will of connected action, here and now, opening up the moment and making the present the place of struggle and listening, without waiting any longer or projecting. The few *femitags* that appeal to the future are imperative statements rather than expressions of utopias: #SeráLey (abortion), #SeVaACaer (patriarchy); sometimes they anticipate the future to call for action in the now and avoid the paradox they announce: #SiMeMatan (...it will be because I liked to go out at night, for example) or #MisSeñasParticulares (...in case they find my corpse). Denouncing through paradoxes is a complex rhetorical device; it involves putting oneself in the place of a future victim to ensure that this should not happen and to fight right now to prevent it.

5.5. The ephemeral life of a hashtag

In terms of duration, some hashtags can jump continents and stretch for years, such as #MeToo. Others become trending and then disappear, but can be resurrected at another time. Their duration does not serve as an indicator of success or failure, as some are one-off campaigns that fulfil their function, such as #UnVioladorNoSeráGobernador or even #MiPrimerAcoso, which trended in Mexico with the call for a march on 24 April 2016. Or #SiMeMatan, which accompanied the anger at the university after the feminicide of Lesvy Berlin.

Other hashtags remain beyond their initial purpose, which was often to call for a spatially situated mobilisation. For example, #NiUnaMenos, which emerged to call to the streets in June 2015, became a global indicator of the current transnational feminist wave. At the same time, it has become an Argentinean movement with a broad agenda, ranging from abortion to denouncing the debt. In the same sense, #MeToo has been reiterated, translated and appropriated in different variants and in multiple campaigns in more than 90 countries. There are *femitags* that are repeated intermittently over the years and localities, as they indicate agendas that are not resolved, such as #AbortoLegal, while global dates are repeated such as #8M for *International women's day*, or for calls that are repeated such as #HuelgaDeMujeres, or slogans that function as motivational frameworks, extending across Latin American time and space, such as #SeVaACaer ("It Is Going To Fall", referring to patriarchy) or #NoesNo (No is No).

6. Conclusions

The hashtags of the connected Mexican feminist crowds make up a series of metacommunicative elements characteristic of this wave of protests and show the profound overlap between street mobilisations and online activity throughout the Spanish-speaking world. In reviewing the arsenal of feminist hashtags from Mexico (which are not Mexican), we have found that they construct a common agenda of feminisms between countries: Argentina, Mexico, Ecuador, Spain, Peru, Paraguay... These *femitags* transcend their indicative function to become performative meta-discursive elements, which sustain the weak ties of feminist multitudes, as they allow for coordinated action and the extension of frames. *Femitags* in their diversity thus constitute a toolbox for convening, denouncing and narrating grievances. Some *femitags*, because of their intertextual quality, show the historical density of protests, invoke previous struggles and weave a global dimension between local mobilisations and sometimes distant causes. They practice a memory of free appropriation, which rescues some legacies (such as the struggle against the feminicides in Ciudad Juárez, the disappeared of Ayotzinapa or the *Mothers of the May Square* in Argentina).

As performatives, *femitags* acquire a strong affective connotation that vindicates life, against feminicides. For example, #VivasNosQueremos, with its emphasis on the present, is not only an invocation to take to the streets but the affirmation of a challenge that implies recognising the importance of each one, denying the possibility that there are lives that do not matter. The questioning of the state by some *femitags* are serious denunciations and markers of impunity. They reveal and make visible

“ VivasNosQueremos, with its emphasis on the present, is not only an invocation to take to the streets but the affirmation of a challenge that implies recognising the importance of each one, denying the possibility that there are lives that do not matter ”

how patriarchal violence is sustained by institutional omission, which proves ineffective in defending life and which normalises sexual violence and re-victimises the victims.

The personalisation of politics, typical of the digital era, is manifested in the abundance of *femitags* in the first (or second) person, calling for participation from the singular and embodied experience: one's own voice, rejecting mediation or representation.

Feminist multitudes have obtained in the common space of the *femitag* the possibility of eluding the abstract character of universals and appealing to a "multiple generalisation": the patient and unexpected creation of an expansive "we", as **Gutiérrez** (2014) would say. The distribution of voices and the absence of strong or permanent leadership constitutes an experience of political power. It also makes it susceptible to backlash, to the use of the same hashtags against activists, with cybermisogyny and hate speech.

The *femitags* of stories and counts that are used to break the silence construct a constellation of affections and voices that manages to put together a jigsaw puzzle: an unprecedented archive of sexual violence. These *femitags* generate conversations that return epistemic authority to the victims themselves about what constitutes harassment or rape. The pedagogical work among peers through hashtags deserves a separate investigation, as well as the feminist consciousness-raising that reading and publishing their own grievances on social media triggers among young women.

Femitags not only sustain networks and oil protests, but also have the function of violence archive and cases indexing. The fleeting nature of networks and the corporate ownership of large platforms such as *Twitter*, *Instagram* and *Facebook* mean that the duration and permanence of this body of discursive evidence is not guaranteed. Only in the case of #Cuéntalo is there a public and accessible repository of the testimonies collected on *Twitter*, thanks to the initiative of the *Association of Archivists of Catalonia* together with the journalists Cristina Fallarás and Karma Peiró, which can be consulted at

<https://proyectocuentalo.org>

Documenting and preserving these testimonies posted on the networks via hashtags is essential for thinking about and diagnosing the state of the world and the efforts to transform it.

7. Notes

1. "Hacking is doing things with code, with machines, with words, with connections, with interventions, with bodies. Hacking is an attitude that brings the hand and politics into play in any space and with any machine or social gadget" (**Rovira**, 2017, p. 111).

2. In the context of the brutal sexual murders of young women in Ciudad Juarez on the US-Mexico border, anthropologist Marcela Lagarde (**Lagarde**, 2005) made the distinction, now incorporated into law in 16 countries, between any murder of women and girls (femicide) and hate crimes as a culmination of gender-based violence and impunity (femicide). According to Lagarde, feminicides are

"the result of misogynist violence taken to the extreme and are thus the most visible sign of multiple previous forms of harassment, mistreatment, harm, repudiation, harassment and abandonment" (2005, p. 1).

3. <https://www.facebook.com/mujeresgrabando/photos/convocatoria-de-grabado3era-edici%C3%B3n-de-la-campa%C3%B1a-gr%C3%A1fica-vivasnosqueremosen-jul/1191818207564841>

4. Originally in Spanish, "contar y contarnos" means at the same time to tell a story and count how many we are.

5. As opposed to the strategy of the powerful, the tactic corresponds to the weak and their ways of trying to get their own way. It is a matter of

"clever tricks of the 'weak' within the order established by the 'strong,' an art of putting one over on the adversary on his own turf, hunter's tricks, maneuverable, poly-morph mobilities, jubilant, poetic, and warlike discoveries" (**De-Certeau**, 1984).

6. It is a graphic and didactic material that circulates in all formats and contains a classification of the various manifestations of intimate partner violence.

7. On 26 April 2018, the La Manada sentence was announced in Spain, which cleared 3 of the 5 men who assaulted an 18-year-old girl in Pamplona in 2016 of rape (sexual assault) and only convicted them of sexual abuse.

8. The *escrache*, which originated in Argentina, consists of going in groups to the homes or workplaces of alleged perpetrators of genocide to publicly point them out to the public.

9. Among the top hashtags of the nearly 40 that emerged during the period were:

#MeTooEscritoresMexicanos ("Me Too Mexican Writers")

#MeTooAcadémicosMx ("Me Too Scholars Mx")

The fleeting nature of networks and the corporate ownership of large platforms such as *Twitter*, *Instagram* and *Facebook* mean that the duration and permanence of this body of discursive evidence is not guaranteed

#MeTooActivistasMexicanos (“Me Too Mexican Activists”)
 #MeTooArtesMx (“Me Too Arts Mx”)
 #MeTooCreativosMexicanos (“Me Too Mexican Creative workers”)
 #MeTooMúsicosMexicanos (“Me Too Mexican Musicians”)
 #MeTooTeatroMexicano (“Me Too Mexican Theatre”)
 #MeTooPeriodistasMexicanos (“Me Too Mexican Journalists”)
 #MeTooUNAM (“Me Too National Autonomous University of Mexico”)
 #MeTooAbogadosMx (“Me Too Lawyers Mx”)
 #MeTooEmpresarios (“Me Too Business Men”)
 #MeTooFotografos (“Me Too Photographers”)
 #Metoolbero (“Me Too Ibero-American University of Mexico-City”)
 #MeTooPolíticos (“Me Too Politicians”).

8. References

- Annunziata, Rocío** (2020). “La acción conectiva de las mujeres. Análisis de la movilización del #ParoInternacionalDeMujeres del 08 de marzo de 2017”. *Revista dígitos*, v. 1, n. 6, pp. 159-180.
<https://doi.org/10.7203/rd.v1i6.177>
- Austin, John-Langshaw** (1982). *Cómo hacer cosas con palabras*. Palabras y acciones. Barcelona: Paidós. ISBN: 978 84 75091419
- Baer, Hester** (2016). “Redoing feminism: Digital activism, body politics, and neoliberalism”. *Feminist media studies*, v. 16, n. 1, pp. 17-34.
<https://doi.org/10.1080/14680777.2015.1093070>
- Banet-Weiser, Sarah; Miltner, Kate** (2016). “#MasculinitySoFragile: Culture, structure, and networked misogyny”, *Feminist media studies*, v. 16, n. 1, pp. 171-174.
<https://doi.org/10.1080/14680777.2016.1120490>
- Barker-Plummer, Bernadette; Barker-Plummer, Dave** (2017), “Twitter as a feminist resource: #YesAllWomen, digital platforms, and discursive social change”. In: Earl, Jennifer; Rohlinger, Deana A. (eds.). *Social movements and media* (Studies in media and communications), v. 14, Emerald Publishing Limited, Bingley, pp. 91-118.
<https://doi.org/10.1108/S2050-206020170000014010>
- Benjamin, Walter** (2008). *Tesis sobre la historia y otros fragmentos*. Edición y traducción de Bolívar Echeverría. México: UNAM-Itaca. ISBN: 970 94353 0 21
- Bennett, W. Lance** (2012). “The personalization of politics: Political identity, social media, and changing patterns of participation”. *The annals of the American Academy of Political and Social Science*, v. 644, n. 1, pp. 20-39.
<https://doi.org/10.1177/0002716212451428>
- Bennett, W. Lance; Segerberg, Alexandra** (2012). “The logic of connective action. Digital media and the personalization of contentious politics”. *Information, communication & society*, v. 15, n. 5, pp. 739-768.
<https://doi.org/10.1080/1369118X.2012.670661>
- Bonila, Yarimar; Rosa, Jonathan** (2015). “#Ferguson: Digital protest, hashtag ethnography, and the racial politics of social media in the United States”. *American ethnologist*, v. 42, n. 1, pp. 4-17.
<https://doi.org/10.1111/amet.12112>
- Briones-Medina, Patricia-Fernanda** (2022). *Hagámoslo juntas (DIT): Hackfeminismos online en México*. Tesis doctoral, Universidad Autónoma Metropolitana.
- Brock, André** (2012). “From the blackhand side: Twitter as a cultural conversation”. *Journal of broadcasting & electronic media*, v. 54, pp. 529-549.
<https://doi.org/10.1080/08838151.2012.732147>
- Candón-Mena, José** (2019). “Movilización cultural y artística en los movimientos tecnopolíticos contemporáneos”. *Teknokultura*, v. 16, n. 1, pp. 11-31.
<https://doi.org/10.5209/TEKN.63162>
- Castells, Manuel** (2012). *Redes de indignación y esperanza*. Madrid: Alianza Editorial. ISBN: 978 84 20609607
- Chávez-Rodríguez, Josefina** (2017). “La revuelta mundial de las mujeres: Un nuevo movimiento”. *Cuadernos feministas*, v. 20, n. 34, pp. 23-27.

- Chen, Boyu; Liao, Da-chi; Wu, Hsin-che; Hwan, San-yih** (2014). "The logic of communitive action: A case study of Taiwan's Sunflower movement". *The internet, politics and policy conference*.
http://ipp.oii.ox.ac.uk/sites/ipp/files/documents/IPP2014_Chen.pdf (no accessible on June 2023)
- Clark, Rosemary** (2016). "Hope in a hashtag: The discursive activism of #WhyIStayed". *Feminist media studies*, v. 16, n. 5, pp. 788-804.
<https://doi.org/10.1080/14680777.2016.1138235>
- Cochrane, Kira** (2014). *All the rebel women. The rise of the fourth wave of feminism*. London: Guardian Books. ISBN: 978 1 783560363
- Conley, Tara L.** (2017). "Decoding black feminist hashtags as becoming". *The black scholar, journal of black studies and research*, v. 47, n. 3, pp. 22-32.
<https://doi.org/10.1080/00064246.2017.1330107>
- Crenshaw, Kimberle** (1991). "Mapping the margins: Intersectionality, identity politics, and violence against women of color". *Stanford law review*, v. 43, n. 6, pp. 1241-1299.
<https://doi.org/10.2307/1229039>
- Daer, Alice R.; Hoffman, Rebecca; Goodman, Seth** (2014). "Rhetorical functions of hashtag forms across social media applications". In: *Sigdoc'14: Proceedings of the 32nd ACM International conference on the design of communication cd-rom*, article n. 16.
<https://doi.org/10.1145/2666216.2666231>
- De-Certeau, Michel** (1984). *The practice of everyday life*. University of California Press, Berkeley. ISBN: 0 520 23699 8
<https://chisineu.files.wordpress.com/2012/10/certeau-michel-de-the-practice-of-everyday-life.pdf>
- De-Kosnik, Abigail; Feldman, Keith P.** (eds.) (2019). *#Identity hashtagging race, gender, sexuality and nation*. USA: University of Michigan Press.
<https://doi.org/10.3998/mpub.9697041>
- Earl, Jennifer; Kimport, Katrina** (2011). *Digitally enabled social change: Activism in the internet age*. Cambridge, MA: MIT.
<https://doi.org/10.7551/mitpress/9780262015103.001.0001>
- Esquivel-Domínguez, Daniela-Carolina** (2019). "Construcción de la protesta feminista en hashtags: Aproximaciones desde el análisis de redes sociales". *Comunicación y medios*, v. 28, n. 40, pp. 184-198.
<https://doi.org/10.5354/0719-1529.2019.53836>
- Foucault, Michel** (1997). *Nietzsche, la genealogía, la historia*. Madrid: Pre-Textos. ISBN: 978 84 85081974
- Fuentes, Marcela** (2019). "#NiUnaMenos. Hashtag performativity, memory and direct action against gender violence in Argentina". In: Gül Altınay, Ayşe; Contreras, María-José; Hirsch, Marianne; Howard, Jean; Karaca, Banu; Solomon, Alisa (eds.). *Women mobilizing memory*, pp. 171-191. New York: Columbia University Press. ISBN: 978 0 231191852
- Ganzer, Miranda** (2014). "In bed with the trolls". *Feminist media studies*, v. 14, n. 6, pp. 1098-1100.
<https://doi.org/10.1080/14680777.2014.975441>
- García-González, Lidia-Ángeles** (2021). "Movimientos feministas en México: prácticas comunicativas digitales y riesgos". *Virtualis: Revista de cultura digital*, v. 12, n. 23, pp. 44-66.
<https://doi.org/10.46530/virtualis.v12i23.382>
- Garibotti, María-Cecilia; Hopp, Cecilia-Marcela** (2019). "Substitution activism: The impact of #MeToo in Argentina". In: B. Flilegorn, Blanca; Loney-Howes, Rachel (eds.). *#MeToo and the politics of social change*. Palgrave Macmillan: pp. 185-199.
https://doi.org/10.1007/978-3-030-15213-0_12
- Garrido-Ortolá, Anabel** (2022). "Reivindicaciones feministas de la cuarta ola: La transnacionalización de la protesta". *Asparkia*, n. 40, pp. 191-216.
<https://doi.org/10.6035/asparkia.6184>
- Gill, Rosalind** (2016). "Post-postfeminism? New feminist visibilities in postfeminist times". *Feminist media studies*, v. 16, n. 4, pp. 610-630.
<https://doi.org/10.1080/14680777.2016.1193293>
- Granovetter, Mark** (1983). "The strength of weak ties: A network theory revisited". *Sociological theory*, v. 1, pp. 201-233.
<https://doi.org/10.2307/202051>
- Gutiérrez, Raquel** (2014). "Políticas en femenino. Reflexiones acerca de lo femenino moderno y del significado de sus políticas". En: Millán, Mária (ed.). *Más allá del feminismo*. México: Red de Feminismos Descoloniales.

- Hochschild, Arlie-Russell** (1979). "Emotion work, feeling rules, and social structure". *American journal of sociology*, v. 85, n. 3, pp. 551-575.
<https://doi.org/10.1086/227049>
- Jones, Quentin** (1997). "Virtual-communities, virtual settlements and cyber-archeology: A theoretical outline". *Journal of computer-mediated communication*, v. 3, n. 3, pp. 35-49.
<https://doi.org/10.1111/j.1083-6101.1997.tb00075.x>
- Kadic, Neira** (2019). *Twitter resistance and digital testimony in 140 characters: Restoring the complexity of Mexico's hashtag feminism*. Master's thesis, College of International Studies, University of Oklahoma.
- Keller, Jessalynn; Mendes, Kaitlynn; Ringrose, Jessica** (2018). "Speaking 'unspeakable things': Documenting digital feminist responses to rape culture". *Journal of gender studies*, v. 27, n. 1, pp. 22-36.
<https://doi.org/10.1080/09589236.2016.1211511>
- Khoja-Moolji, Shenila** (2015). "Becoming an 'intimate publics': Exploring the affective intensities of hashtag feminism". *Feminist media studies*, v. 15, n. 2, pp. 347-350.
<https://doi.org/10.1080/14680777.2015.1008747>
- Kim, Jinsook** (2017). "#iamafeminist as the 'mother tag': Feminist identification and activism against misogyny on Twitter in South Korea". *Feminist media studies*, v. 17, n. 5, pp. 804-820.
<https://doi.org/10.1080/14680777.2017.1283343>
- Knibbs, Kate** (2013). "Slacktivism, unite! Social media campaigns aren't just feel-good back patting". *Digital trends*.
<http://www.digitaltrends.com/social-media/slacktivism-unite-social-media-campaigns-arent-just-feel-good-back-patting>
- Laclau, Ernesto** (1996). "¿Por qué los significantes vacíos son importantes para la política?". En: Laclau, Ernesto. *Emanipación y diferencia*. Buenos Aires: Ariel, pp. 69-86.
- Lagarde, Marcela** (2005). "A qué llamamos feminicidio". *Comisión especial para conocer y dar seguimiento a las investigaciones relacionadas con los feminicidios en la República Mexicana y la Procuración de Justicia*. LIX Legislatura Cámara de Diputados H. Congreso de la Unión, México.
https://xenero.webs.uvigo.es/profesorado/marcela_lagarde/feminicidio.pdf
- Larrondo, Marina; Ponce, Camila** (2019). *Activismos feministas jóvenes. Emergencias, actrices y luchas en América Latina*. Buenos Aires: Clacso, pp. 21-38.
<https://doi.org/10.2307/j.ctvt6rkfs.5>
- Laudano, Claudia** (2019). "Acerca del uso estratégico de TIC en movilizaciones feministas". En: Rivoir, Ana; Morales, María-Julia (coords.). *Tecnologías digitales: Miradas críticas de la apropiación en América Latina*. Buenos Aires: Clacso, pp. 357-369.
<https://doi.org/10.2307/j.ctvt6rmh6.24>
- Liang, Hai; Lee, Francis L. F.** (2021). "Opinion leadership in a leaderless movement: Discussion of the anti-extradition bill movement in the 'Lihkg' web forum". *Social movement studies*, v. 20, n. 1.
<https://doi.org/10.1080/14742837.2021.1989294>
- Lomborg, Stine** (2011). "Social media as communicative genres". *MedieKultur: Journal of media and communication research*, v. 27, n. 51, pp. 17-31.
<https://doi.org/10.7146/mediekultur.v27i51.4012>
- Mendes, Kaitlynn; Ringrose, Jessica; Keller, Jessalynn** (2019). *Digital feminist activism: Girls and women fight back against rape culture*. Oxford University Press.
<https://doi.org/10.1093/oso/9780190697846.001.0001>
- Morales-Sánchez, Deyanira** (2022). "La emocionalidad del activismo digital feminista en los casos de la Brillanteada y la movilización del 8M2020". *Idónea comunicación de resultados*, Máster en Estudios de la Mujer. México: Universidad Autónoma Metropolitana.
- Munro, Ealasaid** (2013). "Feminism: A fourth wave?". *Political insight*, v. 4, n. 2, pp. 22-25.
<https://doi.org/10.1111/2041-9066.12021>
- Natalucci, Ana; Rey, Julieta** (2018). "¿Una nueva oleada feminista? Agendas de género, repertorios de acción y colectivos de mujeres (Argentina, 2015-2018)". *Revista de estudios políticos y estratégicos*, v. 6, n. 2, pp. 14-34.
<https://revistaepae.utem.cl/articulos/una-nueva-oleada-feminista-agendas-de-genero-repertorios-de-accion-y-colectivos-de-mujeres-argentina-2015-2018>

- Núñez-Puente, Sonia; Fernández-Romero, Diana** (2017). "Narrativas transformadoras y testimonio ético: Las estrategias discursivas de la Plataforma Feminista 7N, contra las violencias machistas". *Index comunicación*, v. 7, n. 3, pp. 269-281. <https://dialnet.unirioja.es/descarga/articulo/6277980.pdf>
- Oz, Mustafa** (2016). "Mainstream media's coverage of the Gezi protests and protesters' perception of mainstream media". *Global media and communication*, v. 12, n. 1. <https://doi.org/10.1177/1742766516653164>
- Papacharissi, Zizi** (2016). "Affective publics and structures of storytelling: Sentiment, events and mediality". *Information, communication & society*, v. 19, n. 3, pp. 307-324. <https://doi.org/10.1080/1369118X.2015.1109697>
- Papacharissi, Zizi; De-Fatima-Oliveira, Maria** 2012. "Affective news and networked publics: The rhythms of news storytelling on #Egypt". *Journal of communication*, v. 62, pp. 266-282. <https://doi.org/10.1111/j.1460-2466.2012.01630.x>
- Pedraza, Claudia-Ivette; Rodríguez-Cano, César-Augusto** (2019). "Resistencias sumergidas. Cartografía de la tecnopolítica feminista en México". *Teknokultura. Revista de cultura digital y movimientos sociales*, v. 16, n. 2, pp. 197-212. <https://doi.org/10.5209/tekn.64163>
- Peroni, Caterina; Rodak, Lidia** (2020). "Introduction. The fourth wave of feminism: From social networking and self-determination to sisterhood". *Onati socio-legal series*, v. 10, n. 15, pp. 15-95. <https://doi.org/10.35295/osls.iisl/0000-0000-0000-1160>
- Pfleger, Sabine** (2021). "Fuertes, libres, rebeldes. Hacia una identidad más agentiva del movimiento feminista en México". *Millcayac - Revista digital de ciencias sociales*, v. 8, n. 14, pp. 325-348. <https://doi.org/10.48162/rev.33.015>
- Piñeiro-Otero, Teresa; Martínez-Rolán, Xabier** (2016). "Los memes en el activismo feminista en la Red. #ViajoSola como ejemplo de movilización transnacional". *Cuadernos.info*, v. 39, pp. 17-37. <https://doi.org/10.7764/cdi.39.1040>
- Plumas Atómicas* (2016) #ViajoSola. *Violencia a las mujeres. Video en Facebook*, 10 de marzo. <https://www.facebook.com/watch/?v=1652065365054748>
- Polletta, Francesca** (2006). *Storytelling in protest and politics*. University of Chicago Press.
- Portillo-Sánchez, Maricela; Cano-Rodríguez, César-Augusto; Pérez-Luque, Citlati-Stephany** (2022). "Tecnopolítica feminista, contienda multicampo y derrumbamiento de muros. Una etnografía en red del #8M2021 en México". *Paakat: Revista de tecnología y sociedad*, v. 12, n. 22. <https://doi.org/10.32870/Pk.a12n22.719>
- Rancière, Jacques** (1996). *El desacuerdo. Filosofía y política*. Buenos Aires: Nueva Visión.
- Reverter-Bañón, Sonia; Medina-Vicent, María** (2020). *El feminismo en 35 hashtags*. Madrid: Catarata. ISBN: 978 84 90979051
- Revilla-Blanco, Marisa** (2019). "Del ¡Ni una más! al #NiUnaMenos: Movimientos de mujeres y feminismos en América Latina". *Política y sociedad*, v. 56, n. 1, pp. 47-67. <https://doi.org/10.5209/poso.60792>
- Rovetto, Laura-Florencia** (2015). "Violencia contra las mujeres: Comunicación visual y acción política en 'Ni Una Menos' y 'Vivas Nos Queremos'". *Contratexto*, v. 24, pp. 13-34. <https://doi.org/10.26439/contratexto2015.n024.585>
- Rovira-Sancho, Guiomar** (2017). *Activismo en red y multitudes conectadas*. Barcelona: Icaria-UAMX. ISBN: 978 6072810440
- Rovira-Sancho, Guiomar** (2018). "El devenir feminista de la acción colectiva: Las redes digitales y la política de prefiguración de las multitudes conectadas". *Teknokultura*, v. 15, n. 2, pp. 223-240.
- Rovira-Sancho, Guiomar; Morales-i-Gras, Jordi** (2022). "Femitags for feminist connected crowds in Latin America and Spain". *Acta psicológica*, v. 230, 103756. <https://doi.org/10.1016/j.actpsy.2022.103756>
- Rovira-Sancho, Guiomar; Morales-i-Gras, Jordi** (2023). "Idus de marzo en México. La acción directa en las redes y en las calles de las multitudes". *Teknokultura*, v. 20, n. 1, pp. 11-24. <https://doi.org/10.5209/tekn.81013>
<https://doi.org/10.5209/TEKN.59367>

Signa_Lab (2020). *Ingrid Escamilla: Apagar el horror*. Signa_Lab. Iteso. 13 de febrero.
<https://signalab.mx/2020/02/14/ingrid-escamilla-apagar-el-horror>

Silva-Reis, Josemira; Natansohn, Graciela (2019). "Del ciberfeminismo al hackfeminismo: Notas para pensar Internet en tiempos de la algoritmia". En: Rivoir, Ana-Laura; Morales, María-Julia (eds.). *Tecnologías digitales: Miradas críticas de la apropiación en América Latina*. Clacso: pp. 391-406.
<https://doi.org/10.2307/j.ctvt6rmh6.26>

Sundén, Jenny; Paasonen, Susanna (2019). "Inappropriate laughter: Affective homophily and the unlikely comedy of #MeToo". *Social media + society*, v. 5, n. 4.
<https://doi.org/10.1177/2056305119883425>

Thrift, Samantha (2014). "#Yesallwomen as feminist meme event". *Feminist media studies*, v. 14, n. 6, pp. 1090-1092.
<https://doi.org/10.1080/14680777.2014.975421>

Van-Laer, Jeroen; Van-Aelst, Peter (2010). "Internet and social movement action repertoires: Opportunities and limitations". *Information, communication & society*, v. 13, n. 8, pp. 1146-1171.
<https://doi.org/10.1080/13691181003628307>

Varela, Nuria (2019). *Feminismo 4.0. La cuarta ola*. Barcelona: Random House. ISBN: 978 84 66664431

Zimmerman, Tegan (2017). "#Intersectionality: The fourth wave feminist Twitter community". *Atlantis: Critical studies in gender, culture, and social justice*, v. 38, n. 1, pp. 54-70.
<https://journals.msvu.ca/index.php/atlantis/article/view/4304>

Zuboff, Shoshana (2015). "Big other: Surveillance capitalism and the prospects of an information civilization". *Journal of information technology*, v. 30, n. 1, pp. 75-89.
<https://doi.org/10.1057/jit.2015.5>

Anuario ThinkEPI

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



“As a rule, I don’t have sex on the first date”: Gender and age differences in motivations, perceptions, and experiences on *Tinder*

Carles Roca-Cuberes; Arnau Roig-Mora; Priscila Álvarez-Cueva

Nota: Este artículo se puede leer en español en:
<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87086>

Recommended citation:

Roca-Cuberes, Carles; Roig-Mora, Arnau; Álvarez-Cueva, Priscila (2023). “As a rule, I don’t have sex on the first date’: Gender and age differences in motivations, perceptions, and experiences on *Tinder*”. *Profesional de la información*, v. 32, n. 3, e320308.

<https://doi.org/10.3145/epi.2023.may.08>

Manuscript received on 29th July 2022
Accepted on 17th February 2023



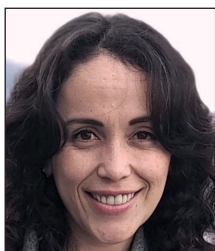
Carles Roca-Cuberes ✉
<https://orcid.org/0000-0002-3604-8722>

Universitat Pompeu Fabra
Departamento de Comunicación
Roc Boronat, 138
08018 Barcelona, Spain
carles.roca@upf.edu



Arnau Roig-Mora
<https://orcid.org/0000-0001-8427-9720>

Universitat Pompeu Fabra
Departamento de Comunicación
Roc Boronat, 138
08018 Barcelona, Spain
arnau.roig@upf.edu



Priscila Álvarez-Cueva
<https://orcid.org/0000-0002-7570-6548>

Universitat Pompeu Fabra
Departamento de Comunicación
Roc Boronat, 138
08018 Barcelona, Spain
priscila.alvarez@upf.edu

Abstract

This research explores the impact that gender and age have on motivations, experiences, and perceptions regarding the use of *Tinder*. Based on an inductive analysis of 37 semi-structured interviews with heterosexual *Tinder* users, we specifically examine gender and age differences in motivations, match selection, and communication management on this mobile dating app. The findings show that age differences have a more significant effect on motivations than gender differences do, whereby older adults use the app to find a stable partner, and young adults use it for sex. Women are more selective when picking matches than men, and when they make these selections, they pay special attention to male attributes that are typically associated with maintaining stable relationships. In contrast, men tend to focus almost exclusively on physical appearance. Between the match and the first date, users need to deal with a considerable volume of communication, which involves the use of different communication media in a series of consecutive stages, toward which matches normatively orient themselves. This transition to new media and stages, in which men tend to take the initiative and women assume the sanctioning role, marks a kind of incremental passage to intimacy. We conclude that, in their courting conduct, *Tinder* users perform conventional gender scripts that are typical of the heteronormative model of intimate relationships.

Keywords

Tinder; Apps; Intimacy; Sex; Gender; Gender differences; Age differences; Heteronormativity; Gender scripts; Youth; Adults; Personal communication; Relations; Motivations; Perceptions; Experiences.



Data availability statement

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Funding

This work was supported by the Spanish *Ministerio de Ciencia e Innovación* under Grant PGC2018-101682-B-I00; and *Departament de Recerca i Universitats de la Generalitat de Catalunya* under Grant 2017SGR53.

1. Introduction

Tinder has become a social phenomenon. It is the most popular of all mobile dating applications, with over 6.5 million downloads a month worldwide in May 2021, practically twice as many as the second most popular app, *Badoo*. With an average of 66 million active users a month, *Tinder* dominates the global dating market with 53.8% of the share. Further data corroborating the *Tinder* phenomenon are the registration of a record 3 billion swipes¹ on May 29, 2020, the over 60 billion matches that have been made since it was launched, and the fact that *Tinder* users go on 1.5 million dates per week (Iqbal, 2021). Therefore, the amount of scholarly attention that it has received since its appearance in 2012 should come as no surprise. Recent studies from different perspectives have explored the motivations for the use of dating apps (Sumter; Vandenbosch; Ligtenberg, 2017; Timmermans; De-Caluwé, 2017), the selection of possible matches (e.g., Timmermans; Courtois, 2018) and (albeit only partially) the management of communication from the match to the first date (Tyson *et al.*, 2016; Sharabi; Dykstra-DeVette, 2019). However, little research has focused on the impact that age and especially gender have on the aforesaid practices and behaviors.

This research addresses this gap in the knowledge base, and specifically investigates the research question: What are the gender and age differences in motivations, match selection, and communication management on *Tinder*?

2. Literature review

Gender and age are crucial dimensions for understanding the practices of online dating, although they have both evolved over time as predictors of online dating usage. As far as gender differences in the use of online dating are concerned, these have been clearly fluctuating. In 2009, the percentage of women compared to men in the United States was 40%-60%, or 38%-62% in Spain (although with exceptions to such a trend in some countries of the former Soviet bloc, like Russia, where the percentage was 65%-35%, or Kazakhstan, 57%-43%) (Kisilevich; Last, 2010). In 2015, statistics showed that in the USA, 45% of online daters were women, compared to 55% men (Smith, 2016). With specific consideration to *Tinder*, the percentage of users in the United States as of 2021 is 24% female versus 75% male (Statista Research Department, 2022).

With regards to age differences in online dating, Stephure *et al.* (2009) found out that older individuals resorted more to online dating to find a prospective partner. This finding was also corroborated by another study from the same period which showed that users aged between 30 and 50 years of 30 were the most active in the online dating domain (Valkenburg; Peter, 2007). A few years later, figures indicated that while the number of 18- to 25-year-old online daters nearly tripled (from 10% in 2013 to 27% in 2015), online dating usage reached its highest among 20- to 40-year-olds, with 39% in 2013 and 43% in 2015 (Smith; Anderson, 2016). By age, the current distribution of *Tinder* users in the United States is (Statista Research Department, 2022):

- 35% in the 18-24 age bracket;
- 25% in the 25-34 age bracket;
- 20% in the 35-44 age bracket;
- 8% in the 45-54 age bracket;and
- 10% in the 55+ age bracket.

Of particular prominence among the studies that have explored the motivations for using *Tinder* is the one by Sumter, Vandenbosch and Ligtenberg (2017), which found that these motivations are love, casual sex, ease of communication, self-worth validation, thrill of excitement, and trendiness. In general, men were more likely than women to describe casual sex as a motivation for using *Tinder*. Men are also more frequently motivated by ease of communication and thrill of excitement (Sumter; Vandenbosch; Ligtenberg, 2017). Interestingly, Sumter, Vandenbosch and Ligtenberg (2017) study also revealed that certain motivations (like love, casual sex, and ease of communication) tended to intensify with age.

Men's sexual motivation for using *Tinder* arises recurrently in the existing literature (Carpenter; McEwan, 2016; Kallis, 2017; Ranzini; Lutz, 2017; Sumter; Vandenbosch; Ligtenberg, 2017; Duncan; March, 2019; Lopes; Vogel, 2019; Palmer, 2020); and this male trend also seems to increase with age (Kallis, 2017). Women, on the other hand, indicate other motivations for using *Tinder*, which include friendship (Ranzini; Lutz, 2017), self-validation or self-esteem (Duncan; March, 2019; Ranzini; Lutz, 2017), and long-term relationships (Palmer, 2020). Even when the use of *Tinder* leads to sex, women (especially younger ones) are reluctant to admit that this may have been their goal (Kallis, 2017). The disparity in motivations for using *Tinder* between men and women can lead to negative experiences. Men looking for sex on *Tinder* can

get frustrated if they cannot find women that are looking for the same thing, and women who are seeking a stable partner may feel unhappy if they continually receive sexual propositions (Carpenter; McEwan, 2016).

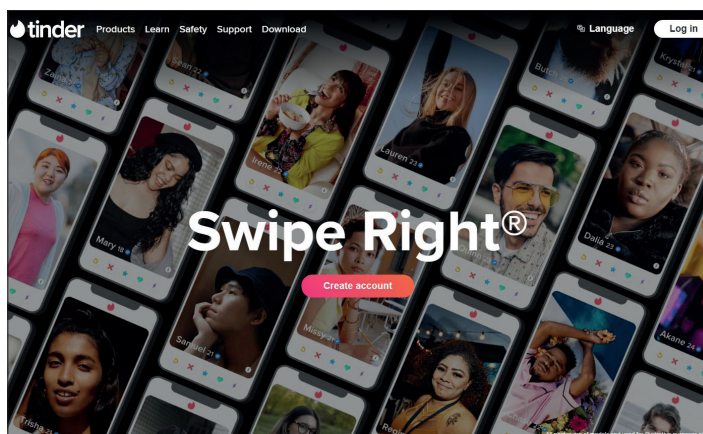
If men and women differ considerably in their approach to sexual permissiveness in *Tinder*, such permissiveness tends to decrease with age (Gatter; Hodkinson, 2016). This online phenomenon mirrors what occurs in the offline world (Le-Gall; Mullet; Shafiqhi, 2002; Mercer *et al.*, 2013).

In a pre-*Tinder* study of online daters over 50 years of age, it was observed that the motivations for trying online dating were considerably different between men and women (McWilliams; Barrett, 2014). Male online daters considered that traditional meeting places were not the most suitable for meeting women of the appropriate age or to start any type of relationship. For these, online dating made up for that inconvenience and could serve the purpose of broadening their options and facilitating the transition to a new relationship. Women, on the other hand, reported a lack of available men in their networks, and online dating was a way to expand the range of possible partners and to gain greater control over the dating process. In general, both men and women were looking for younger partners. However, men paid more attention to the physical attractiveness and women to the social interaction skills of their potential partners. Moreover, while men perceived that women were less beautiful as they get older, women felt that it was men's vivacity that decreased (McWilliams; Barrett, 2014). Older women also found the experiences of online dating stimulating and interesting, on the one hand, as well as stressful and challenging, on the other (Gewirtz-Meydan; Ayalon, 2018). In general, online daters in older age cohorts accepted that they tend to be penalized for their age and that they tried to circumvent this age prejudice by falsifying their real age (i.e., decreasing their age) and exhibiting youthful manners (McWilliams; Barrett, 2014). In addition, in the 50+ age bracket, both men and women agreed that online dating offers them a wider range of people to talk to, without pressure to do so (David, 2012).

The motivations for using online dating (and, in particular, *Tinder*) determine to a large extent actual and perceived dating success, which vary considerably with age (Strugo; Muise, 2019). Thus, for younger users, having approach goals (which refer to the motivation to seek positive outcomes in a relationship, such as growth and intimacy for using *Tinder*) tended to be more firmly related with positive dating outcomes than for older users. Approach goals were also linked with having more romantic partners for older, but not younger *Tinder* users. Overall, the associations between avoidance goals (which manifest the motivation to avoid negative outcomes in a relationship, such as rejection and embarrassment) were most uniform for older users, probably because their reported anxiety might be linked to limited success in *Tinder* (Strugo; Muise, 2019).

The process by which *Tinder* users select other users as potential partners is somewhat similar to that of the offline world, where certain gender differences are also reproduced. Sprecher, Egonie and Treger (2019) found that both online and offline, women are more selective than men, and particularly with regard to resources/success traits and other factors associated with maintaining relationships (such as a pleasant personality). Men, in contrast, are more selective than women with regard to physical appearance and less so concerning resources (e.g., financial security), as corroborated by a large number of previous studies (e.g., Buss; Schmitt, 1993; Hatfield; Sprecher, 1995; Sprecher; Sullivan; Hatfield, 1994). Mate selectivity tends to decrease with age, and particularly in men, and is extended to all traits (including physical appearance). As we get older, our options diminish, and we also adopt a more realistic approach to our mate preferences. Women, however, as they tend to seek long-term partners and invest more in their relationships, maintain a higher degree of mate selectivity than men (Rusbult; Martz; Agnew, 1998). Older users, in comparison to younger participants, also tend to seek long-term relationships and base their choices on factors associated with relationship building (Sprecher; Egonie; Treger, 2019).

The technologically enabled affordances of pre-*Tinder* dating websites allowed their users to present themselves and be selected on the basis of a wide range of socio-demographic characteristics (typically height, weight, ethnicity and education, but also exercise, star sign, drinking, smoking, pets, relationship type, family plans, religion, political leaning, etc.). Such socio-demographic characteristics must be adequately encoded and decoded by the users themselves for adaptation to the environments of those platforms and thus avoid both equivocal self-presentations and the selection of other users with undesirable characteristics. In this type of environment, users overwhelmingly search for similar others (Hardey, 2008), as occurs in offline dating (Harrison; Saeed, 1977; McPherson; Smith-Lovin; Cook, 2001). This phenomenon is known in the literature as homogamy. *Tinder*, unlike those platforms, does not offer so many filtering options beyond a few photographs, a few words of text and (since 2021) a maximum of 5 'interests' (from a list proposed



<https://tinder.com>

by *Tinder*) whereby the user can identify their hobbies, studies or profession. In this context, it has been observed that women are significantly more likely to have matches than men, which suggests that women are more selective in their swiping conduct (Timmermans; Courtois, 2018; Comunello; Parisi; Ieracitano, 2020).

This phenomenon has a simple explanation: as men are

more likely to use *Tinder* for sexual purposes, women tend to try to avoid male users who are only interested in sex. Also, as the *Tinder* interface heavily emphasizes photos and instant appraisals based on limited cues (mainly related to physical appearance) to make swiping decisions, it is not surprising that this trait should take on such special prominence. However, men seem to be more open to making their selections on the basis of looks (Van-Hooff, 2020).

Although physical appearance seems to be the determining criterion when swiping right, age also plays a crucial role. Male users prefer young women while female users prefer partners their age or somewhat older. At older ages, men's tendency to select younger women expands, while women's preferences tend to diversify. Tendencies towards homogeneity are greatest among younger users and women (Šetinová; Topinková, 2021).

The profession of the users is an important factor in the selection process of potential partners. A report issued by the platform itself (*Tinder*, 2016), published a ranking of the professions that men and women found most attractive in 2016. While the most successful female occupations were physical therapist, interior designer, and founder/entrepreneur (in that order), the most popular men were pilots, founders/ entrepreneurs, and fire-fighters (in that order). The 2018 statistics varied slightly, whereby registered nurses, dentists and photographers were the most successful female professions, while for men they were interior designer, pilot and assistant physician (*Tinder*, 2018). What is remarkable about these selection patterns is that while women are interested in prestigious professions or those linked to the eroticism of uniforms, men are more inclined towards professions that are traditionally viewed as feminine and that do not posit competition in professional terms.

When two users both 'like' each other by swiping right, a match occurs. Tyson *et al.* (2016) note that while men gather matches gradually, women gain popularity much more quickly and can even reach as many as 200 matches in the first hour. Once a match has been made, the general expectation is for some kind of conversation to happen via the in-app chat, which should last for a few days (for men) or even weeks (for women) before moving on to another medium (such as *WhatsApp* or *Instagram*) or a face-to-face meeting. According to Licoppe (2019) this conversation should not have sexual connotations, as this is something that women tend to strictly reject. Men typically send the first messages, which are especially flattering ones, while female users are on the receiving end (Comunello; Parisi; Ieracitano, 2020; Timmermans; Courtois, 2018; Tyson *et al.*, 2016; Zytko; Grandhi; Jones, 2014), and were more often the ones being pursued (Kallis, 2017). These behaviors reflect and perpetuate traditional gender roles as already observed at the beginning of offline relationships (Clark; Shaver; Abrahams, 1999). Men and women also differ in the strategies deployed to initiate contact with their matches. While men tend to use cute-flippant, inoffensive pickup lines and directly ask for dates in their messages, women are more likely to focus on the dissimilarities with their matches (Sharabi; Dykstra-DeVette, 2019).

In the transition from online communication to face-to-face encounters, women have been found to be more cautious about meeting strangers (Carpenter; McEwan, 2016). They are generally more selective and swipe for a longer period of time than men (Ward, 2016). Men, on the other hand, are much less demanding when it comes to engaging in face-to-face encounters with their matches (Tyson *et al.*, 2016). Men and women also differ in terms of the reported outcome of their encounters. Contrary to expectations, women were more likely to report a higher number of hookups than men (Timmermans; Courtois, 2018). A possible explanation for this derives from the fact that as women have more matches than men, they are able to have more meetings that end in sex.

The purpose of this study is to investigate the gender and age differences in motivations, match selection and communication management on *Tinder*. In the next sections we present the data and methods employed, the results and discussion, and the conclusion.

3. Method

3.1. Participants

Participants were recruited following the criteria that they needed to be regular *Tinder* users and differed regarding their professional or educational background. We employed snowball and quota sampling, with quotas on age and gender. In total, 37 heterosexual interviewees took part in our study, although the initial aim was to obtain 40 participants. In order to explore possible age differences in uses and practices on *Tinder*, we divided the sample into two age groups: young adults (18-28 years old) and older adults (over 40). The mean age of young adults was 23, while for older adults it was 49. We considered a large enough age gap between young and older adults to attribute plausibility to the possible differences that could be found. The final composition of our sample was: 8 older men, 10 younger men, 10 older women, and 9 younger women. The initial forecast was to find 10 participants for each of the 4 sub-categories.

There are no significant gender differences in terms of motivations for using *Tinder*, except among younger users (especially males), for whom sex is the main motivation

3.2. Instrument for data collection

Our instrument for data collection was semi-structured interviews. For the semi-structured interviews, we used a combination of open and standardized questions, the latter to obtain socio-demographic information about the participants at the beginning of the interviews. The interviews started with a few questions about motivations for using *Tinder* (e.g., What made you decide to download the *Tinder* app? Why and for what do you use *Tinder*? How would you describe *Tinder* to someone who does not know about it?). These were followed by questions about practices when selecting possible matches (e.g., Do you consider yourself selective? In what aspect do you think you are most selective when searching for a possible match?). Finally, there were questions about their communication management (e.g., Did you start conversations or wait for others to do so? How did the conversations flow? Can you give me an example of a positive and a negative *Tinder* date?)

3.3. Study type

We employed a qualitative approach because qualitative methodology offers a greater insight into an individual’s understanding, meaning and experiences using their own language (Aspers; Corte, 2019). Qualitative research provides a framework for discovering new or unexpected findings (White; Cooper, 2022). Such findings allow for ‘information-rich cases’ to acquire a better understanding of the research area (Hamilton, 2020). The goal of qualitative research is not to generalize but instead to offer a rich, contextualized interpretation of people’s experiences through the intensive analysis of particular cases (Lincoln; Guba, 1985).

3.4. Procedure

The interviews were carried out in the metropolitan area of a large Spanish city after we received IRB approval from our university. We conducted the interviews face-to-face in private settings or online depending on the interviewees’ individual preferences and lasted around 45 minutes on average, ranging from 35 to 55 minutes. After obtaining the participants’ permission, the interviews were recorded, transcribed verbatim and anonymized. Pseudonyms were generated for each participant and are used throughout the article.

3.5. Analysis

The analysis of the interviews was done by means of a content analysis, an objective and systematic approach to the study of communication data (Berelson, 1952), performing a frequency analysis of the codes in the interviews. Although the primary use of content analysis has been to analyze media messages, it can also be employed for the study of speech-based data (Prior, 2014).

We combined a deductive and inductive approach in which we first established an initial list of codes derived from prior literature on gender and age differences in motivations, match selection and communication management in online dating (see Appendix). These ‘initial codes’ became sensitizing concepts (Blumer, 1986) that guided the subsequent analysis and led to the development of a preliminary codebook that served to detect new, ‘emerging codes’ (see Appendix), in the interview data. Lastly, the most productive codes and those that were considered most suitable for tracking the motivations, match selection practices, and communication management of our participants (see ‘focused codes’ in Appendix), were incorporated into the final codebook. More codes emerged in the ‘communication management’ category due to the imbalance produced by the larger number of questions in this thematic area. The codes in the final codebook are exhibited in Table 1. For the coding process, we used the *NVivo* software.

Table 1. codes for the analysis of motivations, selection criteria and communication management

Motivations for using <i>Tinder</i>	Selection criteria	Communication management
1. Casual sex 2. Self-worth validation or self-esteem 3. Long-term relationships 4. Shrinking social networks 5. Sentimental breakup 6. To meet people 7. Socializing 8. Entertainment 9. Long-lasting sexual relationship	1. Mate selectivity 2. Factors associated with maintaining relationships (resources, success, a pleasant personality, level of education, profession, hobbies, interests, social status, intelligence) 3. Physical appearance 4. Homogamy/similarity 5. Age 6. Relationship status 7. Ethnicity	1. Who initiates communication 2. Strategies to initiate communication: formal 3. Strategies to initiate communication: personalized 4. Strategies to initiate communication: original 5. Transition to other stages or media: who 6. Transition to other media: only <i>Tinder</i> chat 7. Transition to other media: <i>WhatsApp</i> 8. Transition to other media: <i>Instagram</i> 9. Transition to other media: soon 10. Transition to other media: later stages 11. Thematic agenda: location 12. Thematic agenda: work 13. Thematic agenda: hobbies 14. Thematic agenda: sex 15. Thematic agenda: relationship status (including children and their ages) 16. Thematic agenda: studies 17. Who takes the initiative to propose a date 18. When a date is proposed: soon 19. When a date is proposed: later stages 20. Outcome of the date: sex 21. Outcome of the date: getting to know each other 22. Outcome of the date: friendship

4. Results and discussion

In this section we perform a frequency analysis of the interview data codes showing gender and age differences in motivations for using *Tinder*, in selecting potential matches, and in managing communication with matches.

4.1. Gender and age differences in motivations for using *Tinder*

Table 2 shows the different codes for the category ‘motivations for using *Tinder*’ for the different types of participants in our study (older men, younger men, older women, younger women).

Table 2. Frequency of codes in motivations (raw frequencies and percentage within code)

Gender and age differences in motivations for using <i>Tinder</i>									
Code	Older men		Younger men		Older women		Younger women		Freq. total
	Freq.	%	Freq.	%	Freq.	%	Freq.	%	
Casual sex	12	9%	26	17%	8	5%	18	11%	64
Self-worth validation or self-esteem	3	2%	10	6%	5	3%	15	9%	33
Long-term relationships	29	21%	8	5%	34	23%	12	7%	83
Shrinking social networks	17	12%	4	3%	15	10%	3	2%	39
Sentimental breakup	22	16%	18	12%	24	16%	20	12%	84
To meet people	27	19%	32	21%	33	22%	38	23%	130
Socializing	15	11%	23	15%	17	12%	27	16%	82
Entertainment	12	9%	24	16%	10	7%	28	17%	74
Long-lasting sexual relationship	2	1%	9	6%	1	1%	7	4%	19
Total	139	100%	154	100%	147	100%	168	100%	

Although reasons for using *Tinder* are relatively varied, as shown below, there is a very significant reason for turning to this app (regardless of gender or age), namely a sentimental breakup. This is the second most common code in the category ‘motivations for using *Tinder*’, with 83 frequencies, and it is slightly more prominent among older (46 frequencies) than young (38 frequencies) adults. In older adults, this could be a divorce or separation in recent years, as in the case of Laura, 41:

“I got divorced six years ago and then after a year or so, after seeing how difficult it is to meet other people, they told me about the app and I thought why not give it a go?”

A lot of young adults also end up downloading the app after a more or less recent break-up, instigated by their own social network of people who also use *Tinder*. Luis, 22, said the following:

“Well, a year and a half ago I split up with my partner and so I said OK, when I’d gotten over it and was ready to move on, well I said ‘right, I’ll give this a try’ because I had quite a few friends who were on it and they told me that now and again it can be good and you can meet people and all that.”

Another reason that leads to the use of *Tinder* is, as noted by **McWilliams** and **Barrett** (2014), shrinking social networks. This phenomenon is particularly prominent in older respondents like Antonio, 62, who said:

“In fact, being retired, I have inevitably lost social contact with a lot of people and this made me a bit more inclined to get on *Tinder*, because I didn’t want to be left alone at home, neither at work or socially.”

When interviewees are asked why they use *Tinder*, the most common response is the socially desirable ‘to meet people’ (by far the most frequent code in the category with 130 frequencies). Given that the *Tinder* users in our study have identified themselves as heterosexual, it follows that they can only end up meeting people of the differing sex, and not ‘people’ in general. Daniel, 22, after being asked what he knew about *Tinder* before downloading it, acknowledges that he knew that it was an app used to hook up, and that, in his case, should only be with women. In many cases in our sample, motivations for using *Tinder* only seem to emerge after asking indirect questions or after offering examples of possible uses. What emerges from their responses to questions about motivations for using *Tinder* is that there are no notable gender differences, except among the youngest users. Intra-gender differences, in the form of the age gap, are much more prominent. Thus, young adults in our sample tend to be more inclined to use *Tinder* for casual sex (an average percentage of 14% of all codes for younger adults versus 7% for older adults); this finding is in sharp contrast to **Kallis**’ (2017), who found that sexual motivations increase with age. Older adults, on the other hand, tend to seek a long-term relationship (an average percentage of 22% of all codes in older adults as opposed to 7% in younger adults), which corroborates **Sprecher**, **Econie** and **Treger** (2019) observation. Ana, 49, has been looking for a stable partner after a separation and was advised by her friends to start using *Tinder*. From her own experience and from what she has been told, Ana coincides with our finding with regard to age differences and uses of *Tinder*:

“I think, let’s see, I think that the vast majority of people, in my opinion, use it to find a stable relationship, at least people of my age do. And I think younger people use it more to have sex.”

Andrés, 61, is also hoping to find a long-term relationship:

"I separated ten years ago. The idea is to meet someone, get to know someone, to have a stable partner."

Although having a stable partner is a regular goal of most older male *Tinder* users in our sample, women in the same age group as Monica, 49, refute this assessment. For her,

"Most people of the opposite sex are only looking for sex. I'd say that more than 80 percent are lying when they say, sure, they are looking for a relationship."

Young adults tend to have a more open approach to what they expect to get out of using *Tinder*. It might be for sex, socializing, entertainment, boosting their ego, looking for a partner or the classic 'let's see what happens'. This is succinctly put by Diana, 18:

"Out of boredom. Also to raise self-esteem. Instead of going onto *Instagram*, I go on there and chat to people and it's fun. And for what? Hmm. Nothing specific. I mean, one of the typical questions you ask is 'what are you doing here on *Tinder*?' And it's like: 'Well, I don't know, nothing specific'. I meet people. If I like someone especially then we meet up. If not, well that's as far as it goes. And that's it. But I'm not looking for anything in particular."

Young male adults in our sample, despite what has been shown above, and in line with findings in the existing literature (Carpenter; McEwan, 2016; Kallis, 2017; Ranzini; Lutz, 2017; Sumter; Vandebosch; Ligtenberg, 2017; Duncan; March, 2019; Lopes; Vogel, 2019; Palmer, 2020), are more inclined towards casual sex (the second most frequent code for young male adults in the category with 26 frequencies, after the typical 'to meet people', with 32). This is not a strange coincidence given that these studies, despite generalizing their results to men as a whole, were mostly made with cohorts of young *Tinder* users. Such is the case of David, 23, who has a fairly clear concept of *Tinder*:

"I would say that it is basically for picking up girls quick and easy, and for getting laid."

Yet these motivations might change over time. After a few years on the app, people can end up getting tired of sporadic sex and would prefer to have something more stable. Here's what Javier, 28, says when asked if he's looking for sex on *Tinder*:

"Yes, especially when I was younger, I did. I wasn't getting much... Let's say that with school and work I didn't have that much time. So you'd get chatting and if I liked the girl, then maybe I'd go straight to the point. But right now the idea of having sporadic sex with people, nope."

Young female adults, like Diana above, make more varied and nuanced use of *Tinder*. Even when sex might be one of their priorities, the approach is also much less conventional. Rosa, 21, for example, prioritizes sexual relations, but not sporadic ones. She prefers to have long-lasting sexual relationships (a very infrequent code in the data, with 19 frequencies, and more pronounced among young adults) and does not rule out other forms of intimacy:

"I think a long-lasting sexual relationship would be my priority. Friendship and sporadic sex might be there, on a par, and then comes a stable relationship."

Esther, 23, seeks new sexual experiences, in her case with groups, but also assumes that other forms of relationship might arise:

"To be honest I got into it together with my partner looking for sexual experiences, with other couples or individuals. But I also realized that there are some really nice people, neither of whom is interested in that kind of thing, who make good friends. There is a good connection and we meet up for beers and stuff."

4.2. Gender and age differences in the selection of possible matches

Table 3 shows the various codes employed for the analysis of the category 'selection criteria' for the diverse types of participants in our study.

Table 3. Frequency of codes in selection criteria (raw frequencies within code)

Gender and age differences in the selection of possible matches									
Code	Older men		Younger men		Older women		Younger women		Freq. total
	Freq.	%	Freq.	%	Freq.	%	Freq.	%	
Mate selectivity	18	17%	6	10%	33	30%	29	35%	86
Factors associated with maintaining relationships (resources, success, a pleasant personality, level of education, profession, hobbies, interests, social status, intelligence)	21	20%	3	5%	35	31%	19	23%	78
Physical appearance	28	27%	35	57%	16	14%	22	27%	101
Homogamy/similarity	4	4%	6	10%	5	4%	7	8%	22
Age	25	24%	9	15%	13	12%	4	5%	51
Ethnicity	8	8%	2	3%	9	8%	2	2%	21
Total	104	100%	61	100%	111	100%	83	100%	

In line with the finding by **Sprecher, Sullivan and Hatfield (2019)** and **Comunello, Parisi and Ieracitano (2020)**, we have found that women in our sample are more selective than men. As it may be observed in table 3, the average percentage of 33% of all codes in women versus 13% in men for mate selectivity highlight this fact. For the most part, before swiping another user right, they look at all the photos (if there are more than one), and not only look at their physical appearance but also try to work out or corroborate their various social identity attributes (real age, hobbies, social status, ethnicity, etc.). They also read the description in the bio and study their chosen 'interests' (also with the same purpose as above). This is the case for both young and older women. Mate selectivity, in our sample, is also very high in both female age groups. This code is the second most frequent for older women with 33 frequencies and the most common one for younger female users with 29 frequencies. Raquel, 44, considers herself selective and describes the process she observes before swiping right as follows:

"I read everything they write. [...] I discard anyone showing off their muscles in the gym mirror, out. People with children, out. People with no photo, out. So, when I've eliminated all that, and there's a lot, then I swipe based on what I see."

Although Mati, 21, has not been on the app for long, she has evolved towards more selective swiping behavior:

"Whether I liked the photo or I liked the guy, well, I'd directly give him a like. But now I go a bit further and try to look at his description too, what interests him and I try to also look at more photos, look up his profile on *Instagram*."

Male participants are less selective than women. Unlike what **Rusbult, Martz and Agnew (1998)** propose for the offline world, mate selectivity in men is likely to increase on *Tinder* with age. In our data, the 17% of all codes for older male users contrasts with the 10% for younger men in mate selectivity. It is quite common for young men, like Eduardo, 23, to use up his likes in a single 'swipe session' after swiping right on all the users that the *Tinder* algorithm offers:

"A lot of days I basically go on, swipe everything to use up my likes and then, depending on the people I get matches with, I then decide whether to keep the match or not."

Older men, on the other hand, appreciate that this is a bad swiping strategy, even though they have used it in the past. For instance, Enrique, 57, says:

"I started going pretty full-on. The thing is that experience has taught me that it is absurd to waste time, right? So now I really look for people who fit with what I'm really looking for."

Men generally tend to base their choices of women on *Tinder* on physical appearance, as proposed by **Van-Hooff (2020)** and observed in our data: an average percentage of 42% of all codes for male users versus 21% for women regarding physical appearance. This phenomenon is even more extreme in young men (57% of all codes in physical appearance for younger men versus 27% for older men), who select women first on their physique, and even ignore the social identity cues that can be grasped from their photos. Alberto, 26, admits that looks are his determining criterion:

"I discard, for example, the ones that I don't find physically attractive."

Adult men, like Antonio, 62, select first by looks too, but also pay attention to other things before 'liking' another user:

"But I am more attracted, hmm, by people, women, who are younger than me [...]. The first thing is physical appearance, looks, that's clear. Yeah, yeah. And if I get the chance to read their profile, if they've written one, the content and form also draw my attention."

Selection of younger women, as Antonio does (his search is set in the 50-60 age range), is a phenomenon that increases with age among men. Just like **Šetinová and Topinková, (2021)** revealed in their study, we have found that the older men are, the lower the age range they set in the search engine itself. Age is the second most frequent code for older men with 25 frequencies and also the second most frequent one for younger male users, but only with 9 frequencies. With women, regarding the importance of age, the exact opposite occurs: only 13 and 4 frequencies for older and younger women, respectively.

Female participants, apart from being more selective, tend to base their choices on a wide variety of criteria. Older women, in particular, tend to pay more attention to those cues (discernible in the photos or appreciable in the description and interests) associated with maintaining stable relationships. In effect, the average percentage of 27% of all codes in women versus 12% in men for 'factors associated with maintaining relationships' clearly shows the diverse importance attributed to finding a stable partner for women and men. Laura, 41, despite recognizing the importance of looks in her selections, makes an overall assessment, in which work, level of education (expressed as literacy level) and hobbies are of particular importance:

"It ultimately has to be a bit of everything, but if he has no description and I love the photos, well yeah, if I see someone he might like more things than me and I see him climbing. [...] That there are no spelling mistakes, most of all. [...]. And what he does for work, well, I won't tell a lie, at this stage of the game I'm not going to go with someone who doesn't have two pennies to rub together."

Younger female users, on the other hand, are aware of the relevance of their possible matches' looks in their decisions and primarily value hobbies and (paradoxically in a universe like *Tinder* that is so focused on physical appearance) the

intelligence that can be gleaned from a witty description. What draws the attention of Cristina, 22, in a good description is

“not so much the way he describes himself, but the fact that it’s funny. You know, in fact, there are some I’ve read that you can die laughing, it’s like I’ve pissed myself laughing. And just because of that, just for that, I might like them.”

When picking matches, women are more selective than men; thus, while women focus on those estimable attributes that are usually associated with the maintenance of stable relationships, men pay special or exclusive attention to physical appearance

Older women in the younger group, however, start to appreciate the same things as women in the more senior group. As Gemma, 26, says:

“Well, it’s good, for example, to say what you do for a living, isn’t it? [...] If they’ve got a degree and whatever, well I notice these things because it means we could be more alike, couldn’t we? Or our ideas might be more or less on the same track.”

What we can appreciate as well in the above excerpt is that homogamy has an effect in mate selection in online dating, although with a lower impact (it is the second least frequent code in our ‘selection criteria’ data with 22 frequencies) than that reported by **Hardey** (2008).

4.3. Gender and age differences in managing communication with matches

Table 4 illustrates the codes for the category ‘communication management’ for the groups in our study.

Table 4. Frequency of codes in communication management (raw frequencies within code)

Gender and age differences in managing communication with matches									
Code	Older men		Younger men		Older women		Younger women		Freq.
	Freq.	%	Freq.	%	Freq.	Code	Freq.	%	
Who initiates communication	33	11%	29	10%	7	4%	11	6%	80
Strategies to initiate communication: formal	21	7%	5	2%	4	2%	2	5%	32
Strategies to initiate communication: personalized	9	3%	10	4%	1	1%	4	2%	24
Strategies to initiate communication: original	4	1%	12	4%	1	1%	4	2%	21
Transition to other stages or media: who	32	10%	26	9%	6	4%	9	5%	73
Transition to other media: only <i>Tinder</i> chat	20	11%	3	1%	21	13%	4	2%	48
Transition to other media: <i>WhatsApp</i>	11	4%	13	5%	12	7%	16	9%	52
Transition to other media: <i>Instagram</i>	2	1%	20	7%	2	1%	19	10%	43
Transition to other media: soon	5	2%	34	13%	3	2%	29	16%	71
Transition to other media: later stages	29	9%	3	1%	28	17%	6	3%	66
Thematic agenda: location	7	2%	8	10%	10	6%	10	5%	35
Thematic agenda: work	7	2%	2	1%	8	5%	2	2%	19
Thematic agenda: hobbies	8	3%	7	3%	7	4%	8	4%	30
Thematic agenda: sex	1	1%	2	1%	1	1%	0	0%	4
Thematic agenda: relationship status (including children and their ages)	9	3%	0	0%	11	7%	1	1%	21
Thematic agenda: studies	3	1%	10	4%	1	1%	11	6%	25
Who takes the initiative to propose a date	36	12%	29	10%	8	5%	10	5%	83
When a date is proposed: soon	11	4%	20	7%	1	1%	7	4%	38
When a date is proposed: later stages	22	7%	8	3%	8	4%	4	2%	42
Outcome of the date: sex	14	11%	19	7%	2	1%	5	3%	40
Outcome of the date: getting to know each other	10	3%	9	3%	12	7%	11	6%	42
Outcome of the date: friendship	9	3%	9	3%	14	8%	13	7%	45
Total	305	100%	278	100%	168	100%	186	100%	

From the moment a match occurs and until the potential date, the matches have to manage a sometimes substantial amount of communication. The three consecutive stages and the communication media they have to go through are:

- initial, in which communication is via *Tinder’s* in-app chat;
- pre-date, with communication via *Tinder* chat, *Instagram* or an instant messaging service like *WhatsApp*; and
- date, in which face-to-face communication is employed.

When a match has been made on *Tinder*, someone needs to take the initiative by greeting the other person. As pointed out in recent research (Comunello; Parisi; Ieracitano, 2020; Timmermans; Courtois, 2018; Tyson *et al.*, 2016; Zytko; Grandhi; Jones, 2014), it is men (regardless of their age) who usually initiate communication. Indeed, the average percentage of 11% of all codes in men in contrast to 5% in women for 'who initiates communication' plainly demonstrates the gender differences in taking the initiative to talk to a match. Ricardo, 47, most of the time initiates communication with a greeting lying somewhere between conventional and formal:

"When you get a match, I go in and say 'Hello' or 'Hi, what's up? How are you?' And that's it."

This expectation for men to initiate communication is also shared by women, who (and especially younger ones) also expect the initial communication not to be overly conventional, and instead to particularly reflect interest and personalization. Sara, 45, in awareness, like many other female *Tinder* users, of how unselective men can be, understands that those who have a genuine interest will be the ones who will contact her:

"Because I think, well, maybe there's a match because guys are like that, well ... yeah, yeah, you know? And I think, well, if he's really interested, well, you know?"

That interest also needs to be personalized and to show that there has been substantial inspection of the profile itself. However, that expectation might be unfulfilled, since male users in her age bracket tend to start communication formally (the most frequent strategy with 21 frequencies), rather than personalized (9 frequencies) or with some originality (4 frequencies). Susana, 21, anticipates original messages. In particular, she expects cute-flippant—in line with Sharabi and Dykstra-DeVette's (2019) finding—and personalized first messages from her matches:

"If it's original, so much the better, like, I mean, if you send me a 'hello, what's up, how are you?', well that's not what I find original. If you send me a little message like the other that said 'I'll see you next weekend', well I think that's funny, because I say, hey, at least you've made an effort, and are not just here sending hearts to everybody."

Her expectations are likely to be met, as male users in her age group tend to initiate communication with original messages (the most typical strategy with 12 frequencies) and to a lesser extent personalized (10 frequencies), and even less commonly, formally (5 frequencies).

Whether the conversation progresses to the following stage(s) and (eventually) other media depends on how successful this initial communication is (which in the case of the 'hello, what's up, how are you?' that Susana sometimes receives, would be very low). This progress to other stages and media typically happens at the initiative of men (an average percentage of 10% of all codes for men in contrast to 5% for women), but it is (also typically) women who sanction this initiative. And as we shall see below, this progress also reflects a kind of incremental passage to intimacy.

The pre-date thematic agenda is relatively varied. In the case of our sampled older adults, it usually revolves around two core areas: work and hobbies (both codes with 15 frequencies). As well as these two, there is also, at first, location (17 frequencies). Andrés, 61, describes the things he usually asks about thus:

"Mainly that. Where do you live? What do you do for a living? What do you like to do in your spare time? To find out if there's a connection, right?"

Ricardo, 47, also chats about these subjects and feels that there is one in particular (sex) that should never be touched:

"I don't go directly into sex. That seems highly inappropriate to me."

The expectation of the absence of sex as a topic of conversation (only 4 frequencies for the code 'thematic agenda: sex' in total for all age and gender groups), at least for a 'reasonable' period of time and generally attributable to women, is one of many expectations regarding communication management described by Licoppe (2019). Older adults are inclined to introduce a topic to the thematic agenda that seems to be of major interest to them: their counterparts' past and present relationship status, and also ask about the number and ages of their children (if they have any). This topic is brought up with a frequency of 11 by older women and 9 by older men. Raúl, 44, however, does not think that this is particularly appropriate:

"They sure are more direct sometimes and ask questions that, at first, I don't think belong: if you've been married, how many children you have ... They're things I don't ask straight away. I tend to be more discreet."

Vanesa, 45 (although her profile says 40), searches her matches' profiles for information about their relationship statuses and children. If it does not say anything, it is the first thing she asks:

"I only look to see if they have children or not. If not, the standard questions: 'are you single', 'do you have children'. If it doesn't say anything there, for me it's one of the first things I ask, because it's one of the things that I don't want."

For men, a positive date is one that ends in sex, while for women it is usually one in which the two get along well and leads to some kind of relationship, including friendship

Younger participants’ conversations tend to revolve around three main themes: their location, studies, or hobbies. Within ‘thematic agenda’, these three topics constitute the most common codes with 28, 21, and 15 frequencies, respectively, for this age group. Juan, 28, introduces the topic of hobbies to engage with his match:

“I try to bring up a subject that I know about to see if the other person is interested too and opens up and tells me things about her hobbies.”

There is also sometimes talk of work (4 frequencies), sex (2 frequencies), or relationship status (only 1 frequency), but these topics are rather unusual.

Conversations (as commented by Mati, 21) need to progress from more general to more personal topics:

“So then, maybe, a lot of days might go by before you start talking about kind of more personal stuff. At first, I tell you, it’s like ‘what do you study’, ‘what do you do’, ‘what do you like’, most of all.”

Mati also penalizes those matches that do not normatively adhere to the consecutiveness of the stages and that try to skip some by making no further communication. She gets annoyed when matches try to skip the pre-date stage, the one that is about ‘getting to know each other’:

“There are people that kind of like, that I don’t like either because they are very direct in the sense that they kind of skip a bit the part where you get to know the person, you know? It’s like, hey ‘let’s just meet up now’ and stuff. And you say, ‘hey, we don’t know each other, I hardly know who you are.’”

Interaction in the pre-date phase starts with the *Tinder* chat. At some point, before the date, matches usually switch to other digital media. Older adult interviewees often switch to *WhatsApp* (within ‘transition to other media’, switching to *WhatsApp* is the second most common code with 23 frequencies, while using only the *Tinder* chat is the most frequent with 41 for older users; transiting to *Instagram* is rather unusual, with only 4 frequencies), which can take between 1 and 3 months, to make the necessary arrangements shortly before the date. Before switching to *WhatsApp*, Carmen, 48, needs there to have been an appropriate volume and quality of communication on the *Tinder* chat, just as **Licoppe** (2019) suggested. In those cases, at the men’s request, she gives them her phone number so they can text using *WhatsApp*:

“When you see in the initial chat that they’re asking normal questions, and perhaps write back again the next day, and politely, so well, when you see that they ... Well, I think they’re nice, make comments that get my attention, right? That I like what they say, you know? Well, I end up giving them my, well for example, my *WhatsApp*, okay.”

In this same phase between the *Tinder* chat and *WhatsApp*, young adults in our sample most often also use *Instagram* (within ‘transition to other media’, switching to *Instagram* is the most common code with 39 frequencies and turning to *WhatsApp* the second most common with 29 in young adults, while using only the *Tinder* chat is the least frequent with only 7 frequencies), where in addition to chatting they can obtain a wealth of information of their matches and perform the opportune identity checks. Gemma, 26, describes this transition to other platforms as follows:

“*Instagram* to see, like, more photos because on *Tinder* you ultimately get three or four photos that are the best, but on *Instagram* you can see tags and whatever, and then *WhatsApp* to talk more.”

Each step towards a new medium (normally, also, at the men’s initiative and sanctioned by women) also implies gradually drawing closer together and a higher degree of intimacy with the other person. This is how Lucas, 21, sees it:

“It’s like *Tinder* is for talking. It’s like the first step. And then, if the thing’s flowing, you usually go to other platforms. And creating this bond could be seen as getting closer, to put it one way. It’s like ‘I’m letting you more into my life.’”

The coveted date regularly happens at the man’s initiative (an average percentage of 11% of all codes in ‘who takes the initiative to propose a date’ for men versus 5% for women) and comes after a ‘reasonable’ time, which for older adults might be between 1 and 3 months after the original match, and between 1 and 2 weeks for younger participants. In our data, a date is proposed soon with an average percentage of 3% of all codes for older adults versus 6% for our younger participants, while a date is proposed in later stages in exactly the opposite proportions. The meeting usually happens somewhere central and public, usually a bar.

As it can be seen from Table 4, sex is the most important outcome of a date for men with an average percentage of 9%, versus 2% for women. Women, on the other hand, tend to better appreciate getting to know their dates (7% versus 3% men) or becoming friends with them (8% versus 3% men) as the result of their first date. For example, Agustín, 43, considers that

“a positive date is to meet up with her, go for dinner, go to the movies and have sex on the same day.”

For women, like Carmen, 48, positive dates are those where they have got on well and that have led to some kind of relationship, even if it is only friendship:

“The initial communication is usually led by men, while women mostly assume the role of passive recipients of their messages”

“When I get to meet the person, well, you’ve had an evening out, well, and for me everything has always been positive, and with two or three, well, like, you might carry on writing to each other on *WhatsApp*.”

For there to be sex, there might be one or two further dates. It also generally tends to be women (like Vanesa, 45) who put a stop to things before they get too intimate:

“As a rule, I don’t have sex on the first date. There are people who come and ‘bang’. Not me. Neither the first date, nor the second. I’m very clear about that.”

5. Conclusion

In this paper, we have explored gender (and in a subsidiary manner, age) differences in motivations, practices, and communication management regarding *Tinder* usage. We have observed that there are no appreciable gender differences when it comes to motivations for using *Tinder*, except among the youngest users. Although these motivations may vary over time, the main current reason for our older adult interviewees to use the app is to search for a stable partner. In contrast, the main motivation for young adults is sex, especially in the case of males. Females in the same age group do admit that sex can be a common outcome of their dates but tend to use *Tinder* for a broader range of purposes, which include entertainment, socializing, ego-boosting, and finding a partner, and not just sex.

In terms of selecting matches, we corroborate that women are more selective than men. While women tend to inspect the profiles of their potential matches fairly meticulously, men tend not to go much further than the first photo. Women look at the estimable attributes (such as work, level of education, hobbies, and intelligence) that are typically associated with maintaining stable relationships. However, when men are choosing, they pay special or exclusive attention to the physical appearance observable in the photos. These results reveal, on the one hand, an adequate fit between the selection practices and reported uses of *Tinder* among women and young men. If women are looking for a stable partner, it seems reasonable that their selections will be based on the male attributes that are usually linked to this purpose, or that young men will focus on physical appearance when their goal is sex. On the other hand, if older men’s declared goal is to find a partner, their potential matches’ physical appearance does not seem to be the best criterion that will lead them to a stable relationship. These results also show that, as women are more selective, they are more discerning in their use of *Tinder*, to the extent that they better select how and with whom they establish relationships, and thus practice elective intimacy.

From the moment the match is produced until the first date (if there is one), users who have decided to “like” each other need to handle a sometimes considerable amount of communication, which goes through different media and through a series of consecutive stages, toward which the matches normatively orient themselves. In the transition to new media and stages, men (also normatively) take the initiative, and women assume a sanctioning role. Initial communication usually occurs via the *Tinder* chat feature and is led by men, while women mostly take the role of passive recipients of their messages. In the pre-date phase, the talking points for young people often revolve around their hobbies. Older adults also discuss their work. One topic that is generally dismissed as inappropriate is that of sex, since women may practice “ghosting” if they think a match is coming on too strong too soon. This phase normally involves switching to other media. Older adults usually switch to *WhatsApp* after 1-3 months, but before doing this, younger people usually, and relatively quickly after the initial communication, switch to *Instagram*, where they can find out about or corroborate their matches’ various social identity attributes or continue chatting with them. Older adults can take between 1 and 3 months before going on their first date, while younger users take less time: 1 to 2 weeks. For men, a positive date is one that ends in sex, while one where the two get on well and that leads to some type of relationship usually constitutes a positive date for women. Sex, if it occurs, takes a long time, for one or two further dates are usually needed. The aforesaid progress to other stages and media outlines a kind of incremental passage to intimacy. Each transition to a new medium or stage represents, in a performative fashion, an increase in the matches’ degree of intimacy.

In their uses and practices regarding *Tinder*, interviewees display their awareness of, and a normative orientation towards, the current rules of both offline courtship and those that are adapted to the *Tinder* domain. In their courting conduct, both men and women perform the conventional gender scripts that are typical of the heteronormative model of intimate relationships. These gender scripts, as we have seen above, highlight:

- men’s agency in taking the initiative to make the courting moves, and
- both women’s passiveness, as recipients of men’s initiatives, and sanctioning power over those initiatives.

In general, all these uses and practices are taken by users to express the characteristic traits of masculinity and femininity. We, as analysts, suggest that this is how gender is “performed into being.”

It will be, and indeed already is, interesting to observe what happens with the global implantation and massive use of *Tinder* and similar apps across the globe. In the cases in which our interviewees mention interactions with users living in other countries, they all seem to be aware of, and share, *Tinder*’s courting rules. An increasing global standardization of the processes for starting, developing, and consolidating intimate relationships might seem inevitable.

“ In their courting conduct, both men and women perform the conventional gender scripts that are typical of the heteronormative model of intimate relationships ”

6. Note

1. To *swipe* is the action of sliding left or right on a person in the application. Sliding to the left rejects the person who has not attracted attention, and sliding to the right indicates that that person is interesting for making a *match*.

7. References

- Aspers, Patrik; Corte, Ugo** (2019). "What is qualitative in qualitative research". *Qualitative sociology*, v. 42, n. 2, pp. 139-160. <https://doi.org/10.1007/s11133-019-9413-7>
- Berelson, Bernard** (1952). *Content analysis in communication research*. Glencoe, IL: Free Press. ISBN: 978 0 028412108
- Blumer, Herbert** (1986). "The methodological position of symbolic interactionism". In: Blumer, Herbert (ed.). *Symbolic interactionism: perspective and method*. Berkeley: University of California Press.
- Buss, David M.; Schmitt, David P.** (1993). "Sexual strategies theory: an evolutionary perspective on human mating". *Psychological review*, v. 100, n. 2, pp. 204-232. <https://doi.org/10.1037/0033-295X.100.2.204>
- Carpenter, Christopher J.; McEwan, Bree** (2016). "The players of micro-dating: individual and gender differences in goal orientations toward micro-dating apps". *First Monday*, v. 21, n. 5. <https://doi.org/10.5210/fm.v21i5.6187>
- Clark, Catherine L.; Shaver, Phillip R.; Abrahams, Matthew F.** (1999). "Strategic behaviors in romantic relationship initiation". *Personality and social psychology bulletin*, v. 25, n. 6, pp. 709-722. <https://doi.org/10.1177/0146167299025006006>
- Comunello, Francesca; Parisi, Lorenza; Ieracitano, Francesca** (2021). "Negotiating gender scripts in mobile dating apps: between affordances, usage norms and practices". *Information, communication & society*, v. 24, n. 8, pp. 1140-1156. <https://doi.org/10.1080/1369118X.2020.1787485>
- David, Patricia** (2012). *AARP online dating survey [AARP Research]*. <https://www.aarp.org/research/topics/life/info-2014/online-dating-sites-survey.html>
- Duncan, Zoe; March, Evita** (2019). "Using *Tinder*® to start a fire: predicting antisocial use of *Tinder*® with gender and the dark tetrad". *Personality and individual differences*, v. 145, pp. 9-14. <https://doi.org/10.1016/j.paid.2019.03.014>
- Gatter, Karoline; Hodkinson, Kathleen** (2016). "On the differences between *Tinder*™ versus online dating agencies: questioning a myth. An exploratory study". *Cogent psychology*, v. 3, n. 1, e1162414. <https://doi.org/10.1080/23311908.2016.1162414>
- Gewirtz-Meydan, Ateret; Ayalon, Liat** (2018). "Forever young: visual representations of gender and age in online dating sites for older adults". *Journal of women & aging*, v. 30, n. 6, pp. 484-502. <https://doi.org/10.1080/08952841.2017.1330586>
- Hamilton, Jill B.** (2020). "Rigor in qualitative methods: an evaluation of strategies among underrepresented rural communities". *Qualitative health research*, v. 30, n. 2, pp. 196-204. <https://doi.org/10.1177/1049732319860267>
- Hardey, Mariann** (2008). "The formation of social rules for digital interactions". *Information, communication & society*, v. 11, n. 8, pp. 1111-1131. <https://doi.org/10.1080/13691180802109048>
- Harrison, Albert A.; Saeed, Laila** (1977). "Let's make a deal: an analysis of revelations and stipulations in lonely hearts advertisements". *Journal of personality and social psychology*, v. 35, n. 4, pp. 257-264. <https://doi.org/10.1037/0022-3514.35.4.257>
- Hatfield, Elaine; Sprecher, Susan** (1995). "Men's and women's preferences in marital partners in the United States, Russia, and Japan". *Journal of cross-cultural psychology*, v. 26, n. 6, pp. 728-750. <https://doi.org/10.1177/002202219502600613>
- Iqbal, Mansoor** (2021). "Tinder revenue and usage statistics". *Business of apps*, February 1. <https://www.businessofapps.com/data/tinder-statistics/#2>
- Kallis, Rhiannon-Briana** (2017). *Swipe left or right but what happens for the rest of the night? A qualitative approach to understanding the life cycle of Tinder relationships*. Indiana University of Pennsylvania.
- Kisilevich, Slava; Last, Mark** (2010). "Exploring gender differences in member profiles of an online dating site across 35 countries". In: Atzmueller, Martin; Hotho, Andreas; Strohmaier, Markus; Chin, Alvin. *Analysis of social media and ubiquitous data*. Berlin, Heidelberg: Springer, pp. 57-78. ISBN: 978 3 642 23599 3 https://doi.org/10.1007/978-3-642-23599-3_4

- Le-Gall, Armelle; Mullet, Etienne; Shafighi, Sheila-Riviere** (2002). "Age, religious beliefs, and sexual attitudes". *Journal of sex research*, v. 39, n. 3, pp. 207-216.
<https://doi.org/10.1080/00224490209552143>
- Licoppe, Christian** (2020). "Liquidity and attachment in the mobile hookup culture. A comparative study of contrasted interactional patterns in the main uses of *Grindr* and *Tinder*". *Journal of cultural economy*, v. 13, n. 1, pp. 73-90.
<https://doi.org/10.1080/17530350.2019.1607530>
- Lincoln, Yvonna S.; Guba, Egon G.** (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage. ISBN: 978 0803924314
- Lopes, Milena R.; Vogel, Carl** (2019). "Gender differences in online dating experiences". In: Hetsrone, Amir; Tuncez, Meriç. *It happened on Tinder: reflections and studies on internet-infused dating*, 1st ed. Institute of Network Cultures: Amsterdam, The Netherlands, pp. 31-47. ISBN: 978 94 92302 44 1
<http://www.tara.tcd.ie/bitstream/handle/2262/91343/Lopes-Vogel-IHoT-2019.pdf>
- McPherson, Miller; Smith-Lovin, Lynn; Cook, James M.** (2001). "Birds of a feather: homophily in social networks". *Annual review of sociology*, v. 27, n. 1, pp. 415-444.
<https://doi.org/10.1146/annurev.soc.27.1.415>
- McWilliams, Summer; Barrett, Anne E.** (2014). "Online dating in middle and later life: gendered expectations and experiences". *Journal of family issues*, v. 35, n. 3, pp. 411-436.
<https://doi.org/10.1177/0192513X12468437>
- Mercer, Catherine H.; Tanton, Clare; Prah, Philip; Erens, Bob; Sonnenberg, Pam; Clifton, Soazig; Macdowall, Wendy; Lewis, Ruth; Field, Nigel; Datta, Jessica; Copas, Andrew J.; Phelps, Andrew; Wellings, Kaye; Johnson, Anne M.** (2013). "Changes in sexual attitudes and lifestyles in Britain through the life course and over time: findings from the national surveys of sexual attitudes and lifestyles (Natsal)". *The lancet*, v. 382, n. 9907, pp. 1781-1794.
[https://doi.org/10.1016/S0140-6736\(13\)62035-8](https://doi.org/10.1016/S0140-6736(13)62035-8)
- Palmer, Lauren** (2020). "Dating in the age of Tinder: swiping for love?". In: Carter, Julia; Arocha, Lorena. *Romantic relationships in a time of "cold intimacies"*. Palgrave Macmillan Studies in Family and Intimate Life, Cham., pp. 129-149. ISBN: 978 3 030 29256 0
https://doi.org/10.1007/978-3-030-29256-0_7
- Prior, Lindsay** (2014). "Content analysis". In: Leavy, Patricia (ed.). *The Oxford handbook of qualitative research*. Oxford University Press, pp. 359-379. ISBN: 978 0 190847388
- Ranzini, Giulia; Lutz, Christoph** (2017). "Love at first swipe? Explaining *Tinder* self-presentation and motives". *Mobile media & communication*, v. 5, n. 1, pp. 80-101.
<https://doi.org/10.1177/2050157916664559>
- Rusbult, Caryl E.; Martz, John M.; Agnew, Christopher R.** (1998). "The investment model scale: measuring commitment level, satisfaction level, quality of alternatives, and investment size". *Personal relationships*, v. 5, n. 4, pp. 357-387.
<https://doi.org/10.1111/j.1475-6811.1998.tb00177.x>
- Šetinová, Markéta; Topinková, Renáta** (2021). "Partner preference and age: user's mating behavior in online dating". *Journal of family research*, v. 33, n. 3, pp. 566-591.
<https://doi.org/10.20377/jfr-540>
- Sharabi, Liesel L.; Dykstra-DeVette, Tiffany A.** (2019). "From first email to first date: strategies for initiating relationships in online dating". *Journal of social and personal relationships*, v. 36, n. 11-12, pp. 3389-3407.
<https://doi.org/10.1177/0265407518822780>
- Smith, Aaron** (2016). "15% of American adults have used online dating sites or mobile dating apps". *Pew Research Center*, February 11.
<http://www.pewinternet.org/2016/02/11/15-percent-of-american-adults-have-used-online-dating-sites-or-mobile-dating-apps>
- Smith, Aaron; Anderson, Monica** (2016). "5 facts about online dating". *Pew Research Center*, February 29.
- Sprecher, Susan; Econie, Alexis; Treger, Stanislav** (2019). "Mate preferences in emerging adulthood and beyond: age variations in mate preferences and beliefs about change in mate preferences". *Journal of social and personal relationships*, v. 36, n. 10, pp. 3139-3158.
<https://doi.org/10.1177/0265407518816880>
- Sprecher, Susan; Sullivan, Quintin; Hatfield, Elaine** (1994). "Mate selection preferences: gender differences examined in a national sample". *Journal of personality and social psychology*, v. 66, n. 6, pp. 1074-1080.
<https://doi.org/10.1037/0022-3514.66.6.1074>

Statista Research Department (2022). "Distribution of *Tinder* monthly active users in the United States as of March 2021, by gender". Statista.

<https://www.statista.com/statistics/975925/us-tinder-user-ratio-gender>

Stephure, Robert J.; Boon, Susan D.; MacKinnon, Stacey L.; Deveau, Vicki L. (2009). "Internet initiated relationships: associations between age and involvement in online dating". *Journal of computer-mediated communication*, v. 14, n. 3, pp. 658-681.

<https://doi.org/10.1111/j.1083-6101.2009.01457.x>

Strugo, Julieta; Muise, Amy (2019). "Swiping for the right reasons: approach and avoidance goals are associated with actual and perceived dating success on *Tinder*". *The Canadian journal of human sexuality*, v. 28, n. 2, pp. 93-104.

<https://doi.org/10.3138/cjhs.2019-0010>

Sumter, Sindy R.; Vandenbosch, Laura; Ligtenberg, Loes (2017). "Love me *Tinder*: untangling emerging adults' motivations for using the dating application *Tinder*". *Telematics and informatics*, v. 34, n. 1, pp. 67-78.

<https://doi.org/10.1016/j.tele.2016.04.009>

Timmermans, Elisabeth; Courtois, Cédric (2018). "From swiping to casual sex and/or committed relationships: exploring the experiences of *Tinder* users". *The information society*, v. 34, n. 2, pp. 59-70.

<https://doi.org/10.1080/01972243.2017.1414093>

Timmermans, Elisabeth; De-Caluwé, Elien (2017). "Development and validation of the *Tinder Motives Scale* (TMS)". *Computers in human behavior*, v. 70, pp. 341-350.

<https://doi.org/10.1016/j.chb.2017.01.028>

Tinder (2016). *Tinder's most right-swiped jobs*.

<https://www.tinderpressroom.com/tinders-most-right-swiped-jobs>

Tinder (2018). *Tinder's most right-swiped jobs*.

<https://blog.gotinder.com/post/139924789366/tinders-most-right-swiped-jobs>

[this link no longer works]

Tyson, Gareth; Perta, Vasile C.; Haddadi, Hamed; Seto, Michael C. (2016). "A first look at user activity on *tinder*". In: 2016 IEEE/ACM international conference on advances in social networks analysis and mining (Asonam). IEEE, pp. 461-466.

<https://dl.acm.org/doi/abs/10.5555/3192424.3192510>

Valkenburg, Patti M.; Peter, Jochen (2007). "Who visits online dating sites? Exploring some characteristics of online daters". *Cyber psychology & behavior*, v. 10, n. 6, pp. 849-852.

<https://doi.org/10.1089/cpb.2007.9941>

Van-Hooff, Jenny (2020). "Swipe right? *Tinder*, commitment and the commercialisation of intimate life". In: Carter, Julia; Arocha, Lorena. *Romantic relationships in a time of "cold intimacies"*. Cham: Palgrave Macmillan, pp. 109-127. ISBN: 978 3 030 29256 0

https://doi.org/10.1007/978-3-030-29256-0_6

Ward, Janelle (2017). "What are you doing on *Tinder*? Impression management on a matchmaking mobile app". *Information, communication & society*, v. 20, n. 11, pp. 1644-1659.

<https://doi.org/10.1080/1369118X.2016.1252412>

White, Robert E.; Cooper, Karyn (2022). "Grounded theory". In: Cooper, Karyn; White, Robert E. *Qualitative research in the: post-modern era. Context of qualitative research*. Cham: Springer. ISBN: 978 94 007 2338 2

https://doi.org/10.1007/978-3-030-85124-8_9

Zytka, Douglas; Grandhi, Sukeshini A.; Jones, Quentin (2014). "Impression management struggles in online dating". In: *Proceedings of the 18th international conference on supporting group work*. ACM, pp. 53-62.

<https://doi.org/10.1145/2660398.2660410>

8. Annex. Data-coding process

Categories	Initial codes	Emerging codes	Focused codes
Motivations for using <i>Tinder</i>	<ul style="list-style-type: none"> -Love -Casual sex -Ease of communication -Self-worth validation or self-esteem -Thrill of excitement -Trendiness -Friendship -Long-term relationships -Broadening their options -Shrinking social networks 	<ul style="list-style-type: none"> -Sentimental breakup -To meet people -Socializing -Entertainment -Long-lasting sexual relationship 	<ul style="list-style-type: none"> -Casual sex -Self-worth validation or self-esteem -Long-term relationships -Shrinking social networks -Sentimental breakup -To meet people -Socializing -Entertainment -Long-lasting sexual relationship
Selection criteria	<ul style="list-style-type: none"> -mate selectivity -factors associated with maintaining relationships (resources, success, a pleasant personality) -physical appearance -homogamy/similarity -age -profession 	<ul style="list-style-type: none"> -Factors associated with maintaining relationships (Level of education, profession, interests, social status, intelligence) -Ethnicity 	<ul style="list-style-type: none"> -Mate selectivity -Factors associated with maintaining relationships (resources, success, a pleasant personality, level of education, profession, hobbies, interests, social status, intelligence) -Physical appearance -Homogamy/similarity -Age -Ethnicity
Communication management	<ul style="list-style-type: none"> -Who initiates communication -Strategies to initiate communication (original) -Thematic agenda (sex) -Outcome of the date (sex) 	<ul style="list-style-type: none"> -Strategies to initiate communication (formal, personalized) -Transition to other media: who -Transition to other media: types (Tinder chat, <i>WhatsApp</i>, <i>Instagram</i>) -Transition to other media: when -Thematic agenda: location, work, hobbies, relationship status (including children and their ages), studies -Who takes the initiative to propose a date -When a date is proposed -Outcome of the date (getting to know each other, friendship) 	<ul style="list-style-type: none"> -Who initiates communication -Strategies to initiate communication: formal -Strategies to initiate communication: personalized -Strategies to initiate communication: original -Transition to other stages or media: who -Transition to other media: only Tinder chat -Transition to other media: <i>WhatsApp</i> -Transition to other media: <i>Instagram</i> -Transition to other media: soon -Transition to other media: later stages -Thematic agenda: location -Thematic agenda: work -Thematic agenda: hobbies -Thematic agenda: sex -Thematic agenda: relationship status (including children and their ages) -Thematic agenda: studies -Who takes the initiative to propose a date -When a date is proposed: soon -When a date is proposed: later stages -Outcome of the date: sex -Outcome of the date: getting to know each other -Outcome of the date: friendship

Profesional de la



<http://www.profesionaldelainformacion.com>

Bienvenido a EPI
Revista científica internacional

e-ISSN: 1699-2407
<https://doi.org/10.3145/EPI>

Revista internacional de **Información y Comunicación**
indexada por WoS Social Sciences Citation Index (Q2), Scopus (Q1) y otras bases de datos

Factor de impacto JCR:
JIF 2021=3,596

Scopus/SCImago Journal Rank:
SJR 2021=0,831

Spanish research on Communication in WoS: thematic, methodological, and intellectual comparison between SSCI and ESCI

Francisco Segado-Boj; Valeriano Piñeiro-Naval; Tamara Antona-Jimeno

Nota: Este artículo se puede leer en español en:
<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87184>

Recommended citation:

Segado-Boj, Francisco; Piñeiro-Naval, Valeriano; Antona-Jimeno, Tamara (2023). "Spanish research on Communication in WoS: thematic, methodological, and intellectual comparison between SSCI and ESCI". *Profesional de la información*, v. 32, n. 3, e320309.
<https://doi.org/10.3145/epi.2023.may.09>

Article received on November 11th 2022
Approved on February 2nd 2023



Francisco Segado-Boj ✉
<https://orcid.org/0000-0001-7750-3755>
Universidad Complutense de Madrid
Facultad de Ciencias de la Información
Avda. Complutense, 3
28040 Madrid, Spain
fsegado@ucm.es



Valeriano Piñeiro-Naval
<https://orcid.org/0000-0001-9521-3364>
Universidad de Salamanca
Facultad de Ciencias Sociales
Campus Miguel de Unamuno, Edificio FES
Av. Francisco Tomás y Valiente, s/n
37007 Salamanca, Spain
vale.naval@usal.es



Tamara Antona-Jimeno
<https://orcid.org/0000-0002-8941-1708>
Universidad Complutense de Madrid
Facultad de Ciencias de la Información
Avda. Complutense, 3
28040 Madrid, Spain
tamaanto@ucm.es

Abstract

This study analyses the thematic and conceptual structure of the Spanish scientific production published in *Emerging Sources Citation Index (Web of Science)* journals. In this way, the aim is to identify the closest links between concepts and terms based on the co-occurrence of keywords used by the authors of the papers analysed, and also to point out the theoretical foundations that exist in the discipline through the co-citation relationships of articles in the bibliography of the documents in the sample. Finally, these results were compared with those obtained from the analysis of Spanish scientific production in *Social Sciences Citation Index (SSCI)*. To achieve these objectives, a network analysis of the co-occurrence of keywords and co-citation of references in articles published in Communication journals between 2015 and 2021 in *ESCI* (N = 3,559) and *SSCI* (N = 1,738) with at least one author linked to a Spanish institution was carried out. The results point to similar structural cohesion values and to a thematic and methodological similarity between both sets observed. There is a marked tendency towards quantitative studies on new technologies. While in *SSCI* there is an almost absolute dominance of Journalism studies, in *ESCI* there is a greater diversity of other disciplines such as Audiovisual Communication or Advertising. However, the intellectual structure of the production in *SSCI* reflects a more specialised character than in *ESCI*.

Keywords

Articles; Bibliometrics; Co-citation; Co-words; Communication; Topics; Methods; *Emerging Sources Citation Index*; *ESCI*; *Social Sciences Citation Index*; *SSCI*; *Web of Science*; *WoS*; Meta-research; Social network analysis; Spain.



Acknowledgments

We express our gratitude to the anonymous reviewers for their comments and suggestions that significantly improved the quality of this manuscript.

1. Introduction

According to **Ioannidis** (2018), meta-research is the study of research itself, encompassing its methods, reports, reproducibility, evaluation, and incentives. **Saperas** and **Carrasco-Campos** (2019) define this discipline as

“a descriptive quantitative method linked to content analysis, specifically designed to analyze how the format of a scientific article is organized as a means of communication and dissemination among specialised audiences’ (p. 222).

Its purpose, therefore, is

“to provide a general overview of the features characterizing research in a given field’ (**Igartua et al.**, 2022, p. 59).

In recent years, a body of meta-research focused on Communication Studies has emerged and been consolidated in Spain (**Caffarel**, 2018; **Carrasco-Campos et al.**, 2018; **Goyanes**, 2020; **Lozano-Ascencio et al.**, 2020; **Martínez-Nicolás et al.**, 2019). This research has focused on various features, including predominant theories (**Carrasco-Campos**; **Saperas**, 2022; **Piñeiro-Naval**; **Morais**, 2019), commonly used methods (**Gómez-Escalonilla**, 2021; **Martínez-Nicolás**; **Saperas**, 2011), analysed topics (**Segado-Boj et al.**, 2022a; **Trillo-Domínguez**; **De-Moya-Anegón**, 2022), and funding sources (**Gai-tán-Moya et al.**, 2021; **Piñeiro-Naval et al.**, 2021), among others.

Aligned with this perspective, bibliometric analysis enables, by drawing on a vast range of documents, the examination of various dimensions of a discipline, such as authorship (**Cuocolo et al.**, 2020), productivity (**Carrillo-Vera et al.**, 2018), performance –measured by impact or citation counts– (**Repiso**; **Moreno-Delgado**, 2022), social structure (**Eckl et al.**, 2019), citation structures (**Rauchfleisch**; **Schäfer**, 2018), or preferred topics (**Montero-Díaz et al.**, 2018), among others. This study aims to explore the intellectual and thematic structure of the Spanish scientific output on Communication Studies between 2015 and 2021. To accomplish this goal, the study compares articles of journals indexed in the database *Social Sciences Citation Index (SSCI)* and those featured in *Emerging Sources Citation Index (ESCI)*, both of which are included in the *Web of Science (WoS) Core Collection*.

Clarivate created *ESCI* in 2015 to expand its collection of documents in languages other than English and to compete with the *Scopus* database. It includes journals that are not indexed in the *SSCI* but are considered potential candidates for this index (**De-Filippo**; **Gorraiz**, 2020). Compared with *SSCI*, *ESCI* increases the visibility of regions such as the ‘Global South’, which includes Latin America. Moreover, *ESCI* is more inclusive and provides coverage for fields with fewer documents indexed in *SSCI*, as seen in Social Sciences (**Huang et al.**, 2017). In the 2021 edition of *SSCI*, the ‘Communication’ category features 95 journals, of which only two (*Comunicar* and *Profesional de la información*) are Spanish. In comparison, *ESCI* features 122 Communication journals in 2021, including 24 Spanish and 12 Latin American journals.

In this sense, Spain is one of the countries whose *WoS* coverage has increased the most since *ESCI* was created (**Ruiz-Pérez**; **Jiménez-Contreras**, 2019). 67% of Spain’s production in the ‘Communication’ category of the *WoS* is included in the *ESCI* database (**Repiso**; **Moreno-Delgado**, 2022). *ESCI*’s inclusiveness has enabled it to better represent the Iberoamerican communication studies publishing ecosystem, which is characterized by a majority of open-access journals, published by non-commercial editors, and by a lesser dependence on big publishing houses. Also, the articles published by Iberoamerican Communication journals show publication and citation features distinct from those featured in other international Communication journals (**Demeter et al.**, 2022).

ESCI includes a broader range of documents and reflects diverse streams of research, including particular or alternative perspectives, whereas *SSCI* represents the international standards of Communication research –in other words, the ‘dominant perspective’–. While *SSCI* indexes highly recognized and visible journals, *ESCI* serves as a pre-selection for *SSCI*, indexing journals with lower or insufficient impact by *SSCI* standards. These differences in bibliometric measurements can affect the acceptance rate of articles (**Sugimoto et al.**, 2013). Moreover, given that the editorial boards of journals with high impact may differ in composition from those in lower quartiles (**Dhanani**; **Jones**, 2017; **Goyanes**, 2019), there may be disparities in the editorial and scientific policies among the journals included in the databases (*SSCI* or *ESCI*).

Previous comparisons between *ESCI* and *SSCI* in other disciplines have revealed differences in authorship and other characteristics. For example, in Statistics (**Butt**, 2021), *ESCI* contains a higher proportion of publications from the Global South. In Social Work, differences are not only observed in the language of articles –with a greater percentage of Spanish documents in *ESCI*– but also in the topics covered, as certain issues, such as social representation, are more prevalent in *ESCI* than in *SSCI* (cfr. **Chaves-Montero**; **Vázquez-Aguado**, 2021; **Martínez et al.**, 2015). In Communication studies, **Miquel-Segarra** (2021) notes that *ESCI* publications exhibit a lower level of co-authorship –which is relatively more frequent in *SSCI* (**Castillo-Esparcia et al.**, 2012)– as well as a higher presence of Spanish-language documents. However, according

to this work, the methods used in ESCI publications are of lower quality, contradicting the general trend observed in other studies (Goyanes *et al.*, 2018; Martínez-Nicolás *et al.*, 2019), which suggest that Spanish Communication research is improving the rigor of its methods.

This paper aims to compare articles published in SSCI and ESCI, with the following specific objectives (SOs):

SO1: To analyze the co-word and co-citation networks in the Spanish scientific publications on Communication in ESCI and SSCI and characterize their structures.

SO2: To identify the primary thematic and conceptual relationships in the Spanish scientific publications on Communication in ESCI and SSCI.

SO3: To recognize the main intellectual and theoretical connections in the Spanish scientific publications on Communication in ESCI and SSCI.

In the following section, we provide a review of previous studies that have mapped scientific production in Communication studies, encompassing not only Spanish publications, but also those at the international level.

2. Previous research

Chronologically, the first study to be mentioned is Barnett *et al.* (2011) research on citation patterns in 45 SSCI-indexed journals from 1998 to 2007. Their findings indicate that Communication category journals which are indexed solely in this category are more central in the citation network compared to those attributed simultaneously to other social sciences. Additionally, their clustering analysis reveals the presence of three distinct communities: a psychological micro-cluster, a socio-political macro-cluster, and a minor cluster focused on gender studies. On a separate note, Chung *et al.* (2013) conducted a study on the most applied theories between 2000 and 2009 in four flagship journals in the field, namely *Communication Monographs*, *Communication Research*, *Human Communication Research*, and *Journal of Communication*. They identified 89 conceptual paradigms, with framing, priming, cultivation theory, and agenda-setting being the most prominent.

Taking a different perspective, Günther and Domahidi (2017) analysed a sample of 15,172 manuscripts published in 19 journals from WoS and Scopus over 80 years (from 1935 to 2014). Using topic modeling, they identified 145 different topics and observed the longitudinal evolution of the 15 most prominent topics. They found that in the most recent period (2005-2014), topics such as digital media, health, and media effects emerged as the most popular. Similarly, Montero-Díaz *et al.* (2018) mapped 33,627 articles indexed in SSCI journals from 1980 to 2013 and found that in the closest period (2010-2013), the most frequent topics were news, internet, gender, and advertising. Likewise, Vizoso *et al.* (2019) observed, after analyzing 5,291 texts published in SSCI and Scopus between 2008 and 2018, that content analysis and discourse analysis were the most commonly used methods, that media uses, effects, and reception were the most researched issues, and that Information and Communication Technologies (ICTs) were the most common objects of study.

In another work, Chan and Grill (2020) used topic modeling to analyze 12,990 articles published in 18 international journals from 2000 to 2017. The authors highlighted the relevance of two topics: social media and persuasion. Meanwhile, Demeter and Goyanes (2020) compared the output of two years, 1997 and 2017, through the analysis of a probabilistic sample of articles indexed in SSCI (263 and 283, respectively, a total of 526). Their findings were consistent with those of Vizoso *et al.* (2019), as they noted the current relevance of quantitative methods (content analysis, survey, and experiment), while in 1997, the methods were mostly qualitative (discourse analysis and participant observation).

In another direction, recent analyses of the Spanish scientific output on communication have utilized social network analysis of SSCI-indexed documents Aguado-Guadalupe *et al.*, 2022; Segado-Boj *et al.*, 2021a; 2021b), while other studies have conducted thematic and intellectual analysis using Scopus as their source (Costa-Sánchez, 2017; Heras-Pedrosa *et al.*, 2018; Segado-Boj *et al.*, 2022; Trillo-Domínguez; De-Moya-Anegón, 2022). However, although bibliometric analysis using ESCI has been employed in fields such as Statistics and Probability (Butt *et al.*, 2021) or Education (Repiso *et al.*, 2017), its use in Communication has been limited to Spanish journals within that database (Miquel-Segarra, 2021). This sampling approach excludes works published in international journals by authors affiliated with Spanish institutions. As a significant contribution, this article expands its scope to include this broader universe of study.

3. Methods

Our sampling strategy involved selecting articles from WoS, included in the 'Communication' category, published between 2015 and 2021 (both included) and authored by at least one researcher affiliated with a Spanish institution. We further distinguished between articles published in journals indexed in ESCI (N = 3,559) and journals indexed in SSCI (N = 1,738). This information was downloaded on March 1st, 2022. We converted the bibliographic data to network data using VOSviewer (Van-Eck; Waltman, 2010) and subsequently analysed and visualized it using Pajek (Batagelj; Mrvar, 1998).

Network analysis examines the relationships (known as 'edges' or 'links') between elements (known as 'nodes' or 'vertices'). In the co-word network, the vertices represent the author keywords used to identify the articles (we excluded those keywords generated automatically by WoS). When two keywords appear together in the same article, they are

linked by an edge. Similarly, in the co-citation network, the nodes represent the references cited by an article and are connected by an edge when they appear in the same reference list.

We manually harmonized the keywords and references to homogenize synonyms (e.g., 'youth' and 'young people') and variations in the citation of the same piece of work, such as translations and different editions of the same book. After the harmonization process, the final database contained 12,289 keywords and 156,826 references, compared to the original database of 12,678 keywords and 157,372 references.

We generated separate co-word and co-citation networks for each database (*WoS* and *ESCI*) and calculated the main structural indices of each network, including:

- *Number of nodes*: the total number of vertices in the network.
- *Components*: a component is a group of nodes that are connected to each other but not connected to other groups. The higher the number of components, the more fragmented the network is.
- *Size of the main component*: the percentage of nodes included in the largest component.
- *Density*: measures the level of cohesion among the nodes in the network. A value of "1" would mean that all vertices are directly connected to each other, while a value of "0" would indicate that no nodes are connected to each other, resulting in no arcs in the network.
- *Centrality values*: degree centrality is defined as the number of connections between a given node and other vertices in the network. The nodes with higher centrality degree are considered more relevant and better connected. We provide a ranking of nodes with the highest centrality degree in each network (Tables 2-4), as well as the average centrality degree of the entire network (Table 1). Table 1 also includes the Centrality value of the network, which can be interpreted as the extent to which each network is organized around specific nodes. A centrality value of "1" would mean that the entire network revolves around just one node, and the rest of the vertices show no edges among them. On the contrary, a value of "0" would mean that all nodes in the network are equally connected to each other. We also provide the Betweenness value, which indicates the existence of nodes that, if excluded from the network, would result in fragmentation into detached components. A value of "1" would mean that the components in the network are linked by only one vertex, while a value of "0" would mean that the disappearance of any node would not cause fragmentation at all.
- *Clustering coefficients*: measure the likelihood that two connected nodes share common connections with other vertices. In other words, it examines whether nodes form communities or groups rather than isolated links. This clustering can be calculated using two measures: the Watts-Strogatz coefficient, which is based on low centrality nodes and measures the presence of internal connections within specialised communities, and transitivity, which is based on high centrality nodes and measures the interconnections between different groups or communities.
- *Distance*: the average distance calculates the number of nodes that separate two given vertices. The *longest distance* identifies the pair of nodes that are the farthest apart from each other.

In terms of visualization, since the goal of this work is to identify the primary relationships within each network, we have applied a reduction criterion to each graph. Therefore, the networks only display the arcs (co-occurrences of keywords and references) that exceed a minimum threshold, which is specified as a note in each figure.

The color of each node in Graphs 1-4 represents a cluster identified by the Louvain algorithm (multi-level thickening, single refinement, resolution parameter = 1, number of random restarts = 1, maximum number of levels in each iteration = 20, maximum number of repetitions at each level = 50). These clusters represent groups of vertices that are more likely to be connected to each other compared to the other nodes in the network.

The size of each node in Graphs 1 and 2 is proportional to the frequency of mentions of each keyword, while in Graphs 3 and 4 it is proportional to the number of citations of each reference. The number displayed inside each node represents the sum of citations or mentions, depending on the context.

To simplify the presentation, Graphs 1 and 2 only include the author's last name and the publication year. The full referencing information can be found in Annex I.

4. Results

The co-words and co-citation networks in *ESCI* contain a higher number of nodes (keywords and references) than those in *SSCI*, as shown in Table 1. This can be explained by the higher number of documents found in *ESCI*. Overall, the networks generated from both databases exhibit similar structural values, with relatively low density, decentralization, and a low number of arcs. The networks are well-connected, as indicated by their average and longest distances, and

“ The core of the *ESCI* co-citation network is formed by theoretical references focused on the impact of technologies (Jenkins, 2008; Castells, 2009; Scolari, 2013) and handbooks of research methods (Krippendorff, 1980; Igartua, 2006; Berelson, 1952; Bardin, 1996) ”

the size of the main components in each case approaches 100%. However, the co-word network in *ESCI* exhibit more fragmentation in secondary components compared to *SSCI*, and the average centrality of the co-word network is higher in *SSCI* than in *ESCI*. The most central nodes tend to be more interconnected with each other (Watts-Strogatz coefficient), particularly in the co-word networks, compared to the relationships among peripheral nodes (transitivity).

4.1. Keywords

In both *ESCI* and *SSCI*, the co-keyword networks revolve around the broad concept of “social media”, which exhibits the highest centrality degree (Table 2). Terms associated with specific social media platforms, particularly *Twitter*, as well as *Facebook* and *YouTube*, are also among the top positions. Additionally, both networks demonstrate a preference for new technologies, as indicated by the prominence of keywords such as “internet”, “digital media”, and “Information and Communication Technologies (ICTs)”.

As we analyse Spanish scientific output, it is not surprising that “Spain” appears as the second most central keyword in the ranking. No other geographic scenario or location is mentioned among the most central keywords. Additionally, the impact of the Sars-CoV-2 pandemic is evident in both databases, as indicated by the relevance granted to keywords such as “COVID-19” or “pandemic”.

The ranking of centrality degree suggests that Spanish research on Communication in *SSCI* is dominated by concepts related to Journalism. The *ESCI* network features this trend as well, albeit with less intensity. In *ESCI*, other disciplines such as advertising or television hold more central positions than they do in *SSCI*. Furthermore, the thematic core of *ESCI* features other media, such as “cinema”, which is absent in *SSCI*. Additionally, *ESCI* places greater importance on terms related to Transmedia and framing, as well as education.

The co-word network in *ESCI* reveals six clusters (Graph 1). The one with the most nodes (represented in **green**) centres around the term “social media”, which is linked to other keywords associated with digital technologies (such as “internet”, “digital communication”, “information and communication technologies”, and “web 2.0”), specific platforms (“Instagram” and “Facebook”), aspects of audience profiling (“adolescence”), and forms of interaction and engagement (“engagement”).

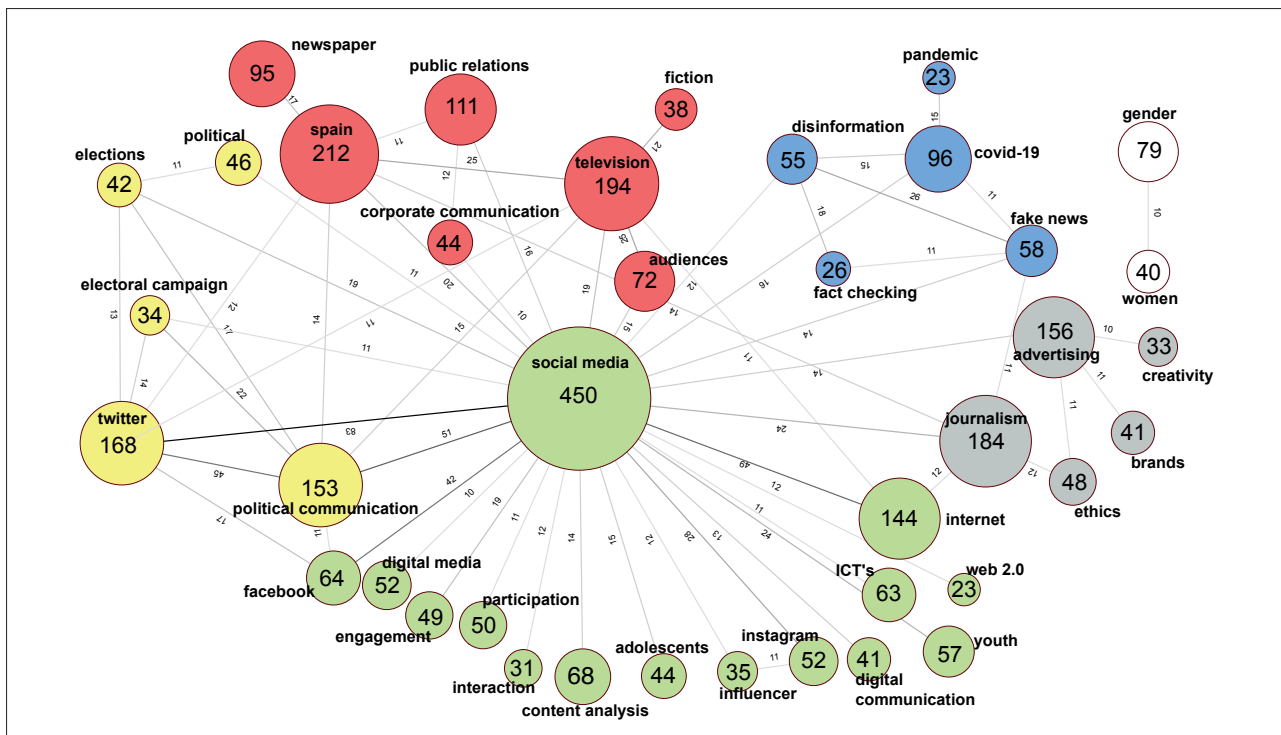
This **green** cluster is closely connected to the **yellow** one, which also features a central node related to social media (“Twitter”). This core is linked to concepts about political communication and electoral campaigns.

Table 1. Co-words and co-citation networks. Source: *Web of Science*.

	Co-words		Co-citation	
	SSCI	ESCI	SSCI	ESCI
Number of nodes	5,983	8,172	8,511	24,065
Density (no loops)	0.002	0.001	0.006	0.004
Average Degree Centrality	14.72	9.67	54.40	58.65
Number of components	49	86	3	6
Size of the main component (%)	95.85	94.80	99.87	99.95
Centrality	0.177	0.126	0.136	0.190
Betweenness	0.158	0.142	0.083	0.109
Watts-Strogatz	0.860	0.853	0.581	0.568
Transitivity	0.131	0.083	0.327	0.241
Average distance	3.208	3.431	3.029	2.898
Longest distance	8	8	7	7

Table 2. Keywords exhibiting the highest centrality degree (CD) in *SSCI* and *ESCI*. Source: *Web of Science*.

SSCI		ESCI	
Keyword	CD	Keyword	CD
social media	1,072	social media	1,039
spain	947	spain	694
journalism	582	journalism	550
twitter	469	television	549
covid-19	414	advertising	484
digital media	400	twitter	446
internet	393	internet	433
political communication	362	political communication	370
digital journalism	358	public relations	340
television	347	newspaper	311
newspaper	346	covid-19	306
audiences	338	cinema	288
pandemic	331	gender	271
content analysis	312	audiences	244
advertising	265	ICTs	242
facebook	245	content analysis	235
public relations	226	educommunication	233
university	221	frames	222
journalist	213	youtube	208
survey	213	digital journalism	192
gender	208	education	191
educommunication	206	transmedia	190
adolescents	205	university	189
university libraries	205	facebook	188
youtube	203	identities	187
		youth	187

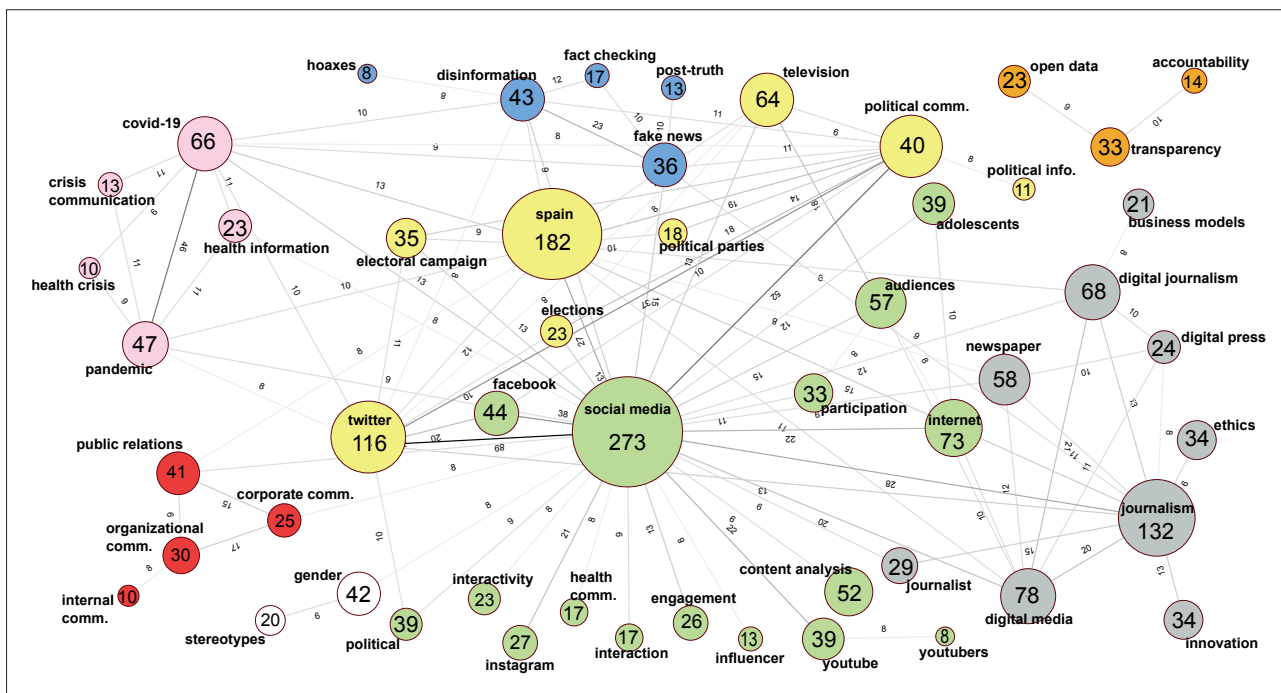


Graph 1. Co-word network of the Spanish scientific output on Communication in ESCI (2015-2021). Source: *Web of Science*. Note. Featuring links equal or above 10.

In more peripheral positions, the graph depicts communities with lower specialization and greater heterogeneity in their nodes. Thus, the grey cluster merges advertising keywords (such as “creativity” and “brands”) with others related to journalism, particularly “ethics”. Similarly, the red cluster combines keywords on public relations (such as “corporate communication”), television (“fiction” and “audiences”), and once again, journalism (“newspapers”). In another sense, a specialised core (deep blue) is centred around disinformation, which also encompasses Covid-19. Finally, a white cluster, disconnected from the rest of the network, is devoted to gender issues.

Comparing this co-word network with the one generated from the SSCI database (Graph 2) reveals several similarities.

As in the previous network, terms and concepts related to social media and digital technologies form a specialised cluster (green) that shares several features with the green cluster in Graph 1. Nevertheless, the green cluster in Gra-



Graph 2. Co-word network of the Spanish scientific output on Communication in SSCI (2015-2021). Source: *Web of Science*. Note. Featuring links equal or above 8.

ph 2 includes specific particularities such as being linked to health communication. A more specialised cluster on journalism (grey) features keywords related to 'digital technologies', 'innovation', 'business models', and 'ethics'. The yellow cluster contains concepts of political communication, such as 'electoral campaigns', 'elections', and 'political participation'.

Just like in ESCI, the SSCI network also shows a cluster (blue) that features keywords related to 'post-truth' and 'disinformation'. However, while in Graph 1 this disinformation cluster gathered concepts about Covid-19, in Graph 2 keywords about the pandemic are organized in a separate community that includes concepts about crisis communication or health information.

The cluster of Public Relations (represented in red) is situated on the periphery and is formed by keywords such as 'organizational communication', 'internal communication', and 'corporate communication'. Similarly to ESCI, SSCI has a cluster (white) dedicated to gender issues. However, in SSCI, this cluster is connected to the main network. Additionally, keywords related to transparency and open data are gathered in a minor cluster (orange).

4.2. References

The core of the ESCI co-citation network is formed by theoretical references focused on the impact of technologies (Jenkins, 2008; Castells, 2009; Scolari, 2013) and handbooks of research methods (Krippendorff, 1980; Igartua, 2006; Berelson, 1952; Bardin, 1996) (Table 3).

Table 3. References exhibiting the highest centrality degree (CD) in ESCI. Source: *Web of Science*.

Reference	CD
Jenkins, Henry (2003). <i>Convergence culture: where old and new media collide</i> . New York: New York University Press. ISBN: 978 0 8147 4368 3	2737
Krippendorff, Klaus (1980). <i>Content analysis: an introduction to its methodology</i> . Beverly Hills: Sage. ISBN: 978 0 8039 1498 8	2195
Castells, Manuel (2009). <i>Communication power</i> . Oxford: Oxford University Press. ISBN: 978 0 19 157063 6	2133
Igartua, Juan-José (2006). <i>Métodos cuantitativos de investigación en comunicación</i> . Barcelona: Editorial Bosch. ISBN: 84 9790 271 8	1437
Chadwick, Andrew (2013). <i>The hybrid media system: politics and power</i> . Oxford: Oxford University Press. ISBN: 978 0 19 975947 7 https://doi.org/10.1093/acprof:oso/9780199759477.001.0001	1286
Scolari, Carlos (2013). <i>Narrativas transmedia: cuando todos los medios cuentan</i> . Barcelona: Deusto. ISBN: 978 84 234 1336 2	1268
Hallin, Daniel; Mancini, Paolo (2007). <i>Comparing media systems: three models of media and politics</i> . Cambridge: Cambridge University Press. ISBN: 978 0 521 54308 8	1249
Bardin, Laurence (1996). <i>El análisis de contenido</i> . Madrid: Akal Ediciones. ISBN: 978 84 7600 093 9	1135
Wimmer, Roger D.; Dominick, Joseph R. (1996). <i>La investigación científica de los medios de comunicación: una introducción a sus métodos</i> . Barcelona: Editorial Bosch. ISBN: 978 84 7676 359 9	1073
Toffler, Alvin (1980). <i>Future shock: the third wave</i> . New York: Bantam Books. ISBN: 0 553 24698 4	1071
Entman, Robert M. (1993). "Framing: toward clarification of a fractured paradigm". <i>Journal of communication</i> , v. 43, n. 4. https://doi.org/10.1111/j.1460-2466.1993.tb01304.x	1009
Berelson, Bernard (1952). <i>Content analysis in communication research</i> . New York: Hafner.	946
Castells, Manuel (2012). "Networks of outrage and hope - Social movements in the Internet age". <i>International journal of public opinion research</i> , v. 25, n. 3, pp. 398-402. https://doi.org/10.1093/ijpor/edt020	924
Campos-Domínguez, Eva (2017). "Twitter y la comunicación política". <i>Profesional de la información</i> , v. 26, n. 5. https://doi.org/10.3145/epi.2017.sep.01	843
López-García, Guillermo (2016). "Nuevos y viejos liderazgos: la campaña de las elecciones generales españolas de 2015 en Twitter". <i>Communication & society</i> , v. 29, n. 3. https://doi.org/10.15581/003.29.35829	779
Bauman, Zygmund (1999). <i>La modernidad líquida</i> . Madrid: Fondo de Cultura Económica de España. ISBN: 950 557 513 0	753
Castells, Manuel (1997). <i>La era de la información: economía, sociedad y cultura</i> . Madrid: Alianza. ISBN: 84 206 4246 0	740
Casero-Ripollés, Andreu; Feenstra, Ramón A.; Tormey, Simon (2016). "Old and new media logics in an electoral campaign: The case of Podemos and the two-way street mediatization of politics". <i>The international journal of press/politics</i> , v. 21, n. 3. https://doi.org/10.1177/1940161216645340	697

A similar pattern can be observed at the core of the SSCI co-citation network, as ten of the most central references in ESCI also appear in the centrality ranking in SSCI. A notable difference between both networks is that Casero-Ripollés (2016) jumps from the twentieth to the eighth position in the SSCI ranking (Table 4).

In both ESCI and SSCI, the co-keyword networks revolve around the broad concept of "social media", which exhibits the highest centrality degree

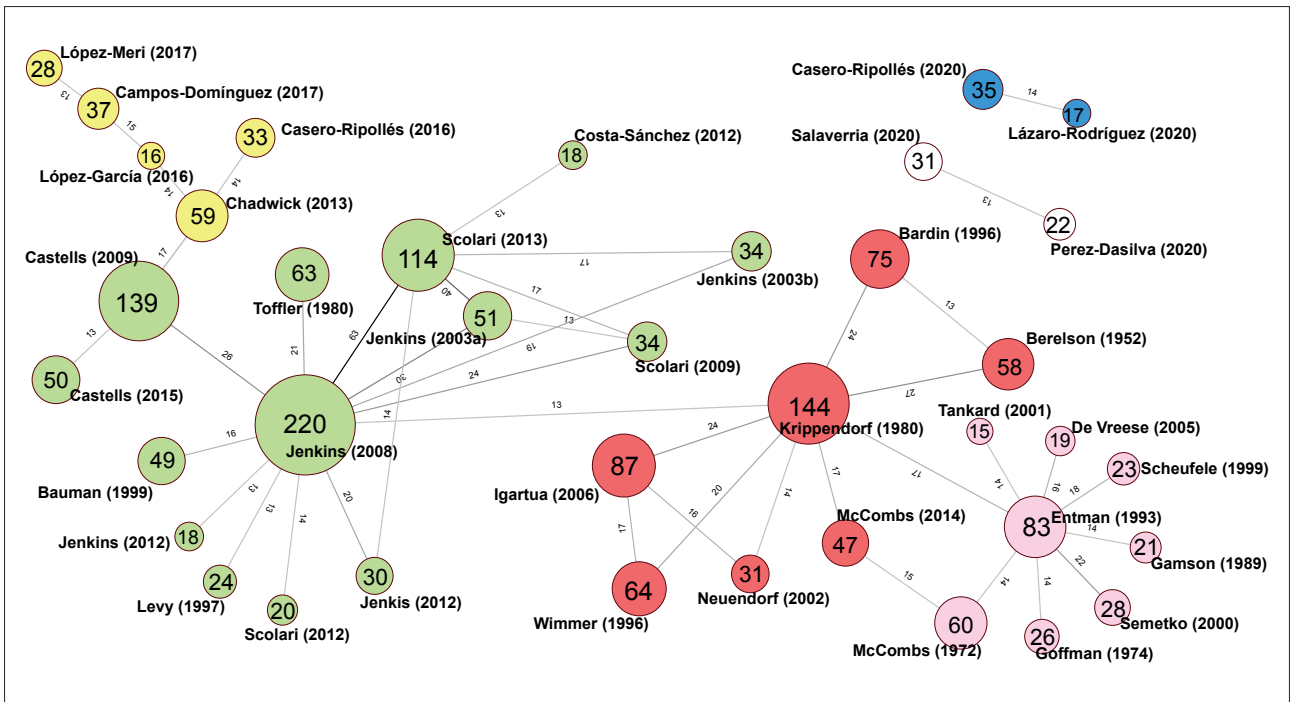
Table 4. References exhibiting the highest centrality degree (CD) in SSCI. Source: *Web of Science*.

Reference	CD
Jenkins, Henry (2003). <i>Convergence culture: where old and new media collide</i> . New York: New York University Press. ISBN: 978 0 8147 4368 3	1210
Hallin, Daniel; Mancini, Paolo (2007). <i>Comparing media systems: three models of media and politics</i> . Cambridge: Cambridge University Press. ISBN: 978 0 521 54308 8	1166
Chadwick, Andrew (2013). <i>The hybrid media system: politics and power</i> . Oxford: Oxford University Press. ISBN: 978 0 19 975947 7 https://doi.org/10.1093/acprof:oso/9780199759477.001.0001	1002
Castells, Manuel (2009). <i>Communication power</i> . Oxford: Oxford University Press. ISBN: 978 0 19 157063 6	847
Krippendorff, Klaus (1980). <i>Content analysis: an introduction to its methodology</i> . Beverly Hills: Sage. ISBN: 978 0 8039 1498 8	756
Krippendorff, Klaus (1980). <i>Content analysis: an introduction to its methodology</i> , Beverly Hills: Sage. ISBN: 978 0 8039 1498 8	552
Kovac, Bill; Rosenstiel, Tom (2001). <i>The elements of journalism: what newspeople should know and the public should expect</i> . New York: Crown Publishers. ISBN: 978 0 60960783 1	536
Pariser, Eli (2011). <i>The filter bubble: what the internet is hiding from you</i> . New York: Penguin. ISBN: 978 0 670 92038 9	529
Entman, Robert M. (1993). "Framing: toward clarification of a fractured paradigm". <i>Journal of communication</i> , v. 43, n. 4, pp. 51-58. https://doi.org/10.1111/j.1460-2466.1993.tb01304.x	527
Igartua, Juan-José (2006). <i>Métodos cuantitativos de investigación en comunicación</i> . Barcelona: Editorial Bosch. ISBN: 84 9790 271 8	502
Gil de Zúñiga, Homero; Weeks, Brian; Ardèvol-Abreu, Alberto (2017). "Effects of the news-finds-me perception in communication: social media use implications for news seeking and learning about politics". <i>Journal of computer-mediated communication</i> , v. 22, n. 3, pp. 105-123. https://doi.org/10.1111/jcc4.12185	466
McCombs, Maxwell E.; Shaw, Donald L. (1972). "The agenda setting function of mass media". <i>Public opinion quarterly</i> , v. 36, n. 2, pp. 176-187. https://doi.org/10.1086/267990	457
Enli, Gunn-Sara; Skogerbø, Eli (2013). "Personalized campaigns in party-centred politics information". <i>Communication & society</i> , v. 16, n. 5, pp. 757-74. https://doi.org/10.1080/1369118X.2013.782330	423
López-Meri, Amparo (2017). "Contribución ciudadana al debate electoral y su cobertura periodística en Twitter". <i>Prisma social: revista de investigación social</i> , n. 18. https://revistaprismasocial.es/article/view/1466	416
Jenkins, Henry; Ford, Sam; Green, Joshua (2013). <i>Spreadable media: creating value and meaning in a networked culture</i> . New York: NYU Press. ISBN: 978 0 8147 4350 8	409
Castells, Manuel (2012). "Networks of outrage and hope - Social movements in the Internet age". <i>International journal of public opinion research</i> , v. 25, n. 3, pp. 398-402. https://doi.org/10.1093/ijpor/edt020	409
Fletcher, Richard; Nielsen, Rasmus-Kleis (2018). "Are people incidentally exposed to news on social media? A comparative analysis". <i>New media & society</i> , v. 20, n. 7. https://doi.org/10.1177/1461444817724170	407
Allcott, Hunt; Gentzkow, Matthew (2017). "Social media and fake news in the 2016 election". <i>Journal of economic perspectives</i> , v. 31, n. 2, pp. 211-236. https://doi.org/10.1257/jep.31.2.211	399
Hermida, Alfred; Fletcher, Fred; Korell, Darryl; Logan, Donna (2012). "Share, like, recommend". <i>Journalism studies</i> , v. 13, n. 5-6, pp. 815-824. https://doi.org/10.1080/1461670X.2012.664430	397
Colleoni, Elanor; Rozza, Alessandro; Arvidsson, Adam (2014). "Echo chamber or public sphere? Predicting political orientation and measuring political homophily in Twitter using big data". <i>Journal of communication</i> , v. 64, n. 2, pp. 317-332. https://doi.org/10.1111/jcom.12084	388

The analysis of the most frequent relationships in *ESCI* (Graph 3) identifies a **green** cluster that focuses on the theoretical extensions of the concept of Transmedia (**Jenkins**, 2008; **Scolari**, 2009, 2013). This cluster is peripherally linked to Toffler through his contributions to the 'Third Wave' and the idea of 'prosumer' (**Toffler**, 1981), which are placed within a broader context of technological changes. This community also includes references about the social changes brought about by digital technologies (**Castells**, 2009). This cluster is connected to another one (**yellow**) that focuses on the impact of technologies in the structures of media systems (**Chadwick**, 2017) or specific areas such as political communication (**Campos-Domínguez**, 2017).

Another community (**red**) features references about content analysis and is linked to another group (**pink**) about media theories. Within this cluster, framing references (**Entman**, 1993; **Semetko**; **Valkenburg**, 2000; **Tankard**, 2001) are more common and placed at the centre, while agenda-setting (**McCombs**; **Shaw**, 1972) is less frequent and located in the periphery.

“ The analysis of the co-word networks provides additional evidence supporting the aforementioned focus of Spanish Communication research on Mass Communication ”



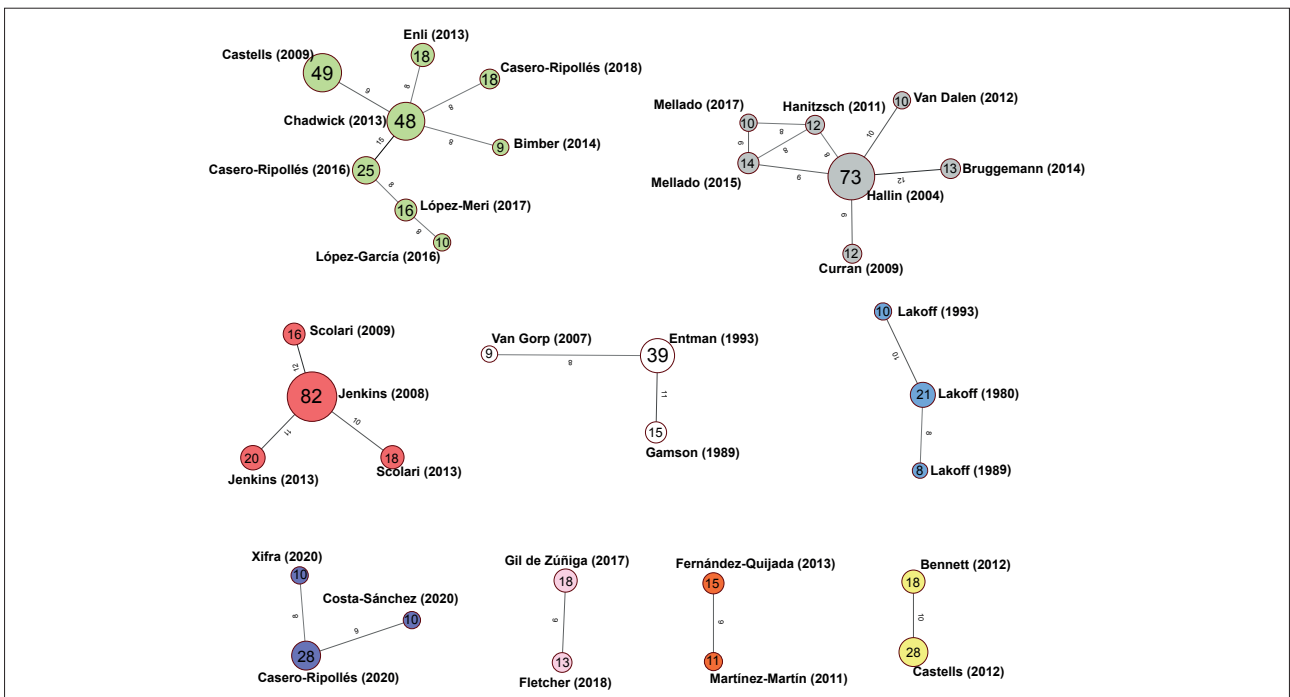
Graph 3. Co-citation network of the Spanish scientific output on Communication in ESCI (2015-2021). Source: Web of Science. Note. Featuring links equal or above 13.

Last, two different dyads appear detached from the rest of the communities. One of them (blue) focuses on Covid-19, and the other (white) revolves around fake news.

The theoretical structure of Spanish scientific production in SSCI is more fragmented, with clusters of references that are more disconnected from each other (Graph 4).

Similar to ESCI, the co-citation network in SSCI contains multiple clusters of references centered on the impact of new technologies. In this case, the Transmedia cluster (red) is disconnected from the community (green) that focuses on the socio-political effects of digital media. Another group of references (gray) is dedicated to journalistic cultures (Hanitzsch et al., 2011; Mellado et al., 2017).

The blue cluster comprises references from the field of Linguistics tied by George Lakoff’s (Lakoff, 1993) concept of ‘metaphor’. Another group of citations (white) centers on framing (Entman, 1993). A triad (deep blue) includes references



Graph 4. Co-citation network of the Spanish scientific output on Communication in SSCI (2015-2021). Source: Web of Science. Note. Featuring links equal or above 8.

that examine the impact of Covid-19 on various phenomena such as corporate communication (Xifra, 2020) or news consumption (Casero-Ripollés, 2020).

Finally, several dyads focus on specific issues such as incidental news consumption (Fletcher; Nielsen, 2018; Gil-de-Zúñiga *et al.*, 2017), digital activism (pink cluster), or meta-research on Communication (Fernández-Quijada; Masip, 2013) and Education (Martínez-Martín *et al.*, 2011).

Our data remarks that Spanish researchers in both *ESCI* and *SSCI* adopt quantitative paradigms, with qualitative analysis remaining a marginal aspect of the field

5. Discussion and conclusions

Our findings highlight certain trends in Spanish scientific publications on Communication, published in *ESCI*, that distinguish them from the characteristics observed in *SSCI* articles.

Spanish Research on Communication exhibits similar structural patterns in both *SSCI* and *ESCI*. The discipline appears to be relatively homogeneous and not highly fragmented, with main components of large size and reduced distances between them. There are no heavily dominant centers (as indicated by low centrality values), and the most central nodes tend to be connected to each other (as indicated by a relatively high Watts-Strogatz value). However, there is lower cohesion among the peripheral nodes (as indicated by a low Transitivity value).

These results are consistent with previous structural analyses of Spanish and Latin American research on Communication published in *Scopus* (Segado-Boj *et al.*, 2021a). The evidence suggests that structural values in Spanish communication research are consistent across various databases, including *Scopus*, *SSCI*, and *ESCI*. However, differences arise when examining the specific placement of individual elements within this overall structure.

Thus, the analysis of the co-citation networks reveals that Spanish *SSCI* research on Communication tends to focus on more specific consequences of new technologies, rather than solely on generic issues. Conversely, research on Communication in *ESCI* tends to analyze more generic topics than in *SSCI*. The theoretical communities in *SSCI* concentrate on specific features, such as incidental news consumption or activism, while in *ESCI*, clusters are organized mainly around the main theories in the discipline, such as framing or agenda setting. In addition, within *ESCI*, the cluster of references related to new technologies primarily comprises works on transmedia narratives, with the socio-political impact of these technologies being a peripheral part of the group. In contrast, in *SSCI*, these two conceptual domains remain separate and disconnected from each other.

While there may be some differences in how issues or concepts are linked, the thematic structures (co-word networks) in *SSCI* and *ESCI* are mostly similar and share many of the same keywords. It is worth noting that the Spanish academic community in Communication has not been immune to the sudden surge in the number of articles related to the health crisis caused by the pandemic. This trend has also been observed in other fields, representing an unprecedented increase in attention paid to a phenomenon in the entire history of science (Odone *et al.*, 2020). This trend has been identified in both *SSCI* and *ESCI*.

The analysis of the co-word networks provides additional evidence supporting the aforementioned focus of Spanish Communication research on Mass Communication (Caffarel-Serra *et al.*, 2017; Martínez-Nicolás *et al.*, 2019). Other areas of Communication research, such as interpersonal communication, are absent from both *ESCI* and *SSCI*.

Disaggregated analysis shows that the thematic core of *ESCI*'s production includes terms such as cinema or advertising, which are absent from *SSCI*. These terms did not appear in previous analyses of Spanish Communication research indexed in *Scopus* (Segado-Boj *et al.*, 2022a). This fact suggests that future meta-studies on Spanish Communication aimed at exploring fields beyond Journalism should utilize *ESCI* as a source, or else a significant portion of the scientific production on such topics will be overlooked.

Furthermore, the analysis of the most central keywords (Table 2) highlights the importance given to new technologies, which is consistent with the findings of Montero *et al.* (2018) at the international level and with Fernández-Fernández *et al.* (2020) at the national level in Spain. The co-word cluster with the highest number of nodes in *SSCI* (Graph 1) consists of socio-political concepts (green cluster), which is consistent with international research on Communication (Barnett *et al.*, 2011). However, in *ESCI* (Graph 1), such terms form smaller, more peripheral clusters (yellow and pink). This suggests that *ESCI* articles cover a wider range of themes compared to *SSCI* documents.

Another noteworthy finding is the isolated and tangential position of gender studies in both *ESCI* and *SSCI*. Gender-related keywords are detached from other specialised communities, and we did not find any specific clusters of references dedicated to gender studies in the co-citation network. This suggests a certain theoretical dispersion, which is consistent with previous evidence of its relatively weak presence in Spanish Communication

Future meta-studies on Spanish Communication aimed at exploring fields beyond Journalism should utilize *ESCI* as a source, or else a significant portion of the scientific production on such topics will be overlooked

studies (Zurbano-Berenguer *et al.*, 2018). Hence, it may be necessary to incorporate a gender perspective into studies on political communication or media effects, to name just two relevant examples.

On another note, while international research is primarily structured around framing and priming (Chung *et al.*, 2013), Spanish research centres on convergence culture (Jenkins, 2008). Specific works on framing are not prominent in the core of the co-citation networks (Tables 3-4). However, the analysis of the most frequent relationships does indicate the significance of the framing theoretical perspective in both ESCI (pink cluster, Graph 3) and SSCI (white cluster, Graph 4). This trend could be linked to the theoretical dispersion identified in articles published by Spanish scholarly journals (Martínez-Nicolás *et al.*, 2019; Piñeiro-Naval; Morais, 2019; Carrasco-Campos; Saperas, 2022) and in Scopus articles authored by Spanish researchers (Segado-Boj *et al.*, 2022b).

Similarly, the presence of content analysis as one of the most central keywords in both ESCI and SSCI (Table 2) reflects the importance of this method in Spanish Communication research (Gómez-Escalonilla, 2021; Piñeiro-Naval, 2020; Vizoso *et al.*, 2019). This relevance is particularly significant in ESCI, where a specific community of references (red cluster, Graph 3) revolves around this method. These findings emphasize the use of content analysis as one of the most common research tools in Communication (Berger, 2016). Additionally, our data remarks that Spanish researchers in both ESCI and SSCI adopt quantitative paradigms, with qualitative analysis remaining a marginal aspect of the field.

In contrast, our findings indicate that media effects, which constitute a fundamental aspect of international Communication research (Günther; Domahidi, 2017), have not been a specific focus of Spanish researchers, as evidenced by the absence of particular communities (Graphs 1 and 2) and keywords related to this area among the most central concepts (Table 2). Instead, the research conducted by Spanish scholars continues to be primarily focused on media messages and content analysis (Caffarel-Serra *et al.*, 2017; Martínez-Nicolás *et al.*, 2019), while other areas (audience research and media context) have yet to be fully explored.

Co-citation and co-word analysis have been extensively employed in prior research to investigate particular themes and subjects in the Communication field (cfr. McGowan *et al.*, 2022; Quevedo-Redondo *et al.*, 2022; Segado-Boj *et al.*, 2022a), showcasing its ability to automatically handle large data sets. Methodologically, we believe that the primary contribution of this study is emphasizing the importance of utilizing ESCI as a resource, and specifically, comparing it to SSCI. Thus, we suggest separating two spheres within Spanish communication research: one that aligns with and conforms to Anglo-Saxon models, and another that is more distinctive and particular, less integrated into these paradigms.

Several possible explanations (not necessarily mutually exclusive) could be proposed to explain the differences between ESCI and SSCI, such as those listed below:

- ESCI displays emerging patterns that have the potential to be consolidated in SSCI in the future.
- Difference in focus: The majority of Communication journals included in SSCI are international publications, with only two journals published in Spain. In contrast, ESCI includes a greater number of Spanish and Latin American journals. It could be argued that the perspective of research published in ESCI may be too local to garner interest from international journals.
- Quality differences cannot be disregarded. It is possible that the articles published in ESCI do not meet the standards necessary to surpass the high rejection rates of SSCI journals.

Limitations

Despite the valuable contributions of this study, it is important to acknowledge its limitations when interpreting the findings. Our analysis relied primarily on a bibliometric approach, which enabled us to automatically process large volumes of data, but it was limited to superficial document features such as author keywords and references cited in the sampled articles. Additionally, the automated analysis used to detect clusters generated heterogeneous communities. This finding should not be taken at face value as proof of the existence of a consolidated subgroup of topics or references, but rather as evidence of smaller communities that lack coherence to be identified as autonomous clusters by the algorithm. This happens, for example, in Graph 1, where the red and grey clusters represent the combination of issues from Journalism and Public Relations and from Journalism and Television, respectively.

Another noteworthy finding is the isolated and tangential position of gender studies in both ESCI and SSCI. Gender-related keywords are detached from other specialised communities, and we did not find any specific clusters of references dedicated to gender studies in the co-citation network

Possible explanations for the differences between ESCI and SSCI: 1) ESCI displays emerging patterns that have the potential to be consolidated in SSCI in the future. 2) The journals indexed in SSCI are more international than in ESCI, where they are more Spanish and Latin American. 3) Quality differences.

Prospective

As a prospect for further research, content analyses could complement our findings by manually analysing a significant random sample of articles. This approach would provide a deeper look at the frames, theories, methods, and other features of Spanish communication research. Similarly, we propose investigating whether the differences identified between *ESCI* and *SSCI* are also present in Latin American research or if these differences are unique to the Spanish field. Finally, based on the hypothesis that *ESCI* articles may be of lower quality compared to those indexed in *SSCI*, future studies could explore whether there is a significant difference (measured, for example, in citations or other altmetrics) between the works indexed in one database versus the other.

“The Spanish *SSCI* research on Communication tends to focus on more specific consequences of new technologies, rather than solely on generic issues. Conversely, research on Communication in *ESCI* tends to analyze more generic topics than in *SSCI*”

6. References

- Aguado-Guadalupe, Guadalupe; Herrero-Curiel, Eva; De-Oliveira-Lucas, Elaine-Rosangela** (2022). “Dinámicas de la producción científica española en las revistas de Comunicación en WoS”. *Revista española de documentación científica*, v. 45, n. 2, e326.
<https://doi.org/10.3989/redc.2022.2.1854>
- Barnett, George A.; Huh, Catherine; Kim, Youngju; Park, Han Woo** (2011). “Citations among communication journals and other disciplines: a network analysis”. *Scientometrics*, v. 88, n. 2.
<https://doi.org/10.1007/s11192-011-0381-2>
- Batagelj, Vladimir; Mrvar, Andrej** (2004). “Pajek - Analysis and visualization of large networks”. In: *Graph drawing software*, Berlin: Springer, pp. 77-103. ISBN: 978 3 642 18638 7
https://doi.org/10.1007/978-3-642-18638-7_4
- Berger, Arthur-Asa** (2000). *Media and communication research methods: an introduction to qualitative and quantitative approaches*. Thousand Oaks: Sage Publications. ISBN: 978 0 7619 1852 3
- 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.
<https://doi.org/10.1111/jcom.12127>
- Butt, Nadeem-Shafique** (2021). “Bibliomining and comparison of Q4 and ESCI indexed journals under statistics and probability category”. *Pakistan journal of statistics and operation research*, v. 17, n. 1.
<https://doi.org/10.18187/pjsor.v17i1.3243>
- Butt, Nadeem-Shafique; Malik, Ahmad-Azam; Shahbaz, Muhammad-Qaiser** (2021). “Bibliometric analysis of statistics journals indexed in Web of Science under Emerging Source Citation Index”. *SAGE Open*, v. 11, n. 1.
<https://doi.org/10.1177/2158244020988870>
- Caffarel-Serra, María-Carmen** (2018). “La metainvestigación en comunicación, una necesidad y una oportunidad”. *adComunica*, v. 15, pp. 293-295.
<https://doi.org/10.6035/2174-0992.2018.15.16>
- Campos-Domínguez, Eva** (2017). “Twitter y la comunicación política”. *Profesional de la información*, v. 26, n. 5.
<https://doi.org/10.3145/epi.2017.sep.01>
- Carrasco-Campos, Ángel; Saperas, Enric** (2022). “Between centrality and theoretical dispersion: theoretical frameworks in Spanish peer-reviewed communication journals. Analysis of research papers published by five top-ranked journals 2011-2020”. *Communication & society*, v. 35, n. 2, pp. 155-168.
<https://doi.org/10.15581/003.35.2.155-168>
- Carrasco-Campos, Ángel; Saperas, Enric; Martínez-Nicolás, Manuel** (2018). “¿Cómo investigamos la comunicación en España? Universidades públicas y privadas en las publicaciones científicas de comunicación españolas (1990-2014)”. *adComunica*, v. 15, pp. 45-63.
<https://doi.org/10.6035/2174-0992.2018.15.4>
- Carrillo-Vera, José-Agustín; Aguado-Terrón, Juan-Miguel; Gómez-García, Salvador** (2021). “Following the trail of esports: the multidisciplinary boom of research on the competitive practice of video games”. *International journal of gaming and computer-mediated simulations*, v. 10, n. 4, pp. 42-61.
<https://doi.org/10.4018/IJGCMS.2018100103>

- 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.
<https://doi.org/10.3145/epi.2020.mar.23>
- Casero-Ripollés, Andreu; Feenstra, Ramón A.; Tormey, Simón** (2016). "Old and new media logics in an electoral campaign: the case of podemos and the two-way street mediatization of politics". *The international journal of press/politics*, v. 21, n. 3.
<https://doi.org/10.1177/1940161216645340>
- Castillo-Esparcia, Antonio; Rubio-Moraga, Ángel; Almansa-Martínez, Ana** (2012). "La investigación en Comunicación. Análisis bibliométrico de las revistas de mayor impacto del ISI". *Revista latina de comunicación social*, n. 67, pp. 248-270.
<https://doi.org/10.4185/RLCS-067-955-248-270>
- Chadwick, Andrew** (2013). *The hybrid media system: Politics and power*. Oxford University Press. ISBN: 978 0 19 975947 7
<https://doi.org/10.1093/acprof:oso/9780199759477.001.0001>
- Chan, Chung-hong; Grill, Christiane** (2022). "The highs in communication research: research topics with high supply, high popularity, and high prestige in high-impact journals". *Communication research*, v. 49, n. 5.
<https://doi.org/10.1177/0093650220944790>
- Chaves-Montero, Alfonso; Vázquez-Aguado, Octavio** (2021). "La calidad de la producción de trabajo social de autores/as españoles/as indexada en ESCI". *Alternativas cuadernos de trabajo social*, v. 28, n. 2, pp. 189-213.
<https://doi.org/10.14198/ALTERN2021.28.2.03>
- Chung, Chung-Joo; Barnett, George A.; Kim, Kitae; Lackaff, Derek** (2013). "An analysis on communication theory and discipline". *Scientometrics*, v. 95, n. 3.
<https://doi.org/10.1007/s11192-012-0869-4>
- Costa-Sánchez, Carmen** (2017). "Análisis de la productividad y visibilidad en Scopus de los investigadores españoles en comunicación". *Observatorio (OBS*)*, v. 11, n. 3.
<https://doi.org/10.15847/obsOBS11320171030>
- Costa-Sánchez, Carmen; López-García, Xosé** (2020). "Comunicación y crisis del coronavirus en España. Primeras lecciones". *Profesional de la información*, v. 29, n. 3.
<https://doi.org/10.3145/epi.2020.may.04>
- Cuocolo, Renato; Ponsiglione, Andrea; Dell'Aversana, Serena; D'Acierno, Ludovica; Lassandro, Giulia; Ugga, Lorenzo; Romeo, Valeria; Vola, Elena-Augusta; Stanzione, Arnaldo; Verde, Francesco; Picariello, Valentina; Capaldo, Iolanda; Pontillo, Giuseppe; Cantoni, Valeria; Green, Roberta; Petretta, Mario; Cuocolo, Alberto; Imbriaco, Massimo** (2020). "The cardiac conundrum: a systematic review and bibliometric analysis of authorship in cardiac magnetic resonance imaging studies". *Insights into imaging*, v. 11, n. 1.
<https://doi.org/10.1186/s13244-020-00850-1>
- De-Filippo, Daniela; Gorraiz, Juan** (2020). "Is the Emerging Source Citation Index an aid to assess the citation impact in social science and humanities?". *Journal of informetrics*, v. 14, n. 4, 101088.
<https://doi.org/10.1016/j.joi.2020.101088>
- De-las-Heras-Pedrosa, Carlos; Martel-Casado, Telva; Jambrino-Maldonado, Carmen** (2018). "Análisis de las redes académicas y tendencias científicas de la comunicación en las universidades españolas". *Revista prisma social*, n. 22, pp. 229-246.
<https://revistaprismasocial.es/article/view/2584>
- Dhanani, Alpa; Jones, Michael-John** (2017). "Editorial boards of accounting journals: gender diversity and internationalization". *Accounting, auditing & accountability journal*, v. 30, n. 5.
<https://doi.org/10.1108/AAAJ-08-2014-1785>
- Entman, Robert M.** (1993). "Framing: toward clarification of a fractured paradigm". *Journal of communication*, v. 43, n. 4.
<https://doi.org/10.1111/j.1460-2466.1993.tb01304.x>
- Fernández-Fernández, Maximiliano; Díaz-Cano, Eduardo; Fernández-Alameda, Carlos M.; Tardivo, Giuliano** (2020). "Temas de investigación predominantes en los congresos de la Asociación Española de Investigación de la Comunicación (AE-IC)". En: *Comunicación y diversidad. Selección de comunicaciones del VII Congreso Internacional de la Asociación Española de Investigación de la Comunicación (AE-IC)*. Valencia, España, 28-30 de octubre, pp. 11-24. EPI SL. ISBN: 978 84 120239 5 4
<https://doi.org/10.3145/AE-IC-epi.2020>

- Gaitán-Moya, Juan-Antonio; Lozano-Ascencio, Carlos-Horacio; Caffarel-Serra, Carmen; Piñuel-Raigada, José-Luis** (2021). "La investigación en comunicación en proyectos I+D en España de 2007 a 2018". *Revista latina de comunicación social*, n. 79.
<https://doi.org/10.4185/RLCS-2020-1486>
- Gómez-Escalonilla, Gloria** (2021). "Métodos y técnicas de investigación utilizados en los estudios sobre comunicación en España". *Revista mediterránea de comunicación*, v. 12, n. 1.
<https://doi.org/10.14198/MEDCOM000018>
- Goyanes, Manuel** (2019). "Diversidad geográfica y de género en los comités editoriales de las revistas en comunicación". *Estudios sobre el mensaje periodístico*, v. 25, n. 1.
<https://doi.org/10.5209/ESMP.63727>
- Goyanes, Manuel** (2020). "Editorial. Meta-investigación en comunicación: antecedentes, efectos y retos de una investigación y gobernanza estandarizada". *Profesional de la información*, v. 29, n. 4.
<https://doi.org/10.3145/epi.2020.jul.06>
- Goyanes, Manuel; Rodríguez-Gómez, Eduardo-Francisco; Rosique-Cedillo, Gloria** (2018). "Investigación en comunicación en revistas científicas en España (2005-2015): de disquisiciones teóricas a investigación basada en evidencias". *Profesional de la información*, v. 27, n. 6.
<https://doi.org/10.3145/epi.2018.nov.11>
- Günther, Elisabeth; Domahidi, Emese** (2017). "What communication scholars write about: an analysis of 80 years of research in high-impact journals". *International journal of communication*, v. 11, n. 0.
<https://ijoc.org/index.php/ijoc/article/view/6989>
- Hallin, Daniel; Mancini, Paolo** (2007). *Comparing media systems: three models of media and politics*. Cambridge University Press. ISBN: 978 0 521 54308 8
- Hanitzsch, Thomas; Hanusch, Folker; Mellado, Claudia; Anikina, Maria; Berganza, Rosa; Cangoz, Incilay; Coman, Mihai; Hamada, Basyouni; Hernández, María-Elena; Karadjov, Christopher D.; Virginia Moreira, Sonia; Mwesige, Peter G.; Plaisance, Patrick Lee; Reich, Zvi; Seethaler, Josef; Skewes, Elizabeth A.; Vardiansyah-Noor, Dani; Kee-Wang-Yuen, Edgar** (2011). "Mapping journalism cultures across nations: a comparative study of 18 countries". *Journalism studies*, v. 12, n. 3.
<https://doi.org/10.1080/1461670X.2010.512502>
- Huang, Ying; Zhu, Donghua; Lv, Qi; Porter, Alan L.; Robinson, Douglas K. R.; Wang, Xuefeng** (2017). "Early insights on the Emerging Sources Citation Index (ESCI): an overlay map-based bibliometric study". *Scientometrics*, v. 111, n. 3.
<https://doi.org/10.1007/s11192-017-2349-3>
- Igartua, Juan-José; Arcila-Calderón, Carlos; Piñeiro-Naval, Valeriano; González, Alejandro; Blanco-Herrero, David** (2022). "La teoría del framing en la investigación iberoamericana sobre comunicación política". En: Muñiz, Carlos. *Framing y política. Aportaciones empíricas desde Iberoamérica*. Ciudad de México: Tirant Humanidades, pp. 57-98. ISBN: 978 84 19071 71 2
- Ioannidis, John P. A.** (2018). "Meta-research: Why research on research matters". *PLoS biology*, v. 16, n. 3.
<https://doi.org/10.1371/journal.pbio.2005468>
- Jenkins, Henry** (2008). *Convergence culture: la cultura de la convergencia de los medios de comunicación*. Barcelona: Paidós. ISBN: 978 84 493 2153 5
- Krippendorff, Klaus** (1980). *Content analysis: an introduction to its methodology*. Sage. ISBN: 978 0 8039 1498 8
- Lakoff, George** (1993). "The contemporary theory of metaphor". In: *Metaphor and thought*, Cambridge: Cambridge University Press, pp. 202-251. ISBN: 978 0 521 40561 4
- López-García, Guillermo** (2016). "Nuevos y viejos liderazgos: la campaña de las elecciones generales españolas de 2015 en Twitter". *Communication & society*, v. 29, n. 3.
<https://doi.org/10.15581/003.29.35829>
- López-Meri, Amparo; Marcos-García, Silvia; Casero-Ripollés, Andreu** (2017). "What do politicians do on Twitter? Functions and communication strategies in the Spanish electoral campaign of 2016". *Profesional de la información*, v. 26, n. 5.
<https://doi.org/10.3145/epi.2017.sep.02>
- Lozano-Ascencio, Carlos; Gaitán-Moya, Juan-Antonio; Caffarel-Serra, Carmen; Piñuel-Raigada, José-Luis** (2020). "Una década de investigación universitaria sobre Comunicación en España, 2007-2018". *Profesional de la información*, v. 29, n. 4.
<https://doi.org/10.3145/epi.2020.jul.12>

- Martínez, María-Ángeles; Cobo, Manuel-Jesús; Herrera, Manuel; Herrera-Viedma, Enrique** (2015). "Analyzing the scientific evolution of social work using science mapping". *Research on social work practice*, v. 25, n. 2.
<https://doi.org/10.1177/1049731514522101>
- Martínez-Martín, Miquel; Esteban-Bara, Francisco; Buxarrais-Estrada, María-Rosa** (2011). "Escuela, profesorado y valores". *Revista de educación*, n. extra, pp. 95-113.
<http://hdl.handle.net/11162/92222>
- Martínez-Nicolás, Manuel; Saperas, Enric; Carrasco-Campos, Ángel** (2019). "La investigación sobre comunicación en España en los últimos 25 años (1990-2014). Objetos de estudio y métodos aplicados en los trabajos publicados en revistas españolas especializadas". *Empiria*, n. 42, pp. 37-69.
<https://doi.org/10.5944/empiria.42.2019.23250>
- McCombs, Maxwell E.; Shaw, Donald L.** (1972). "The agenda setting function of mass media". *Public opinion quarterly*, v. 36, n. 2.
<https://doi.org/10.1086/267990>
- McGowan, Nadia; Repiso, Rafael; Montero-Díaz, Julio** (2022). "Las obras de mayor relevancia en el estudio científico del Cine y Fotografía". *Fotocinema*, n. 25, pp. 381-410.
<https://doi.org/10.24310/Fotocinema.2022.vi25.14428>
- Mellado, Claudia; Hellmueller, Lea; Márquez-Ramírez, Mireya; Humanes, María-Luisa; Sparks, Colin; Stepinska, Agnieszka; Pasti, Svetlana; Schielicke, Anna-Maria; Tandoc, Edson; Wang, Haiyan** (2017). "The hybridization of journalistic cultures: a comparative study of journalistic role performance". *Journal of communication*, v. 67, n. 6.
<https://doi.org/10.1111/jcom.12339>
- Miquel-Segarra, Susana** (2021). "Capítulo 4.2. Análisis de la investigación publicada en las revistas de comunicación con índice ESCI (Emerging Sources Citation Index)". *Espejo de monografías de comunicación social*, n. 2.
<https://doi.org/10.52495/c4.2.emcs.2.mic6>
- Montero-Díaz, Julio; Cobo, Manuel-Jesús; Gutiérrez-Salcedo, María; Segado-Boj, Francisco; Herrera-Viedma, Enrique** (2018). "Mapeo científico de la categoría comunicación en WoS (1980-2013)". *Comunicar*, v. 26, n. 55.
<https://doi.org/10.3916/C55-2018-08>
- Odone, Anna; Salvati, Stefano; Bellini, Lorenzo; Bucci, Daria; Capraro, Michele; Gaetti, Giovanni; Amerio, Andrea; Signorelli, Carlo** (2020). "The runaway science: a bibliometric analysis of the COVID-19 scientific literature: How COVID-19 has changed academic publishing". *Acta biomedica Atenei Parmensis*, v. 91, n. 9, pp. 34-39.
<https://doi.org/10.23750/abm.v91i9-5.10121>
- Piñeiro-Naval, Valeriano** (2020). "The content analysis methodology. Uses and applications in communication research on Spanish-speaking countries". *Communication & society*, v. 33, n. 3.
<https://doi.org/10.15581/003.33.3.1-15>
- Piñeiro-Naval, Valeriano; Morais, Ricardo** (2019). "Estudio de la producción académica sobre comunicación en España e Hispanoamérica". *Comunicar*, v. 27, n. 61.
<https://doi.org/10.3916/C61-2019-10>
- Piñeiro-Naval, Valeriano; Morais, Ricardo; Baptista, João-Pedro** (2021). "Una perspectiva hispánica sobre la actual investigación en comunicación (2014-2019)". *Revista general de información y documentación*, v. 31, n. 2.
<https://doi.org/10.5209/rgid.79464>
- Repiso, Rafael; Jiménez-Contreras, Evaristo; Aguaded, Ignacio** (2017). "Revistas iberoamericanas de educación en ScIELO Citation Index y Emerging Source Citation Index". *Revista española de documentación científica*, v. 40, n. 4, e186.
<https://doi.org/10.3989/redc.2017.4.1445>
- Repiso, Rafael; Moreno-Delgado, Alicia** (2022). "Producción científica española en comunicación indexada en Web of Science: contextualización y presencia en el ranking de Shanghai". *Profesional de la información*, v. 31, n. 1.
<https://doi.org/10.3145/epi.2022.ene.19>
- Ruiz-Pérez, Rafael; Jiménez-Contreras, Evaristo** (2019). "Los Emerging Sources Citation Index y la internacionalización de las revistas científicas españolas, con especial mención a las de Psicología". *Psicothema*, v. 31, n. 4, pp. 376-383.
<https://doi.org/10.7334/psicothema2019.59>
- Saperas, Enric; Carrasco-Campos, Ángel** (2019). "¿Cómo investigamos la comunicación? La meta-investigación como método para el estudio de las prácticas de investigación en los artículos publicados en revistas científicas". En: Francisco Sierra-Caballero; Jordi Alberich-Pascual (eds.). *Epistemología de la comunicación y cultura digital: retos emergentes*, pp. 217-230. Granada: Editorial Universidad de Granada. ISBN: 978 84 338 6527 4

- Scolari, Carlos A.** (2009). "Transmedia storytelling: Implicit consumers, narrative worlds, and branding in contemporary media production". *International journal of communication*, v. 3, pp. 586-606.
- Scolari, Carlos A.** (2013). *Narrativas transmedia: cuando todos los medios cuentan*. Barcelona: Deusto. ISBN: 978 84 234 1336 2
- Segado-Boj, Francisco; Chaparro, María-Ángeles; Ramírez-Daza, Óscar-Israel** (2022a). "Periodismo de investigación: análisis bibliométrico, conceptual e intelectual (2000-2020)". *Transinformação*, v. 34. <https://doi.org/10.1590/2318-0889202234e210080>
- Segado-Boj, Francisco; Gómez-García, Salvador; Díaz-Campo, Jesús** (2022b). "Intellectual and thematic structure of Communication research in *Scopus* (1980-2020). A comparative perspective among Spain, Europe, and Latin America". *Profesional de la información*, v. 31, n. 1. <https://doi.org/10.3145/epi.2022.ene.10>
- Segado-Boj, Francisco; Prieto-Gutiérrez, Juan-José; Díaz-Campo, Jesús** (2021a). "Redes de coautorías de la investigación española y latinoamericana en Comunicación (2000-2019): cohesión interna y aislamiento transcontinental". *Profesional de la información*, v. 30, n. 3. <https://doi.org/10.3145/epi.2021.may.05>
- Segado-Boj, Francisco; Prieto-Gutiérrez, Juan-José; Quevedo-Redondo, Raquel** (2021b). "El Efecto Matilda en la red de coautorías Hispanoamericana en Comunicación". *Revista mediterránea de comunicación*, v. 12, n. 2. <https://doi.org/10.14198/MEDCOM.18971>
- Semetko, Holli A.; Valkenburg-Valkenburg, Patti M.** (2000). "Framing european politics: a content analysis of press and television news". *Journal of communication*, v. 50, n. 2. <https://doi.org/10.1111/j.1460-2466.2000.tb02843.x>
- Sugimoto, Cassidy R.; Larivière, Vincent; Ni, Chaoqun; Cronin, Blaise** (2013). "Journal acceptance rates: A cross-disciplinary analysis of variability and relationships with journal measures". *Journal of informetrics*, v. 7. <https://doi.org/10.1016/j.joi.2013.08.007>
- Toffler, Alvin** (1980). *Future shock: the third wave*. New York: Bantam Books. ISBN: 0 553 24698 4
- Trillo-Domínguez, Magdalena; De-Moya-Anegón, Félix** (2022). "Map of scientific research on Communication in Spain: study fronts and rankings of authors, publications and institutions". *Profesional de la información*, v. 31, n. 1. <https://doi.org/10.3145/epi.2022.ene.12>
- Van-Eck, Nees-Jan; Waltman, Ludo** (2010). "Software survey: VOSviewer, a computer program for bibliometric mapping". *Scientometrics*, v. 84. <https://doi.org/10.1007/s11192-009-0146-3>
- Vizoso, Ángel; Pérez-Seijo, Sara; López-García, Xosé** (2019). "Diez años de investigación en las principales revistas científicas de comunicación. Tendencias y temáticas en las publicaciones de mayor impacto en JCR y SJR". *adComunica*, v. 18, pp. 245-270. <https://doi.org/10.6035/2174-0992.2019.18.12>
- Tankard, James W.** (2001). "The empirical approach to the study of media framing". In: Stephen D. Reese; Oscar H. Gandy; August E. Grant (eds.). *Framing public life*. Routledge. ISBN: 978 1 4106 0568 9
- Xifra, Jordi** (2020). "Comunicación corporativa, relaciones públicas y gestión del riesgo reputacional en tiempos del Covid-19". *Profesional de la información*, v. 29, n. 2. <https://doi.org/10.3145/epi.2020.mar.20>
- Zurbano-Berenguer, Belén; Cano-Orón, Lorena; Liberia-Vayá, Irene** (2018). "Gender studies in communication research: A longitudinal analysis of scientific papers published in Spanish journals indexed in the Journal Citation Reports (JCR) and the SCImago Journal Rank (SJR) (1988-2017)". *Journal of research in gender studies*, v. 8, n. 2, pp. 169-200. <https://doi.org/10.22381/JRGS82201810>

7. Annexes

Appendix 1. References of Graph 3

Bardin, Laurence (1996). <i>El análisis de contenido</i> . Madrid: Akal Ediciones. ISBN: 978 84 7600 093 9
Bauman, Zygmund (1999). <i>La modernidad líquida</i> . Madrid: Fondo de Cultura Económica de España. ISBN: 950 557 513 0 https://doi.org/10.3145/epi.2017.sep.01
Berelson, Bernard (1952). <i>Content analysis in communication research</i> . New York: Hafner.
Campos-Domínguez, Eva (2017). "Twitter y la comunicación política". <i>Profesional de la información</i> , v. 26, n. 5. https://doi.org/10.3145/epi.2017.sep.01
Casero-Ripollés, Andreu (2020). "Impact of Covid-19 on the media system. Communicative and democratic consequences of news consumption during the outbreak". <i>Profesional de la información</i> , v. 29, n. 2. https://doi.org/10.3145/epi.2020.mar.23
Casero-Ripollés, Andreu; Feenstra, Ramón A.; Tormey, Simon (2016). "Old and new media logics in an electoral campaign: The case of Podemos and the two-way street mediatization of politics". <i>The international journal of press/politics</i> , v. 21, n. 3. https://doi.org/10.1177/1940161216645340
Castells, Manuel (2009). <i>Communication power</i> . Oxford: Oxford University Press. ISBN: 978 0 19 157063 6
Castells, Manuel (2012). "Networks of outrage and hope - Social movements in the Internet age". <i>International journal of public opinion research</i> , v. 25, n. 3, pp. 398-402 https://doi.org/10.1093/ijpor/edt020
Chadwick, Andrew (2013). <i>The hybrid media system: politics and power</i> . Oxford: Oxford University Press. ISBN: 978 0 19 975947 7 https://doi.org/10.1093/acprof:oso/9780199759477.001.0001
Costa-Sánchez, Carmen; Piñero-Otero, Teresa (2012). "Nuevas narrativas audiovisuales: multiplataforma, crossmedia y transmedia: El caso de Águila Roja (RTVE)". <i>Icono 14</i> , v. 10. https://doi.org/10.7195/ri14.v10i2.156
Entman, Robert M. (1993). "Framing: toward clarification of a fractured paradigm". <i>Journal of communication</i> , v. 43, n. 4. https://doi.org/10.1111/j.1460-2466.1993.tb01304.x
Gamson, William A.; Modigliani, Andre (1989). "Media discourse and public opinion on nuclear power: a constructionist approach". <i>American journal of sociology</i> , v. 95, n. 1. https://doi.org/10.1086/229213
Goffman, Erving (1974). <i>Frame analysis: an essay on the organization of experience</i> . Cambridge: Harvard University Press. ISBN: 978 0 674 31656 0
Igartua, Juan-José (2006). <i>Métodos cuantitativos de investigación en comunicación</i> . Barcelona: Editorial Bosch. ISBN: 84 9790 271 8
Jenkins, Henry (2003a). <i>Convergence culture: where old and new media collide</i> . New York: New York University Press. ISBN: 978 0 8147 4368 3
Jenkins, Henry (2003b). "Transmedia storytelling". <i>MIT technology review</i> .
Jenkins, Henry (2012). <i>Textual poachers: television fans and participatory culture</i> . Routledge. ISBN: 978 0 415533294
Jenkins, Henry; Ford, Sam; Green, Joshua (2013). <i>Spreadable media: creating value and meaning in a networked culture</i> . New York: NYU Press. ISBN: 978 0 8147 4350 8
Krippendorff, Klaus (1980). <i>Content analysis: an introduction to its methodology</i> . Beverly Hills: Sage. ISBN: 978 0 8039 1498 8
Lázaro-Rodríguez, Pedro; Herrera-Viedma, Enrique (2020). "Noticias sobre Covid-19 y 2019-nCoV en medios de comunicación de España: el papel de los medios digitales en tiempos de confinamiento". <i>Profesional de la información</i> , v. 29, n. 3. https://doi.org/10.3145/epi.2020.may.02
López-García, Guillermo (2016). "Nuevos y viejos liderazgos: la campaña de las elecciones generales españolas de 2015 en Twitter". <i>Communication & society</i> , v. 29, n. 3. https://doi.org/10.15581/003.29.35829
López-Meri, Amparo (2017). "Contribución ciudadana al debate electoral y su cobertura periodística en Twitter". <i>Prisma social: revista de investigación social</i> , n. 18. https://revistaprismasocial.es/article/view/1466
McCombs, Maxwell E.; Shaw, Donald L. (1972). "The agenda setting function of mass media". <i>Public opinion quarterly</i> , v. 36, n. 2, pp. 176-187. https://doi.org/10.1086/267990
Neuendorf, Kimberly A. (2002). <i>The content analysis guidebook</i> . Sage. ISBN: 978 0 7619 1977 3
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". <i>Profesional de la información</i> , v. 29, n. 3. https://doi.org/10.3145/epi.2020.may.08
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". <i>Profesional de la información</i> , v. 29, n. 3. https://doi.org/10.3145/epi.2020.may.15
Scheufele, Dietram A. (1999). "Framing as a theory of media effects". <i>Journal of communication</i> , v. 49, n. 1. https://doi.org/10.1111/j.1460-2466.1999.tb02784.x
Scolari, Carlos A. (2009). "Transmedia storytelling: implicit consumers, narrative worlds, and branding in contemporary media production". <i>International journal of communication</i> , v. 3.
Scolari, Carlos A. (2012). "Comunicación digital. Recuerdos del futuro". <i>Profesional de la información</i> , v. 21, n. 4. https://doi.org/10.3145/epi.2012.jul.01
Scolari, Carlos A. (2013). <i>Narrativas transmedia: cuando todos los medios cuentan</i> . Barcelona: Deusto. ISBN: 978 84 234 1336 2
Semetko, Holli A.; Valkenburg, Patti M. Valkenburg-Valkenburg (2000). "Framing European politics: a content analysis of press and television news". <i>Journal of communication</i> , v. 50, n. 2. https://doi.org/10.1111/j.1460-2466.2000.tb02843.x
Tankard, James W. (2001). "The empirical approach to the study of media framing". In: Stephen D. Reese; Oscar H. Gandy; August E. Grant (eds.). <i>Framing public life</i> . Routledge. ISBN: 978 1 4106 0568 9
Toffler, Alvin (1980). <i>Future shock: the third wave</i> . New York: Bantam Books. ISBN: 0 553 24698 4
Vreese, Claes H. (2005). "News framing: theory and typology". <i>Information design journal</i> , v. 13, n. 1. https://doi.org/10.1075/idjd.13.1.06vre
Wimmer, Roger D.; Dominick, Joseph R. (1996). <i>La investigación científica de los medios de comunicación: una introducción a sus métodos</i> . Barcelona: Editorial Bosch. ISBN: 978 84 7676 359 9

Appendix 2. References of Graph 4

Bennett, Lance; Segerberg, Alexandra (2012). <i>The logic of connective action: digital media and the personalization of contentious politics</i> . Cambridge University Press. ISBN: 978 1 139198752 https://doi.org/10.1017/CBO9781139198752
Bimber, Bruce (2014). "Digital media in the Obama campaigns of 2008 and 2012: Adaptation to the personalized political communication environment". <i>Journal of information technology & politics</i> , v. 11, n. 2, pp. 130-150. https://doi.org/10.1080/19331681.2014.895691
Brüggemann, Michael; Engesser, Sven; Büchel, Florin; Humprecht, Edda; Castro, Laia (2014). "Hallin and Mancini revisited: four empirical types of Western media systems". <i>Journal of communication</i> , v. 64, n. 6, pp. 1037-1065. https://doi.org/10.1111/jcom.12127
Casero-Ripollés, Andreu (2018). "Research on political information and social media: Key points and challenges for the future". <i>Profesional de la información</i> , v. 27, n. 5. https://doi.org/10.3145/epi.2018.sep.01
Casero-Ripollés, Andreu (2020). "Impact of Covid-19 on the media system. Communicative and democratic consequences of news consumption during the outbreak". <i>Profesional de la información</i> , v. 29, n. 2. https://doi.org/10.3145/epi.2020.mar.23
Casero-Ripollés, Andreu; Feenstra, Ramón A.; Tormey, Simon (2016). "Old and new media logics in an electoral campaign: The case of Podemos and the two-way street mediatization of politics". <i>The international journal of press/politics</i> , v. 21, n. 3. https://doi.org/10.1177/1940161216645340
Castells, Manuel (2009). <i>Communication power</i> . Oxford: Oxford University Press. ISBN: 978 0 19 157063 6
Castells, Manuel (2012). "Networks of outrage and hope - Social movements in the Internet age". <i>International journal of public opinion research</i> , v. 25, n. 3, pp. 398-402. https://doi.org/10.1093/ijpor/edt020
Chadwick, Andrew (2013). <i>The hybrid media system: politics and power</i> . Oxford: Oxford University Press. ISBN: 978 0 19 975947 7 https://doi.org/10.1093/acprof:oso/9780199759477.001.0001
Costa-Sánchez, Carmen; López-García, Xosé (2020). "Comunicación y crisis del coronavirus en España. Primeras lecciones". <i>Profesional de la información</i> , v. 29, n. 3. https://doi.org/10.3145/epi.2020.may.04
Curran, James; Iyengar, Shanto; Lund, Anker-Brink; Salovaara-Moring, Inka (2009). "Media system, public knowledge and democracy: a comparative study". <i>European journal of communication</i> , v. 24. https://doi.org/10.1177/0267323108098943
Enli, Gunn-Sara; Skogerboe, Eli (2013). "Personalised campaigns in party-centered politics: Twitter and Facebook as arenas for political communication (2015)". <i>Information, communication and society</i> , v. 16, n. 5, pp. 757-774. https://doi.org/10.1080/1369118X.2013.782330
Entman, Robert M. (1993). "Framing: toward clarification of a fractured paradigm". <i>Journal of communication</i> , v. 43, n. 4. https://doi.org/10.1111/j.1460-2466.1993.tb01304.x
Fernández-Quijada, David; Masip-Masip, Pere (2013). "Tres décadas de investigación española en comunicación: hacia la mayoría de edad". <i>Comunicar</i> , v. 21, n. 41. https://doi.org/10.3916/C41-2013-01
Fletcher, Richard; Nielsen, Rasmus-Kleis (2018). "Are people incidentally exposed to news on social media? A comparative analysis". <i>New media & society</i> , v. 20, n. 7. https://doi.org/10.1177/1461444817724170
Gamson, William A.; Modigliani, Andre (1989). "Media discourse and public opinion on nuclear power: a constructionist approach". <i>American journal of sociology</i> , v. 95, n. 1. https://doi.org/10.1086/229213
Gil de Zúñiga, Homero; Weeks, Brian; Ardèvol-Abreu, Alberto (2017). "Effects of the news-finds-me perception in communication: social media use implications for news seeking and learning about politics". <i>Journal of computer-mediated communication</i> , v. 22, n. 3, pp. 105-123. https://doi.org/10.1111/jcc4.12185
Hallin, Daniel; Mancini, Paolo (2007). <i>Comparing media systems: three models of media and politics</i> . Cambridge: Cambridge University Press. ISBN: 978 0 521 54308 8
Hanitzsch, Thomas; Hanusch, Folker; Mellado, Claudia; Anikina, Maria; Berganza, Rosa; Cangoz, Incilay; Coman, Mihai; Hamada, Basyouni; Hernández, María-Elena; Karadjov, Christopher D.; Virginia-Moreira, Sonia; Mwesige, Peter G.; Plaisance, Patrick-Lee; Reich, Zvi; Seethaler, Josef; Skewes, Elizabeth A.; Vardiansyah-Noor, Dani; Kee-Wang-Yuen, Edgar (2011). "Mapping journalism cultures across nations: A comparative study of 18 countries". <i>Journalism studies</i> , v. 12, n. 3. https://doi.org/10.1080/1461670X.2010.512502
Jenkins, Henry (2008). <i>Convergence culture: la cultura de la convergencia de los medios de comunicación</i> . Barcelona: Paidós. ISBN: 978 84 493 2153 5
Jenkins, Henry; Ford, Sam; Green, Joshua (2013). <i>Spreadable media: creating value and meaning in a networked culture</i> . New York: NYU Press. ISBN: 978 0 8147 4350 8
Lakoff, George (1993). "The contemporary theory of metaphor". In: <i>Metaphor and thought</i> . Cambridge: Cambridge University Press, pp. 202-251. ISBN: 978 0 521 40561 4 https://doi.org/10.1017/CBO9781139173865.013
Lakoff, George; Turner, Mark (1989). <i>More than cool reason: A field guide to poetic metaphor</i> . University of Chicago Press. ISBN: 978 0 226 46812 9 https://doi.org/10.7208/chicago/9780226470986.001.0001
Lakoff, Sanford A. (1980). "Moral responsibility and the Galilean imperative". <i>Ethics</i> , v. 91, n. 1. https://doi.org/10.1086/292206
López-García, Guillermo (2016). "Nuevos y viejos liderazgos: la campaña de las elecciones generales españolas de 2015 en Twitter". <i>Communication & society</i> , v. 29, n. 3. https://doi.org/10.15581/003.29.35829
López-Meri, Amparo (2017). "Contribución ciudadana al debate electoral y su cobertura periodística en Twitter". <i>Prisma social: revista de investigación social</i> , n. 18. https://revistaprimasocial.es/article/view/1466
Martínez-Martín, Miguel (2011). "Educación, valores y democracia". <i>Revista de educación</i> , n. 1, pp. 15-19.
Mellado, Claudia (2015). "Professional roles in news content". <i>Journalism studies</i> , v. 16, n. 4, pp. 596-614. https://doi.org/10.1080/1461670X.2014.922276
Scolari, Carlos A. (2009). "Transmedia storytelling: implicit consumers, narrative worlds, and branding in contemporary media production". <i>International journal of communication</i> , v. 3.
Scolari, Carlos A. (2013). <i>Narrativas transmedia: cuando todos los medios cuentan</i> . Barcelona: Deusto. ISBN: 978 84 234 1336 2
Van-Dalen, Arjen; De-Vreese, Claes H.; Albæk, Erik (2012). "Different roles, different content? A four-country comparison of the role conceptions and reporting style of political journalists". <i>Journalism</i> , v. 13, n. 7. https://doi.org/10.1177/1464884911431538
Van-Gorp, Baldwin (2007). "The constructionist approach to framing: Bringing culture back in". <i>Journal of communication</i> , v. 57, n. 1, pp. 60-78. https://doi.org/10.1111/j.1460-2466.2006.00329_3.x
Xifra, Jordi (2020). "Comunicación corporativa, relaciones públicas y gestión del riesgo reputacional en tiempos del Covid-19". <i>Profesional de la información</i> , v. 29, n. 2. https://doi.org/10.3145/epi.2020.mar.20

Public health communication and the Covid-19: A review of the literature during the first wave

Alfonso Méndiz-Noguero; Laia Wennberg-Capellades; Elisa Regadera-González; Blanca Goni-Fuste

Nota: Este artículo se puede leer en español en:
<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87233>

Recommended citation:

Méndiz-Noguero, Alfonso; Wennberg-Capellades, Laia; Regadera-González, Elisa; Goni-Fuste, Blanca (2023). "Public health communication and the Covid-19: A review of the literature during the first wave". *Profesional de la información*, v. 32, n. 3, e320313.
<https://doi.org/10.3145/epi.2023.may.13>

Article received on December 26th 2022
Approved on February 20th 2023



Alfonso Méndiz-Noguero
<https://orcid.org/0000-0003-4450-1272>

Universitat Internacional de Catalunya
Immaculada, 22
08017 Barcelona, Spain
amendiz@uic.es



Laia Wennberg-Capellades ✉
<https://orcid.org/0000-0002-6811-1343>

Universitat Internacional de Catalunya
Immaculada, 22
08017 Barcelona, Spain
lwennberg@uic.es



Elisa Regadera-González
<https://orcid.org/0000-0002-1793-8105>

Universitat Internacional de Catalunya
Immaculada, 22
08017 Barcelona, Spain
eregadera@uic.es



Blanca Goni-Fuste
<https://orcid.org/0000-0002-1475-4466>

Universitat Internacional de Catalunya
Immaculada, 22
08017 Barcelona, Spain
bgoni@uic.es

Abstract

The expansion of the Covid-19 virus in early 2020 grew in parallel with the spread of rumours, false or unverified news and even contradictions between information sources and health sources. It has been the first pandemic to be broadcast live on social media and has generated disinformation which was described by the *WHO* as an "infodemic", a pandemic as serious as the virus itself. The aim was to identify and analyse the impact generated by the first wave of Covid-19 (January-June 2020) on public health communication. The review was carried out under the *Prisma* guidelines. A systematic search was performed in *PubMed*, *Scopus* and *Web of Science* databases, which yielded a figure of 1.157 papers. Using seven keywords as a filter a corpus of 193 articles was reached. Four main themes were identified: 1) Need for massive public health literacy; 2) Social networks as an information and disinformation during pandemic; (3) The uncertain response of institutional communication; and (4) Media coverage of the pandemic. The authors propose large-scale health literacy and point out the need to work on health information together -governments, health institutions and the media-.

Keywords

Systematic review; Covid-19; Coronavirus; Infodemics; Health communication; Public health; Ehealth literacy; Social media; Social networks; Disinformation; Health information; Pandemics.



1. Introduction

The Covid-19 pandemic is a Public Health Emergency of International Concern (PHEIC). Worldwide, 634.5 million infections and more than 6.5 million deaths were counted as of November 21, 2022 (*World Health Organization, 2022*). This is the fifth PHEIC declared by the *WHO*, the four previous ones being

- the severe acute syndrome (SARS) pandemic in 2002-2004 in China and Southeast Asia;
- H1N1 swine influenza in 2009;
- the Middle Eastern respiratory syndrome (MERS) in 2013 in Saudi Arabia; and
- Ebola in 2014 in Africa.

Covid-19 was a true health crisis situation on a global scale, affecting the whole population of the world. This pandemic can be described as an unexpected crisis in which critical elements coincided such as:

“the surprise factor, destabilization, stress and the reduction of time that determine decision-making, the emotion of the affected public, media pressure, the challenge for the image of the authorities and the danger to the position of power the governing officials” (**Crespo-Martínez; Garrido, 2020**).

The different communication models were disrupted the pandemic. There was no clear responsible beyond those pointed out by the political *framing*. Moreover, this crisis has generated infinite economic and social crises that have led to a very alarming situation in all countries (**Xifra, 2020**). From a strategic point of view, it has affected different sectors and publics –governments, institutions, media, citizens, companies, among others– (**De-la-Hoz, 2014**), and has required global health coordination between many countries to solve it.

During the Covid-19 pandemic, digital channels and, above all, social networks have had a clear preeminence as tools at the service of institutional communication to inform the population in a fast and accessible way, and to establish effective relationships with the public (**Castro-Martínez; Morán-Urdiales; Díaz-Morilla, 2021**). In previous pandemics, the value of social networks as a tool to quickly inform the population was already highlighted (**Vijaykumar; Nowak; Himelboim; Jin, 2018**). However, in this pandemic they have gained a special role, generating such an incredible volume of news which made impossible to discern between reliable and non-reliable information. The speed with which information circulated, and the mixture of erroneous and contradictory news has made containment measures difficult and has become an additional threat. The *WHO* has denounced this situation, which constitutes a great challenge for public health communication, and has pointed out that this deluge of inaccurate and contradictory information is the main obstacle to public health policies and, consequently, to preventive actions and interventions (**Venegas-Vera; Colbert; Lerma, 2020**).

All this has highlighted the need for digital literacy to be able to search, understand and evaluate the health information conveyed in digital channels in relation to pandemic prevention and control measures, since low literacy is directly related to misuse of information (**Choukou et al., 2022**). Health literacy encompasses people’s knowledge, motivation and skills to find, understand, value and apply health information, as well as to make informed judgments and decisions about health care, disease prevention and health promotion to act accordingly (**Juvinyà-Canal; Bertran-Noguer; Suñer-Soler, 2020**). The proper use of communication channels contributes powerfully to the promotion of health and public awareness of the causes and remedies in the face of a pandemic, which generates a positive impact on the behavior of the population in relation to their health (**Mheidly; Fares, 2020**).

One of the key elements in dealing with the so-called PHEIC is how communication is managed in public health (**Mirbaiae et al., 2020**). **Bernhardt (2004)** defines it as

“the scientific development, strategic dissemination and critical evaluation of relevant and accurate information, in an accessible and understandable way, in order to improve population health.”

Its main objective is to generate behavioral changes in the population which must be done with a strategy that puts health objectives before the satisfaction of mere informational curiosity. As **Glick (2007)** states, communication management is crucial in the administration of a public health crisis, which is why the *Centers for Disease Control and Prevention (CDC)* consider it one of the 15 key points in preparing for a health emergency (*Centers for Disease Control and Prevention, 2018*).

Public health communication aims to instruct the population in their personal health management (**Paakkari; Okan, 2020**), helping to mitigate the risks of contagion, supporting official protection measures and reducing the negative impact on citizens’ mental health as a result of a pandemic. Since the beginning of Covid-19, local and international public health agencies have used the media to inform the population immediately about the evolution of the pandemic, with the aim of raising awareness and reducing the impact of the spread of the virus on the population (**Mheidly; Fares, 2020**).

All stakeholders (Ministries of Health, Public Health Agencies, etc.) should join efforts to conduct health education campaigns on the pandemic and preventive behaviors, ensuring that information is clear, truthful and credible (**Frieden, 2014**). The use of the media as channels of health information has highlighted the importance of health literacy and, in particular, the importance of digital literacy. The *WHO* defines health literacy as

“the personal characteristics and social resources needed for individuals and communities to access, understand, appraise and use information and services to make decisions about health” (*World Health Organization*, 2015).

As **Naeem** and **Boulos** (2021) point out, such literacy is essential to deal with a global health crisis, as it exponentially increases the resilience of the population.

A retrospective analysis of the first PHEIC has shown the importance of media campaigns to inform the population and educate them about preventive measures (**Savoia; Lin; Viswanath**, 2013; **Basch et al.**, 2020; **Liu, Q. et al.**, 2020). Likewise, a systematic review on the use of social networks during the Ebola epidemic showed a strong breakdown in health communication that led to widespread panic among the population due to the misuse of new technologies (**Roberts et al.**, 2017; **Wong et al.**, 2017).

The Covid-19 pandemic, and specifically this first wave, introduced important challenges for health communication for governments and authorities. The analysis of this communication during this phase of the health crisis has generated not only valuable knowledge, but also relevant conclusions to improve communication strategies for future health emergencies.

2.1. Search

The first step was to implement an adequate search strategy in the main databases. Due to the theme of our study, we used *PubMed*, and for the scientific relevance of the databases, *Web of Science* (core collection) and *Scopus*. The literature review was carried out in these databases as the outcome of our study was communication in public health. The dates indicated by the *WHO* were established as a time frame: January to June 2020. Table 1 shows the search strategy implemented in *PubMed*. This same strategy has been implemented in *Web of Science* and *Scopus*, adapting the terms to the specific language of each of these bases. The results obtained were imported into the *Rayyan* screening tool, to extract duplicates and select the articles.

2.2. Eligibility criteria

The second step was to rigorously determine the criteria for inclusion and exclusion of the articles. Articles that analyzed public health communication during the first wave of the Covid-19 pandemic in any communicative context (media, social networks, advertising campaigns, institutional communications, etc.), with any type of format and design, and published in English or Spanish, were included.

The search strategy in the three databases yielded 2,206 references (*PubMed*: 720, *Web of Science*: 316, *Scopus*: 1170; search date: July 2020). A total of 1,049 duplicates were removed and further 954 that did not specifically address the issue raised were discarded. Finally, a total of 193 articles were selected (Figure 1).

2.3. Synthesis

The study review process was carried out by all four researchers. It was done independently and in three sub-phases following the *Prisma* guidelines (**Moher et al.**, 2009): 1) Review by title, 2) Review by abstract, and 3) Review by full text based on the inclusion criteria. The reasons for exclusion of full text were recorded (Figure 1). Any discrepancies in the selection process were resolved by discussion among the members of the research team. An ad hoc data extraction matrix was designed with the following list: first author, year of publication, country of the PI, country under study, study design, aim, main results and conclusions.

2.4. Analysis

The data in the extraction table, as well as the main objectives and results, were analysed by means of a thematic analysis

Table 1. Search strategy in the *PubMed* database

	Terms used in <i>PubMed</i> (10/07/2020)	Nº of references
#1	covid-19	30,423
#2	coronavirus	35,044
#3	#1 OR #2	47,134
#4	public health campaign*	1,036
#5	media	497,391
#6	health communication	7,182
#7	health information	33,186
#8	health literacy	10,636
#9	#4 OR #5 OR #6 OR #7 OR #8	542,503
#10	#3 AND #9	720

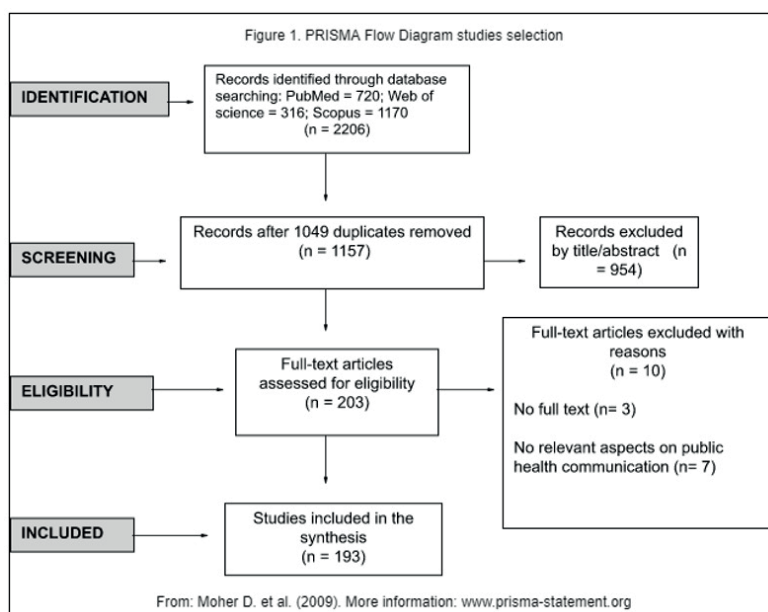


Figure 1. Scheme of the bibliographic search

that took into account the key concepts alluded to in the title, abstract and body of the article. Because of the nature of this study, approval by an ethics committee was not required.

3. Analysis and results

The analysis of the articles shows that most of the published works have studied the pandemic as a global issue, with reference to its incidence in the general population ($n=91$). In terms of geographical overview, the countries that were the subject of a more specific study were: China ($n=19$); United States ($n=17$) and United Kingdom ($n=12$), in first place; followed by India ($n=4$), Pakistan ($n=3$) and South Korea ($n=3$). The other countries have had a residual frequency.

Regarding the type of article, the following articles fall from highest to lowest in: original studies (119), commentaries (32), reviews (17), letter to the editor (14), editorial (7) and gray literature (4). Regarding the design of the original studies, 91 of them were cross-sectional quantitative studies and 28 qualitative studies. Of these, 13 were descriptive and 15 case studies.

The literature review allowed us to identify four main themes:

- (1) The need for mass public health literacy ($n=63$);
- (2) Social networks as a source of information and misinformation ($n=105$);
- (3) The uncertain response of institutional communication ($n=27$); and
- (4) Media coverage of the pandemic ($n=31$).

The total of the 4 themes adds up to 226 articles because several of them address more than one storyline. We describe each of the themes below.

3.1. The need for mass public health literacy

Some authors point to the importance of an urgent improvement in public health literacy (**Abdel-Latif**, 2020; **Abel and McQueen**, 2020; **Finset et al.**, 2020; **Gray Back**, 2020; **Liao et al.**, 2020; **Lin, Y. H. et al.**, 2020). A total of 63 works affirm this exhaustively and an editorial by **Van-den-Broucke** (2020) highlights the desire and determined will of the population to be well informed. For the authors, responding to this social need is unavoidable in the face of upcoming health emergencies (**Li, S. et al.**, 2020; **Liu, P. et al.**, 2020; **Paakkari and Okan**, 2020).

Quite a few authors also see the need to involve the population in the adoption of preventive measures, changing the common tendency to leave all responsibility in government hands. Awareness needs to be raised so that they have greater control over their health (**Habersaat et al.**, 2020; **Lazcano-Ponce and Alpuche-Aranda**, 2020). **Hashemi-Shahri et al.**, (2020) demonstrate with data how poor literacy leads to a greater spread of the pandemic. **Chan et al.** (2020) also point out that the lack of information has endangered health professionals, increasing their chances of contagion. Other authors have highlighted that, although internet searches have sometimes facilitated the adoption of preventive behaviors (**De Coninck et al.**, 2020; **Kamenidou; D'Haenens; Matthijs**, 2020; **Liu, Q.**, 2020; **Yusof et al.**, 2020), citizens have often failed to interpret or transmit information with due prudence (**Harnett**, 2020). Also the saturation of content has made it difficult to decide which sources deserved confidence and which did not (**Lima; Albanez; Brito**, 2020).

The importance of informing in a language accessible to all is also noted (**Airhihenbuwa et al.**, 2020; **Karamouzián; Johnson; Kerr**, 2020), especially for those most vulnerable (**Rudd; Baur**, 2020). **Kluger and Scrivener** (2020) highlight that people use more informal language when searching the internet; therefore, others advice offering visual and educational content (**Adam; Bärnighausen; McMahan**, 2020; **Ali; Bhatti**, 2020; **Chiodini**, 2020) or manifest the need to improve digital literacy (**Katapally**, 2020; **López**, 2020; **Sharma et al.**, 2020; **Szmuda et al.**, 2020). Some authors highlight the need to achieve an emotional connection with the public, giving prominence to the population (**Martínez-Estrella**, 2020).

To build trust, it is key that the information is transmitted by doctors publicly recognized as experts in the field. Messages should not be based on fear, nor should they use a paternalistic tone; rather, they must build trust in the community, relying on standards recommended by health workers and adopted by all (**Bilal et al.**, 2020; **St. Amant**, 2020; **Wang et al.**, 2020). Finally, health workers should not be portrayed as 'heroes', as this puts on them the psychological pressure to feel obliged to work beyond their duty, while reducing the population's awareness of their responsibility for their own health and the obligation to comply with health measures (**Cox**, 2020).

“ The main objective of public health communication is to generate behavioral changes in the population, which is why health objectives must be put before the satisfaction of mere informational curiosity ”

3.2. Social networks as a source of information and misinformation about the pandemic

Covid-19 has been the first pandemic broadcast live through social networks, which have been hyper-consulted before each turn of the health crisis (**Bento et al.**, 2020). Unlike what happened in other crises –such as Ebola or avian flu– the governments and health workers found themselves unable to control the flow of information, outpaced in time and credibility by social networks and overwhelmed by the saturation of unverified information. A total of 105 articles analyze this phenomenon.

3.2.1. Widespread disinformation

The authors harshly criticize the role of social media in the early days of the pandemic. They are accused of having massively disseminated fake news and unverified news (**Ahmed-Siddiqui et al., 2020; Armitage et al., 2020; Romanò; Majerova; Machová, 2020**) with extraordinary ease to replicate rumors (**Bastani; Bahrami, 2020**). This has generated a new concept in public health: “infodemic” (information epidemic). This term, created in the networks themselves and officially adopted by the *WHO*, refers to the high and uncontrolled flow of news about the pandemic (**Kulkarni et al., 2020; Lu, Y.; Zhang, 2020**), which has caused overexposure to the media, making it impossible to discern false information from true information (**Sasaki et al., 2020; Sharov, 2020**). Within this concept, some authors (**Looi et al., 2020; Mirbabaie et al., 2020**) point out that continued over-information has also generated a negative and fatalistic impression of public health. In fact, the large amount of fake or exaggerated news has given rise to a climate of skepticism, along with complete disorientation (**Ippolito et al., 2020**).

Many authors see the need to involve the population in the adoption of preventive measures, changing the common trend of leaving all responsibility in government hands.

This disorientation and misinformation has generated great concern among citizens (**Chong et al., 2020; Ratzan; Sommariva; Rauh, 2020**). Hence, “the first measure to control the pandemic” is the need to verify any information before sharing it; instead, it is strongly advised to refer to official websites.

3.2.2. Beneficial effects of social networks

On the other hand, several authors highlight the effectiveness of social networks in the rapid transmission of protocols between medical staff and health authorities (**González-Padilla; Tortolero-Blanco, 2020; Gottlieb; Dyer, 2020**). They have also been shown to be effective in training and informing doctors and health workers (**Chan et al., 2020**), emergency responders in health centers (**Merchant; Lurie, 2020**) and those who should manage public health information (**Amin, 2020; O’Brien; Moore; McNicholas, 2020**). Their potential use for rapid and effective communication with citizens is noted (**Bao et al., 2020; Leng; Phua, 2020; Nazir et al., 2020**), as they positively increase risk awareness (**Luu, 2020; Malecki; Keating; Safdar, 2020**), contribute to the adoption of preventive measures (**Abd-Alrazaq et al., 2020; Su et al., 2020**) and are useful for soliciting volunteer collaboration (**Ekzayez et al., 2020; Chen; Lerman; Ferrara, 2020**). Their systematic and protocolized use by public health officials is proposed (**Balhara; Chandiok, 2020**), as there is evidence that medical personnel are, in general, a reliable source for citizens (**Hunter, 2020; Lohiniva et al., 2020; Topf, 2020**). Some articles appeal to health organizations, media and all stakeholders to create common platforms to disseminate quality information on the Covid-19 (**Jain; Sinha, 2020; Tasnim; Hossain; Mazumder, 2020**).

Other authors highlight the help that social networks have provided to medical students to act as health agents (**Huddart et al., 2020**), as well as the intensive and beneficial participation of young people in digital channels (**Hashim et al., 2020; Mohamad, 2020; Olaimat et al., 2020**). In healthcare, social networks have contributed to the rediscovery of telehealth, which includes medical care via the internet and social networks (**Massaad; Cherfan, 2020; Mulrennan; Colt, 2020**).

3.2.3. Medical use of specific social networks

Several studies have focused on the analysis of a particular social network. Regarding *YouTube*, studies show that videos produced by health professionals or university professors have been more efficient than those disseminated by citizens or other information channels (**Fitz-Maurice; Di-Tommaso; Baranchuk, 2020; Kocyigit; Akaltun; Sahin, 2020; D’Souza et al., 2020; Vervoort et al., 2020**). This is why one author laments the fact that the US health authorities have hardly used this medium, a missed opportunity (**Basch et al., 2020**).

Regarding *Twitter*, some authors point out that the information disseminated here by health professionals is, with exceptions, mostly adequate (**Shah; Kim; Mian, 2020**). Some good practices have been analysed (**Park; Park; Chong, 2020**), such as the creation of an account for physicians’ communication with patients with diabetes (**Iacobucci, 2020**), the #GetMePPE movement to raise awareness of protective measures (**He et al., 2020**) and its use for monitoring people with sclerosis (**Nesbitt et al., 2020**). In contrast, it is clearly noted that the G7 world leaders made controversial use of this network during the pandemic (**Rufai; Bunce, 2020**).

Some authors have tested the effectiveness of specific networks in providing information to specific groups of patients: *Facebook* for communication with patients with diabetes (**Isip-Tan et al., 2020**); *WhatsApp* as an alerting network and medical surveillance system (**Ekzayez et al., 2020**); *Weibo* as a source of information for patients with coronavirus (**Huang et al., 2020**); and in some cases the use of *Facebook* live to disseminate medical information in real time has also been proposed (**Kauffman et al., 2020**).

3.3. The uncertain response of institutional communication

Institutional communication has proved uncertain for many authors and a lack of coordination between the various institutions at national and international level has been observed, which has increased the existing disinformation. A total of 27 papers refer to this topic, which we have divided into two thematic areas.

3.3.1. Criticisms of government and institutional communication

The Covid-19 crisis has also been a political and health communication crisis. Several factors have contributed to this: partisan use, disparity of health criteria, minimisation of risk by governments and lack of accountability in some media (Gollust; Nagler; Fowler, 2020). The inaction of authorities to combat fake news (Pennycook *et al.*, 2020), especially concerning magical cures or racist information (Rathore; Farooq, 2020), has been criticized in general, leading to many people ignoring health recommendations (Brivio; Oliveri; Pravettoni, 2020).

Several authors point out that governments have not adequately coordinated the dissemination of information, which explains why they have generated so little trust in society. Some criticize the scant information on protective measures (Hu *et al.*, 2020; Ruiu, 2020), or the self-sufficient way in which some leaders have managed the transmission of this information, prioritizing economic development over pandemic containment (Singer, 2020). Two striking examples were Trump's promotion of hydroxychloroquine without *Health Department* approval (Samy; Ahmed; Kelada, 2020) and the British government's changing attitude towards the virus: first it spoke of "containing it", then of "delaying it" and finally it proposed "herd immunity" (Cowper, 2020).

3.3.2. The need to rethink institutional public health communication

Several articles propose strengthening and improving the response of governments and health institutions towards public opinion: to prevent epidemics or to cope with them and to foster government-society-citizen collaboration to jointly combat them (Han *et al.*, 2020; Harnett, 2020; Zhao *et al.*, 2020; Zhu *et al.*, 2020).

A study has been conducted on how various institutional groups (government leaders, social media companies and healthcare providers) could jointly respond to the communication challenges of Covid-19 and their role in preventing relevant but uninformed voices from generating mistrust or endangering public health (Limaye *et al.*, 2020; Yin *et al.*, 2020). Some outline digital initiatives that can help institutions respond more proactively (Dheeraj, 2020; El-Jardali; Bou-Karroum; Fadlallah, 2020; Huang *et al.*, 2020; Jayawardena *et al.*, 2020; Raamkumar; Tan; Wee, 2020; Rashid; Wang, 2020; Schillinger; Chittamuru; Ramírez, 2020). Bilbatua *et al.* (2020) point out that the strategic public health communication response should have three objectives: to address citizens' need for information; to pre-empt misinformation; and to build trust in health institutions. Eysenbach (2020) identifies four pillars for infodemic management: 1) information monitoring (info-surveillance); 2) fostering health literacy; 3) better data verification and peer review; and 4) knowledge sharing, minimizing political or commercial interference.

Finally, Guest, Del-Río and Sánchez (2020) indicates three key steps to end Covid-19: health leadership, rapid innovation and political will. A key element on which several authors agree is that there should be a strategy of international, national and EU coordination, through the creation of partnerships and common platforms to disseminate quality information (Sahoo; Sahu; Kankaria, 2020; Tasnim; Hossain; Mazumder, 2020).

3.4. Media coverage of the pandemic

In contrast to the negative view of social media, the 31 articles that analyse media coverage offer a generally positive view; perhaps this is why they have experienced an increase in readership during the pandemic (Casero-Ripollés, 2020; La *et al.*, 2020; Mahima *et al.*, 2020). Some attribute this increase to their empathy with people's suffering (Barile; Bovolino, 2020); others to their ability to mitigate the tendency to scaremongering (Bilal *et al.*, 2020); and others to their ability to raise awareness about preventive measures (Chang *et al.*, 2020; Segura, 2020; Sezgin; Karaaslam; Ersoy, 2020), key factor in mitigating the spread of the disease (Bilal *et al.*, 2020; Yan *et al.*, 2020).

For health professionals the media has generally been seen as a reliable source (Ko *et al.*, 2020; Walker; Sulyok, 2020; Wang *et al.*, 2020), providing them with the means to assess the evolution of the pandemic (Dkhar *et al.*, 2020; Karasneh *et al.*, 2021; Liu, 2020). Hence, some have encouraged the presence of health professionals and academics in the media, to support the measures taken by health authorities (Laufer, 2020).

To a much lesser extent, other authors have criticised the "infodemics" that the media itself has generated (Oh *et al.*, 2020), by continuously and falsely disseminating negative news (Cuan-Baltazar *et al.*, 2020; Ippolito, 2020; Rommer; Majerova; Machová, 2020) and allowing a wave of fake news that has fostered racism (Rovetta; Bhagavathula, 2020), increased public anxiety (Arora; Grey, 2020; Wen *et al.*, 2020) or generated unease in financial markets (Haroon; Rizvi, 2020).

However, the main complaint has been in the reverse direction: government pressure on the media to control information on Covid-19 has been criticized. This has been particularly notable in India (Singh, 2020), where the authorities refused to listen to voices critical of their management; and in China, where official media reports were highly inconsistent with those issued in the rest of the world (Bento *et al.*, 2020; Fu; Zhu, 2020).

“ The role of social networks at the beginning of the pandemic is harshly criticized; they are accused of having massively disseminated hoaxes and unverified news, with extraordinary ease in replicating rumors ”

4. Discussion and conclusions

Below we point out the main conclusions of the results obtained from the general review of the literature.

1) The need for broad public health literacy is confirmed. It has also been noted that the current lack of public health literacy education severely limits the adoption of preventive behaviours (**Hashemi-Shahri et al.**, 2020). A cross-sectional study conducted in Vietnam at the beginning of the pandemic with 3,947 participants concludes that literacy has a protective effect against depression and improves quality of life (**Nguyen et al.**, 2020). Along these lines, there has been discussion in the health system in recent years on how to empower the population in public health decisions, as research shows how literacy and community participation in health are clearly interconnected (**McCormack et al.**, 2017). It is also recommended to involve expert health professionals to lead public literacy in the communication channels to which they should be given access.

There is also a need to inform by adapting the language to the general public. Some health organizations recommend writing information at a reading level equivalent to 2nd grade secondary education (13-14 years old) (**Badarudeen; Sa-bharwal**, 2010). Several studies highlight the usefulness of visual material for health education. The results of this review also suggest the importance of providing culturally segmented messages for each group of individuals, a strategy previously used in poor literacy contexts (**Luque**, 2018). **Airhihenbuwa et al.** (2020) propose tailoring information to a culturally appropriate framework to engage the community.

2) It highlights the relevance of social networks in the transmission of information and as potentially effective tools for communicating health emergencies, something already pointed out in previous epidemics (**Tang et al.**, 2018). On the other hand, the articles also point out the enormous proliferation of fake news in social networks, and the confusion and misinformation generated during the first wave; This confirms the need to take measures at the highest level to control the “infodemic” –the rapid spread of false or unverified news– and to incentivise the use of reliable sources of expert professionals. Previous studies had already shown this high prevalence of disinformation on social networks, as well as its capacity to induce fear, anxiety and distrust in institutions (**McDougall et al.**, 2019).

It underlines the desirability of health professionals and health and government institutions to better use their own social networks to carry out this task, giving voice to experts and using social networks as allies to disseminate immediate and truthful information.

3) It is also clear the need to strengthen institutional communication (health and government), leaving aside political and / or partisan interests in the search for culprits, as many analysts have criticized the biased use of information by some politicians and rulers, the excess of non-relevant information in official communications and the lack of measures to combat fake news. Similarly, it has been denounced that the authorities would not have been able to take advantage of a digitalized society, informing through traditional channels, which have often been of little relevance to the population.

It appears as a fundamental need to achieve adequate coordination in the management of communication in public health. This coordination should involve governments and health institutions -both internationally and at the national and local levels-, without leaving in government hands all the responsibility for the management of health communication. Prior to Covid-19, a study on social media information during the Zika epidemic (**Gui et al.**, 2018) revealed notable discrepancies between what the population is interested in or concerned about and the information provided by health authorities.

4) In the articles analyzed, there is a generally positive view of the coverage of the journalistic media during Covid-19, since they were empathetic with the suffering of the population, knew how to attenuate the tendency to alarmism and favored awareness of preventive measures. In addition, in the face of a health catastrophe such as the one that occurred, the need to raise morale and encourage the population to counteract the panic and anxiety generated by the pandemic became clear.

The authors point out the need for joint work between health authorities and the media to share reliable information and use social networks in a coordinated way, as well as to actively counter misinformation (**Harnett**, 2020).

The present systematic review highlights as a strength the use of a comprehensive search strategy (*PubMed*, *Scopus* and *Web of Science*) that allowed the exhaustive review of the articles published during the first wave of Covid-19. The study selection process was conducted considering *Prisma* guidelines. It also highlights the high number of articles included in the review, which has allowed us to understand the practical consequences in relation to public health communication for future pandemics. As limitations of the research, we must point out the inclusion of articles of different nature, and even of very different quality. However, our objective was to review all publications of scientific journals that examined communication in public health, so no type of articles was excluded. Another limitation of this review has been the time frame in which it was carried out, taking into account that some publications referring to the first wave had to be published later. Finally, we must note the limited methodological design, which had to be simpli-

“ The climate of social alarm was increased by the appearance of conspiracy theories, which led to distrust of sanitary measures ”

fied due to the exceptional production analyzed. In this sense, the large volume of articles included in this review has not allowed an integrative analysis, leaving the results obtained at a more descriptive level.

As a summary, we can conclude that the analysis of the articles on the first wave of Covid-19 has provided interesting proposals that could serve to improve public health communication in the face of future health crises. These include the presence and leadership of health professionals in the media in the face of a pandemic, and the need for adequate public health literacy among the population. It also shows the need for coordination between governments, health authorities and mass media to, from the beginning and unanimously, counteract possible social alarm and misinformation. Likewise, it is committed to the use of social networks as a habitual channel for general and frequent information about the state of the pandemic and for communication between doctors and health authorities. The networks, which in inexperienced hands caused confusion and anxiety during the first wave, must transmit clarity and serenity in authorized and reliable hands, taking the lead in communicating any incident and clearly explaining the ways to prevent contagion. For this, it will be essential to segment the audiences to which to direct the communication, and adapt the messages to an informative level, prioritizing visual communication over verbal communication. We believe that the findings reported in this review can contribute to improving the design and coordination of communication strategies in the face of future pandemics.

“The extraordinary facility of social networks to replicate rumors and unverified news generated a new concept in public health: the “infodemic” (information epidemic). The WHO adopted this term to refer to the confusing, contradictory and uncontrolled flow of news about the pandemic.”

5. References

- Abd-Alrazaq, Alaa; Alhuwail, Dari; Househ, Mowafa; Hamdi, Mounir; Shah, Zubair** (2020). “Top concerns of tweeters during the Covid-19 pandemic: a surveillance study”. *Journal of medical Internet research*, v. 22, n. 4. <https://doi.org/10.2196/19016>
- Abdel-Latif, Mohammad M. M.** (2020). “The enigma of health literacy and Covid-19 pandemic”. *Public health*, v. 185, pp. 95-96. <https://doi.org/10.1016/j.puhe.2020.06.030>
- Abel, Thomas; McQueen, David** (2020). “Critical health literacy and the Covid-19 crisis”. *Health promotion international*, v. 35, n. 6, pp. 1612-1613. <https://doi.org/10.1093/heapro/daaa040>
- Adam, Maya; Bärnighausen, Till; McMahon, Shannon A.** (2020). “Design for extreme scalability: A wordless, globally scalable Covid-19 prevention animation for rapid public health communication”. *Journal of global health*, v. 10, n. 1, 010343. <https://doi.org/10.7189/JOGH.10.010343>
- Ahmed-Siddiqui, Mohammed-Yaseen; Mushtaq, Kamran; Mohamed, Mouhand; Al Soub, Hussam; Hussein-Mohamedali, Mohamed-Gaafar; Yousaf, Zohaib** (2020). “Social media misinformation - An epidemic within the Covid-19 pandemic”. *American journal of tropical medicine and hygiene*, v. 103, n. 2, pp. 920-921. <https://doi.org/10.4269/ajtmh.20-0592>
- Airhienbuwa, Collins O.; Iwelunmor, Juliet; Munodawafa, Davison; Ford, Chandra; Oni, T.; Ebenezer; Agyemang, Charles; Castaño-De-la-Mota, C.; Ikuomola, B. S.; Simbayi, Leickness; Fallah, Mosoka-Papa; Quian, Zhengmin; Ma-kinwa, B.; Niang, Cheikh-Ibrahima; Okosun, Ike** (2020). “Culture matters in communicating the global response to Covid-19”. *Preventing chronic disease*, v. 17, E60. <https://doi.org/10.5888/PCD17.200245>
- Ali, Muhammad-Yousuf; Bhatti, Rubina** (2020). “Covid-19 (Coronavirus) pandemic: Information sources channels for the public health awareness”. *Asia-pacific journal of public health*, v. 32, n. 4, pp. 168-169. <https://doi.org/10.1177/1010539520927261>
- Amin, Saquib** (2020). “Why ignore the dark side of social media? A role of social media in spreading corona-phobia and psychological well-being”. *International journal of mental health promotion*, v. 22, n. 1, pp. 29-37. <https://doi.org/10.32604/IJMHP.2020.011115>
- Armitage, Laura; Lawson, Beth K.; Whelan, Maxin E.; Newhouse, Nikki** (2020). “Paying special consideration to the digital sharing of information during the Covid-19 pandemic and beyond”. *British journal of general practice*, v. 4, n. 2. <https://doi.org/10.3399/BJGPOPEN20X101072>
- Arora, Teresa; Grey, Ian** (2020). “Health behaviour changes during Covid-19 and the potential consequences: A mini-review”. *Journal of health psychology*, v. 25, n. 9, pp. 1155-1163. <https://doi.org/10.1177/1359105320937053>

- Badarudeen, Sameer; Sabharwal, Sanjeev** (2010). "Assessing readability of patient education materials: current role in orthopaedics". *Clinical orthopaedics and related research*, v. 468, n. 10, pp. 2572-2580.
<https://doi.org/10.1007/s11999-010-1380-y>
- Balhara, Yatan-Pal-Singh; Chandiok, Ketaki** (2020). "Can #PlayOurpartTogether help prevent miscommunication?". *Asian journal of psychiatry*, v. 52, 102123.
<https://doi.org/10.1016/j.ajp.2020.102123>
- Bao, Huanyu; Cao, Bolin; Xiong, Yuan; Tang, Weiming** (2020). "Digital media 's role in the Covid-19 pandemic". *Journal of medical internet research, mhealth uhealth*, v. 8, n. 9, pp. 16-18.
<https://doi.org/10.2196/20156>
- Barile, Nello; Bovalino, Guerino-Nuccio** (2020). "Empathic media, emotions and gamification: from the fourth industrial revolution to the postpandemic society". *H-Ermes-journal of communication*, n. 16, pp. 7-26.
<https://doi.org/10.1285/i22840753n16p7>
- Basch, Corey H.; Hillyer, Grace C.; Meleo-Erwin, Zoe C.; Jaime, Christi; Mohlman, Jan; Basch, Charles E.** (2020a). "Preventive behaviors conveyed on YouTube to mitigate transmission of Covid-19: cross-sectional study". *Journal of medical internet research, public health and surveillance*, v. 6, n. 2, e18807.
<https://doi.org/10.2196/18807>
- Basch, Corey H.; Mohlman, Jan; Hillyer, Grace C.; Garcia, Philip** (2020b). "Public health communication in time of crisis: readability of on-line Covid-19 information". *Disaster medicine and public health preparedness*, v. 14, n. 5, pp. 635-637.
<https://doi.org/10.1017/dmp.2020.151>
- Bastani, Peivand; Bahrami, Mohamed Amin** (2020). "Covid-19 related misinformation on social media: a qualitative study from Iran". *Journal of medical internet research*. PMID: 32250961.
<https://doi.org/10.2196/18932>
- Bento, Ana; Nguyen, Thuy; Wing, Coady; Lozano-Rojas, Felipe; Ahn, Yong-Yeol; Simon, Kosali** (2020). "Evidence from internet search data shows information-seeking responses to news of local Covid-19 cases". *Proceedings of the National Academy of Sciences*, v. 117, n. 21, pp. 11220-11222.
<https://doi.org/10.1073/pnas.2005335117>
- Bernhardt, Jay M.** (2004). "Communication at the core of effective public health". *American journal of public health*, v. 94, n. 12, pp. 2051-2053.
<https://doi.org/10.2105/ajph.94.12.2051>
- Bilal, Sharif; Latif, Faiza; Bashir, Muhammad-Farhan; Komal, Bushra; Tan, Duoqiao** (2020). "Role of electronic media in mitigating the psychological impacts of novel coronavirus (Covid-19)". *Psychiatry research*, n. 289, 113041.
<https://doi.org/10.1016/j.psychres.2020.113041>
- Bilbatua-Andrés, Jon; Mudaber, Jawad; Jiménez-Vargas, Dennis; Arrieta-Angulo, Brandon** (2020). "Sobre las estrategias y campañas de comunicación en salud pública: ¿Qué sabes del 2019-nCoV?". *Revista española de salud pública*, v. 94, e202005052.
- Brivio, Eleonora; Oliveri, Serena; Pravettoni, Gabriela** (2020). "Empowering communication in emergency contexts: Reflections from the Italian coronavirus outbreak". *Mayo Clinic proceedings*, v. 95, n. 5, pp. 849-851.
<https://doi.org/10.1016/j.jmayocp.2020.03.021>
- 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>
- Castro-Martínez, Andrea; Morán-Urdiales, Ignacio; Díaz-Morilla, Pablo** (2021). "Comunicación institucional de crisis y redes sociales en la Covid-19: los casos de Yunquera y Gaucín". *Ámbitos. Revista internacional de comunicación*, n. 52, pp. 102-123.
<https://doi.org/10.12795/Ambitos.2021.i52.07>
- Choukou, Mohamed-Amine; Sánchez-Ramírez, Diana C.; Pol, Margriet; Uddin, Mohy; Monnin, Caroline; Syed-Abdul, Shabbir** (2022). "Covid-19 infodemic and digital health literacy in vulnerable populations: A scoping review". *Digital health*, v. 8.
<https://doi.org/10.1177/20552076221076927>
- Cox, Caitríona L.** (2020). "Healthcare heroes': Problems with media focus on heroism from healthcare workers during the Covid-19 pandemic". *Journal of medical ethics*, v. 46, n. 8, pp. 510-513.
<https://doi.org/10.1136/medethics-2020-106398>
- Cowper, Andy** (2020). "Budgeting for Covid-19: changing the narrative and narrating the change". *The British Medical Association, medicine and the media*, n. 368, m1053.
<https://doi.org/10.1136/bmj.m1053>

- Crespo-Martínez, Ismael; Garrido, Antonio** (2020). "La pandemia del coronavirus: estrategias de comunicación de crisis". *Más poder local*, v. 41, pp. 12-19.
- Cuan-Baltazar, José-Yunam; Muñoz-Pérez, María-José; Robledo-Vega, Carolina; Pérez-Zepeda, María-Fernanda; Soto-Vega, Elena** (2020). "Misinformation of Covid-19 on the internet: infodemiology study". *Journal of medical internet research, public health and surveillance*, v. 6, n. 2, e18444.
<https://doi.org/10.2196/18444>
- Chan, Albert K. M.; Nickson, Christopher P.; Rudolph, Jenny W.; Lee, Anna; Joynt, Gavin M.** (2020). "Social media for rapid knowledge dissemination: early experience from the Covid-19 pandemic". *Anaesthesia*, v. 75, n. 12, pp. 1579-1582.
<https://doi.org/10.1111/anae.15057>
- Chang, Xinghua; Liu, Maoxing; Jin, Zhen; Wang, Jianrong** (2020). "Studying the impact of media coverage on the spread of Covid-19 in Hubei province, China". *Mathematical biosciences and engineering*, v. 17, n. 4, pp. 3147-3159.
<https://doi.org/10.3934/MBE.2020178>
- Chen, Emily; Lerman, Kristina; Ferrara, Emilio** (2020). "Tracking social media discourse about the Covid-19 pandemic: Development of a public coronavirus Twitter data set". *Journal of medical internet research, public health and surveillance*, v. 6, n. 2, e19273.
<https://doi.org/10.2196/19273>
- Chiodini, Jane** (2020). "Maps, masks and media - Traveller and practitioner resources for 2019 novel coronavirus (2019-nCoV) acute respiratory virus". *Travel medicine and infectious disease*, n. 33, 101574.
<https://doi.org/10.1016/j.tmaid.2020.101574>
- Chong, Yuen-Yu; Cheng, Ho-yu; Chan, Helen-Yue-Lai; Chien, Wai-Tong; Wong, Samuel-Yeung-Shan** (2020). "Covid-19 pandemic, infodemic and the role of eHealth literacy". *International journal of nursing studies*, n. 108, 103644.
<https://doi.org/10.1016/j.ijnurstu.2020.103644>
- De-Coninck, David; D'Haenens, Leen; Matthijs, Koen** (2020). "Perceived vulnerability to disease and attitudes towards public health measures: Covid-19 in Flanders, Belgium". *Personality and individual differences*, v. 166, 110220.
<https://doi.org/10.1016/j.paid.2020.110220>
- De-la-Hoz, Karoline F.** (2014). "La comunicación en crisis sanitarias: perspectiva de las administraciones públicas". *Revista española de comunicación en salud*, v. 5, n. 1, pp. 14-19.
- Dheeraj, Khatod** (2020). "Analysing Covid-19 news impact on social media aggregation". *International Journal of advanced trends in computer science and engineering*, v. 9, n. 3, pp. 2848-2855.
<https://doi.org/10.30534/ijatcse/2020/56932020>
- Dkhar, Sabira-Aalia; Quansar, Ruqia; Saleem, Sheikh-Mohd; Khan, Muhammad-Salim** (2020). "Knowledge, attitude, and practices related to Covid-19 pandemic among social media users in J & K, India". *Indian journal of public health*, v. 64, n. 6, pp. 205-210.
https://doi.org/10.4103/ijph.IJPH_469_20
- D'Souza, Ryan S.; D'Souza, Sawn; Strand, Natalie; Anderson, Alexandra; Vogt, Matthew N. P.; Olatoye, Oludare** (2020). "YouTube as a source of medical information on the novel coronavirus 2019 disease (Covid-19) pandemic". *Global public health*, v. 15, n. 7, pp. 935-942.
<https://doi.org/10.1080/17441692.2020.1761426>
- Ekzayez, Abdulkarim; Al-Khalil, Munzer; Jasiem, Mohamad; Al-Saleh, Raed; Alzoubi, Zedoun; Meagher, Kristen; Patel, Preeti** (2020). "Covid-19 response in Northwest Syria: Innovation and community engagement in a complex conflict". *Journal of public health (United Kingdom)*, v. 42, n. 3, pp. 504- 509.
<https://doi.org/10.1093/pubmed/fdaa068>
- El-Jardali, Fadi; Bou-Karroum, Lama; Fadlallah, Racha** (2020). "Amplifying the role of knowledge translation platforms in the Covid-19 pandemic response". *Health research policy and systems*, v. 18, n. 58.
<https://doi.org/10.1186/s12961-020-00576-y>
- Eysenbach, Gunther** (2020). "How to fight an infodemic: the four pillars of infodemic management". *Journal of medical internet research*, v. 22, n. 6, e21820.
<https://doi.org/10.2196/21820>
- Finset, Arnstein; Bosworth, Hayden; Butow, Philip; Gulbrandsen, Pål; Hulsman, Robert L.; Pieterse, Arwen H.; Robin; Street, Richard; Tschoetschel, Rubin; Van-Weert, Julia** (2020). "Effective health communication - A key factor in fighting the Covid-19 pandemic. *Patient education and counseling*, v. 103, n. 5, pp. 873-876.
<https://doi.org/10.1016/j.pec.2020.03.027>

- Fitz-Maurice, Mario; Di-Tommaso, Fernando; Baranchuk, Adrián** (2020). "Efecto de una intervención en las redes sociales en tiempos del Covid-19". *Medicina (Buenos Aires)*, v. 80, n. 3, pp. 310.
https://www.researchgate.net/publication/345770369_Efecto_de_una_intervencion_en_las_redes_sociales_en_tiempos_del_COVID-19
- Frieden, Thomas R.**; *Centers for Disease Control and Prevention (CDC)* (2014). "Strategies for reducing health disparities - Selected CDC-sponsored interventions, United States, 2014. Foreword". *MMWR supplements*, v. 63, n. 1.
- Fu, King-Wa; Zhu, Yuner** (2020). "Did the world overlook the media's early warning of Covid-19?". *Journal of risk research*, v. 23, n. 7-8, pp. 1047-1051.
<https://doi.org/10.1080/13669877.2020.1756380>
- Glick, Michael** (2007). "The health of the nation. Why you should care". *JADA*, v. 138, n. 2.
<https://doi.org/10.14219/jada.archive.2007.0115>
- Gollust, Sara E.; Nagler, Rebekah H.; Fowler, Erika-Franklin** (2020). "The emergence of Covid-19 in the US: a public health and political communication crisis". *Journal of health politics, policy and law*, v. 45, n. 6, pp. 967-981.
<https://doi.org/10.1215/03616878-8641506>
- González-Padilla, Daniel A.; Tortolero-Blanco, Leonardo** (2020). "Social media influence in the Covid-19 pandemic". *International Brazilian journal of urology*, v. 46, Supl. 1, pp. 120-124.
<https://doi.org/10.1590/S1677-5538.IBJU.2020.S121>
- Gottlieb, Michael; Dyer, Sean** (2020). "Information and disinformation: social media in the Covid-19 crisis". *Academic emergency medicine*, v. 27, n. 7, pp. 640-641.
<https://doi.org/10.1111/acem.14036>
- Gray, Natan A.; Back, Anthony L.** (2020). "Covid-19 communication aids". *The British Medical Association*, n. 369, m2255.
<https://doi.org/10.1136/bmj.m2255>
- Guest, Jodie L.; Del-Río, Carlos; Sánchez, Travis** (2020). "The three steps needed to end the Covid-19 pandemic: bold public health leadership, rapid innovations, and courageous political will". *Journal of medical internet research, public health and surveillance*, v. 6, n. 2, e19043.
<https://doi.org/10.2196/19043>
- Gui, Xinning; Wang, Yue; Kou, Yubi; Reynolds, Tera-Leigh; Chen, Yunan; Mei, Qiaozhu; Zheng, Kai** (2018). "Understanding the patterns of health information dissemination on social media during the Zika outbreak". In: *AMIA annual symposium proceedings, eCollection 2017*, pp. 820-829. PMID: 29854148.
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5977662/>
- Habersaat, Katrine; Betsch, Cornelia; Danchin, Margie; Sunstein, Cass R.; Böhm, Robert; Falk, Armin; Brewer, Noel T.; Omer, Saad B.; Scherzer, Martha; Sah, Sunita; Fischer, Edward F.; Scheel, Andrea E. et al.** (2020). "Ten considerations for effectively managing the Covid-19 transition". *Nature human behaviour*, v. 4, pp. 677-687.
<https://doi.org/10.1038/s41562-020-0906-x>
- Han, Xuehua; Wang, Juanle; Zhang, Min; Wang, Xiaojie** (2020). "Using social media to mine and analyze public opinion related to Covid-19 in China". *International journal of environmental research and public health*, v. 17, n. 8.
<https://doi.org/10.3390/ijerph17082788>
- Harnett, Susan** (2020). "Health literacy, social media and pandemic planning". *Journal of consumer health on the internet*, v. 24, n. 2, pp. 157-162.
<https://doi.org/10.1080/15398285.2020.1756677>
- Haroon, Omair; Rizvi, Syed-Aun** (2020). "Covid-19: media coverage and financial markets behavior - A sectoral inquiry". *Journal of behavioral and experimental finance*, v. 27.
<https://doi.org/10.1016/j.jbef.2020.100343>
- Hashemi-Shahri, Seyed-Mohammad; Khammarnia, Mohammad; Ansari-Moghaddam, Alireza; Setoodehzadeh, Fate-meh; Okati-Aliabad, Hassan; Peyvand, Mostafa** (2020). "Sources of news as a necessity for improving community health literacy about Covid-19". *Medical journal of the Islamic Republic of Iran*, v. 34, n. 63.
<https://doi.org/10.34171/mjiri.34.63>
- Hashim, Suhaizal; Masek, Alias; Abdullah, Nurhaim-Saadah; Paimin, Aini-Nazura; Muda, Wan-Hanim** (2020). "Students' intention to share information via social media: a case study of Covid-19 pandemic". *Indonesian journal of science and technology*, v. 5, n. 2, pp. 236-245.
<https://doi.org/10.17509/ijost.v5i2.24586>

- He, Susan; Ojo, Ayatomiwa; Beckman, Adam L.; Gondi, Suhas; Ranney, Megan; Betz, Marian; Faust, Jeremy S.; Choo, Esther; Kass, Dara; Raja, Ali S.** (2020). "The story of #GetMePPE and #GetUsPPE.org to mobilize health care response to Covid-19 : rapidly deploying digital tools for better health care". *Journal of medical internet research*, v. 22, n. 7, e20469. <https://doi.org/10.2196/20469>
- Hu, Dingtao; Lou, Xiaoqi; Xu, Zhiwei, Xu; Nana, Meng; Qiaomei, Xie; Man, Zhang; Yanfeng, Zou; Jiatao, Liu; Guoping, Sun; Wang, Fang** (2020). "More effective strategies are required to strengthen public awareness of Covid-19: evidence from Google Trends". *Journal of global health*, v. 10, n. 1. <https://doi.org/10.7189/jogh.10.011003>
- Huang, Hao; Peng, Zongchao; Wu, Hongtao; Xie, Qihui** (2020). "A big data analysis on the five dimensions of emergency management information in the early stage of Covid-19 in China". *Journal of Chinese governance*, v. 5, n. 2, pp. 213-233. <https://doi.org/10.1080/23812346.2020.1744923>
- Huddart, Daniel; Hirniak, Johnathan; Sethi, Rajiv; Hayer, Gursharun; Dibblin, Connor; Meghna-Rao, Biyyam; Zaman, Mohammed-Ehsaanuz; Jenkins, Cameron; Hueso, Benjamin; Sethi, Sunil** (2020). "#MedStudentCovid: how social media is supporting students during Covid-19". *Medical education*, v. 54, n. 10, pp. 951-952. <https://doi.org/10.1111/medu.14215>
- Hunter, Philip** (2020). "The growth of social media in science". *European Molecular Biology Organization, reports*, v. 21, n. 5, e50550. <https://doi.org/10.15252/embr.202050550>
- Iacobucci, Gareth** (2020). "Covid-19: diabetes clinicians set up social media account to help alleviate patients' fears". *British Medical Association (Clinical research ed.)*, n. 368, m1262. <https://doi.org/10.1136/bmj.m1262>
- Ippolito, Giuseppe; Hui, David S.; Ntoumi, Francine; Maeurer, Markus; Zumla, Alimuddin** (2020). "Toning down the 2019-nCoV media hype - and restoring hope". *The lancet. Respiratory medicine*, v. 8, n. 3, pp. 230-231. [https://doi.org/10.1016/S2213-2600\(20\)30070-9](https://doi.org/10.1016/S2213-2600(20)30070-9)
- Isip-Tan, Iris-Thiele; Gutiérrez, Jerico; Bernardo, Diane-Carla** (2020). "Use of Facebook to serve information needs of persons with diabetes in the Philippines amid the covid-19 pandemic". *Journal of the Asean Federation of Endocrine Societies*, v. 35, n. 1, pp. 32-37. <https://doi.org/10.15605/jafes.035.01.09>
- Jain, Somya; Sinha, Adwitiya** (2020). "Identification of influential users on Twitter: a novel weighted correlated influence measure for Covid-19". *Chaos, solitons and fractals*, v. 139. <https://doi.org/10.1016/j.chaos.2020.110037>
- Jayawardena, Asitha D. L.; Romano, Sarah-Romano; Callans, Kevin; Fracchia, M. Shannon; Hartnick, Christopher J.** (2020). "Family-centered information dissemination: a multidisciplinary virtual Covid-19 'Town Hall'". *Otolaryngology-head and neck surgery (United States)*, v. 163, n. 5, pp. 929-930. <https://doi.org/10.1177/0194599820935419>
- Juvinyà-Canal, Dolors; Bertran-Noguer, Carme; Suñer-Soler, Rosa** (2018). "Alfabetización para la salud, más que información". *Gaceta sanitaria*, v. 32, n. 1, pp. 8-10. <https://doi.org/10.1016/j.gaceta.2017.07.005>
- Kamenidou, Irene; Stavrianea, Aikaterini; Liava, Chistina** (2020). "Achieving a Covid-19 free country: citizens preventive measures and communication pathways". *International journal of environmental research and public health*, v. 17, n. 13. <https://doi.org/10.3390/ijerph17134633>
- Karamouzian, Mohammad; Johnson, Cheyenne; Kerr, Thomas** (2020). "Public health messaging and harm reduction in the time of Covid-19". *The lancet psychiatry*, v. 7, n. 5, pp. 390-391. [https://doi.org/10.1016/S2215-0366\(20\)30144-9](https://doi.org/10.1016/S2215-0366(20)30144-9)
- Karasneh, Reema; Al-Azzam, Sayer; Muflih, Suhaib; Soudah, Ola; Hawamdeh, Sahar; Khader, Yousef** (2021). "Media's effect on shaping knowledge, awareness, risk perceptions and communication practices of pandemic Covid-19 among pharmacists". *Research in social and administrative pharmacy*, v. 17, n. 1, pp. 1897-1902. <https://doi.org/10.1016/j.sapharm.2020.04.027>
- Katapally, Tarun R.** (2020). "A global digital citizen science policy to tackle pandemics like Covid-19". *Journal of medical internet research*, v. 22, n. 5. <https://doi.org/10.2196/19357>

- Kauffman, Lilly; Fishman, Elliot K.; Mahesh, Mahadevappa; Weisberg, Edmund M.** (2020). "Breaking news: using Facebook Live to transmit radiologic information quickly on a global scale". *Journal of the American College of Radiology*, v. 17, n. 7, pp. 899-902.
<https://doi.org/10.1016/j.jacr.2020.05.013>
- Kluger, Nicolas; Scrivener, Jean-Nicolas** (2020). "The use of Google Trends for acral symptoms during Covid-19 outbreak in France". *Journal of the European Academy of Dermatology and Venereology*, v. 34, n. 8, pp. e358-e360.
<https://doi.org/10.1111/jdv.16572>
- Ko, Nai-Ying; Lu, Wei-Hsin; Chen, Yi-Lung; Li, Dian-Jeng; Wang, Peng-Wei; Hsu, Su-Ting; Chen, Chang-Chun; Lin, Yi-Hsuan; Chang, Yu-Ping; Yen, Cheng-Fang** (2020). "Covid-19-related information sources and psychological well-being: an online survey study in Taiwan". *Brain, behavior, and immunity*, v. 87, pp. 153-154.
<https://doi.org/10.1016/j.bbi.2020.05.019>
- Kocigit, Burhan-Fatih; Akaltun, Mazlum-Serdar; Sahin, Ahmet-Riza** (2020). "YouTube as a source of information on Covid-19 and rheumatic disease link". *Clinical rheumatology*, v. 39, pp. 2049-2054.
<https://doi.org/10.1007/s10067-020-05176-3>
- Kulkarni, Praveen; Prabhu, Sudhir; Kumar, Sunil; Ramraj, Balaji** (2020). "Covid-19-infodemic overtaking pandemic?. Time to disseminate facts over fear". *Indian journal of community health*, v. 32, n. 2 (suppl.), pp. 264-268.
<https://doi.org/10.47203/ijch.2020.v32i02supp.018>
- La, Viet-Phuong; Pham, Thanh-Hang; Ho, Manh-Toan; Nguyen, Minh-Hoang; Nguyen, Khanh-Linh P.; Vuong, Thu-Trang; Nguyen, Hong-Kong T.; Tran, Trung; Khuc, Quy; Ho, Manh-Tung; Vuong, Quan-Hoang** (2020). "Policy response, social media and science journalism for the sustainability of the public health system amid the Covid-19 outbreak: the Vietnam lessons". *Sustainability*, v. 12, n. 7.
<https://doi.org/10.3390/su12072931>
- Laufer, Daniel** (2020). "Academics engaging through the media - Insights from creating a monthly column on crisis management". *Public relations review*, v. 47, n. 1, 101932.
<https://doi.org/10.1016/j.pubrev.2020.101932>
- Lazcano-Ponce, Eduardo; Alpuche-Aranda, Celia** (2020). "Public health literacy in the face of the Covid-19 pandemic emergency". *Salud pública de México*, v. 62, n. 3, pp. 331-340.
<https://doi.org/10.21149/11408>
- Leng, Ho-Keat; Phua, Yi-Xian-Philip** (2020). "Athletes as role models during the Covid-19 pandemic". *Managing sport and leisure*, v. 27, n. 1-2, pp. 163-167.
<https://doi.org/10.1080/23750472.2020.1762330>
- Liao, Qiuyan; Yuan, Jiehu; Dong, Meihong; Yang, Lin; Fielding, Richard; Tak-Lam, Wendy-Wing** (2020). "Public engagement and government responsiveness in the communications about Covid-19 during the early epidemic stage in China: infodemiology study on social media data". *Journal of medical internet research*, v. 22, n. 5.
<https://doi.org/10.2196/18796>
- Limaye, Rupali-Jayant; Sauer, Molly; Ali, Joseph; Bernstein, Justin; Wahl, Brian; Barnhill, Anne; Labrique, Alain** (2020). "Building trust while influencing online Covid-19 content in the social media world". *The lancet digital health*, v. 2, n. 6, e277-e278.
[https://doi.org/10.1016/S2589-7500\(20\)30084-4](https://doi.org/10.1016/S2589-7500(20)30084-4)
- Lima, Diego-Laurentino; Albanex-A-De-Medeiros-Lopes, Maria-Antonieta; Brito, Ana-María** (2020). "Social media: friend or foe in the Covid-19 pandemic?". *Clinics (São Paulo)*, n. 75, e1953.
<https://doi.org/10.6061/clinics/2020/e1953>
- Lin, Yu-Hsuan; Liu, Chun-Hao; Chiu, Yu-Chuan** (2020). "Google searches for the keywords of "wash hands" predict the speed of national spread of Covid-19 outbreak among 21 countries". *Brain, behavior, and immunity*, v. 87, pp. 30-32.
<https://doi.org/10.1016/j.bbi.2020.04.020>
- Liu, Piper-Liping** (2020). "Covid-19 information seeking on digital media and preventive behaviors: the mediation role of worry". *Cyberpsychology, behavior, and social networking*, v. 23, n. 1, pp. 677-682.
<https://doi.org/10.1089/cyber.2020.0250>
- Liu, Qian; Zheng, Zequan; Zheng, Jiabin; Chen, Qiuyi; Liu, Guan; Chen, Sihan; Chu, Bojia; Zhu, Hongyu; Akinwunmi, Babatunde; Huang, Jian; Zhang, Casper J. P.; Ming, Wai-Kit** (2020). "Health communication through news media during the early stage of the Covid-19 outbreak in China: digital topic modeling approach". *Journal of medical internet research*, v. 22, n. 4, e19118.
<https://doi.org/10.2196/19118>

- Liu, Xi; Luo, Wen-Tao; Li, Ying; Li, Chun-Na; Hong, Zhong-Si; Chen, Hui-Li; Xiao, Fei; Xia, Jin-Yu** (2020). "Psychological status and behavior changes of the public during the Covid-19 epidemic in China". *Infectious diseases of poverty*, v. 9, n. 1. <https://doi.org/10.1186/s40249-020-00678-3>
- Lohiniva, Anna-Leena; Sane, Jussi; Sibenberg, Katja; Puumalainen, Taneli; Salminen, Mika** (2020). "Understanding coronavirus disease (Covid-19) risk perceptions among the public to enhance risk communication efforts: a practical approach for outbreaks, Finland, february 2020". *Eurosurveillance*, v. 25, n. 13, pp. 3-6. <https://doi.org/10.2807/1560-7917.ES.2020.25.13.2000317>
- Looi, Jeffrey C. L.; Allison, Stephen; Bastiampillai, Tarun; Maguire, Paul** (2020). "Fire, disease and fear: effects of the media coverage of 2019-2020 Australian bushfires and novel coronavirus 2019 on population mental health". *Australian and New Zealand journal of psychiatry*, v. 54, n. 9, pp. 938-939. <https://doi.org/10.1177/0004867420931163>
- López, Taylarr** (2020). "Building local health department Covid-19 emergency and risk communications capacity". *Journal of public health management and practice*, v. 26, n. 4, pp. 384-386. <https://doi.org/10.1097/PHH.0000000000001196>
- Lu, Yue; Zhang, Leiliang** (2020). "Social media WeChat infers the development trend of Covid-19". *Journal of infection*, v. 81, n. 1, e82-e83. <https://doi.org/10.1016/j.jinf.2020.03.050>
- Luque, John S.** (2018). "The faces on our educational materials: real stories behind the messages". *Journal of cancer education*, v. 33, pp. 242-244. <https://doi.org/10.1007/s13187-016-1018-1>
- Luu, Toan** (2020). "The Covid-19 risk perception: a survey on socioeconomics and media attention". *Economics bulletin*, v. 40, n. 1, pp. 758-764. <https://EconPapers.repec.org/RePEc:ebl:ecbull:eb-20-00175>
- Mahima, B. N.; Tiwari, Herman Kr; Mahapatra, Payel; Amudhan, Senthil; Rao, Girish N.** (2020). "Covid-19 epidemiology: through the eyes of vernacular newspapers". *Indian journal of public health*, v. 64, n. 6, pp. 217-220. https://doi.org/10.4103/ijph.IJPH_485_20
- Malecki, Kristen M. C.; Keating, Julie A.; Safdar, Nasia** (2020). "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>
- Martínez-Estrella, Eva-Citlali** (2020). "Uso de personajes y metáforas en la gestión de la crisis sanitaria del Covid-19. Revisión de la comunicación de sanidad pública en España y México". *Revista española de comunicación en salud*, Supl. 1, pp. 319-327.
- Massaad, Elie; Cherfan, Patrick** (2020). "Social media data analytics on telehealth during the Covid-19 pandemic". *Cureus*, v. 12, n. 4, e7838. <https://doi.org/10.7759/cureus.7838>
- McCormack, Lauren; Thomas, Verónica; Lewis, Megan A.; Rudd, Rima** (2017). "Improving low health literacy and patient engagement: a social ecological approach". *Patient education and counseling*, v. 100, n. 1, pp. 8-13. <https://doi.org/10.1016/j.pec.2016.07.007>
- McDougall, Julian; Brites, Maria-José; Couto, Maria-João; Lucas, Catarina** (2019). "Digital literacy, fake news and education". *Cultura y educación*, v. 31, n. 2, pp. 203-212. <https://doi.org/10.1080/11356405.2019.1603632>
- Merchant, Raina M.; Lurie, Nicole** (2020). "Social media and emergency preparedness in response to novel coronavirus". *JAMA - Journal of the American Medical Association*, v. 323, n. 20, pp. 2011-2012. <https://doi.org/10.1001/jama.2020.4469>
- Mheidly, Nour; Fares, Jawad** (2020). "Leveraging media and health communication strategies to overcome the Covid-19 infodemic". *Journal of public health policy*, v. 41, n. 4, pp. 410-420. <https://doi.org/10.1057/s41271-020-00247-w>
- Mirbabaie, Milad; Bunker, Deborah; Stieglitz, Stefan; Marx, Julian; Ehnis, Christian** (2020). "Social media in times of crisis: learning from hurricane Harvey for the coronavirus disease 2019 pandemic response". *Journal of information technology*, v. 35, n. 3, pp. 195-213. <https://doi.org/10.1177/026839622092925>
- Mohamad, Siti-Mazidah** (2020). Creative production of 'Covid-19 social distancing' narratives on social media. *Tijdschrift voor economische en sociale geografie*, v. 111, n. 3, pp. 347-359. <https://doi.org/10.1111/tesg.12430>

- Moher, David; Liberati, Alessandro; Tetzlaff, Jennifer; Altman, Douglas G. (2009).** "Preferred reporting items for systematic reviews and meta-analyses: The Prisma statement". *PLoS medicine*, v. 339, b2535.
<https://doi.org/10.1136/bmj.b2535>
- Mulrennan, Siobhain; Colt, Henri (2020).** "Medical information and social media in the time of Covid-19". *Respirology*, v. 25, n. 6, pp. 578-579.
<https://doi.org/10.1111/resp.13832>
- Naeem, Salman-Bin; Boulos, Maged N. Kamel (2021)** "Covid-19 misinformation online and health literacy: a brief overview". *International journal of environmental research and public health*, v. 18, n. 15, 8091.
<https://doi.org/10.3390/ijerph18158091>
- Nazir, Mehrab; Hussain, Iftikhar; Tian, Jian; Akram, Sabahat; Tshiaba, Sidney-Mangenda; Mushtaq, Shahrukh; Shad, Muhammad-Afzal (2020).** "A multidimensional model of public health approaches against Covid-19". *International journal of environmental research and public health*, v. 17, n. 11.
<https://doi.org/10.3390/ijerph17113780>
- Nesbitt, Cassie; Rath, Louise; Yeh, Wei Zhen; Zhong, Michael; Wesselingh, Robb; Monif, Mastura; Richards, Janene; Minh, Viet-Bui; Jokubaitis, Vilija G.; Skibina, Olga; Butzkueven, Helmut; Van-der-Walt, Anneke (2020).** "MsCovid19: using social media to achieve rapid dissemination of health information". *Multiple sclerosis and related disorders*, v. 45, n. 1, 102338.
<https://doi.org/10.1016/j.msard.2020.102338>
- Nguyen, Hoang C.; Nguyen, Minh H.; Do, Binh N.; Tran, Cuong Q.; Nguyen, Thao T. P.; Pham, Khue M.; Pham, Linh V.; Tran, Khanh V.; Duong, Trang T.; Tran, Tien V.; Duong, Thai H.; Nguyen, Tham T. et al. (2020).** "People with suspected Covid-19 symptoms were more likely depressed and had lower health-related quality of life: the potential benefit of health literacy". *Journal of clinical medicine*, v. 9, n. 4.
<https://doi.org/10.3390/jcm9040965>
- O'Brien, Michael; Moore, Kathleen; McNicholas, Fiona (2020).** "Social media spread during Covid-19: the pros and cons of likes and shares". *Irish medical journal*, v. 113, n. 4, pp. 52-54.
<https://imj.ie/social-media-spread-during-covid-19-the-pros-and-cons-of-likes-and-shares>
- Oh, Sang-Hwa; Lee, Seo-Yoon; Han, Changhyun (2020).** "The effects of social media use on preventive behaviors during infectious disease outbreaks: the mediating role of self-relevant emotions and public risk perception". *Health communication*, v. 36, n. 8, pp. 972-981.
<https://doi.org/10.1080/10410236.2020.1724639>
- Olaimat, Amin; Aolymat, Ilman; Shahbaz, Hafiz; Holley, Richard (2020).** "Knowledge and information sources about Covid-19 among university students in Jordan: A cross-sectional study. *Frontiers in public health*, v. 8, n. 254.
<https://doi.org/10.3389/fpubh.2020.00254>
- Paakkari, Leena; Okan, Orkan (2020).** "Covid-19: health literacy is an underestimated problem". *The lancet public health*, v. 5, n. 5, e249-e250.
[https://doi.org/10.1016/S2468-2667\(20\)30086-4](https://doi.org/10.1016/S2468-2667(20)30086-4)
- Park, Han-Woo; Park, Sejung; Chong, Miyoung (2020).** "Conversations and medical news frames on Twitter: Infodemiological study on Covid-19 in South Korea". *Journal of medical internet research*, v. 22, n. 5.
<https://doi.org/10.2196/18897>
- Pennycook, Gordon; McPhetres, Jonathon; Zhang, Yunhao; Lu, Jackson G.; Rand, David G. (2020).** "Fighting Covid-19 misinformation on social media: experimental evidence for a scalable accuracy-nudge intervention". *Psychological science*, v. 31, n. 7, pp. 770-780.
<https://doi.org/10.1177/0956797620939054>
- Raamkumar, Aravind-Sesagiri; Tan, Soon-Guan; Wee, Hwee-Lin (2020).** "Use of health belief model-based deep learning classifiers for Covid-19 social media content to examine public perceptions of physical distancing: model development and case study". *Journal of medical internet research, public health and surveillance*, v. 6, n. 3.
<https://doi.org/10.2196/20493>
- Rashid, Md-Tahmid; Wang, Dong (2021).** "CovidSens: a vision on reliable social sensing for Covid-19". *Artificial intelligence review*, v. 54.
<https://doi.org/10.1007/s10462-020-09852-3>
- Rathore, Farooq-Azam; Farooq, Fareeha (2020).** "Information overload and infodemic in the Covid-19 pandemic". *Journal of the Pakistan Medical Association*, v. 70, suppl. 3, n. 5, S162-S165.
<https://doi.org/10.5455/JPMA.38>

- Ratzan, Scott C.; Sommariva, Silvia; Rauh, Lauren** (2020). "Enhancing global health communication during a crisis: lessons from the Covid-19 pandemic". *Public health research & practice*, v. 30, n. 2, e3022010.
<https://doi.org/10.17061/phrp3022010>
- Roberts, Hal; Seymour, Brittany; Fish, Sands-Alden; Robinson, Emily; Zuckerman, Ethan** (2017). "Digital health communication and global public influence: a study of the ebola epidemic". *Journal of health communication*, v. 22, sup1, pp. 51-58.
<https://doi.org/10.1080/10810730.2016.1209598>
- Romanò, Carlo-Luca; Drago, Lorenzo; Del-Sel, Hernán; Johari, Ashok; Lob, Guenter; Mavrogenis, Andreas F.; Benzakour, Thami; Waiot** (2020). "Loud and silent epidemics in the third millennium: tuning-up the volume". *International orthopaedics*, v. 44, n. 6, pp. 1019-1022.
<https://doi.org/10.1007/s00264-020-04608-8>
- Rommer, Derek; Majerova, Jana; Machová, Veronika** (2020). "Repeated Covid-19 pandemic-related media consumption: minimizing sharing of nonsensical misinformation through health literacy and critical thinking". *Linguistic and philosophical investigations*, v. 19, pp. 107-113.
<https://doi.org/10.22381/LPI1920207>
- Rovetta, Alessandro; Bhagavathula, Akshaya-Srikanth** (2020). "Covid-19-related web search behaviors and infodemic attitudes in Italy: infodemiological study". *The journal of medical internet research, health and surveillance*, v. 6, n. 2, e19374.
<https://doi.org/10.2196/19374>
- Rudd, Rima; Baur, Cynthia** (2020). "Health literacy and early insights during a pandemic". *Journal of communication in healthcare*, v. 13, n. 1, pp. 13-16.
<https://doi.org/10.1080/17538068.2020.1760622>
- Rufai, Sohaib R.; Bunce, Catey** (2020). "World leaders' usage of Twitter in response to the Covid-19 pandemic: a content analysis". *Journal of public health (United Kingdom)*, v. 42, n. 3, pp. 510-516.
<https://doi.org/10.1093/pubmed/fdaa049>
- Ruiu, Maria-Laura** (2020). "Mismanagement of Covid-19: lessons learned from Italy". *Journal of risk research*, v. 23, n. 7-8, pp. 1007-1020.
<https://doi.org/10.1080/13669877.2020.1758755>
- Sahoo, Soumya-Swaroop; Sahu, Dinesh-Prasad; Kankaria, Ankita** (2020). "Mis-infodemic: the Achilles' heel in combating the Covid-19 pandemic in an Indian perspective". *Monaldi archives for chest disease*, v. 90, n. 2, pp. 330-332.
<https://doi.org/10.4081/MONALDI.2020.1405>
- Samy, Micahel-Abdelmalak; Ahmed A., Rebecca; Kelada, Mary** (2020). "Social media as a source of medical information during Covid-19". *Medical education online*, v. 25, 1791467.
<https://doi.org/10.1080/10872981.2020.1791467>
- Sasaki, Natsu; Kuroda, Reiko; Tsuno, Kanami; Kawakami, Norito** (2020). "Exposure to media and fear and worry about Covid-19". *Psychiatry and clinical neurosciences*, v. 74, n. 9, pp. 501-502.
<https://doi.org/10.1111/pcn.13095>
- Savoia, Elena; Lin, Leesa; Viswanath, Kasisomayajula** (2013). "Communications in public health emergency preparedness: a systematic review of the literature". *Biosecurity and bioterrorism: biodefense strategy, practice, and science*, v. 11, n. 3, pp. 170-184.
<https://doi.org/10.1089/bsp.2013.0038>
- Schillinger, Dean; Chittamuru, Deepti; Ramírez, Susana A.** (2020). "From 'infodemics' to health promotion: a novel framework for the role of social media in public health". *American journal of public health*, v. 110, n. 9, pp. 1393-1396.
<https://doi.org/10.2105/AJPH.2020.305746>
- Segura, María-Soledad** (2020). "Con alerta pero sin pánico. El rol de los medios durante la pandemia". *Revista de la Facultad de Ciencias Médicas de Córdoba*, v. 77, n. 1, pp. 55-58.
<https://doi.org/10.31053/1853.0605.v77.n1.28066>
- Sezgin, Deniz; Karaaslan, Yesim-Sert; Ersoy, Ibrahim** (2020). "The pandemic infodemic: the role of risk communication and media in a pandemic". *Gazi medical journal*, v. 31, n. 2, pp. 325-327.
<https://doi.org/10.12996/gmj.2020.79>
- Shah, Paras; Kim, F. J.; Mian, Badar M.** (2020). "Genitourinary cancer management during a severe pandemic: utility of rapid communication tools and evidence-based guidelines". *BJUI Compass*, v. 1, n. 2, pp. 45-59.
<https://doi.org/10.1002/bco2.18>

- Sharma, Manoj-Kumar; Anand, Nitin; Vishwakarma, Akash; Murthy, Kashava D.** (2020). "Mental health issues mediate social media use in rumors: Implication for media based mental health literacy". *Asian journal of psychiatry*, n. 53, 102132.
<https://doi.org/10.1016/j.ajp.2020.102132>
- Sharov, Konstantin S.** (2020). "Adaptation to SARS-CoV-2 under stress: role of distorted information". *European journal of clinical investigation*, v. 50, n. 9.
<https://doi.org/10.1111/eci.13294>
- Singer, Nermeen** (2020). "Coronavirus media discourse and current situation (Qca for different responses to combat covid-19)". *Humanities and social sciences reviews*, v. 8, n. 3, pp. 665-675.
<https://doi.org/10.18510/hssr.2020.8371>
- Singh, Bhupen** (2020). "Media in the time of Covid-19". *Economic & political weekly*, v. 55, n. 16.
- St. Amant, Kirk** (2020). "Communicating about Covid-19: practices for today, planning for tomorrow". *Journal of technical writing and communication*, v. 50, n. 3, pp. 211-223.
<https://doi.org/10.1177/0047281620923589>
- Su, Yue; Xue, Jia; Liu, Xiaoqian; Wu, Peijing; Chen, Jungxiang; Chen, Chen; Liu, Tianli; Gong, Weigang; Zhu, Thu** (2020). "Examining the impact of covid-19 lockdown in Wuhan and Lombardy: a psycholinguistic analysis on Weibo and Twitter". *International journal of environmental research and public health*, v. 17, n. 12.
<https://doi.org/10.3390/ijerph17124552>
- Szmuda, Tomasz; Özdemir, Cevdet; Ali, Siraj M.; Singh, Anumeha; Syed, Maryam-Raza; Słoniewski, Pawel** (2020). "Reability of online patient education material for the novel coronavirus disease (Covid-19): a cross-sectional health literacy study". *Public health*, n. 190, pp. 145-146.
<https://doi.org/10.1016/j.puhe.2020.09.001>
- Tang, Lu; Bie, Bijie; Park, Sung-Eun; Zhi, Degui** (2018). "Social media and outbreaks of emerging infectious diseases: a systematic review of literature". *American journal of infection control*, n. 46, pp. 962-972.
<https://doi.org/10.1016/j.ajic.2018.02.010>
- Tasnim, Samia; Hossain, Md-Mahbub; Mazumder, Hoimonty** (2020). "Impact of rumors and misinformation on Covid-19 in social media". *Journal of preventive medicine and public health*, v. 53, n. 3, pp. 171-174.
<https://doi.org/10.3961/JPMPH.20.094>
- Topf, Joel M.** (2020). "Introduction: social media and medical education come of age". *Seminars in nephrology*, v. 40, n. 3, pp. 247-248.
<https://doi.org/10.1016/j.semnephrol.2020.04.002>
- Van-den-Broucke, Stephan** (2020). "Why health promotion matters to the Covid-19 pandemic, and vice versa". *Health promotion international*, v. 35, n. 2, pp. 181-186.
<https://doi.org/10.1093/heapro/daaa042>
- Venegas-Vera, Averter; Colbert, Gates B.; Lerma, Edgar V.** (2020). "Positive and negative impact of social media in the Covid-19 era". *Reviews in cardiovascular medicine*, v. 21, n. 4, pp. 561-564.
<https://doi.org/10.31083/j.rcm.2020.04.195>
- Vervoort, Dominique; Ma, Xiya; Luc, Jessica G. Y.; Zieroth, Shelley** (2020). "Rapid scholarly dissemination and cardiovascular community engagement to combat the infodemic of the Covid-19 pandemic". *Canadian journal of cardiology*, v. 6, n. 36, 969.e1-969.e2.
<https://doi.org/10.1016/j.cjca.2020.03.042>
- Vijaykumar, Santosh; Nowak, Glen; Himelboim, Itai; Jin, Yan** (2018). "Managing social media rumors and misinformation during outbreaks". *American journal of infection control*, v. 46, n. 7, pp. 850.
<https://doi.org/10.1016/j.ajic.2018.03.014>
- Walker, Mark D.; Sulyok, Mihály** (2020). "Online behavioral patterns for coronavirus disease 2019 (Covid-19) in the United Kingdom". *Epidemiology and infection*, n. 5, pp.148, e110.
<https://doi.org/10.1017/S0950268820001193>
- Wang, Peng-Wei; Lu, Wei-Hsin; N.Y. Chen, Yi-Lu; 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.
<https://doi.org/10.2196/20021>
- Wen, Jun; Aston, Joshua; Liu, Xinyi; Ying, Tianyu** (2020). "Effects of misleading media coverage on public health crisis: A case of the 2019 novel coronavirus outbreak in China". *Anatolia*, v. 31, n. 2, pp. 331-336.
<https://doi.org/10.1080/13032917.2020.1730621>

Wong, Roger; Harris, Jenine K.; Staub, Mackenzie; Bernhardt, Jay M. (2017). "Local health departments tweeting about ebola: characteristics and messaging". *Journal of public health management and practice*, v. 23, n. 2, e16-e24.
<https://doi.org/10.1097/PHH.0000000000000342>

World Health Organisation (2015). *Health literacy toolkit for low- and middle-income countries. A series of information sheets to empower communities and strengthen health systems*. World Health Organization, Geneva.
<https://apps.who.int/iris/handle/10665/205244>

World Health Organization (2022). *WHO coronavirus disease (Covid-19) dashboard overview*. World Health Organization.
<https://covid19.who.int>

Xifra, Jordi (2020). "Comunicación corporativa, relaciones públicas y gestión del riesgo reputacional en tiempos del Covid-19". *Profesional de la información*, v. 29, n. 2, e290220.
<https://doi.org/10.3145/epi.2020.mar.20>

Yan, Qinling; Tang, Yingling; Yan, Dingding; Wang, Jiaying; Yang, Linqian; Yang, Xinpei; Tang, Sanyi (2020). "Impact of media reports on the early spread of Covid-19 epidemic". *Journal of theoretical biology*, n. 502, 110385.
<https://doi.org/10.1016/j.jtbi.2020.110385>

Yin, Fulian; Lv, Jiahui; Zhang, Xiaojian; Xia, Xinyu; Wu, Jianhong (2020a). "Covid-19 information propagation dynamics in the Chinese Sina-microblog". *Mathematical biosciences and engineering*, v. 17, n. 3, pp. 2676-2692.
<https://doi.org/10.3934/mbe.2020146>

Yin, Fulian; Xia, Xinyu; Song, Nan; Zhu, Lingyao; Wu, Jianhong (2020b). "Quantify the role of superspreaders -opinion leaders- on Covid-19 information propagation in the Chinese Sina-microblog". *PLoS one*, v. 15, n. 6, e0234023.
<https://doi.org/10.1371/journal.pone.0234023>

Yusof, Aimi-Nadia-Mohd; Muuti, Muhamad-Zaid; Ariffin, Lydia-Aiseach; Tan, Mark-Kiak-Min (2020). "Sharing information on Covid-19: the ethical challenges in the Malaysian setting". *Asian bioethics review*, v. 12, n. 3, pp. 349-361.
<https://doi.org/10.1007/s41649-020-00132-4>

Zhao, Yuxin; Cheng, Sixuang; Yu, Xiaoyan; Xu, Huilan (2020). "Chinese public's attention to the Covid-19 epidemic on social media: observational descriptive study". *Journal of medical internet research*, v. 22, n. 5.
<https://doi.org/10.2196/18825>

Zhu, Yun'er; Fu, King-Wa; Grépin, Karen A.; Liang, Hai; Fung, Isaac-Chun-Hai (2020). "Limited early warnings and public attention to coronavirus disease 2019 in China, january-february, 2020: a longitudinal cohort of randomly sampled Weibo users". *Disaster medicine and public health preparedness*, v. 14, n. 5, e24-e27.
<https://doi.org/10.1017/dmp.2020.68>

Dialnet | métricas
Nueva edición 2020

Fundación Dialnet

dialnet.unirioja.es/metricas

Data-driven scientific research based on public statistics: a bibliometric perspective

Jorge-Eusebio Velasco-López; Ramón-Alberto Carrasco; Manuel J. Cobo; Gema Fernández-Avilés

Nota: Este artículo se puede leer en español en:
<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87085>

Recommended citation:

Velasco-López, Jorge-Eusebio; Carrasco, Ramón-Alberto; Cobo, Manuel J.; Fernández-Avilés, Gema (2023). "Data-driven scientific research based on public statistics: a bibliometric perspective". *Profesional de la información*, v. 32, n. 3, e320314.

<https://doi.org/10.3145/epi.2023.may.14>

Manuscript received on 4th August 2022
Accepted on 8th March 2023



José-Eusebio Velasco-López ✉
<https://orcid.org/0000-0003-2392-5690>

Universidad Complutense de Madrid
Facultad de Estudios Estadísticos
Avda. Puerta de Hierro, s/n
28040 Madrid, Spain
jorge.velasco.lopez@ine.es



Ramón-Alberto Carrasco
<https://orcid.org/0000-0001-7365-349X>

Universidad Complutense de Madrid
Facultad de Estudios Estadísticos
Avda. Puerta de Hierro, s/n
28040 Madrid, Spain
ramoncar@ucm.es



Manuel J. Cobo
<https://orcid.org/0000-0001-6575-803X>

Instituto Andaluz Interuniversitario
en Ciencia de Datos e Inteligencia
Computacional (DaSCI)
Universidad de Granada
Periodista Daniel Saucedo Aranda, s/n
18071 Granada, Spain
mjcobo@ugr.es



Gema Fernández-Avilés
<https://orcid.org/0000-0001-5934-1916>

Universidad de Castilla-La Mancha
Facultad de Ciencias Jurídicas y Sociales
de Toledo
San Pedro Mártir, s/n
45071 Toledo, Spain
gema.faviles@uclm.es

Abstract

Official statistics provide information on different areas of citizens' lives and are widely used in scientific research as a source of data due to their open data nature and quality assurance. In this context, a bibliometric analysis is carried out using all *Scopus* publications from 1960 to 2020 that use official statistics as data sources. Thus, 10,777 publications are analyzed using the *SciMAT* bibliometric analysis software, providing a complete conceptual analysis of the main research topics in the literature through the quantification of the main bibliometric performance indicators, identifying the most important authors, organizations, countries, sources, and intellectual structures corresponding to the main fields of research and bringing classification by subject area as an innovation to the methodology.

Keywords

Official statistics; Co-word analysis; Strategic diagram; Science mapping analysis; Bibliometric analysis; *SciMAT*.

Funding

This paper has been partially supported by:

- Spanish State Research Agency through project PID2019-105381GA-I00/AEI/10.13039/501100011033 (*iScience*).
- University of Castilla-La Mancha, Faculty of Law and Social Science at Toledo, Code: 00440710.



1. Introduction

Today's society is undergoing an accelerated process of digital transformation, or datafication. big data and open data are the most important pillars of this transformation. While big data is defined by volume, open data is defined by its public accessibility for research, new ventures, patterns and trends analysis, data-driven decisions (**Moreno; Carrasco; Herrera-Viedma**, 2019), and complex problem-solving. A key source of open data is official statistics, which are statistical operations published as a public good by government agencies or other public bodies, such as international organizations (*United Nations*, 2021).

Scientific researchers are among the potential users that may benefit from the data provided by official statistics. They are usually familiar with the statistical methods used and may want to delve into statistical facts and observations for analytical purposes, studying the cause-and-effect interrelationships among different phenomena. In this context, we focus on examining the areas in which there has been scientific research published that employs official statistics as data sources (in whole or in part).

Advancing the knowledge of any scientific discipline requires an examination of the theoretical and empirical contributions of academics and professionals over time. As such, a periodic analysis of the accumulated knowledge is necessary to understand the current state of a discipline (**Rodríguez-López et al.**, 2020). After conducting a literature review, we have found no studies similar to the one proposed here. The present article thus seeks to fill this gap and provide a better understanding of the state of the question through an exhaustive bibliometric study involving an analysis of the co-occurrence of keywords and the mapping of all published articles related to official statistics and indexed in the *Scopus* bibliographic database (**Mukherjee et al.**, 2022).

The expected result is to discover which themes can be defined as "central" for research on the basis of official statistics in the period examined, which themes play a minor, or "non-central" role, and what changes can be observed during the different periods for all the themes. A performance analysis will be carried out.

In the rest of the article, section 2 provides a description of the methodology applied for the bibliometric analysis, the dataset, and the main tools used. Section 3 presents the bibliometric analysis and section 4 presents a discussion of the relationship detected between official statistics, open data, and scientific research. Finally, section 5 summarizes the most relevant results of the performance and content analysis conducted and outlines possible lines of future research.

2. Methodology and dataset

Below, we detail the methodology employed in our study, with a description of the general context, the specific methodology, the computer tool, and the dataset used.

2.1. Science mapping analysis

Science mapping or bibliometric mapping is a spatial representation of how disciplines, fields, specialties, documents, or authors are related to one another (**Small**, 1999). It has been widely used to reveal the hidden relationships among key elements (documents, authors, institutions, topics, etc.) in different research fields (**Cobo et al.**, 2011a; **Rodríguez-Ledesma et al.**, 2015; **Rodríguez-López et al.**, 2020; **Galán; Carrasco; LaTorre**, 2022)

Science mapping analysis can be performed with several software tools (**Cobo et al.**, 2011b). *SciMAT* was presented by **Cobo et al.** (2012) as a powerful tool that integrates most of the advantages of the available science mapping software tools and was designed according to the science mapping analysis approach presented in **Cobo et al.** (2011c). It combines both performance analysis tools and science mapping tools to analyze a research field and detect and visualize its conceptual subdomains (particular topics/themes or general thematic areas) and its thematic evolution.

Therefore, in this contribution, *SciMAT* was employed to develop a co-word science mapping analysis (**Callon et al.**, 1983; **Batagelj; Cerinšek**, 2013; **Börner; Chen; Boyack**, 2003). In line with **Cobo et al.** (2011c), the analysis is performed following a four-stage methodology:

1) Detection of the research themes. For each of the periods analyzed, research themes are detected by applying a clustering algorithm (**Coulter; Monarch; Konda**, 1998). over a normalized co-word network (**Callon et al.**, 1983). The similarity between the keywords is assessed using the equivalence index (**Callon; Courtial; Laville**, 1991).

2) Visualization of the research themes and thematic networks. The detected themes are visualized through different visualization instruments (Figure 1): the strategic diagram (**He**, 1999) and the thematic network (**Cobo et al.**, 2011c). Each theme can be characterized by two measures (**Callon; Courtial; Laville**, 1991): centrality, which measures the degree of interaction of a network with other networks; and density, which measures the internal strength of the network. Given both measures, a research field can be visualized as a set of research themes, mapped in a two-dimensional strategic diagram, Figure 1(a), and classified into four groups:

- Motor themes (quadrant Q1): They present strong values of centrality and density. These are well-developed themes that are important or central to structuring a research field.
- Highly developed and isolated subjects (quadrant Q2): These are strongly related, highly specialized, and peripheral, but they do not have the appropriate background or importance for the field.

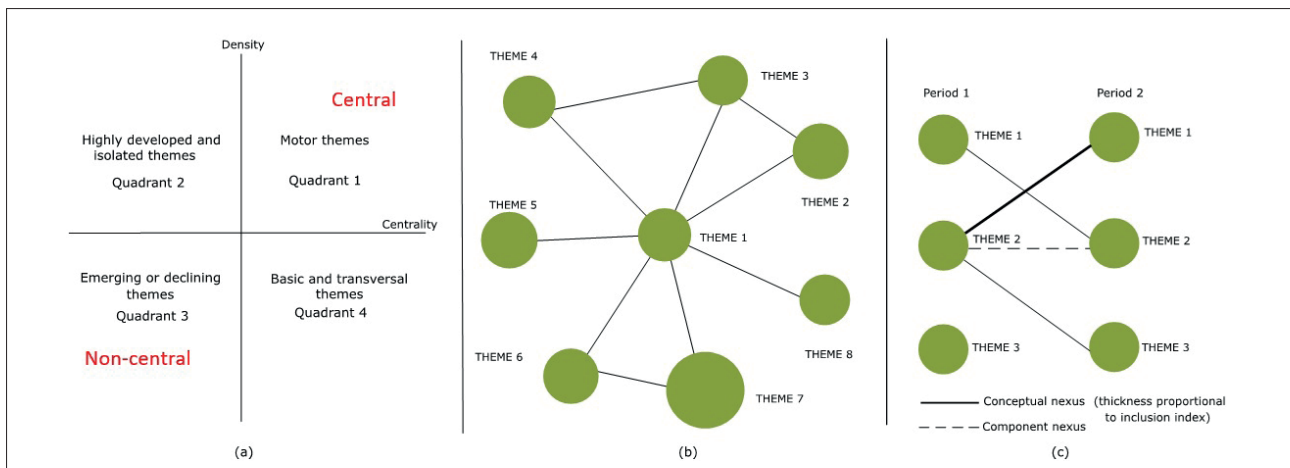


Figure 1. (a) Strategic diagram, (b) thematic network, and (c) thematic evolution

- Emerging or declining themes (niche) (quadrant Q3): They present both low density and low centrality. They are underdeveloped and marginal to the field of research.
- Basic and transversal themes (quadrant Q4): These themes are important for the research field, but internally they are not extensively developed.

In this study, we refer to the themes on the right-hand side (motor and basic) as central and those on the left-hand side (niche and peripheral) as non-central.

3) Discovery of thematic areas. In this stage, the evolution of the research themes over a set of periods is first detected and then analyzed to identify the main general areas of evolution in the research field, their origins, and their interrelationships. Their evolution over the whole period is then measured as the overlapping of clusters from two consecutive periods. For this purpose, the inclusion index (Sternitzke; Bergmann, 2009) is used to detect conceptual nexuses between research themes in different periods, thus allowing the identification of the thematic areas in a research field. The thematic evolution map, Figure 1(c), depicts the current development of research that uses official statistics, applying a longitudinal perspective covering the entire period examined. In the visualization, the solid lines mean that the linked cluster shares the main item, while a dotted line means that the themes share elements that are not the main item. The thickness of the lines is proportional to the inclusion index value, and the size of the circles is proportional to the number of published documents associated with each cluster. Section 3.3 will show the thematic fluctuations.

4) Performance analysis. In this stage, the relative contribution of the research themes to the whole research field is measured (quantitatively and qualitatively) and used to establish the most prominent, productive, and highest-impact subfields. Some of the bibliometric indicators to use are the number of published documents, the number of citations, and different types of h-index (Alonso *et al.*, 2009; Hirsch, 2005; Martínez *et al.*, 2004). For each theme, the performance measures are computed based on the documents associated with it.

The methodology described above is extended by introducing a new step in which each theme detected in the analysis phase is assigned to a certain main area of knowledge, through a classification procedure based on a cluster network analysis (thematic network, Figure 1b). After determining this area of knowledge, we then add this new information to our maps from the previous stage. A theme can cut across different fields of knowledge, although there is always one field that predominates, to which it is therefore assigned. Given that the visualization techniques applied already make use of the positioning and the size of the object to graphically represent the properties of the objects (circles that represent a theme), in this study we assign colors to these objects according to the main field of knowledge to which they belong. The colors associated with each knowledge area are defined in Table 1.

Table 1. Subject areas with their associated colors

Medicine	Psychology	Sociology	Economy
Microbiology	Arts	Environment	Business and management

2.2. Dataset

Scopus is one of the world's largest abstract and citation databases of peer-reviewed research literature, with over 25,000 titles from more than 7,000 international publishers and extensively used in bibliometric analyses. To identify possible changes in the themes identified in research based on official statistics, a global search by title, abstract, and keywords on articles in the above-mentioned Scopus database is carried out. The results are then analyzed using a co-word analysis in the following steps. Note that no filter by starting year is applied, since the purpose of the study is to analyze the entire Scopus database, which starts in 1960.

It is empirically observed that entering the term GOVERNMENT DATA distorts the search, yielding many results about E-GOVERNMENT, which is not directly related to official statistical information. Likewise, the term PUBLIC DATA does not adequately capture the use of official statistical data sources. Therefore, the search eventually implemented is as follows:

“ A key provider of open data is official statistics, which are statistical operations published by government agencies or other public bodies such as international organizations ”

TITLE-ABS-KEY (“OFFICIAL STATISTICS”) OR TITLE-ABS-KEY (“OFFICIAL DATA”) OR TITLE-ABS-KEY (“GOVERNMENT STATISTICS”) OR TITLE-ABS-KEY (“PUBLIC STATISTICS”) OR TITLE-ABS-KEY (“NATIONAL STATISTICS”).

A total of 10,995 results are obtained. By conducting a manual review, we discard 218 of them that relate to E-GOVERNMENT and other non-related topics, considering the remaining 10,777 of interest and suitable for the study.

This selection is exported in Research Information Systems (RIS) format, and then imported into *SciMAT* (Cobo *et al.*, 2012). It includes all the citation information, the abstract along with the keywords, as well as the references, all of which are used to build the knowledge base for a subsequent science mapping analysis. Approximately 10% of the documents are found to be lacking any keywords (and therefore will not be analyzed). In addition, some keywords without meaning in this context, such as stop words or words with a very broad and general meaning, are removed to improve the analysis. Considering that our focus is on understanding the development of the themes, we exclude methodological terms, codes, years, geographic information, and dates, although some information could also be learned from those terms. The stop words are selected following the procedure proposed by Jacob and Brust (2019). The next step is time slicing, whereby the entire time frame of the sample (1960-2020) is divided into the following four periods:

- 1960-2000 (1,885 documents);
- 2001-2010 (2,527);
- 2011-2015 (2,486);
- 2016-2020 (3,878).

In this study, we configure the keyword frequency threshold and the co-occurrence frequency threshold for each period: (9, 2), (13, 2), (11, 2), and (17, 2), respectively.

3. Bibliometric analysis

Next, the bibliometric analysis is carried out using the methodology and dataset presented in section 2.

3.2. Performance statistical analysis

From the documents retrieved for this study, a total of 33,078 keywords have been identified, with the most frequent ones presented in Table 2.

Table 2. Keywords with corpus frequency higher than 200

Keywords	Frequency	Keywords	Frequency
Mortality	1,219	Morbidity	286
Epidemiology	677	Health	278
Sex-Difference	530	Geography	277
Cause-of-Death	502	Suicide	245
Population	499	Pregnancy	244
Economics	466	Health-Care-Policy	230
Demography	389	Sex-Factors	229
Public-Health	382	Economic-Analysis	228
Socioeconomics	353	Cancer	225
Health-Survey	347	Hospitalization	222
Age-Distribution	333	Time-Factors	215
Migration	330	Psychology	210
Socioeconomic-Factors	329	Gender	210
Health-Care-System	307	Survival-Rate	207
Age-Factors	291	Employment	206

Documents with more than 500 citations are shown in Table 3, along with their respective authors, citations, and year of publication. The table is ordered by subject area and the Field-Weighted Citation Impact (FWCI) metric (Purkayastha *et al.*, 2019), which is the ratio of the total citations received by the denominator output and the total citations expected based on the average of the subject field.

Table 3. Documents with more than 500 citations, sorted by subject and FWCI

Subject	FWCI	Document	Citations
Business and Management	22.75	(Gustavsson <i>et al.</i> , 2011)	970
	18.22	(Olesen <i>et al.</i> , 2012).	828
Economy	47.55	(Henderson; Storeygard; Weil, 2012)	553
Environment	27.89	(Giri <i>et al.</i> , 2011)	1,271
	9.20	(Liu <i>et al.</i> , 2015)	627
	8.23	(Chen, 2007)	564
Medicine	129.00	(Allison <i>et al.</i> , 1999)	1,463
	58.41	(Hippisley-Cox <i>et al.</i> , 2008)	861
	43.95	(Llewelyn <i>et al.</i> , 2004)	907
	29.88	(Peto <i>et al.</i> , 2000)	1,006
	17.19	(Unal; Critchley; Capewell, 2004)	539
	10.27	(Multicentre Aneurysm Screening Study Group, 2002)	909
	7.79	(Bernstein <i>et al.</i> , 2000)	575
Sociology	49.41	(Gillborn, 2008)	520
	44.32	(Van-Dijk; Hacker, 2003)	635
	25.01	(Grantham-McGregor <i>et al.</i> , 2007)	1,599
	13.96	(Abraido-Lanza <i>et al.</i> , 1999)	692
	6.87	(Savage <i>et al.</i> , 2013)	546

The most prolific countries in terms of document output can be seen in Table 4, together with the affiliation organizations accounting for the highest number of documents in the sample (all of them public institutions in the United Kingdom).

Table 4. Countries with more than 300 documents published and affiliation organizations with more than 75 documents

Country	Documents	Affiliation	Documents
United Kingdom	2,292	University College London	163
United States	1,579	Office for National Statistics London	163
Germany	636	University of Oxford	128
Spain	630	University of Manchester	123
Russian Federation	625	King's College London	112
Italy	587	London School of Hygiene-Tropical Medicine	106
China	420	Imperial College London	103
Sweden	339	University of Cambridge	90
Australia	333	University of Oxford-Medical Science Division	81
Brazil	328	University of Bristol	81
France	319	University of Southampton	78

3.2. Thematic analysis

To analyze for each period the most prominent themes in research based on official statistics, a strategic diagram is provided for each period. Performance analysis is used to improve the results.

3.2.1. First period (1960-2000)

According to the strategic diagram presented in Figure 2a, during this period the research activity was focused on 22 themes, with the following 12 central themes (motor themes plus basic themes): *Economic-Factors*, *Developed-Countries*, *Developing-Countries*, *Employment*, *Contraceptive-Usage*, *Geography*, *Social-Policy*, *Pregnancy*, *Cause-of-Death*, *Forensic-Medicine*, *Life-Table*, *Health-Survey*.

Note that for all strategic diagrams in this section, the color of the circles corresponds to the classification of the themes obtained from the cluster network analysis, the size of the circles reflects the number of documents found for each theme, and the quadrant where they are located indicates the centrality and density of the theme, as outlined in Figure 1. The performance measures of the themes are shown separately for each of the four periods in Tables 5, 6, 7, and 8, sorted by the number of documents. Each table shows the number of documents assigned to the themes with more than 100 documents, the citations, and the h-index of the documents associated with the theme. Also, a summary of the central (basic and motor) and non-central themes and their subject area can be found.

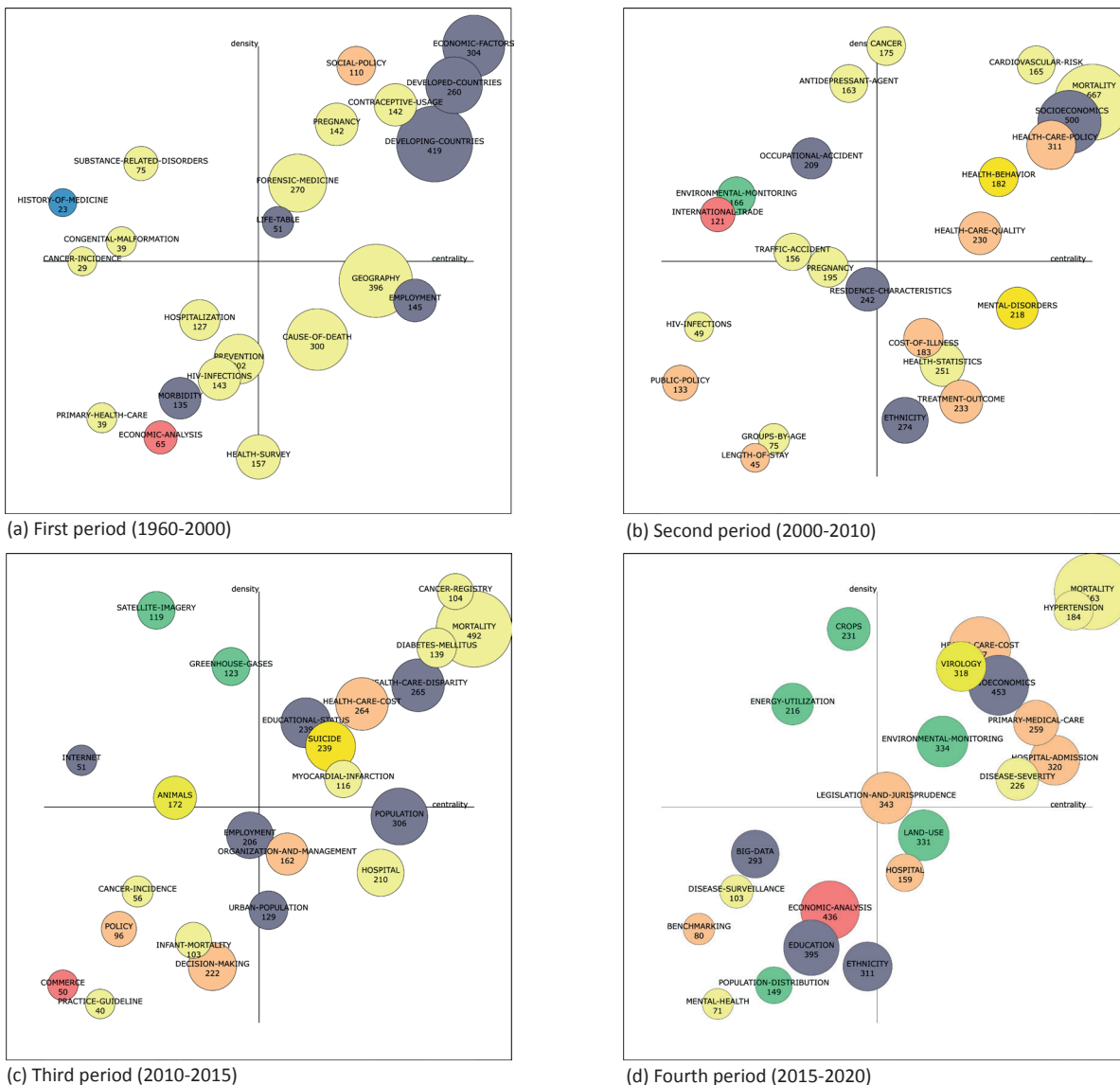


Figure 2. Strategic diagram of the four periods

According to the performance measures, the following themes stand out with an h-index greater than 28 and a number of citations greater than 3,000: *Developing-Countries*, *Geography*, *Economic-Factors*, *Cause-Of-Death*, *Forensic-Medicine*, *Developed-Countries*, *Prevention*, *Health-Survey*, *Employment*, and *HIV-Infections*. Most of these fall within the central themes.

Note that themes related to Medicine and Sociology predominate, both in motor and basic themes as well as in non-central ones (niche and peripheral)

Non-central themes (i.e., peripheral or niche) within the research are shown in the left section of Figure 2a. The themes *Congenital-Malformation*, *Cancer-Incidence*, *History-Of-Medicine*, and *Substance-Related-Disorders* (all health-related) in the upper left quadrant can be classified as isolated and developed themes. The themes of *Prevention*, *HIV-Infections*, *Hospitalization*, *Morbidity*, *Economic-Analysis*, and *Primary-Health-Care* in the lower left quadrant are related to emerging or declining themes.

3.2.2. Second period (2001-2010)

In this period, the research delved into 23 themes (see Figure 2b). In this case, according to the strategic diagram, 11 central themes can be identified (motor themes plus basic themes): *Mortality*, *Socioeconomics*, *Health-Care-Policy*, *Cardiovascular-Risk*, *Mental-Disorders*, *Health-Behavior*, *Health-Care-Quality*, *Treatment-Outcome*, *Health-Statistics*, *Cost-Of-Illness*, *Ethnicity*, and *Cancer*. However, of the motor themes, *Mortality*, *Socioeconomics*, *Health-Care-Policy* and *Cardiovascular-Risk* are the most influential as they are well developed and central to research based on official statistics during this period.

“ We investigate which themes can be defined as “central” and “peripheral” of the research based on official statistics in the period examined and, if a change in trend is observed in these themes, which may be the reasons for such changes ”

Table 5. Performance measures of the themes (central and non-central) of the first period (1960-2000), sorted by the number of documents. Note: *Non-central themes

Name	Docs.	h index	Cita-tions	Subject area	Summary
<i>Developing-Countries</i>	419	42	9,881	Sociology	Topics related to mortality and health, the status of women, and education
<i>Geography</i>	396	41	6,333	Medicine	Epidemiology related to gender, ethnicity, age, and geographic factors
<i>Economic-Factors</i>	304	33	5,015	Sociology	Economic factors that influence the socioeconomic status and working conditions
<i>Cause-Of-Death</i>	300	46	10,131	Medicine	Different types of diseases (coronary, respiratory, cancer...) and their associated cause of death
<i>Forensic-Medicine</i>	270	36	5,370	Medicine	Ethological and psychological aspects that can lead to suicide, etc.
<i>Developed-Countries</i>	260	28	2,923	Sociology	Sociodemographic characteristics in developed countries (type of family, migration, marriage)
<i>Prevention</i>	202	31	3,646	Medicine*	Prevention of work accidents and related aspects such as diagnosis and injuries
<i>Health-Survey</i>	157	32	4,819	Medicine	Uses of health statistics in demographic and medical studies
<i>Employment</i>	145	29	3,097	Sociology	Sociological and psychological aspects of employment and unemployment
<i>HIV-Infections</i>	143	30	3,365	Medicine*	Determinants of HIV infections according to cultural, educational, and racial characteristics
<i>Pregnancy</i>	142	26	2,334	Medicine	Aspects related to infant mortality and characteristics of the pregnant mother and newborn
<i>Contraceptive-Usage</i>	142	17	1,861	Medicine	Use of contraceptive measures and sexual behavior and related aspects such as abortions, fertility, and access to healthcare
<i>Morbidity</i>	135	26	3,687	Sociology*	Hospitalized illnesses according to lifestyles, race, and cost of stay
<i>Hospitalization</i>	127	22	1,896	Medicine*	Medical issues related to the length of stay in hospitals
<i>Social-Policy</i>	110	17	1,918	Business and Management	Social, population, health, and public policies

The *Mortality* theme presents the highest performance indicators (h-index of 87), with the highest number of documents and citations. *Socioeconomics*, *Health-Care-Policy*, *Ethnicity*, *Health-Statistics*, *Residence-Characteristics*, *Treatment-Outcome*, *Cardiovascular-Risk* are also important themes, all with more than 9,000 citations, h-index values greater than 45, and a significant number of documents. All these themes are framed within the central themes.

According to the number of documents, it is observed that themes related to Medicine predominate, as in the first period. Now, however, themes related to Business and Management (mostly public policies) emerge as the second most important group. Sociology moves to third place, followed by themes related to Psychology, Environment, and Economy. Medicine and Sociology are distributed among all the quadrants, while Business and Management does not appear in the second one.

Non-central themes within the research are shown in the left section of Figure 2b. The *Antidepressant-Agent*, *Occupational-Accident*, *Traffic-Accident*, *Environmental-Monitoring*, and *International-Trade* themes in the upper left quadrant can be classified as isolated and developed themes. We observe that only the first two are related to health, but the first theme is related to the Environment and two more related to the Economy appear. The themes *Residence-Characteristics*, *Pregnancy*, *Groups-By-Age*, *Length-Of-Stay*, *HIV-Infections*, and *Public-Policy* in the lower left quadrant relate to emerging or declining themes.

3.2.3. Third period (2011-2015)

In this period, the research delved into 23 themes (Figure 2c). In this case, according to the strategic diagram, 12 central themes can be identified (motor themes plus basic themes): *Mortality*, *Cancer-Registry*, *Diabetes-Mellitus*, *Health-Care-Disparity*, *Population*, *Hospital*, *Health-Care-Cost*, *Myocardial-Infarction*, *Suicide*, *Educational-Status*, *Organization-And-Management*, and *Urban-Population*. However, of the motor themes, *Cancer-Registry*, *Diabetes-Mellitus*, and *Health-Care-Disparity* are the most influential as they are well developed and central to the research based on official statistics during this period.

The *Mortality* theme registers the highest performance indicator (h-index of 56) with the highest number of documents and citations. *Population*, *Health-Care-Disparity*, *Health-Care-Cost*, *Suicide*, *Educational-Status*, and *Hospital* are also important themes, with more than 6,000 citations, an h-index greater than 38, and a notable number of documents. All these themes are framed within the central themes.

The themes identified are classified into large thematic areas, where those related to Medicine and Sociology predominate in terms of the number of documents, as in the first period. Themes related to Business and Management (mostly

Table 6. Performance measures of the themes (central and non-central) of the second period (2000-2010), sorted by the number of documents. Note: *Non-central themes

Name	Docs.	h index	Cita-tions	Subject area	Summary
<i>Mortality</i>	667	87	30,032	Medicine	Topics related to causes of death
<i>Socioeconomics</i>	500	68	19,156	Sociology	Topics involving socio-cultural and economic aspects of the population
<i>Health-Care-Policy</i>	311	49	8,823	Business and Management	Topics related to health policies
<i>Ethnicity</i>	274	48	9,442	Sociology	Social and demographic aspects related to ethnicity
<i>Health-Statistics</i>	251	50	8,977	Medicine	Topics related to health quality
<i>Residence-Characteristics</i>	242	47	9,859	Sociology*	Socioeconomic characteristics of dwellings according to whether they are rural or urban
<i>Treatment-Outcome</i>	233	54	12,102	Business and Management	Results of medical treatment
<i>Health-Care-Quality</i>	230	48	7,199	Business and Management	Health quality in terms of accessibility and other factors
<i>Mental-Disorders</i>	218	45	7,813	Psychology*	Social, legal, and economic aspects of mental disorders
<i>Occupational-Accident</i>	209	39	5,723	Sociology*	Accidents by type of employment
<i>Pregnancy</i>	195	40	6,833	Medicine*	Medical and social aspects related to motherhood
<i>Cost-Of-Illness</i>	183	42	5,977	Business and Management	Cost of diseases by type
<i>Health-Behavior</i>	182	45	8,083	Psychology	Health behavior according to socioeconomic and mental health factors
<i>Cancer</i>	175	47	8,181	Medicine	Cancer-related aspects (risk, mortality, incidence)
<i>Environmental-Monitoring</i>	166	41	5,721	Environmental*	Environmental supervision concerning contamination and protection
<i>Cardiovascular-Risk</i>	165	47	9,826	Medicine	Cardiovascular risk factors and their consequences
<i>Antidepressant-Agent</i>	163	45	6,898	Medicine*	Consequences and risks of using antidepressant medications
<i>Traffic-Accident</i>	156	36	4,406	Medicine*	Topics related to traffic accidents
<i>International-Trade</i>	121	24	2,261	Economy*	Topics related to international trade
<i>Public-Policy</i>	133	27	2,358	Business and Management*	Public policies based on economic analysis

public policies) are still important, but not as much as in the previous period. Economic and Environmental themes are less important but continue to appear, while it is notable that those related to Psychology disappear altogether. Themes are distributed throughout all the quadrants, except the Environment, which remains a peripheral and developed theme, as in the previous period. Likewise, as was the case in the previous period, Management themes do not appear in the second quadrant.

Non-central themes within the research are shown in the left section of Figure 2c. The *Greenhouse-Gases*, *Animals*, *Satellite-Imagery*, and *Internet* themes in the upper left quadrant can be classified as isolated and developed themes. We note that the themes related to the Environment are located here, as well as one related to the use of the internet. The themes *Residence-Characteristics*, *Pregnancy*, *Groups-By-Age*, *Length-Of-Stay*, *HIV-Infections*, and *Public-Policy* in the lower left quadrant relate to emerging or declining themes.

3.2.4. Fourth period (2016-2020)

In this period, the research delved into 22 themes (see Figure 2d). In this case, according to the strategic diagram, 12 central themes can be identified (motor themes plus basic themes): *Mortality*, *Hypertension*, *Hospital-Admission*, *Primary-Medical-Care*, *Disease-Severity*, *Socioeconomics*, *Health-Care-Cost*, *Virology*, *Environmental-Monitoring*, *Land-Use*, *Hospital*, and *Legislation-And-Jurisprudence*. Of the motor themes, *Mortality* and *Hypertension*, are the most influential because they are well-developed and central to research based on official statistics during this period.

The *Mortality* theme registers the highest performance indicator (h-index of 38), while *Health-Care-Cost*, *Economic-Analysis*, *Socioeconomics*, *Land-Use*, *Hospital-Admission*, and *Virology* are also important themes, with

“ A new step is introduced in the traditional mapping process, consisting of each theme detected in the analysis phase of the process being assigned to a certain main area of knowledge ”

Table 7. Performance measures of the themes (central and non-central) of the third period (2010-2015), sorted by the number of documents.
Note: *Non-central themes

Name	Docs.	h index	Cita-tions	Subject area	Summary
<i>Mortality</i>	492	56	13,283	Medicine	Morbidity, epidemiology, and causes of death by sex and age
<i>Population</i>	306	43	7,842	Sociology	Studies on cultural and social issues affecting the population
<i>Health-Care-Disparity</i>	265	43	7,396	Sociology	Socioeconomic issues that affect access to healthcare facilities
<i>Health-Care-Cost</i>	264	41	8,501	Business and Management	Cost of healthcare related to socioeconomic factors
<i>Suicide</i>	239	39	6,008	Psychology	Social, medical, and psychological aspects of suicide
<i>Educational-Status</i>	239	38	6,553	Sociology	Pregnancy and access to healthcare according to educational status
<i>Decision-Making</i>	222	34	4,073	Business and* Management	Political aspects of the Sustainable Development Goals (SDG)
<i>Hospital</i>	210	41	6,296	Medicine	Access to healthcare according to sociodemographic aspects
<i>Employment</i>	206	30	3,387	Sociology*	Economic aspects of employment
<i>Animals</i>	172	34	5,459	Microbiology*	Environmental protection and animal epidemics
<i>Organization-And- Management</i>	162	28	3,363	Business and Management	The political organization of public health
<i>Diabetes-Mellitus</i>	139	31	3,515	Medicine	Various diseases related to diabetes-mellitus together with risk factors
<i>Urban-Population</i>	129	31	3,493	Sociology	Social aspects related to the urban population
<i>Greenhouse-Gases</i>	123	34	4,298	Environmental*	Climate aspects of greenhouse gases
<i>Satellite-Imagery</i>	119	35	4,743	Environmental*	Use of satellite images for monitoring of agriculture and ecosystems
<i>Myocardial- Infarction</i>	116	29	3,432	Medicine	Topics related to myocardial infarction
<i>Cancer-Registry</i>	104	31	3,090	Medicine	Treatment, survival, and other aspects related to cancer registries
<i>Infant-Mortality</i>	103	25	1,877	Medicine*	Infant mortality related to vaccinations and cardiovascular risk

more than 3,000 citations, an h-index greater than 26, and a notable number of documents. Most of these themes fall within the central themes.

For this last period, the themes obtained are classified into large thematic areas, showing that the areas of Medicine and Sociology are still important in terms of number of documents, as in all previous periods, but now the most important field is related to Business and Management (mostly public policy). In addition, the Environment gains prominence while the Economy loses it. Finally, new areas appear, notably Microbiology. Sociology and Medicine mainly lie in the first and third quadrants, whereas Environment is now distributed throughout all quadrants. As in the previous period, Business and Management themes do not appear in the second quadrant.

Non-central themes within the research are shown in the left section of Figure 2d. The *Crops* and *Energy-Utilization* themes in the upper left quadrant can be classified as isolated and developed themes. We note that this is the quadrant that contains themes related to the Environment and agriculture. The *Ethnicity*, *Economic-Analysis*, *Education*, *Population-Distribution*, *Big-Data*, *Disease-Surveillance*, *Mental-Health*, and *Benchmarking* themes in the lower left quadrant are related to emerging or declining themes.

3.3. Conceptual evolution map

Using *SciMAT*, an analysis of the themes detected in each period is carried out, considering their keywords and their evolution over time.

The evolution map (Figure 3) depicts the current development of research using official statistics, applying a longitudinal perspective covering the entire period examined. Few thematic fluctuations can be found here, although there is more variation in themes that are not basic or motor. In Figure 3, only thick dashed lines (connecting groups in adjacent periods that share individual keywords) have been left, while we have removed the themes that are not basic or motor and that did not show any connections with adjacent periods, with themes assigned to their corresponding main thematic area through the color of the circle.

Table 8. Performance measures of the themes (central and non-central) of the fourth period (2016-2020), sorted by the number of documents. Note: *Non-central themes.

Name	Docs.	h index	Citations	Subject area	Summary
<i>Mortality</i>	663	38	6,776	Medicine	Causes of death
<i>Health-Care-Cost</i>	477	30	4,217	Business and Management	Cost of healthcare
<i>Economic-Analysis</i>	436	27	3,578	Economy*	The financial crisis and its consequences
<i>Socioeconomics</i>	453	27	3,863	Sociology	Socioeconomic aspects (poverty, health) according to population profiles
<i>Education</i>	395	26	2,949	Sociology*	Influence of education and other factors such as age and sex on well-being and employment
<i>Legislation-And- Jurisprudence</i>	343	25	2,599	Business and Management	Public policies related to trade, health, and crime
<i>Environmental-Monitoring</i>	334	28	2,923	Environmental	Environmental issues
<i>Land-Use</i>	331	28	3,134	Environmental	Land use and Environmental aspects
<i>Hospital-Admission</i>	320	31	3,886	Business and Management	Management aspects of public health and the treatments used
<i>Virology</i>	318	26	3,114	Microbiology	COVID-19 related issues
<i>Ethnicity</i>	311	20	2,124	Sociology*	Relationship between ethnicity and health and social aspects such as social class and migration
<i>Big-Data</i>	293	17	1,457	Sociology*	Alternative data sources to obtain statistical indicators
<i>Primary-Medical-Care</i>	259	25	3,117	Business and Management	Aspects of primary health care system by types of disease
<i>Crops</i>	231	23	2,144	Environmental*	Aspects related to agriculture and means of analysis
<i>Disease-Severity</i>	226	25	2,969	Medicine	The severity of various diseases
<i>Energy-Utilization</i>	216	24	1,867	Environmental*	Energy efficiency
<i>Hypertension</i>	184	23	1,844	Medicine	Hypertension as a cause of cardiovascular events, diabetes, and other diseases
<i>Hospital</i>	159	18	1,563	Business and Management	Aspects related to accidents and illnesses and hospital stays
<i>Population-Distribution</i>	149	15	982	Environmental*	Lifestyle influence on climate change
<i>Disease-Surveillance</i>	103	17	1,081	Medicine*	Follow-up of some diseases

Considering the above, we can establish some of the major trends in the themes according to the related subject areas.

Concerning the field of knowledge to which themes belong, we make the following observations:

- For the field of Medicine, an especially relevant timeline is *Cause-of-Death* (period 1)/ *Mortality* (period 2)/ *Mortality* (period 3)/ *Mortality* (period 4), dealing with aspects such as causes of death, epidemiology, cancer mortality, hospitalization, and treatment outcomes.
- For the field of Sociology, we observe the timeline *Economic-Factors* (period 1)/ *Socioeconomics* (period 2)/ *Health-Care-Disparity* (period 3)/ *Socioeconomics* (period 4). It covers aspects related to the social sciences, especially sociology, and features interconnections with other fields such as Economy and Medicine.
- For the field of Business and Management, we identify the timeline *Social-Policy* (period 1)/*Health-Care-Policy, Cost-Of-Illness* (period 2)/*Health-Care-Cost* (period 3)/ *Health-Care-Cost* (period 4), covering aspects related to healthcare management.
- There are also timelines related to the Environment starting in period 2: *Environmental-Monitoring* (period 2) / *Greenhouse-Gases* (period 3)/ *Energy- Utilization* (period 4). It covers aspects related to environmental protection.
- The relationships between different periods are very solid, although there are also themes that have a different assigned field of knowledge (subject area) in different periods. This is because, as has been pointed out before, themes do not necessarily belong exclusively to a single field of knowledge; in fact, they usually relate to several, although they remain assigned to the main one. The themes with the highest rate of such cross-links are *Mortality*, *Socioeconomics*, and *Health-Care-Cost*, which again represent some of the main central themes within the research based on official statistics over time. Other themes develop quite independently with only a few cross-links throughout the entire period examined, such as *Animals* and *Crops*.

Regarding the themes:

- The cross-links between the themes in the upper half of Figure 3 (central research themes) are stronger than those in the lower half (non-central themes). Furthermore, we can identify which specific central and non-central themes

occur more frequently and which other themes, unusual or not at all related, hardly appear throughout the period examined.

- Some themes appear throughout almost all the periods examined, such as *Mortality*, *Socioeconomics*, or *Health-Care-Cost*, with these generally being central themes. Other themes have developed very recently and are notably central, such as *Virology*, *Environmental-Monitoring*, *Hypertension*, and *Primary-Medical-Care*. They are all themes related to Medicine and the Environment. Specifically, *Environmental-Monitoring* and *International-Trade*, which had previously established hardly any cross-links in the network, emerged at the bottom of the diagram in the second period. In the third, the same happened with *Satellite-Imagery*, *Internet*, and *Animals*.
- It is also possible to identify topics that lose researchers' interest over time, such as *HIV-Infections*, *Developed-Countries* and *Developing-Countries*, *Contraceptive-Usage*, or *Traffic-Accidents*.
- Some themes gain more centrality and density over time, such as the *Cancer* cluster. This is the case until the fourth period, where it disappears and is integrated into that of *Mortality*, albeit with less importance.



Figure 3. Map of the evolution of all the clusters over the four periods

4. Discussion

As indicated by **Giovannini, Martins and Gamba (2009)**, decision-making at any level is increasingly based on the culture of evidence. They also highlight the changes generated by technological developments, big data, and the datafication of society (**Harford, 2014**), which make statistical information more accessible to any potential user. Moreover, globalization means that the need for statistical information is expanding and is not restricted to governmental or political frameworks. Based on such shifts, these authors conclude that we are witnessing a change in the role of National Statistical Offices, from being information providers to facilitators of knowledge, which is why they are considered a public good. Given this public good character, official statistics may affect the decisions of all members of our society, including scientific researchers and the private sector. Since decisions are based on the available evidence (**García-Villar, 2012**), they may be distorted or affected if these statistics are poorly prepared, or people lack confidence in them. Since statistical output revolves around users' perception of trust, which is largely derived from the quality of statistics, official statisticians invest a lot of effort in ensuring it. The quality management system of the National Statistics Offices is aimed at establishing Quality Assurance Frameworks, inspired by the preeminent schemes at the international level, such as that of the *United Nations* or the one defined by *Eurostat* from the Code of Good Practices (**Radermacher, 2014**) of European Statistics.

Data typically needed for an investigation are collected in different ways. In addition to the specific ad-hoc generated or available data for a specific issue researchers may use official statistical sources. It is worth highlighting the value of open data for scientific work (**Lnenicka; Luterek; Nikiforova, 2022**), which can come from any of the abovementioned sources. In this sense, the data provided by official statistics are essentially open, pursuant to the legislation on data protection. The provision of open data is currently being promoted at an international level, with legislation such as *Directive (EU) 2019/1024* (on open data and reuse of public sector information), the *Data Governance Act*, the *Data Act* (linked to the *European Digital Strategy*), the creation of *European Data Spaces* and other legislative initiatives to promote high-value public sector datasets (HVDS), which include official statistics. Initiatives such as the *ODIN* project, by *Open Data Watch*, which seeks to assess the coverage and openness of official statistics, confirm the rise in reusable and free-use statistical data, but also indicate that there are challenges ahead in both availability and quality.

Regarding their relevance, public statistics are constantly being updated, with national forums such as the *Higher Council of Statistics* and its *National Statistical Plans for Spain (España, 2020)*, where these changes are legally approved. At an international level, forums of this kind can also be found in the field of official statistics, such as the *Conference of European Statisticians (CES)* of the *United Nations Economic Commission for Europe (Unece)*.

They analyze the production of statistics in pioneering areas where measurement is crucial but notoriously difficult, such as climate change, international migration, well-being, human capital, gender identity, social exclusion, and advances in the use of new data sources from the big data environment due to the appearance of new phenomena in society, such as social networks. Indeed, in our study, we can identify a series of themes studied using official statistics, which match these aforementioned recently emerging areas. Some of these themes that appear are *Big-Data*, which enters quadrant 4, linked to *Statistical Commission Decision (United Nations, 2014)*, which promotes the practical use of big data sources; *Environmental-Monitoring (Unece, 2020)*, *Sustainable-Development-Goals (Unece, 2021)*, *Crops (Unece, 2019)*, *Gender (Vikat; Jones, 2014)*, *Land-Use (Unece, 2017)*, among others. In addition, we note the capacity of official statistics to provide information on themes that arise unexpectedly; for example, *Virology*, closely related to terms such as *Covid-19*, *Coronavirus*, etc., stands out in the last period for its high density and centrality. Therefore, this study makes clear that societal demand for the availability of new topics leads to thematic shifts and updates over time.

“ Show relationship between the demand and supply of official statistics as a source of information that can be used as a basis for promoting from Public Statistics those areas of work that respond to the needs of academic research ”

This work may have various implications:

- For those responsible for public statistics, it confirms that their efforts to regularly update contents through statistical plans and to meet the unexpected needs that arise, such as the need to measure the effects of the covid pandemic, are reflected in the use of this information by social agents, and specifically by the scientific community. On the other hand, this study makes it possible to quantify the use that has been made of public statistical sources for research purposes in each area of knowledge. Where appropriate, this can enable statisticians to adjust or promote the statistical resources provided for this research. It has also been pointed out that National Statistical Offices are called on to play a leading role when it comes to coordinating open data publication policies that can serve citizens and the scientific community. To achieve this, it is essential to ensure that open data and interoperability are key criteria in the creation, implementation, and execution of national statistical platforms from now on. For this reason, the *United Nations Statistical Commission* is currently discussing how to update the fundamental principles of official statistics (Georgiou, 2017) and its statistical manuals to incorporate the principles of open data.
- For the scientific community, the efforts made by national statistics offices to guarantee quality data and make relevant open data available to society, mean public statistics can be valued as a source of open and reliable data for researchers, enabling them to test their theoretical models. In addition, this study allows researchers to find out whether this source of information and its thematic structure and evolution has been used previously in their research topic.
- It is also important to highlight the advantages that official statistics offer to data-driven public and private organizations. The conceptual architecture required for such enterprises has several layers (Moreno; Carrasco; Herrera-Viedma, 2019; Galán; Carrasco; LaTorre, 2022): data storage, insight generation, action, and outcomes. Due to the important role played by data (the first layer of this architecture), this study, which analyzes the use of official (and therefore high quality) data, can be useful when incorporating new sources into this architecture for decision-making.

5. Conclusion and future work

Applying a bibliometric perspective, this paper analyzes the use of official statistics as data sources, either in whole or in part, for scientific research. To that end, we examine 10,777 related articles from the period 1960-2020, extracted from the *Scopus* database. Due to the heterogeneous nature of the fields of knowledge of the themes found, and to simplify their categorization and interpretation, we propose an original classification of the themes, based on an analysis of their thematic network. The authors of this paper combine experience of official statistics with scientific expertise in bibliometrics, which has been essential for reaching and interpreting the conclusions of this study.

Regarding the thematic analysis, the following results can be highlighted:

- The main themes, those with the highest performance measures, have been described in detail for each period. They are closely related to Medicine, Business and Management and Sociology. Other high performance themes from different areas have also been found in recent periods, related to Psychology, Environment and Microbiology.
- We identify a progressive appearance of basic and motor themes from fields other than Medicine (*Mortality...*) and Sociology (*Socioeconomics...*), especially Business and Management (*Health-Care-Cost...*), which have become especially relevant in the last period. Themes related to Psychology also appear in the second period (*Health-Behavior, Suicide*, etc.), while the fourth period sees the appearance of *Virology* in the field of Microbiology, and other themes related to the Environment (*Land-Use, Environmental-Monitoring*).

- On the other hand, a few themes persist as non-central, with the ones related to the Economy being especially relevant. In addition to being scarce, they do not become central in any of the periods. However, from the second period on, Environmental themes that have been gaining centrality over time appear as non-central.

Capacity of official statistics to provide information on topics that arise spontaneously, as happened in the last period with the theme Virology, closely related to terms such as Covid-19

From the longitudinal study and the analysis of the visualizations, the following findings can be noted:

- Several timelines have been identified that relate themes over time, generally from the same field, but in some cases intertwining fields of knowledge. The themes related to Medicine and Sociology persist in all periods as motor and central themes, especially the themes of *Mortality* and *Socioeconomics*, respectively. Concerning the field of Business and Management, the *Health-Care-Cost* theme stands out.
- We identify themes that lose researchers' interest over time (such as *HIV-Infections*, *Developed-Countries*, *Developing-Countries*, *Contraceptive-Usage*, or *Traffic-Accidents*) as well as those that gain greater centrality and density over time, such as the *Cancer* theme. Therefore, we can affirm that slight changes in trends have been observed in research themes based on official statistics.
- It can be said that there is a greater degree of thematic change within the non-central research area. Conversely, during the period analyzed, we find greater stability in the thematic structure within the central area of research.

We have also used the classification by fields of knowledge (subject areas) to determine the number of documents produced in each. In the first period, the studies on Medicine stand out, and in the second those on Sociology, focusing both on central themes as well as peripheral or niche ones. In the second period, in addition to the themes in the field of Medicine and Sociology, those in Business and Management and to a lesser extent Psychology are prominent. The same pattern occurs in the third period, where Environmental themes also begin to appear. In the last period, topics related to Medicine cease to be the most prolific and are overtaken by Business and Management themes in terms of the number of documents produced, with themes related to the Environment also becoming much more relevant. It is also noteworthy that Business and Management issues, despite their importance during all periods, are never in the second quadrant of the strategic diagram (isolated and developed), which is not the case for the other important areas.

Furthermore, we have summarized the most relevant papers, authors, countries, and institutions. Some of the most relevant results are the following:

- the most prolific countries are the UK and the US, followed by Germany, Spain, and the Russian Federation;
- the organizations with the highest number of affiliations are University College of London, Office for National Statistics and the University of Oxford;
- the studies with the highest number of citations are **Grantham-McGregor et al. (2007)** and **Allison et al. (1999)** from the field of Medicine, and those with the highest FWCI index are **Hippisley-Cox et al. (2008)** also from the field of Medicine and **Gillborn (2008)** from the field of Sociology.

The limitations of this research need to be considered. Word co-occurrence analysis enables a broad examination of a topic to gain initial insights into a research area. Therefore, the results of our analysis only scratch the surface of the underlying structure of the research based on official statistics. Our central objective has been to provide an overall impression of the thematic structure of the research based on official statistics. Another critical aspect is the choice of keywords as the basis of the analysis as there are no general rules for assigning keywords when they are selected by the authors. Also, we noticed that approximately 10% of the documents do not contain keywords. In addition, some keywords that have no meaning in this context, such as stop words or words with a very broad meaning, were removed. Despite all this, it can be seen that the keywords examined have a notable signaling effect, with the common appearance of keywords enabling the identification of thematic networks and an underlying thematic structure.

Future studies could focus on finding more detailed and reliable findings by examining smaller scientific subdivisions, but our analysis takes on the challenge of examining all periods and scientific disciplines together. It would also be worth focusing on methodology, due to recent changes in official statistics (**Buskirk; Kirchner, 2020; Groves, 2011**) that are leading to a "Data Science Sampling Age", marked by triangulation, integration, innovation, and refinement. Similarly, another future research task would be to repeat the bibliometric study sometime later to assess the evolution of the findings of this study.

6. References

Abraído-Lanza, Ana F.; Dohrenwend, Bruce P.; Ng-Mak, Daisy S.; Turner, J. Blake (1999). "The Latino mortality paradox: a test of the 'salmon bias' and healthy migrant hypotheses". *American journal of public health*, v. 89, n. 10, pp. 1543-1548. <https://doi.org/10.2105/AJPH.89.10.1543>

- Allison, David B.; Fontaine, Kevin R.; Manson, JoAnn E.; Stevens, June; VanItallie, Theodore B.** (1999). "Annual deaths attributable to obesity in the United States". *Jama* v. 282, n. 16, pp. 1530-1538.
<https://doi.org/10.1001/jama.282.16.1530>
- Alonso, Sergio; Cabrerizo, Francisco-Javier; Herrera-Viedma, Enrique; Herrera, Francisco** (2009). "h-Index: A review focused in its variants, computation and standardization for different scientific fields". *Journal of informetrics*, v. 3, n. 4, pp. 273-289.
<https://doi.org/10.1016/j.joi.2009.04.001>
- Batagelj, Vladimir; Cerinšek, Monika** (2013). "On bibliographic networks". *Scientometrics*, v. 96, n. 3, pp. 845-864.
<https://doi.org/10.1007/s11192-012-0940-1>
- Bernstein, Ira M.; Horbar, Jeffrey D.; Badger, Gary J.; Ohlsson, Arne; Golan, Agneta** (2000). "Morbidity and mortality among very-low-birth-weight neonates with intrauterine growth restriction". *American journal of obstetrics and gynecology*, v. 182, n. 1, pp. 198-206.
[https://doi.org/10.1016/S0002-9378\(00\)70513-8](https://doi.org/10.1016/S0002-9378(00)70513-8)
- Börner, Katy; Chen, Chaomei; Boyack, Kevin W.** (2003). "Visualizing knowledge domains". *Annual review of information science and technology*, v. 37, n. 1, pp. 179-255.
<https://doi.org/10.1002/aris.1440370106>
- Buskirk, Trent D.; Kirchner, Antje** (2020). "Why machines matter for survey and social science researchers: Exploring applications of machine learning methods for design, data collection, and analysis". In: Hill, Craig A.; Biemer, Paul P.; Buskirk, Trent D.; Japac, Lilli; Kirchner, Antje; Kolenikov, Stas; Lyberg, Lars E. *Big data meets survey science: A collection of innovative methods*, pp. 9-62. ISBN: 978 111 897 635 7
<https://doi.org/10.1002/9781118976357.ch1>
- Callon, Michel; Courtial, Jean-Pierre; Laville, Françoise** (1991). "Co-word analysis as a tool for describing the network of interactions between basic and technological research: The case of polymer chemistry". *Scientometrics*, v. 22, pp. 155-205.
<https://doi.org/10.1007/BF02019280>
- Callon, Michel; Courtial, Jean-Pierre; Turner, William A.; Bauin, Serge** (1983). "From translations to problematic networks: An introduction to co-word analysis". *Social science information*, v. 22, n. 2, pp. 191-235.
<https://doi.org/10.1177/053901883022002003>
- Chen, Jie** (2007). "Rapid urbanization in China: A real challenge to soil protection and food security". *Catena*, v. 69, n. 1.
<https://doi.org/10.1016/j.catena.2006.04.019>
- Cobo, Manuel J.; López-Herrera, Antonio G.; Herrera, Francisco; Herrera-Viedma, Enrique** (2011a). "A note on the ITS topic evolution in the period 2000-2009 at T-ITS". *IEEE transactions on intelligent transportation systems*, v. 13, n. 1, pp. 413-420.
<https://doi.org/10.1109/TITS.2011.2167968>
- Cobo, Manuel J.; López-Herrera, Antonio G.; Herrera-Viedma, Enrique; Herrera, Francisco** (2011b). "Science mapping software tools: Review, analysis, and cooperative study among tools". *Journal of the American Society for Information Science and Technology*, v. 62, n. 7, pp. 1382-1402.
<https://doi.org/10.1002/asi.21525>
- Cobo, Manuel J.; López-Herrera, Antonio G.; Herrera-Viedma, Enrique; Herrera, Francisco** (2011c). "An approach for detecting, quantifying, and visualizing the evolution of a research field: A practical application to the Fuzzy Sets Theory field". *Journal of informetrics*, v. 5, n. 1, pp. 146-166.
<https://doi.org/10.1016/j.joi.2010.10.002>
- Cobo, Manuel J.; López-Herrera, Antonio G.; Herrera-Viedma, Enrique; Herrera, Francisco** (2012). "SciMAT: A new science mapping analysis software tool". *Journal of the American Society for Information Science and Technology*, v. 63, n. 8, pp. 1609-1630.
<https://doi.org/10.1002/asi.22688>
- Coulter, Neal; Monarch, Ira; Konda, Suresh** (1998). "Software engineering as seen through its research literature: A study in co-word analysis". *Journal of the American Society for Information Science*, v. 49, n. 13, pp. 1206-1223.
[https://doi.org/10.1002/\(SICI\)1097-4571\(1998\)49:13<1206::AID-ASIT>3.0.CO;2-F](https://doi.org/10.1002/(SICI)1097-4571(1998)49:13<1206::AID-ASIT>3.0.CO;2-F)
- España** (2020). "Real decreto 1110/2020, de 15 de diciembre, por el que se aprueba el Plan Estadístico Nacional 2021-2024". *BOE*, n. 340, 30 diciembre.
https://www.boe.es/diario_boe/txt.php?id=BOE-A-2020-17283

- EU (2019). "Directiva (UE) 2019/1024 del Parlamento Europeo y del Consejo relativa a los datos abiertos y la reutilización de la información del sector público (versión refundida)". *Diario oficial de la UE*, 26 junio.
<https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:32019L1024&from=ES>
- EU (2022). "Reglamento (UE) 2022/868 del Parlamento Europeo y del Consejo de 30 de mayo de 2022 relativo a la gobernanza europea de datos y por el que se modifica el Reglamento (UE) 2018/1724 (Reglamento de Gobernanza de Datos)". *Diario oficial de la Unión Europea*, n. 152, 3 junio.
<https://www.boe.es/buscar/doc.php?id=DOUE-L-2022-80835>
- Galán, José-Javier; Carrasco, Ramón-Alberto; LaTorre, Antonio** (2022). "Military applications of machine learning: A bibliometric perspective". *Mathematics*, v. 10, n. 9, 1397.
<https://doi.org/10.3390/math10091397>
- García-Villar, Jaume** (2012). "La estadística oficial como bien público: retos del presente". *Revista asturiana de economía*, n. 46, pp. 61-85.
<https://dialnet.unirioja.es/servlet/articulo?codigo=4204191>
- Georgiou, Andreas V.** (2017). "Towards a global system of monitoring the implementation of UN fundamental principles in national official statistics". *Statistical journal of the IAOS*, v. 33, n. 2, pp. 387-397.
<https://doi.org/10.3233/SJI-160335>
- Gillborn, David** (2008). *Racism and education: Coincidence or conspiracy?*. Routledge. ISBN: 978 0 203928424
<https://doi.org/10.4324/9780203928424>
- Giovannini, Enrico; Martins, J. Oliveira; Gamba, Michela** (2009). "Statistics, knowledge and governance". *Statistika*, v. 6, pp. 471-490.
- Giri, Chandra; Ochieng, Edward; Tieszen, Larry L.; Zhu, Zq; Singh, Ashbindu; Loveland, Tomas; Masek, Jeff; Duke, Norman** (2011). "Status and distribution of mangrove forests of the world using earth observation satellite data". *Global ecology and biogeography*, v. 20, n. 1, pp. 154-159.
<https://doi.org/10.1111/j.1466-8238.2010.00584.x>
- Grantham-McGregor, Sally; Cheung, Yin-Bun; Cueto, Santiago; Glewwe, Paul; Richter, Linda; Strupp, Barbara** (2007). "Developmental potential in the first 5 years for children in developing countries". *The lancet*, v. 369, n. 9555, pp. 60-70.
[https://doi.org/10.1016/S0140-6736\(07\)60032-4](https://doi.org/10.1016/S0140-6736(07)60032-4)
- Groves, Robert M.** (2011). "Three eras of survey research". *Public opinion quarterly*, v. 75, n. 5, pp. 861-871.
<https://doi.org/10.1093/poq/nfr057>
- Gustavsson, Anders; Svensson, Mikael; Jacobi, Frank; Allgulander, Christer; Alonso, Jordi; Beghi, Ettore; Dodel, Richard; Ekman, Mattias; Faravelli, Carlo; Fratiglioni, Laura; Gannon, Brenda; Jones, David H.; Jennum, Poul; Jordanova, Albena; Jönsson, Linus; Karampampa, Korinna; Knapp, Martin; Kolbelt, Gisela; Kurth, Tobias; Lieb, Roselind; Linde, Mattias; Ljungcrantz, Christina; Maercker, Andreas; Melin, Beatrice; Moscarilli, Massimo; Musayev, Amir; Norwood, Fiona; Preisig, Martin; Pugliatti, Maura; Rehm, Juergen; Salvador-Carulla, Luis; Schlehofer, Brigitte; Simon, Roland; Steinhausen, Hans-Christoph; Stovner, Lars J.; Vallat, Jean-Michel; Van-den-Bergh, Peter; Van-Os, Jim; Vos, Pieter; Xu, Weili; Wittchen, Hans-Ulrich; Jönsson, Bengt; Olesen, Jes** (2011). "Cost of disorders of the brain in Europe 2010". *European neuropsychopharmacology*, v. 21, n. 10, pp. 718-779.
- Harford, Tim** (2014). "Big data: A big mistake?". *Significance*, v. 14, n. 5, pp. 14-19.
<https://rss.onlinelibrary.wiley.com/doi/full/10.1111/j.1740-9713.2014.00778.x>
- He, Qin** (1999). "Knowledge discovery through co-word analysis". *Library trends*, v. 48, n. 1, p. 133-159.
<https://www.ideals.illinois.edu/items/8226>
- Henderson, J. Vernon; Storeygard, Adam; Weil, David N.** (2012). "Measuring economic growth from outer space". *American economic review*, v. 102, n. 2, pp. 994-1028.
<https://doi.org/10.1257/aer.102.2.994>
- Hippisley-Cox, Julia; Coupland, Carol; Vinogradova, Yana; Robson, John; Minhas, Rubin; Sheikh, Aziz; Brindle, Peter** (2008). "Predicting cardiovascular risk in England and Wales: prospective derivation and validation of QRISK2". *British medical journal*, v. 336, pp. 1475-1482.
<https://doi.org/10.1136/bmj.39609.449676.25>
- Hirsch, Jorge E.** (2005). "An index to quantify an individual's scientific research output". *Proceedings of the National Academy of Sciences*, v. 102, n. 46, pp. 16569-16572.
<https://doi.org/10.1073/pnas.0507655102>

- Jacob, Ulrike; Brust, Oliver A.** (2019). "Confronting the anomaly: directions in (German) economic research after the crisis". *Science in context*, v. 32, n. 4, pp. 449-471.
<https://doi.org/10.1017/S026988972000006X>
- Liu, Zhu; Guan, Dabo; Wei, Wei; Davis, Steven J; Ciais, Philippe; Bai, Jin; Peng, Shushi; Zhang, Qiang; Hubacek, Klaus; Marland, Gregg; Andres, Robert J.; Crawford-Brown, Douglas; Lin, Jintai; Zhao, Hongyan; Hong, Chaopeng; Boden, Thomas A.; Feng, Kuishuang; Peters, Glen P.; Xi, Fengming; Liu, Junguo; Li, Yuan; Zhao, Yu; Zeng, Ning; He, Kebin** (2015). "Reduced carbon emission estimates from fossil fuel combustion and cement production in China". *Nature*, v. 524, n. 7565, pp. 335-338.
<https://doi.org/10.1038/nature14677>
- Llewelyn, Charlotte A.; Hewitt, Patricia E.; Knight, Richard S. G.; Amar, K.; Cousens, S.; Mackenzie, J.; Will, Robert G.** (2004). "Possible transmission of variant Creutzfeldt-Jakob disease by blood transfusion". *The lancet*, v. 363, n. 9407, pp. 417-421.
[https://doi.org/10.1016/S0140-6736\(04\)15486-X](https://doi.org/10.1016/S0140-6736(04)15486-X)
- Lnenicka, Martin; Luterek, Mariusz; Nikiforova, Anastasija** (2022). "Benchmarking open data efforts through indices and rankings: Assessing development and contexts of use". *Telematics and informatics*, v. 66, 101745.
<https://doi.org/10.1016/j.tele.2021.101745>
- Martínez, Miguel-Ángel; Herrera, Manuel; López-Gijón, Javier; Herrera-Viedma, Enrique** (2004). "H-Classics: Characterizing the concept of citation classics through". *Scientometrics*, v. 98, pp. 1971-1983.
<https://doi.org/10.1007/s11192-013-1155-9>
- Moreno, Caio; Carrasco, Ramón-Alberto; Herrera-Viedma, Enrique** (2019). "Data and artificial intelligence strategy: A conceptual enterprise big data architecture to enable market-oriented organizations". *IJIMAI*, v. 5, n. 6, pp. 7-14.
<https://doi.org/10.9781/ijimai.2019.06.003>
- Mukherjee, Debmalya; Lim, Weng-Marc; Kumar, Satish; Donthu, Naveen** (2022). "Guidelines for advancing theory and practice through bibliometric research". *Journal of business research*, v. 148, pp. 101-115.
<https://doi.org/10.1016/j.jbusres.2022.04.042>
- Multicentre Aneurysm Screening Study Group** (2002). "The Multicentre Aneurysm Screening Study (MASS) into the effect of abdominal aortic aneurysm screening on mortality in men: a randomised controlled trial". *The lancet*, v. 360, n. 9345, pp. 1531-1539.
[https://doi.org/10.1016/S0140-6736\(02\)11522-4](https://doi.org/10.1016/S0140-6736(02)11522-4)
- Olesen, Jes; Gustavsson, Anders; Svensson, Mikael; Wittchen, H.-U.; Jönsson, Bengt** (2012). "The economic cost of brain disorders in Europe". *European journal of neurology*, v. 19, n. 1, pp. 155-162.
<https://doi.org/10.1111/j.1468-1331.2011.03590.x>
- Peto, Richard; Darby, Sarah; Deo, Harz; Silcocks, Paul; Whitley, Elise; Doll, Richard** (2000). "Smoking, smoking cessation, and lung cancer in the UK since 1950: combination of national statistics with two case-control studies". *British medical journal*, v. 321, n. 7257, pp. 323-329.
<https://doi.org/10.1136/bmj.321.7257.323>
- Purkayastha, Amrita; Palmaro, Eleonora; Falk-Krzesinski, Holly J.; Baas, Jeroen** (2019). "Comparison of two article-level, field-independent citation metrics: Field-Weighted Citation Impact (FWCI) and Relative Citation Ratio (RCR)". *Journal of informetrics*, v. 13, n. 2, pp. 635-642.
<https://doi.org/10.1016/j.joi.2019.03.012>
- Radermacher, Walter J.** (2014). "The European statistics code of practice as a pillar to strengthen public trust and enhance quality in official statistics". *Statistical and Social Inquiry Society of Ireland*, v. 43, pp. 27-33.
<http://www.tara.tcd.ie/handle/2262/72773>
- Rodríguez-Ledesma, Antonio; Cobo, Manuel J.; López-Pujalte, Cristina; Herrera-Viedma, Enrique** (2015). "An overview of animal science research 1945-2011 through science mapping analysis". *Journal of animal breeding and genetics*, v. 132, n. 6, pp. 475-497.
<https://doi.org/10.1111/jbg.12124>
- Rodríguez-López, María-Eugenia; Alcántara-Pilar, Juan-Miguel; Del-Barrio-García, Salvador; Muñoz-Leiva, Francisco** (2020). "A review of restaurant research in the last two decades: A bibliometric analysis". *International journal of hospitality management*, v. 87, 102387.
<https://doi.org/10.1016/j.ijhm.2019.102387>
- Savage, Mike; Devine, Fiona; Cunningham, Niall; Taylor, Mark; Li, Yaojun; Hjellbrekke, Johs; Le-Roux, Brigitte; Friedman, Sam; Miles, Andrew** (2013). "A new model of social class? Findings from the BBC's great British class survey experiment". *Sociology*, v. 47, n. 2, pp. 219-250.
<https://doi.org/10.1177/0038038513481128>

Small, Henry (1999). "Visualizing science by citation mapping". *Journal of the American Society for Information Science*, v. 50, n. 9, p. 799-813.

[https://doi.org/10.1002/\(SICI\)1097-4571\(1999\)50:9<799::AID-ASI9>3.0.CO;2-G](https://doi.org/10.1002/(SICI)1097-4571(1999)50:9<799::AID-ASI9>3.0.CO;2-G)

Sternitzke, Christian; Bergmann, Isumo (2009). "Similarity measures for document mapping: a comparative study on the level of an individual scientist". *Scientometrics*, v. 78, p. 113-130.

<https://doi.org/10.1007/s11192-007-1961-z>

Unal, Belgin; Critchley, Julia-Alison; Capewell, Simon (2004). "Explaining the decline in coronary heart disease mortality in England and Wales between 1981 and 2000". *Circulation*, v. 109, n. 9, pp. 1101-1107.

<https://doi.org/10.1161/01.CIR.0000118498.35499.B2>

United Nations (2014). "Statistical Commission Decision 45/110 on big data and modernization of statistical systems". *Statistics Division*, Decision code: 45/110, chapter: I, section: B.

<https://unstats.un.org/unsd/statcom/decisions-ref/?code=45/110>

Unece (2017). *Guidance on land-use planning, the siting of hazardous activities and related safety aspects*. Geneva, Switzerland: United Nations Economic Commission for Europe.

https://unece.org/DAM/env/eia/Publications/2017/1735403E_Final_ENG_web.pdf

Unece (2019). *Use of satellite image and Earth observation data in official statistics*. Geneva, Switzerland: United Nations Economic Commission for Europe.

https://unece.org/DAM/stats/documents/ece/ces/2019/ECE_CES_2019_16-1906490E.pdf

Unece (2020). *Guidelines on the shared environmental information system reporting mechanism*. Geneva, Switzerland: United Nations Economic Commission for Europe.

https://unece.org/sites/default/files/2020-12/2014795E_WEB.pdf

Unece (2021). *Measuring and monitoring progress towards the Sustainable Development Goals*. Geneva, Switzerland: United Nations Economic Commission for Europe.

https://unece.org/sites/default/files/2021-04/2012761_E_web.pdf

United Nations (2021). *The handbook on management and organization of National Statistical Systems*.

<https://unstats.un.org/capacity-development/handbook/html/topic.htm#t=Handbook%2FCover%2FCover.htm>

Van-Dijk, Jan; Hacker, Kenneth (2003). "The digital divide as a complex and dynamic phenomenon". *The information society*, v. 19, n. 4, pp. 315-326.

<https://doi.org/10.1080/01972240309487>

Vikat, Andres; Jones, Christopher (2014). *Indicators of gender equality*. Unece: Geneva, Switzerland.

https://unece.org/sites/default/files/2022-02/ECE_CES_37_WEB.pdf



Profesional de la información

Servicio de traducciones al inglés

<https://www.profesionaldelainformacion.com/wp-content/uploads/2023/02/traducciones.pdf>

Información: **Isabel Olea**
epi.iolea@gmail.com

Peer review: the attitudes and behaviours of Covid-19 pandemic-era early career researchers

David Nicholas; Eti Herman; Blanca Rodríguez-Bravo; Anthony Watkinson; Cherifa Boukacem-Zeghmouri; Marzena Świgoń; Abdullah Abrizah; David Sims; Jie Xu; David Clark; Galina Serbina; Hamid R. Jamali; Carol Tenopir; Suzie Allard

Nota: Este artículo se puede leer en español en:
<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87375>

Recommended citation:

Nicholas, David; Herman, Eti; Rodríguez-Bravo, Blanca; Watkinson, Anthony; Boukacem-Zeghmouri, Chérifa; Świgoń, Marzena; Abrizah, Abdullah; Sims, David; Xu, Jie; Clark, David; Serbina, Galina; Jamali, Hamid R.; Tenopir, Carol; Allard, Suzie (2023). "Peer review: the attitudes and behaviours of Covid-19 pandemic-era early career researchers". *Profesional de la información*, v. 32, n. 3, e320306.

<https://doi.org/10.3145/epi.2023.may.06>

Manuscript received on 12th April 2023
Accepted on 30th April 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



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



Anthony Watkinson

<https://orcid.org/0000-0002-2317-6557>

CIBER Research Ltd
Newbury, Berkshire
RG147RU, United Kingdom
anthony.watkinson@btinternet.com



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



Marzena Świgoń

<https://orcid.org/0000-0003-3600-8349>

Uniwersytet Warmińsko-Mazurski
Wydział Humanistyczny
10-719 Olsztyn, Poland
marzena.swigon@uwm.edu.pl



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



David Sims

<https://orcid.org/0000-0003-1916-5617>

University of Tennessee
School of Information Sciences
1345 Circle Park Drive, 423
Knoxville, Tennessee 37996-0341, USA
dsims7@utk.edu





Jie Xu

<https://orcid.org/0000-0002-9820-8066>

Wuhan University
School of Information Management
Hubei 430072, Wuhan, China
xuj@whu.edu.cn



David Clark

<https://orcid.org/0000-0002-5893-4274>

CIBER Research Ltd
Newbury, Berkshire
RG147RU, United Kingdom
david.clark@ciber-research.uk



Galina Serbina

<https://orcid.org/0000-0003-0196-6010>

Tomsk State University
TSU Research Library
Lenin st. 34a, 634050 Tomsk, Russia
serbina@lib.tsu.ru



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



Suzie Allard

<https://orcid.org/0000-0001-9421-3848>

University of Tennessee
School of Information Sciences
1345 Circle Park Drive, 423
Knoxville, TN 37996-0341, USA
sallard@utk.edu

Abstract

Explores science and social science early career researchers' (ECRs) perceptions and experiences of peer review, seeking also to identify their views of any pandemic-associated changes that have taken place. Data are drawn from the *Harbingers-2* project, which investigated the impact of the pandemic on scholarly communications. Peer review, one of the activities covered, is singled out as it proved to be the activity of greatest concern to ECRs. Findings are obtained from interviews, which covered around 167 ECRs from China, France, Malaysia, Poland, Russia, Spain, UK and US, supplemented by an international survey that took the data out to a bigger and wider audience for confirmation and generalisation. Results obtained are enhanced by comparisons with pre-pandemic evidence yielded by *Harbingers-1*, the forerunner of the present study, and anchored in an extensive review of the literature. Main findings are: 1) most ECRs were experienced in peer review, both as reviewers and authors, but few had formal training; 2) half the ECRs had a lot or some reservations as to whether peer review vouches for the trustworthiness of research; 3) inadequate reviewers and slow processes were the main peer review associated problems; 4) there was a strong feeling that some kind of compensation, whether monetary or reputational, could help in dealing with these problems; 5) the pandemic impacted most on the speed of processing, with the majority of ECRs saying it had slowed the process; 6) nearly everyone thought that any pandemic-induced impacts would be temporary.

Keywords

Research; Scholarly communication; Scientific communication; Young researchers; ECRs; Peer review; Reliability in peer review; Pandemics; Covid-19; Harbingers Project; Impacts; Consequences of the pandemic; interviews; Peer review trustworthiness; Surveys; Resilience; Differences between countries; China; France; Malaysia; Poland; Russia; Spain; United Kingdom; USA.

Funding

This paper comes from the *Harbingers-2* project: *Early career researchers and the pandemic*, funded by the Alfred P. Sloan Foundation (<http://www.sloan.org>).

1. Introduction

We have learnt about early career researchers' (ECRs) views and experiences of peer review in the course of the just-concluded, international, six-year long, longitudinal *Harbingers* research project, which investigated the working lives and scholarly communication behavior of junior science and social science researchers. The project had as its basic premise the belief that ECRs –tomorrow's leading professors, influencers, and decision makers– hold the key to understanding where the scholarly communications is heading, indeed, will be instrumental in transforming it. After all, they are the new generation, traditionally held to be trailblazers, and, as such, disruptors of the established order of things, as well as millennials, who characteristically possess generational values of openness to change and community-mindedness (Burstein, 2013; Duffy; Shrimpton; Clemence, 2017; FEPS; ThinkYoung, 2018; Pew Research Center, 2010a; 2010b; Schewe et al., 2013, Sørensen et al., 2017).

In fact, the pandemic might have tipped the scale toward their giving precedence and acting upon their millennial tendencies, for, as Nugin and Kalmus (2022) suggest, it is particularly during profound societal changes that young people can become agents of social transformation, rather than being objects of socialization. With the pandemic creating the kind of changed and challenging societal reality that calls for and incentivizes new ways of coping (Herman et al., 2021), the likelihood that the new generation of researchers will mobilize transformations in the scholarly system has arguably become more pronounced.

Of course, it is only to be expected that in their endeavours to bring about change, ECRs, as newcomers to academe, would focus first on the system's age-old problems, a major one of which, as the literature amply evidences (see the *Background and context* section), is peer review. Not only has it long been a much-debated aspect of the scholarly communication system, but recently it unmistakably came even more to the fore, figuring as highly as it did among the weaknesses that the pandemic exacerbated or at least shed new light on. Indeed, the evidence obtained over the years of the *Harbingers* project has proven the peer review process to be a major reason for concern among ECRs, with the scale, seriousness and centrality of its weaknesses having come top of the list among the 'cracks' they have identified in the pandemic-era scholarly system (Nicholas et al., 2022a; 2022b; 2023).

This paper seeks to follow-up on this finding and present a comprehensive portrayal of ECRs' perceptions and practices of peer review by examining in detail the vast amount of data we have collected, inclusive of the impact that the pandemic might have had on its processes. We are well-placed to do so, as we have been investigating ECRs' attitudes to and practices of peer review in context and in the round and exploring its relationships to other scholarly activities over six-years, the whole duration of the project: in *Harbingers-1*¹, the first leg of the undertaking (2016-2019), which directly preceded the pandemic, and in *Harbingers-2*², the *Alfred P. Sloan Foundation*³ funded extension to the project (2020-2022), conducted during and after the pandemic. The analysis of the peer review process, as ECRs see it, thus covers ECRs' training for the role of referees, their attitudes to and experiences of the peer review process, their suggestions for its improvement as well as their views of the impact of the pandemic on peer review.

2. Aims

The overarching aim of this study is to establish ECRs' perceptions and experiences of peer review, seeking in particular to identify their views of any pandemic-associated, possibly long-term changes that have been taking place. Within this broad aim the paper shall seek to find out:

- How experienced in peer review ECRs are.
- What training for the role of reviewers ECRs obtain.
- What ECRs think of the peer review process: its strengths, weaknesses and future.
- What suggestions ECRs have (if any) for improving peer review.
- What ECRs believe to be the impact of the pandemic on peer review (if any).

3. Background and context

Peer review, the

“social mechanism through which a discipline's ‘experts’ maintain quality control over new knowledge entering the field” (Berkenkotter, 1995, p. 245),

has been described as

“the lynchpin about which the whole business of science is pivoted” (Ziman, 1968, p. 111),

indeed, as

“the institutionalised practice... [that] glues the academy together” (Neylon, 2018).

Justifiably so, of course, for the procedure aims at safeguarding the quality, novelty, reliability, soundness, theoretical and empirical validity, and potential impact of new knowledge produced (Eve et al., 2021; Mulligan; Hall; Raphael, 2013; Nicholas et al., 2015b). Just what a central role peer review plays in the scholarly undertaking has never been more clearly demonstrated than in pandemic times, when the crucial need for rapid dissemination of relevant scientific knowledge shone a strong light on its direct, and very powerful bearing on the publication process (Horbach, 2020). Still,

whilst stakeholders in the academic enterprise agree that peer review, per se, is indispensable (Nicholas *et al.*, 2015b; 2019; *Publishing Research Consortium*, 2016; Tennant; Ross-Hellauer, 2020), it nevertheless seems to be the most debated and charged aspect of the scholarly communication.

Indeed, peer review has been found in study after study to be wanting, manifesting as it does a host of characteristic limitations. Thus, inter alia, it has been criticised for bias and unfair assessment (Demarest; Freeman; Sugimoto, 2014; Haffar, Bazerbachi; Murad, 2019; Lee *et al.*, 2013; Silbiger; Stubler, 2019), arbitrariness (Brezis; Birukou, 2020; Roubanis, 2022), problematic scientific gate-keeping (Bartneck, 2017; Flaherty, 2022; Jubb, 2016; Seeber, 2022), suppression of innovation (Siler; Lee; Bero, 2015), ineffective detecting of error or fraud (Brainard; You, 2018; Horbach; Halffman, 2019), and delays in publishing (Allen *et al.*, 2022; Christie *et al.*, 2021). The flurry of retractions in the past decade or so (Hesselmann *et al.*, 2017; Sharma, 2021; Steen; Casadevall; Fang, 2013), and especially during the pandemic (Kodvanj *et al.*, 2022; Shimray, 2022), which can arguably be traced back, among other reasons, to oversight in peer review and editorial verification, certainly seems to imply that there are inadequacies in the system.

By the same token, so does the unrelenting search for achieving more equitable, fair, efficient and effective ways and means of doing peer review. Indeed, the past few years have seen numerous innovations in peer review and quality assurance in scholarly publishing (Woods *et al.*, 2023). These range from initiatives striving to make the process more open, in order to increase referee accountability and minimize their bias (Ross-Hellauer, 2017), through efforts at de-coupling the procedure from journal publishing and making it more collaborative and community-led (Tennant *et al.*, 2017), to recent technological advances, inclusive of AI-assisted ones, which aim at rendering it more efficient (Barroga, 2020; Horbach; Halffman, 2018).

Plainly then, peer review occupies the hearts and minds of the scholarly community, and with good reason, too: as Squazzoni and Gandelli (2013) suggest, summing it all up in a nutshell, it is peer review that determines, whether directly or indirectly, how the resources of the science system –including funding, positions and reputation– are allocated. Indeed, with the publication of research achievements hinging on peer review, the refereeing process assumes a pivotal role in shaping scholarly careers, for –to reiterate what by now has become a truism– research productivity is held to be a major yardstick, if not *the* major yardstick by which scholarly success is measured (Blankstein; Wolff-Eisenberg, 2019; Herman, 2018; Herman; Nicholas, 2019; Nicholas *et al.* 2015a; Niles *et al.*, 2020; Van-Dalen; Henkens, 2012).

This, of course, is particularly so in the case of novice researchers, yet to prove their suitability to an academic career by producing an impressive publishing record (Jamali *et al.*, 2020; Nicholas *et al.*, 2017a; 2018a; 2020a; 2020b). With the pandemic negatively impacting many aspects of their work-life (Christian *et al.*, 2021; Douglas *et al.*, 2022; Herman *et al.*, 2021; Jamali *et al.*, 2023; Johnson *et al.*, 2021; Levine *et al.*, 2021; Morin *et al.*, 2022), ECRs' ability to successfully navigate the turbulent waters of the publishing process, inclusive of its peer review component, seems bound to be further affected.

4. Scope

The findings reported here come from the aforementioned, Sloan Foundation-funded, *Harbingers-2 - Early Career Researchers and the Pandemic Research Project*, the two-year extension to the original, four-year *Harbingers-1* study (for more detail see the *Methods* section). The data are drawn from both the interview leg of the project, which was at the heart of the research, and the international survey that subsequently extended and generalised it, albeit majoring on the former. There were three rounds of interviews, but for the purposes of this paper we are only taking data from the final round, when ECRs would have had two years familiarity with peer review in pandemic times and would thus be in a good position to comment on it. The findings are enhanced, where relevant, by the pre-pandemic results obtained in *Harbingers-1*.

5. Definitions

Peer review

Peer review is definable, as Ross-Hellauer (2017) suggests, as the formal quality assurance mechanism whereby scholarly manuscripts (e.g., journal articles, books, grant applications and conference papers) are made subject to the scrutiny of others, whose feedback and judgements are then used to improve works and make final decisions regarding selection (for publication, grant allocation or speaking time).

ECRs

With different, conflicting and country-specific definitions of an ECR circulating (Teixeira-Da-Silva, 2021), a pragmatic conceptualization of the term was decided on. Our definition of an ECR focusses on the common denominators of their standing, that is, their being employed in a research position but, being relatively young and in an early phase of their career, not yet established as permanent faculty:

“Researchers who are generally not older than 40, who either have received their doctorate and are currently in a research position or have been in research positions, but are currently doing a doctorate. In neither case are they researchers in established or tenured positions. In the case of academics, some are non-tenure line faculty research employees”.

Subject/discipline

The reference to subject/disciplinary representativeness in this paper builds on **Fanelli & Glänzel's** (2013) findings, which support the 'gradualist' view of scientific knowledge suggested by the *Hypothesis of the Hierarchy of Sciences* –the placing of each field of research, moving from the physical to the social sciences, along a continuum of complexity and softness. Thus, the wide disciplinary area of Social Sciences is divided here into 'hard' disciplines, such as Psychology and 'soft' disciplines, such as Sociology.

6. Methods

The *Harbingers-2* project continues the mixed methodology approach of *Harbingers-1*, as detailed in **Nicholas et al.** (2019; 2020a) and on the project website¹. This provides for a high-degree of data triangulation: an ongoing literature review, semi-structured interviews and a questionnaire survey to test the results on a larger and more diverse population. The data are drawn from both the interview leg of the project, and the international survey that subsequently extended and generalised it, albeit majoring on the former. The findings are enhanced, where relevant and possible, by the pre-pandemic results obtained in *Harbingers-1*, which covered roughly the same disciplines and countries, although the number and composition of the cohort differed a little, and an extensive, analytical literature review.

6.1. Interviews

As noted, the interview stage of *Harbingers-2* consisted of three rounds of interviews, but the results presented here come from the final round, two years into the pandemic (Spring of 2022), when ECRs were best placed to report on their experiences of peer review as things unfolded on the ground.

Recruitment

Interviewees for *Harbingers-2* included both ECRs who participated in *Harbingers-1* and new ones, recruited to fill the ranks of participants who had left research or no longer qualified as ECRs. New (and old) ECRs were recruited by the eight national interviewers, utilizing their local networks and connections, with numbers supplemented by mail-outs from scholarly publisher lists. Each country was given a quota of interviewees (between 20 and 24) to achieve an element of representativeness in terms of age, gender and subject and also to ensure some consistency across countries.

Characteristics of the sample

177 ECRs were recruited for the first round of interviews, of which 167 remained for the third round, mainly because of ECRs' moving out of academe/research. Table 1 provides a country and discipline breakdown of the cohort as of the final round of interviews. The main differences between the rounds of interviews in terms of make-up is that in the final round of interviews there were: a) 3 fewer French ECRs; b) 3 fewer youngest ECRs, aged 24-28. The disciplinary and gender balance remained much the same.

Table 1. Discipline and country breakdown of ECRs (Interview 3)

	Total	CHEM	ENV	LIFE	MATH	MED	PHY	SOCH	SOCS
China	23	0	0	0	5	9	5	1	3
	14%				22%	39%	22%	4%	13%
France	17	2	0	2	3	2	5	0	3
	10%	12%		12%	18%	12%	29%		18%
Malaysia	20	1	0	3	4	2	2	5	3
	12%	5%		15%	20%	10%	10%	25%	15%
Poland	22	2	3	3	3	3	4	1	3
	13%	9%	14%	14%	14%	14%	18%	5%	14%
Russia	20	3	2	2	2	3	4	3	1
	12%	15%	10%	10%	10%	15%	20%	15%	5%
Spain	20	3	3	2	2	2	2	4	2
	12%	15%	15%	10%	10%	10%	10%	20%	10%
UK	24	1	2	5	2	6	2	4	2
	14%	4%	8%	21%	8%	25%	8%	17%	8%
US	21	2	2	3	2	5	3	2	2
	13%	10%	10%	14%	10%	24%	14%	10%	10%
Total	167	14	12	20	23	32	27	20	19
	100%	8%	7%	12%	14%	19%	16%	12%	11%

Interview procedure

Interviews were conducted remotely via *Zoom* or similar video conferencing systems, and in the national language, except Malaysia where English is widely understood. They lasted between 60-100 minutes, were recorded, transcribed and translated (if necessary) and returned to the interviewees for agreement and comment.

The interview protocol contained a mix of closed, open and hybrid questions, covering every aspect of the scholarly system: job, status, career aims/progression, assessment, research directions, working life, reputation, as well as scholarly communications practices –collaboration, searching/finding, networking, ethics, informal communication (preprints, blog posts, posters), social media, publishing, metrics, sharing, outreach and transformations. There were 9 direct questions about peer review and, in addition, a free-text search was conducted for mentions of peer review in responses to other questions. Indeed, ECRs volunteered information about peer review in the context of a further 33 questions and sub-questions, covering more than a dozen scholarly activities, undoubtedly a testimony to the ubiquity and importance of the topic. These mentions of peer review were also included in the data analysed and reported here.

Data analysis

All the interview transcripts were transferred by the national interviewers to a coding sheet, which closely matched the questions of the original interview schedule, but left room for information derived from additional enquiries or clarifications during the interview process. A mapping was maintained so that same, revised, and new questions could be matched between schedule, coding and the three rounds of interviews. The coding sheets were multi-faceted, containing both quantitative and qualitative data, and often a question generated both. For each question, the code sheet captured the interviewee's response in three ways: 1) as a code (e.g., Y/N); 2) as a quotation; 3) and as an explanatory comment from the interviewer. We refer to 2) and 3) as being free-text comments. Not all questions were encoded, as some sought a more extensive explanation, which is recoded only as quotation and commentary. Quotations and commentary were 'at will', they vary in extent and quality and, often for coded responses, were omitted by respondents. Free-text data were analysed using thematic coding, based on themes from the questions and any new themes emerging from the data.

6.2. Survey

A questionnaire, informed by the interview phase of the study, was developed, piloted and sent out towards the end of the study (Summer of 2022) in order to obtain further data on key interview findings from a bigger, more international and disciplinary population of early career researchers. There were just two questions concerning peer review, which asked ECRs whether, based on their own personal experience, the pandemic affected: a) the quality standards of peer review; b) the speed of review.

Recruitment

We did not have a sampling framework because there is no register of ECRs in any of the case study countries. Therefore, a probability sampling approach was not possible, and it was decided to distribute the survey as widely as possible through various channels. Four methods were used for distribution:

- Invitations were sent out by scholarly publishers or relevant institutions to potential ECRs (e.g., *Taylor and Francis*).
- A link to the survey was tweeted by publishers or relevant institutions to researchers (e.g., *Oxford University Press*).
- A banner image with a link to the survey was put on *Wiley Digital Library* and anyone who saw the banner while visiting a journal or article, and was interested, could click and go to the survey.
- Direct invitation emails and texts were sent to ECRs at universities in the case study countries by the national interviewers.

Survey tool

The questionnaire contained 17 questions, two of which, as noted, were on the impact of the pandemic on peer review. The questionnaire was translated by the research team into Chinese, French, Polish and Spanish and was hosted on *Qualtrics* in the third quarter of 2022. For the sake of consistency and to be able to triangulate the data, we used the same definition for ECR used in the other phases of the *Harbingers-2* project (see the *Definitions* section). The survey started with a screening question that asked respondents to self-identify whether they were an ECR based on the definition. Those who said *No* in answer to the screening question exited the survey.

Characteristics of respondents

After data cleaning, 800 responses remained for analysis. Slightly more women (440, 55%) than men (314, 39.3%) participated. The majority were 31 years or older (560, 70.1%). The disciplinary distribution of respondents was biased towards the social sciences (294, 36.8%) and was followed in rank by the life/biological sciences (158, 19.8%). Chemical sciences (21, 2.6%) and mathematical sciences (28, 3.5%) had the lowest numbers of respondents. Respondents came from 71 countries, with those based in the USA accounting for slightly more than a third of responses (285, 35.6%), followed by a large gap to China (61), Spain (48), France (40), Australia (31), Malaysia (30), India (28), UK (27) and Poland (17).

Comparing the make-up of the survey respondents with those who participated in the interview phase of the study, the survey respondents were more international, more American, on average older, and, unlike the interview cohort, some came from the arts and humanities, too. Health/medical sciences were the largest disciplinary group in interviews, whereas they were the third largest group in the survey.

Data analysis

Statistical analysis including descriptive (frequency and percentage) and some inferential (non-parametric Chi-square, Mann Whitney U, and Kruskal-Wallis H tests) were conducted using the *Statistical Package for the Social Sciences (SPSS)*. Non-parametric tests were used because of the nature of variables (some nominal or ordinal) and lack of normality of the data. For six questions with Likert options (strongly disagree to strongly agree, or significant negative impact to significant positive impact) the mean value was also calculated using numeric values of the options (1 being 'strongly disagree' or 'significant negative impact' and 5 being 'strongly agree' or 'significant positive impact'). 'Don't know' or 'not sure' options were excluded in the mean calculation for these questions. Comparisons between countries were only made in the case of the seven countries that were included in the qualitative phase of the study (Russia was also initially part of the project but was missing due to problems resulting from the war in Ukraine), so that we could compare the survey findings with interviews.

7. Results and discussion

7.1. ECRs' involvement in peer review

Experience

Three interview questions sought to find out how experienced ECRs were when it came to peer review, in order to enable us to weigh the depth of their understanding by establishing how informed they were about the process:

- Are they involved in responding to criticisms of their/their groups' publications?
- Have they undertaken peer review themselves?
- Have they had any formal and/or informal training for these roles?

In response to the first question as to whether they were involved in responding to criticisms/suggestions of reviewers on papers that they/their research group authored, the vast majority (89%) of ECRs answered yes (Table 2). This came as no real surprise, as the *Harbingers-1* findings have already indicated that ECRs were keen to participate in responding to reviewers commenting on their work: 85% of the 116 interviewees said so (Rodríguez-Bravo *et al.*, 2017).

Chinese and Malaysian ECRs were the most experienced, with all of the interviewees in both of these groups saying they had been involved in responding to referees. Russian ECRs, with half of the group answering 'no' to the question, were the least experienced.

As to ECRs' extent of experience as reviewers – belying their junior status, well-over three-quarters of the interviewees have undertaken peer reviews of other people's papers (Table 3). Again, not much of a surprise here, as we have already been told by the participants of the *Harbingers-1* interviews, 58% of whom reported having had reviewing experience, that it is quite common for ECRs to stand in for their busy seniors when it comes to performing the task (Rodríguez-Bravo *et al.*, 2017). US ECRs were the most experienced, with all 21 of them having reviewed, but then so had many, if not most of the interviewees in the other countries, the only outlier being France, with only a third of ECRs (5/15) reporting to have reviewed.

7.2. Training

Putting the two tables together, the Chinese, Americans, Spaniards and Malaysians have particularly high levels of experience when it comes to peer review –on both sides of the peer review fence. Is it, though, because ECRs in these coun-

Table 2. Involved in responding to reviewers?

Country	Total	No	Yes
China	23	0	23
France	15	4	11
Malaysia	20	0	20
Poland	22	2	20
Russia	20	7	13
Spain	20	1	19
UK	24	3	21
US	21	1	20
Total	165 (100%)	18 (11%)	147 (89%)

Table 3. Involved in peer reviewing?

Country	Total	No	Yes
China	23	2	21
France	15	10	5
Malaysia	20	1	19
Poland	22	7	15
Russia	20	9	11
Spain	20	1	19
UK	24	7	17
US	21	0	21
Total	165 (100%)	37 (22%)	128 (78%)

tries are better trained for the task? The next question set out therefore to find out whether, in addition to experience, ECRs had any formal and/or informal training for peer review. Only two-fifths had had some kind of training, so most must have learned on the job (Table 4), a finding that comes as no great surprise: for example, in a systematic review of all openly available online training in scholarly peer review, a comprehensive search of the literature identified only 20 openly accessible online training materials (Willis *et al.*, 2022a).

According to our findings, it is the US that stands out as provider of training, with 76% of the interviewees having received training, although in China, UK and Russia there is some training given, too, with around half of the ECRs in each of these countries saying so. However, there does not seem to be a correlation between the training ECRs get and the extent to which they participate in the peer review process. It would seem reasonable to surmise that the more prevalent a task is, the more training would be given to those that do it, which, indeed, is the case in the US, where ECRs are very active in peer review. Still, in Malaysia, for example, with all that, nearly all ECRs take part in the peer review process, none have received any training whatsoever for the task. In the case of Russia, ECRs who are not reviewers do not receive training. And those who are, receive training in one form or another: either official, from the journal, or from senior colleagues or Principal Investigator.

It is certainly not for lack of need or want for training in peer review that there seems to be so little of it given, even where senior researchers are concerned. Rather the contrary, as a survey among biomedical researchers found: 84.2% of the 171 participants (assistant, associate, or full professors) said they had never received formal training in peer review, although most (75.7%) agreed that they should have (Willis *et al.*, 2022b). Indeed, in a roundtable discussion focusing on how publishers and editors can help early career researchers, all agreed that publishers should provide more assistance/training/ support for ECRs with peer review (O'Brien; Graf; McKellar, 2019).

7.3. Peer review as ECRs see it

Moving on to ECRs' perceptions of the peer review, the next five questions concentrated on the extent to which they felt they could trust the system, whether it was in need of improvements, and if so, how.

Trustworthiness of the system

Seeking to establish the extent to which ECRs rely on the system, the interviewees were asked:

- To what extent do they feel that the peer review system vouches for the quality and trustworthiness of formally published research?

Around half of the ECRs who responded to the question (which provided an opportunity to provide free text comments, too) said 'true to a great extent', 43% as 'true to an extent', and 6% as 'not a lot' (Table 5). With 49% of ECRs thus expressing at least some reservation about the system's capability to guarantee the soundness of the scholarly endeavour, there are clearly problems here.

However, there were some very big country differences. Thus, French or Malaysian ECRs were of the opinion that peer review delivered quality/trustworthiness to an extent/to a great extent, so much so, that in fact none of either cohort thought it did not. In all likelihood in both cases the roots of the phenomenon are traceable to widely held cultural attitudes. With mutual trust and respect seen as necessary in order to establish a strong relationship, French academics have a profound respect for their peers (Scroope, 2017). By the same token, so do their Malaysian counterparts, guided as they are by the traditional Malay concept of 'budi', which sees politeness and respect as essential to human interaction (Evason, 2016).

Even bigger supporters of peer review, at least in terms of absolute numbers, were the Chinese ECRs, with 22 out of 23 saying that peer review vouched for quality to a great extent. Indeed, Chinese researchers tend to be greatly appreciative of peer review,

Table 4. Any training for peer review?

Country	Total	No	Yes
China	23	11	12
France	14	13	1
Malaysia	20	20	0
Poland	22	14	8
Russia	20	9	11
Spain	20	16	4
UK	24	12	12
US	21	5	16
Total	164 (100%)	100 (61%)	64 (39%)

Table 5. Peer review - how trustworthy?

Country	Total	To an extent	Great extent	Not a lot
China	23	0	22	1
France	16	3	13	0
Malaysia	18	5	13	0
Poland	21	11	7	3
Russia	20	10	10	0
Spain	20	13	6	1
UK	24	15	7	2
US	17	11	3	3
Total	159	68 (43%)	81 (51%)	10 (6%)

considering it the very basis of scientific communication. Having said that, our Chinese interviewees, needing to navigate foreign waters in their publishing practices (**Jiang; Borg; Borg, 2017**), were in a position to point out that the peer review processes of Chinese journals and English-language journals, published by international publishers, are quite different, with peer review felt to be more trustworthy in English-language journals.

British and Spanish ECRs on the other hand had relatively high levels of reservation or concern in respect to the trustworthiness of the procedure, with only around a third of each cohort saying that peer review could be vouched for to a great extent. US ECRs showed even more concern, with just 3 opting for 'to a great extent' answer when asked about the trustworthiness of peer review. In fact, one of them, a Medical scientist, even raised the possibility of doing altogether without peer review:

"Part of me just thinks that it shouldn't exist and everything should just be pre-prints and the consumer should just think critically about it themselves."

ECRs' main reservations, echoing concerns voiced in *Harbingers-1*, too, (**Jamali et al., 2020; Nicholas et al., 2018b; Rodríguez-Bravo et al., 2017**), were:

- it took too long to get reviews completed during the pandemic;
- reviewers for the task were unsuitable, with some asking questions already answered in the paper and others wanting authors to do more experiments or analyses that were outside the scope of the paper; and
- the quality of reviews was not up to scratch, with some reviewers providing little feedback, others asking questions that suggested that they were not familiar with the field, and, here again, asking for additional work that did not make sense.

Undoubtedly then, ECRs are well-aware of the problems with the system, but, as many ECRs suggested, much along the lines of Churchill's well-known aphorism about democracy –the worst form of government except for all those other forms that have been tried from time to time– even if peer review is not perfect, it is indispensable. Bearing testimony to this are the mentions of peer review in responses to questions not directly related to the topic. Indeed, many interviewees stressed the basic necessity of having and trusting peer review accredited information in a variety of contexts. Thus, for example, a French life scientist expressed no doubts whatsoever as to the centrality of peer review:

"Journals are still the gatekeepers and are still law. They validate ideas exposed in preprint repositories; It's the peer review process that is crucial. There has to be a check."

In the same vein, a US medical scientist explained why, despite their reservations, peer review remains at the heart of research:

"I trust it if it's gone through the peer review process, which is contradictory to what I just said about it. You kind of have to trust the system but I don't have to like the system."

The same notions emerged from ECRs' responses when asked about numerous scholarly practices, which, of course, is yet another proof of the centrality of peer review to research. Thus, for example, a Malaysian hard social scientist, talking about what they call 'pay to publish' (i.e. OA) journals, echoing the concerns encountered in *Harbingers-1* (**Nicholas et al., 2020c**), stressed the importance of peer review for being able to distinguish between acceptable and questionable publications:

"I can say that now I join some group who do research and publish the paper to pay-to-publish journal. But it still goes through peer review process. My suspicion towards the group lessens because I understand these days your papers still need to go through peer review, it's just your chances [to getting published] is higher because of the additional [money]."

By the same token, a British medical scientist, looking at the rise of preprints during the pandemic from the reader's point of view, noted that

"preprints have become more common practice, but still considered to be unreliable sources of evidence because of absence of peer review,"

and a US chemist, speaking mainly as an author, pointed out that with all that preprints have their welcome uses,

"at the end of the day you are evaluated by your publications in peer reviewed journals."

Indeed, discussing whether informal communications (often interpreted as preprints) would have a greater role in scholarly communications, the consensus among interviewees was that it would not, with the barrier to greater use being the absence of peer review. As a French physicist, speaking for others who voiced the same notions, too, explained:

"Only if the peer reviewed [sic] is guaranteed, otherwise it cannot work whatever is the innovation. Journals are deeply integral to research and the way to do research."

Thinking along the same lines, a Chinese soft social scientist, whilst acknowledging the positive role of preprints, which "may allow everyone to share as soon as possible and ensure the author's initial ideas, nevertheless stressed that the dissemination of academic ideas should be screened through specific procedures, such as peer review and formal publishing procedures."

7.4. Need for improvements to the system

Having learned ECRs' views as to the extent to which the peer review is trustworthy, we probed more deeply and asked whether they thought that it needed improving. This, via a question that offered Yes/No/Don't know options, but also allowed for the interviewees to comment freely on the topic:

- Do they feel that the peer review system needs improving in any way?

Nearly two-thirds of the ECRs (66%) opted for 'yes' in response to the question, i.e., they were of the opinion that peer review needed improving (Table 6), a percentage that is a little higher than might have expected from the answers to the previous question. With only 18% of the interviewees saying that the system was good as is, and 16% of them not sure about it, plainly few researchers thought that the system did not require change, strong evidence, perhaps, that it does not always work for them. This was not surprising, having learned from *Harbingers-1* participants of a wide range of concerns about the system (Jamali *et al.*, 2020; Rodríguez-Bravo *et al.*, 2017).

It was ECRs from Spain and the US who felt most strongly that the system was in need of improvement, with almost universal agreement voiced in both cohorts: 19/20 interviewees in the former and 18/19 in the latter. Arguably, though, as there were quite a few assistant professors among them, who were likely to have had more experience with peer review, they might have been simply more knowledgeable about the system and its problematic aspects. Spanish ECRs were especially critical, claiming that there were big and widespread problems with peer review, principally because the best reviewers do not review any longer (*burnt out*), so the quality of reviewing is diminishing. Besides, they said, there are more papers for declining numbers of reviewers, which slows up the whole process at a time when authors want to be reviewed faster than ever. The pace of the process, a long-standing problem (Allen *et al.*, 2022; Christie *et al.*, 2021), which had *Harbingers-1* participants complaining bitterly (Jamali *et al.*, 2019; Rodríguez-Bravo *et al.*, 2017), came even more to the fore, as we are about to see, when the pandemic dictated the quick reviews of COVID-related results and the speeding up of their publication (Horbach, 2021).

The free-text, optional contributions to the question, which came to nearly a hundred, demonstrating how vocal ECRs are in respect to peer review, lent further support to ECRs' awareness of multiple weaknesses in the system. Thus, complaints about the lack of speed in processing and the poor training provided re-emerged here, too, but the biggest problem of them all was thought to be the (in)appropriateness, (poor) quality and bias of reviewers, with more than a fifth of the interviewees saying so. The notion of some sort of compensation for reviewers, be it monetary or reputation-building recognition on the institutional- or the field-level, was brought up again and again by ECRs as a possible way to incentivise peer reviewers to spend as much effort as required on producing quality reports.

Thus, for example, a US Mathematical scientist, musing aloud on the topic, said:

"I think *compensating reviewers* would probably encourage better use of time and a higher quality. You can sometimes take on too many peer reviews and you just don't have time to give it the in-depth review it needs. And compensation may make up for that, but it might not. I think it would have to be tested more."

Another interviewee, a Malaysian Life scientist, was a proponent of reputation-building, rather than financial incentives:

"Peer-reviewing activities are not emphasised in our KPI. If journals publish reviewers' name alongside the published article, reviewers would be more committed to provide quality review. And Universities should support this by including this activity in research assessment. University will also have data on how many of their academics contribute to science through peer-review."

That ECRs should think along these lines certainly came as no surprise, given the proposals in this vein that have been circulating for quite some time now (Bonaccorsi, 2023; García; Rodríguez-Sánchez; Fernández-Valdivia, 2022).

7.5. Possible remedies for peer review ills

Admittedly representing an extremist view of the state of peer review, several ECRs across all countries and disciplines thought the system so bad it was beyond repair, *on the verge of collapse*, as one Spanish environmental scientist put it, indeed, needing *a complete tear-down and rebuild*, as a US medical scientist stated. However, most interviewees held a more positive outlook as to the future of the peer review system, certainly not seeing it as irredeemable, as evidenced by their views on a number of possible solutions or scenarios that were proffered to those among them who thought the

Table 6. Peer review in need of improving?

Country	Total	Don't know	No	Yes
China	22	2	2	18
France	16	7	5	4
Malaysia	18	3	4	11
Poland	22	3	6	13
Russia	20	4	9	7
Spain	20	0	1	19
UK	24	6	1	17
US	19	0	1	18
Total	161	25 (16%)	29 (18%)	107 (66%)

system needed improving or were unsure whether it did or not. Two questions asked about specific solutions and a third asked for their suggestions, with all also allowing for the interviewees to comment freely on the topic:

- Would double blind peer review, defined as author blind to reviewer, and reviewer blind to author, improve things?
- Would the use of Open Reports, where the full content of the reviewer report is published along with the name of the reviewer (Open Identities), improve things?
- Are there any other ways that would improve quality of peer review?

Endorsing the prevalent view in academe, which sees anonymity as the key to fair peer reviewing (Lee *et al.*, 2013; Mulligan *et al.*, 2017; Tomkins; Zhang; Heavlin, 2017), indeed, lending support to the findings of *Harbingers-1* which indicated that ECRs felt very strongly about double-blind reviewing (Jamali *et al.*, 2020; Nicholas *et al.*, 2017b; Rodríguez-Bravo *et al.*, 2017), around two thirds of interviewees said ‘yes’ when asked whether it would improve things, and just a fifth said ‘no’ (Table 7). The greatest support for double-blind came from Poland and the greatest resistance and uncertainty from the UK. The reason for many ECRs’ preference for double-blind reviewing was succinctly put by a British chemist:

“The system is very broken based on gentleman’s club approach...”

The second solution posed was the Open Reports peer review model (Ross-Hellauer, 2017), where the full reviewer report is published along with the name of the reviewers (Open Identities). Just under half of those that answered the question thought the model would be an improvement over the traditional way of reviewing (Table 8), with the greatest endorsement coming from the UK, with 17 ECRs in support, whereas at the other end of the scale just one French ECR opting for the possibility. Not an overwhelming support then of this type of peer review, especially as nearly a quarter of the interviewees chose ‘Don’t know’ when asked whether the model could improve the peer review system, which suggests that even now it is not so well known.

This is all understandable given the risks faced by junior researchers, yet to prove that themselves, as a French ECR put it:

“Open Peer Review is tricky because you engage your own reputation as a reviewer”.

Nevertheless, Open Reports is no longer as unpopular as it was not so long ago –*Harbingers-1* interviewees generally exhibited a suspicious attitude to the model, using terms such as ‘dangerous’, ‘risky’ and ‘unwanted effects’ when referring to it (Rodríguez-Bravo *et al.*, 2017). Indeed, in the *Harbingers-1* survey that followed only around 10% of the participants named Open Report/Open Identities as the peer review model of their choice (Jamali *et al.*, 2019).

ECRs were also asked whether there were any other ways they could think of which could improve peer review, with 75 ECRs making a relevant suggestion. Unsurprisingly, the most frequent suggestion (27), by some margin, focussed on the aforementioned possibility of financial remuneration for reviewers. Relatedly, another 5 mentioned financial incentives other than direct payments, such as certificates or publishing discounts. The second biggest group of comments, made by 12 interviewees, concerned again the quality and appropriateness of reviewers. Stronger editorial control/intervention, raised by 6 ECRs, was another suggestion, as were the need for reputational recognition, raised by 5, the provision of more time for the process, raised by 4, and more detailed reviewing criteria, raised by 3.

Another improvement that was suggested, if only by 3 interviewees, was the right to appeal the reviewers’ ‘verdict’, which is not very surprising, coming as it does from our cohort of millennials, primed by their generational belief that traditional ways of doing things can and should be scrutinised. This is what a Chinese physicist had to say on the topic:

“Authors should be given the right to appeal when a paper is rejected. Because from a historical point of view, many grand theories were not recognized by their peers at the beginning, but they were finally confirmed. Therefore, even if all the comments given by the peer reviewers are rejections, the authors should be given channels and rights to appeal.”

Table 7. Does double blind reviewing improve things?

Country	Total	Don’t know	No	Yes
China	21	1	5	15
France	4	1	2	1
Malaysia	7	0	5	2
Poland	22	0	5	17
Russia	11	4	0	7
Spain	20	2	2	16
UK	24	9	4	11
US	20	3	4	13
Total	129	20 (15%)	27 (21%)	82 (64%)

Table 8. Does the Open Reports model improve things?

Country	Total	Don’t know	No	Yes
China	21	3	5	13
France	10	5	4	1
Malaysia	10	8	1	1
Poland	22	1	11	10
Russia	12	4	4	4
Spain	20	3	8	9
UK	21	3	1	17
US	20	5	6	9
Total	136	31 (23%)	40 (32%)	64 (47%)

Finally, a US life scientist introduced a novel approach to improving the current system –a do-it-yourself one:

“At our [government institution] we have a system where, if a paper hasn’t been peer reviewed, we peer review it. We have two peer reviewers and an editor. The article will be scored, and we can judge if it’s useful or not.”

8. Impact of the pandemic on the peer review

Two inter-related interview questions sought to discover ECRs’ views as to the impact of the pandemic on the peer review, probing for both quantitative and qualitative data:

- Do they think the pandemic is changing the peer review process?
- If so, in what ways and for what reasons?

8.1. Has the pandemic been changing the peer review process?

A third of ECRs thought the pandemic was changing peer review (Table 9), which, although a sizeable proportion, seems to amount to a lesser agreement with the notion than expected at a time when, as a US physicist pointed out, even the general public were much more aware of validation processes because

“everyone was talking about COVID research peer review process even in the general public space.”

Indeed, there was palpable concern at the beginning of the pandemic about the time-consuming nature of the peer review procedure, and the resulting long turnaround times in publishing novel results, as a British ECR, a hard social scientist, put it:

“I think the pandemic has made people think about how quick or slow the peer review process is. Whether some instances require a “quicker” process, e.g., with COVID research.”

Hardly surprisingly, of course: it was, after all, a time when rapid dissemination of scientific knowledge, aimed at battling the virus, was of paramount importance, so much so, that publishers of traditional medical journals, rising up to the challenge, adopted new policies to enable quicker responses, centring upon efforts to fast-track peer review of COVID-19 manuscripts (Horbach, 2020; 2021).

UK researchers most felt that the pandemic was changing peer review, with three-fifths of them thinking so, with, by contrast, no French ECRs thinking this. In pure numbers, perhaps surprisingly, given the alleged big impact that the pandemic had there, China stood out with 15/22 ECRs asserting that the pandemic had made no difference.

Table 9. Pandemic changing peer review?

Country	Total	Don't know	No	Yes
China	22	2	15	5
France	13	4	9	0
Malaysia	19	6	6	7
Poland	21	8	9	4
Russia	19	1	11	7
Spain	19	4	9	6
UK	22	3	6	13
US	20	0	11	9
Total	155	28 (18%)	76 (49%)	51 (33%)

8.2. Peer review during the pandemic

The free-text responses to the specific question on peer review, as well as to the final question of the interview, which asked ECRs to reflect over the last two pandemic years, rendered the picture of the COVID-wrought change to the peer review system more nuanced. With all that most of the interviewees, who thought that there was a change, did mention speed as the big change/impact, many of them said it had actually slowed down the review, as a Russian soft social scientist put it:

“Fewer people available to do the work because of COVID (falling ill) so slowed down.”

A few did argue the opposite, but they spoke of the fast-tracking of COVID-related papers that was taking place. Indeed, coronavirus-related (vs. non-related) articles were considerably more expeditiously processed and published in peer-reviewed journals, largely due to quicker peer review processes (Helliwell et al., 2020; Horbach, 2020; 2021; Jung et al., 2021; Kodvanj et al., 2022). However, according to ECRs, too, on the whole the process became slower, as this US life scientist explained:

“I think it [i.e. the pandemic] has made it slower. As a reviewer, I have accepted fewer reviews and I take longer to do them. As an author, it takes me longer to respond to them so everything is slower.”

Lending further support to the picture emerging from the qualitative data, three quarters of the 800 ECRs participating in the survey felt that the pandemic had an impact on the speed of the peer review process (Graph 1). The majority found the impact to be negative, that is, they were of the opinion that the pandemic had been slowing down the process, with over a quarter even saying that the impact was significantly negative, i.e. slower. Regarding the impact on quality a majority thought there was no impact, but there were more ECRs thinking it had a negative rather than positive impact.

Nevertheless, as Graph 2 indicates, there were considerable country differences: Malaysian ECRs registered the highest impact, with most (60%) saying there had been a positive impact on peer review during the pandemic, with their US

counterparts reporting the lowest level of impact, although the US also had the highest percentage of significant negative impact (46%).

Indeed, some of the interviewees thought that the accelerated speed of the process meant lower quality of reviews and reviewers, echoing a widespread concern characteristic at the time of the scholarly community. Certainly, the possibility that speed may come at the expense of the quality of rushed COVID research has been raised again and again in the literature, along with prognostications of a horror scenario of an infodemic—an epidemic of low-quality information on COVID-19 (De-Araújo-Grisi *et al.*, 2022; Jung *et al.*, 2021). Thus, for example, a British medical scientist, worried about the possibility that the state of the peer review system would only get worse in general, noted the dire effect of the pandemic, too:

“It is going to get bigger and bigger as a problem as more material gets published—it’s linear with quantity of work being done and being published. The pandemic has produced accelerated publishing processes, which translate to ‘light touch’ peer review, also contributing.”

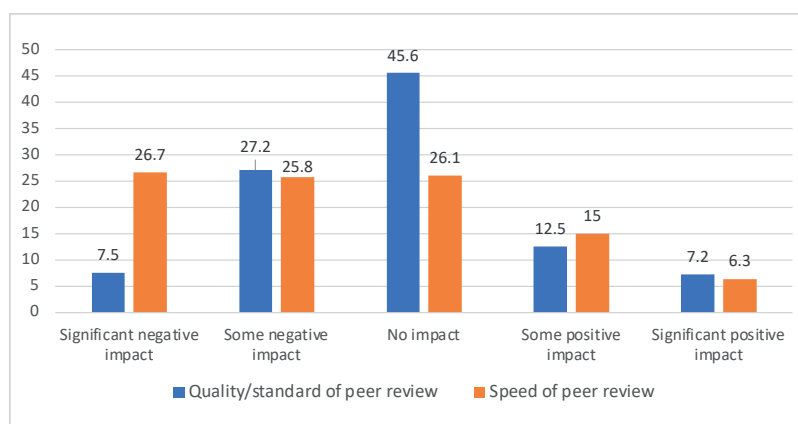
These concerns were not divorced from reality, as the rise in the number of retractions amply proves (Kodvanj *et al.*, 2022), for retractions have found to be traceable back to either unreliable data or to oversight/greater lenience in peer review and editorial verification that have taken place even in journals with an IF and/or high-ranking biomedical journals (Anderson; Nugent; Peterson, 2021; Horbach, 2021; Jung *et al.*, 2021; Teixeira-Da-Silva; Bornemann-Ciment; Tsigaris, 2021; Shimray, 2022). As a British medical scientist’s comments testify, the possibility that during the pandemic quality was sacrificed for the sake of rapid peer review did not escape ECRs’ attention:

“Traditional journals have been ‘shaken’ in a way through the pandemic, they first saw a great increase in submissions with COVID-related research, which decreased the quality of peer review. However, once some of the big scandals occurred with falsified COVID data and the big Lancet retraction, there was new focus on the peer-review process, which I believe it helps refining the system itself.”

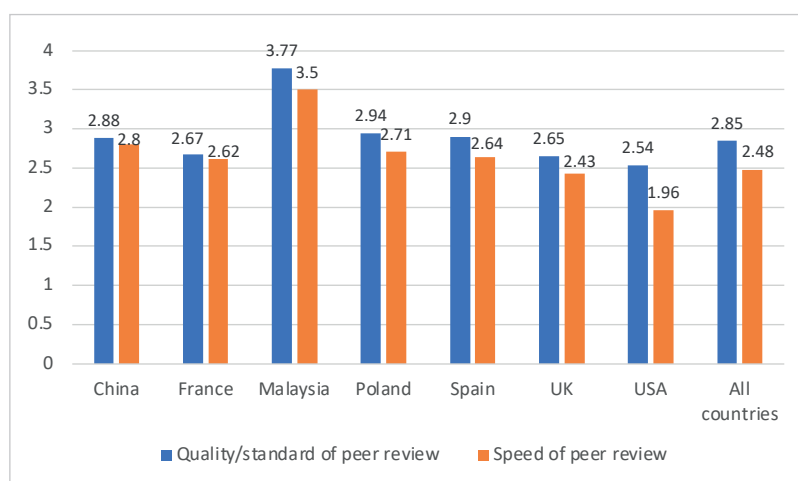
Here again, the qualitative findings are borne out by the survey data: over half of ECRs (54.6%) thought that the pandemic had an impact of some kind on the quality of peer review, with most (34.7%) finding it a negative impact, although only 7.5% said this was significant. Once more, there were significant differences between the case study countries. Again, Malaysians were most likely to say there was an impact (two-thirds though it was positive) and the Americans least likely to say that, although more likely to say it had a negative impact (nearly a half thought this).

8.3. The impact of the pandemic on the future of peer review

Still, with all that the impact of the pandemic on peer review was thus deemed to be rather more negative than positive, virtually all of the comments made by ECRs suggested that it would be temporary. Their prognostications seem to be on the right track: as the pandemic evolved, consistently longer publication delays for COVID-19 manuscripts were found (Fraser *et al.*, 2021; Sevryugina; Dicks, 2021). Thus, as Sevryugina & Dicks (2022) contend, the much shorter submission-to-acceptance times at the beginning of the pandemic might have been simply a manifestation of the early bird effect, associated with any new and ‘hot’ topic. Also, as Homolak, Kodvanj and Virag (2020) argue, the lower quality seen in the case of COVID-19 articles might have been associated with circumstances peculiar to the circumstances at the time: the true experts that could review COVID-19 research must have been far too busy, involved with first-hand fighting with the pandemic, to devote their time to keeping up to date with new developments and/or assessing new contributions.



Graph 1. Level of impact on peer review (%)



Graph 2. Country differences for impact on peer review (mean values)

All this is not to say, though, that the interviewees were not aware of the very real need for accelerated peer reviewing practices; rather the contrary. Thus, for example, a Spanish chemist argued for

“more speed in peer review for the need of generating science faster,”

with another Spaniard, a mathematician, highlighting the role played by the pandemic in bringing the problem to the fore:

“The pandemic has agitated the debate about the speed of science. We are aware that we cannot wait months for the peer review process and that publishing not validated results can be necessary because society asks for them.”

Indeed, a US hard social scientist singled out peer review among the pandemic-induced changes in their practices:

“I think the main one was the change in my peer reviewer approach. That was what happened during the pandemic when I realized just what this kind of bottleneck situation was doing to the publication times for a lot of people where their papers were having these incredibly long publication times.”

Going one step further, another American ECR, a life scientist, was hoping that the pandemic, having drawn attention to the problems with the current publishing system, would bring about change to its peer review component, too:

“The pandemic will be seen as the inflection point of pre-prints. Before, the minority did pre-prints, and soon, the majority will do pre-prints. I hope the pandemic will lead to improvements in peer review, because it has exposed how broken the system is. The problem with pre-prints is we are not yet ready to get rid of peer review. The great thing about pre-prints is [that] they accelerate science. We need something. We’re still waiting for what that is.”

9. Conclusions

Having established that despite their junior status most of our ECRs were experienced in being reviewed and reviewing others, we could be confident that they knew what they were talking about when we asked them about peer review. True, only a minority had received formal training for participating in the peer review process, but with many of them being part of research teams and working with senior researchers, they were mentored and learnt on the job.

A significant proportion of the interviewees –nearly half of them, had a lot or at least some reservations as to the capability of the peer review system to vouch for the trustworthiness of formally published research. However, there were big differences between countries: whilst French or Malaysian ECRs were of the opinion that peer review delivered quality/trustworthiness to an extent and even to a great extent, with none of either cohort thinking it did not, and the Chinese ECRs, bar one, believing that the system was wholly trustworthy, British, Spanish and American researchers were more subdued in their views. Indeed, only around a third of the first two cohorts and only a sixth of the last one said that peer review could be vouched for to a great extent, with the majority in each opting for ‘to an extent’.

In terms of discipline, medical scientists proved to be the greatest believers in peer review, with two-thirds of them saying the system vouched for trustworthiness to a great extent. This is perhaps why the Chinese ECRs were found to be so supportive of the capability of the system to safeguard the trustworthiness of peer review, as the most medical scientists were in their cohort. In answering the trustworthiness question ECRs’ raised three main criticisms of the system: (1) it took too long to obtain reviews; (2) there were too many unsuitable reviewers; (3) partly as a consequence of (2), the quality of reviews was not up to scratch. These criticisms echoed those of pre-pandemic ECRs.

A large majority (two-thirds) of the ECRs felt that peer review needed improving, indeed, in US and Spain nearly all ECRs thought so. In the case of Spain, the short times for reviewing that some journals give, such as those of *MDPI*, may have played a role in the mistrust of the system. ECRs may perceive that reviews are too short and superficial. It was hard social scientists who were less likely to believe the system needed changing. Poor reviewers and slow processes were thought to be the main problems identified. There was a strong feeling that some form of reward (monetary or reputational-building recognition) could help solve these problems. As to the two possible remedies for the ills of the system that the participants were posed –double blind and open reports– the anonymity that double blind afforded was clearly welcomed by ECRs, lending further support to the findings of *Harbingers-1*, with two thirds thinking so. Open reports were thought to be attractive by around a half of all ECRs, with the model, approached with far less apprehension than in *Harbingers-1*, gaining most support in the UK and among environmental scientists. When asked for other remedies, again compensation was mentioned most, adding strength to what we have heard earlier. Financial rewarding of reviewers was also high among recommendations in *Harbingers-1*, indeed, it is a recurrent theme in the literature, too, but it seems the notion is largely falling upon deaf ears.

As to the pandemic impact on peer review, a third of ECRs interviewed felt there was an impact, although no French ECRs thought so, and this is partly explained by the fact that no French ECR was involved in COVID research. The most support of the possibility came from life scientists, with two-thirds of them voicing the opinion, which is perhaps not very surprising, as they were at the forefront of the efforts to combat the virus and witnessed closely how peer review affected research. Speed was thought to be the big impact, a finding that emerged quite clearly from the interviews and from the survey (three quarters of survey respondents thought so), with most saying it has slowed the process down.

However, most ECRs thought that the pandemic impacts would be transitional. But what the pandemic has done without doubt is raised the question of why most peer review is so slow. Of course, ECRs are always in two minds: as reviewers they need more time, but as authors they want more speed.

Comparing pre and post pandemic peer review data is made difficult because the questions were not exactly the same, but there is a sense that ECRs have become more independent and less influenced by their seniors and, also, more critical about the process and the latter is the reason why they are now –especially Spanish ECRs, pressing for some kind of payment, which they believe would lead to better reviews and greater trust in the system.

10. Notes

1. Harbingers-1
<http://ciber-research.com/harbingers.html>
2. Harbingers-2
<http://ciber-research.com/harbingers-2>
3. Alfred P. Sloan Foundation
<https://sloan.org>

11. References

- Allen, Kelly-Ann; Reardon, Jonathan; Lu, Yumin; Smith, David V.; Rainsford, Emily; Walsh, Lucas** (2022). "Towards improving peer review: Crowd-sourced insights from Twitter". *Journal of university teaching & learning practice*, v. 19, n. 3. <https://ro.uow.edu.au/jutlp/vol19/iss3/02>
- Anderson, Caleb; Nugent, Kenneth; Peterson, Christopher** (2021). "Academic journal retractions and the COVID-19 pandemic". *Journal of primary care & community health*, v. 12. <https://doi.org/10.1177/21501327211015592>
- Barroga, Edward** (2020). "Innovative strategies for peer review". *Journal of Korean medical science*, v. 35, n. 20, e138. <https://doi.org/10.3346/jkms.2020.35.e138>
- Bartneck, Christoph** (2017). "Reviewers' scores do not predict impact: Bibliometric analysis of the proceedings of the human-robot interaction conference". *Scientometrics*, n. 110, 179-194. <https://doi.org/10.1007/s11192-016-2176-y>
- Berkenkotter, Carol** (1995). "The power and the perils of peer review". *Rhetoric review*, v. 13, n. 2, pp. 245-248. <https://www.jstor.org/stable/465828>
- Blankstein, Melissa; Wolff-Eisenberg, Christine** (2019). *U.S. faculty survey 2018*. New York, NY: Ithaca S+R. <https://sr.ithaca.org/wp-content/uploads/2019/03/SR-Report-US-Faculty-Survey-2018-04122019.pdf>
- Bonaccorsi, Andrea** (2023). "Towards peer review as a group engagement". *JLIS.it*, v. 14, n. 1, pp. 46-59. <https://doi.org/10.36253/jlis.it-511>
- Brainard, Jeffrey; You, Jia** (2018). "What a massive database of retracted papers reveals about science publishing's 'death penalty'". *Science*. October 25. <https://www.sciencemag.org/news/2018/10/what-massive-database-retracted-papers-reveals-about-science-publishing-s-death-penalty>
- Brezis, Elise S.; Birukou, Aliaksandr** (2020). "Arbitrariness in the peer review process". *Scientometrics*, v. 123, n. 1, pp. 393-411. <https://doi.org/10.1007/s11192-020-03348-1>
- Burstein, David D.** (2013). *Fast future: How the millennial generation is shaping our world*. Boston: Beacon Press. ISBN: 978 0 807033227
- Christian, Katherine; Johnstone, Carolyn; Larkins, Jo-ann; Wright, Wendy; Doran, Michael R.** (2021). "Research culture: A survey of early-career researchers in Australia". *ELife*, n. 10, e60613. <https://doi.org/10.7554/eLife.60613>
- Christie, Alec P.; White, Thomas B.; Martin, Philip A.; Petrovan, Silviu O.; Bladon, Andrew J.; Bowkett, Andrew E.; Littlewood, Nick A.; Mupepele, Anne-Christine; Rocha, Ricardo; Sainsbury, Katherine A.; Smith, Rebecca K.; Taylor, Nigel G.; Sutherland, William J.** (2021). "Reducing publication delay to improve the efficiency and impact of conservation science". *PeerJ*, n. 9, e12245. <https://doi.org/10.7717/peerj.12245>
- De-Araújo-Grisi, Gabriel; De-Deus-Barreto-Segundo, João; Freire, Camila-Verônica-Souza; Matias, Denise-Silva; Cruz, Mariana-Correia-Moreira; Laporte, Larrie-Rabelo; Medina-da-Silva, Daniel-Oliveira; Taniguchi, Thiago-Masashi; Requião, Leticia-Escorse; Goes, Bruno-Teixeira; Correia, Luis-Claudio-Lemos** (2022). *Evidence on the role of journal editors in the COVID19 infodemic - metascientific study analyzing COVID19 publication rates and patterns*, medRxiv, 2022-01. <https://doi.org/10.1101/2022.01.23.22269716>

- Demarest, Bradford; Freeman, Guo; Sugimoto, Cassidy R.** (2014). "The reviewer in the mirror: Examining gendered and ethnicized notions of reciprocity in peer review". *Scientometrics*, v. 101, pp. 717-735.
<https://doi.org/10.1007/s11192-014-1354-z>
- Douglas, Hannah M.; Settles, Isis H.; Cech, Erin A.; Montgomery, Georgina M.; Nadolsky, Lexi R.; Hawkins, Arika K.; Ma, Guizhen; Davis, Tangier M.; Elliott, Kevin C.; Cheruvelil, Kendra-Spence** (2022). "Disproportionate impacts of COVID-19 on marginalized and minoritized early-career academic scientists". *PloS one*, v. 17, n. 9, e0274278.
<https://doi.org/10.1371/journal.pone.0274278>
- Duffy, Bobby; Shrimpton, Hannah; Clemence Michael** (2017). *Millennial myths and realities*. London: Ipsos Mori.
<https://www.ipsos.com/ipsos-mori/en-uk/millennial-myths-and-realities>
- Evason, Nina** (2016). *Malaysian culture: Core concepts*. Cultural Atlas.
<https://culturalatlas.sbs.com.au/malaysian-culture/malaysian-culture-core-concepts>
- Eve, Martin-Paul; Neylon, Cameron; O'Donnell, Daniel-Paul; Moore, Samuel; Gadie, Robert; Odeniyi, Victoria; Parvin, Shahina** (2021). *Reading peer review: PLoS one and institutional change in academia*. Cambridge University Press.
<https://doi.org/10.1017/9781108783521>
- Fanelli, Daniele; Glänzel, Wolfgang** (2013). Bibliometric evidence for a hierarchy of the sciences. *PLoS one*, v. 8, n. 6, e66938.
<https://doi.org/10.1371/journal.pone.0066938>
- FEPS (Foundation for European Progressive Studies); ThinkYoung*, (2018). *The millennial dialogue: Engaging and creating a better understanding of the priorities and values of millennials*.
<https://www.millennialdialogue.com/#download>
- Flaherty, Colleen** (2022). "The peer review crisis". *Inside higher ed.*, June 13.
<https://www.insidehighered.com/news/2022/06/13/peer-review-crisis-creates-problems-journals-and-scholars>
- Fraser, Nicholas; Brierley, Liam; Dey, Gautam; Polka, Jessica K.; Pálffy, Máté; Nanni, Federico; Coates, Jonathon-Alexis** (2021). "The evolving role of preprints in the dissemination of COVID-19 research and their impact on the science communication landscape". *PLoS biology*, v. 19, n. 4, e3000959.
<https://doi.org/10.1371/journal.pbio.3000959>
- García, José A.; Rodríguez-Sánchez, Rosa-María; Fernández-Valdivia, Joaquín** (2022). "The cross-subsidy and buy-one-give-one models of compensated peer review: A comparative study for mission-driven journals". *Journal of information science*, Online first.
<https://doi.org/10.1177/01655515221125321>
- Haffar, Samir; Bazerbachi, Fateh; Murad, M. Hassan** (2019). "Peer review bias: A critical review". *Mayo Clinic proceedings*, v. 94, n. 4, pp. 670-676.
<https://doi.org/10.1016/j.mayocp.2018.09.004>
- Helliwell, Jack A.; Bolton, William S.; Burke, Joshua R.; Tiernan, Jim P.; Jayne, David G.; Chapman, Stephen J.** (2020). "Global academic response to COVID-19: cross-sectional study". *Learned publishing*, v. 33, n. 4, pp. 385-393.
<https://doi.org/10.1002/leap.1317>
- Herman, Eti** (2018). "Scholarly reputation". *FEMS microbiology letters*, v. 365, n. 18, fny200.
<https://doi.org/10.1093/femsle/fny200>
- Herman, Eti; Nicholas, David** (2019). "Scholarly reputation building in the digital age: An activity-specific approach. Review article". *El profesional de la información*, v. 28, n. 1, e280102.
<https://doi.org/10.3145/epi.2019.ene.02>
- Herman, Eti; Nicholas, David; Watkinson, Anthony; Rodríguez-Bravo, Blanca; Abdullah, Abrizah; Boukacem-Zeghmouri, Chérifa; Jamali, Hamid R.; Sims, David; Allard, Suzie; Tenopir, Carol; Xu, Jie; Świgoń, Marzena; Serbina, Galina; Parke-Cannon, Leah** (2021). "The impact of the pandemic on early career researchers: what we already know from the internationally published literature". *Profesional de la información*, v. 30, n. 2, e300208.
<https://doi.org/10.3145/epi.2021.mar.08>
- Hesselmann, Felicitas; Graf, Verena; Schmidt, Marion; Reinhart, Martin** (2017). "The visibility of scientific misconduct: A review of the literature on retracted journal articles". *Current sociology*, v. 65, n. 6, pp. 814-845.
<https://doi.org/10.1177/0011392116663807>
- Homolak, Jan; Kodvanj, Ian; Virag, Davor** (2020). "Preliminary analysis of COVID-19 academic information patterns: A call for open science in the times of closed borders". *Scientometrics*, n. 124, pp. 2687-2701.
<https://doi.org/10.1007/s11192-020-03587-2>

- Horbach, Serge P. J. M.; Halffman, Willem** (2018). "The changing forms and expectations of peer review". *Research integrity and peer review*, v. 3, n. 1.
<https://doi.org/10.1186/s41073-018-0051-5>
- Horbach, Serge P. J. M.; Halffman, Willem** (2019). "The ability of different peer review procedures to flag problematic publications". *Scientometrics*, n. 118, pp. 339-373.
<https://doi.org/10.1007/s11192-018-2969-2>
- Horbach, Serge P. J. M.** (2020). "Pandemic publishing: Medical journals strongly speed up their publication process for COVID-19". *Quantitative science studies*, v. 1, n. 3, pp. 1056-1067.
https://doi.org/10.1162/qss_a_00076
- Horbach, Serge P. J. M.** (2021). "No time for that now! Qualitative changes in manuscript peer review during the COVID-19 pandemic". *Research evaluation*, v. 30, n. 3, pp. 231-239.
<https://doi.org/10.1093/reseval/rvaa037>
- Jamali, Hamid R.; Nicholas, David; Watkinson, Anthony; Abrizah, Abdullah; Rodríguez-Bravo, Blanca; Boukacem-Zeghmouri, Chérifa; Xu, Jie; Polezhaeva, Tatiana; Herman, Eti; Swigon, Marzena** (2020). "Early career researchers and their authorship and peer review beliefs and practices: An international study". *Learned publishing*, v. 33, n. 2, pp. 142-152.
<https://doi.org/10.1002/leap.1286>
- Jamali, Hamid R.; Nicholas, David; Sims, David; Watkinson, Anthony; Herman, Eti; Boukacem-Zeghmouri, Chérifa; Rodríguez-Bravo, Blanca; Świgon, Marzena; Abrizah, Abdullah; Xu, Jie; Tenopir, Carol; Allard, Susie** (2023a). "The pandemic and changes in early career researchers' career prospects, research and publishing practices". *PLoS one*.
<https://doi.org/10.1371/journal.pone.0281058>
- Jiang, Xiaoli; Borg, Erik; Borg, Michaela** (2017). "Challenges and coping strategies for international publication: Perceptions of young scholars in China". *Studies in higher education*, v. 42, n. 3, pp. 428-444.
<https://doi.org/10.1080/03075079.2015.1049144>
- Johnson, Timothy P.; Feeney, Mary K.; Jung, Heyjie; Frandell, Ashlee; Caldarulo, Mattia; Michalegko, Lesley; Islam, Shaika; Welch, Eric W.** (2021). "COVID-19 and the academy: opinions and experiences of university-based scientists in the US". *Humanities and social sciences communications*, v. 8, n. 1.
<https://doi.org/10.1057/s41599-021-00823-9>
- Jubb, Michael** (2016). "Peer review: The current landscape and future trends". *Learned publishing*, v. 29, n. 1, pp. 13-21.
<https://doi.org/10.1002/leap.1008>
- Jung, Richard G.; Di-Santo, Pietro; Clifford, Cole; Prospero-Porta, Graeme; Skanes, Stephanie; Hung, Annie; Parlow Simon; Visintini, Sarah; Ramírez F. Daniel; Simard, Trevor; Hibbert, Benjamin** (2021). "Methodological quality of COVID-19 clinical research". *Nature communications*, v. 12, n. 1, 943.
<https://doi.org/10.1038/s41467-021-21220-5>
- Kodvanj, Ivan; Homolak, Jan; Virag, Davor; Trkulja, Vladimir** (2022). "Publishing of COVID-19 preprints in peer-reviewed journals, preprinting trends, public discussion and quality issues". *Scientometrics*, n. 127, pp. 1339-1352.
<https://doi.org/10.1007/s11192-021-04249-7>
- Lee, Carole J.; Sugimoto, Cassidy R.; Zhang, Guo; Cronin, Blaise** (2013). "Bias in peer review". *Journal of the American Society for Information Science and Technology*, v. 64, n. 1, pp. 2-17.
<https://doi.org/10.1002/asi.22784>
- Levine, Felice J.; Nasir, Na'ilah-Suad; Ríos-Aguilar, Cecilia; Gildersleeve, Ryan-Evely; Rosich, Katherine J.; Bang, Megan; Bell, Nathan E.; Holsapple, Matthew A.** (2021). *Voices from the field: The impact of COVID-19 on early career scholars and doctoral students* [Focus group study report]. American Educational Research Association; Spencer Foundation.
<https://doi.org/10.3102/aera20211>
- Morin, Andréanne; Helling, Britney A.; Krishnan, Seetha; Risner, Laurie E.; Walker, Nykia D.; Schwartz, Nancy B.** (2022). "Research culture: Surveying the experience of postdocs in the United States before and during the COVID-19 pandemic". *Elife*, v. 11, e75705.
<https://doi.org/10.7554/eLife.75705>
- Mulligan, Adrian; Hall, Louise; Raphael, Ellen** (2013). "Peer review in a changing world: An international study measuring the attitudes of researchers". *Journal of the American Society for Information Science and Technology*, v. 64, n. 1, pp. 132-161.
<https://doi.org/10.1002/asi.22798>
- Neylon, Cameron** (2018). *Arenas of productive conflict: Universities, peer review, conflict and knowledge*.
<https://doi.org/10.17613/fkrd-ef72>
<https://hcommons.org/deposits/item/hc:22483>

Nicholas, David; Boukacem-Zeghmouri, Chérifa; Rodríguez-Bravo, Blanca; Herman, Eti; Abrizah, Abdullah; Clark, David; Serbina, Galina; Sims, David; Świgoń, Marzena; Xu, Jie; Watkinson, Anthony; Jamali, Hamid R.; Tenopir, Carol; Allard, Susie (2023). “‘Cracks’ in the scholarly communications system: Insights from a longitudinal international study of early career researchers”. *Learned publishing*, Early view.

<https://doi.org/10.1002/leap.1539>

Nicholas, David; Herman, Eti; Boukacem-Zeghmouri, Chérifa; Watkinson, Anthony; Sims, David; Rodríguez-Bravo, Blanca; Świgoń, Marzena; Abrizah, Abdullah; Xu, Jie; Serbina, Galina; Jamali, Hamid R. (2022a). Early career researchers in the pandemic-fashioned ‘new scholarly normality’: Voices from the research frontline. *Profesional de la información*, v. 31, n. 5.

<https://doi.org/10.3145/epi.2022.sep.10>

Nicholas, David; Herman, Eti; Boukacem-Zeghmouri, Chérifa; Watkinson, Anthony; Sims, David; Rodríguez-Bravo, Blanca; Świgoń, Marzena; Abrizah, Abdullah; Xu, Jie; Serbina, Galina; Jamali, Hamid R.; Tenopir, Carol; Allard, Suzie (2022b). “Early career researchers in the pandemic-fashioned ‘new scholarly normality’: a first look into the big changes and long-lasting impacts (international analysis)”. *Profesional de la información*, v. 31, n. 4, e310418.

<https://doi.org/10.3145/epi.2022.jul.18>

Nicholas, David; Herman, Eti; Jamali, Hamid R.; Abrizah, Abdullah; Boukacem-Zeghmouri, Chérifa; Xu, Jie, Rodríguez-Bravo, Blanca; Watkinson, Anthony; Polezhaeva, Tatiana; Świgoń, Marzena (2020a). “Millennial researchers in a metric-driven scholarly world: An international study”. *Research evaluation*, v. 29, n. 3, pp. 263-274.

<https://doi.org/10.1093/reseval/rvaa004>

Nicholas, David; Herman, Eti; Jamali, Hamid R.; Rodríguez-Bravo, Blanca; Boukacem-Zeghmouri, Chérifa; Dobrowolski, Tom; Pouchot, Stephanie (2015a). “New ways of building, showcasing, and measuring scholarly reputation”. *Learned publishing*, v. 28, n. 3, pp. 169-183.

<https://doi.org/10.1087/20150303>

Nicholas, David; Herman, Eti; Xu, Jie; Boukacem-Zeghmouri, Chérifa; Abrizah, Abdullah; Watkinson, Anthony; Świgoń, Marzena; Rodríguez-Bravo, Blanca (2018a). “Early career researchers’ quest for reputation in the digital age”. *Journal of scholarly publishing*, v. 49, n. 4, pp. 375-396.

<https://doi.org/10.3138/jsp.49.4.01>

Nicholas, David; Jamali, Hamid R.; Herman, Eti; Watkinson, Anthony; Abrizah, Abdullah; Rodríguez-Bravo, Blanca; Boukacem-Zeghmouri, Chérifa; Xu, Jie; Świgoń, Marzena; Polezhaeva, Tatiana (2020b). “A global questionnaire survey of the scholarly communication attitudes and behaviours of early career researchers”. *Learned publishing*, v. 33, n. 3, pp. 198-211.

<https://doi.org/10.1002/leap.1286>

Nicholas, David; Jamali, Hamid R.; Herman, Eti; Xu, Jie; Boukacem-Zeghmouri, Chérifa; Watkinson, Anthony; Rodríguez-Bravo, Blanca; Abrizah, Abdullah; Świgoń, Marzena; Polezhaeva, Tatiana (2020c). How is open access publishing going down with early career researchers? An international, multi-disciplinary study. *Profesional de la información*, v. 29, n. 6.

<https://doi.org/10.3145/epi.2020.nov.14>

Nicholas, David; Rodríguez-Bravo, Blanca; Watkinson, Anthony; Boukacem-Zeghmouri, Chérifa; Herman, Eti; Xu, Jie; Abrizah, Abdullah; Świgoń, Marzena (2017a). “Early career researchers and their publishing and authorship practices”. *Learned publishing*, v. 30, n. 3, pp. 205-217.

<https://doi.org/10.1002/leap.1102>

Nicholas, David; Watkinson, Anthony; Abrizah, Abdullah; Boukacem-Zeghmouri, Chérifa; Xu, Jie; Rodríguez-Bravo, Blanca; Świgoń, Marzena; Herman, Eti (2018b). “What publishers can take away from the latest early career researcher research”. *Learned publishing*, v. 31, n. 3.

<https://doi.org/10.1002/leap.1165>

Nicholas, David; Watkinson, Anthony; Abrizah, Abdullah; Rodríguez-Bravo, Blanca; Boukacem-Zeghmouri, Chérifa; Xu, Jie; Świgoń, Marzena; Herman, Eti (2020d). “Does the scholarly communication system satisfy the beliefs and aspirations of new researchers? Summarizing the Harbingers research”. *Learned publishing*, v. 33, n. 2, pp. 132-141.

<https://doi.org/10.1002/leap.1284>

Nicholas, David; Watkinson, Anthony; Boukacem-Zeghmouri, Chérifa; Rodríguez-Bravo, Blanca; Xu, Jie; Abrizah, Abdullah; Świgoń, Marzena; Clark, David; Herman, Eti (2019). “So, are early career researchers the harbingers of change?”. *Learned publishing*, v. 32, n. 3, pp. 237-247.

<https://doi.org/10.1002/leap.1232>

- Nicholas, David; Watkinson, Anthony; Boukacem-Zeghmouri, Chérifa; Rodríguez-Bravo, Blanca; Xu, Jie; Abrizah, Abdullah; Świgoń, Marzena; Herman, Eti** (2017b). "Early career researchers: Scholarly behaviour and the prospect of change". *Learned publishing*, v. 30, n. 2, pp. 157-166.
<https://doi.org/10.1002/leap.1098>
- Nicholas, David; Watkinson, Anthony; Jamali, Hamid R.; Herman, Eti; Tenopir, Carol; Volentine, Rachel; Allard, Suzie; Levine, Kenneth** (2015b). "Peer review: Still king in the digital age". *Learned publishing*, v. 28, n. 1, pp. 15-21.
<https://doi.org/10.1087/20150104>
- Niles, Meredith T.; Schimanski, Lesley A.; McKiernan, Erin C.; Alperin, Juan-Pablo** (2020). "Why we publish where we do: Faculty publishing values and their relationship to review, promotion and tenure expectations". *Plos one*, v. 15, n. 3.
<https://doi.org/10.1371/journal.pone.0228914>
- O'Brien, Anna; Graf, Chris; McKellar, Kate** (2019). "How publishers and editors can help early career researchers: Recommendations from a roundtable discussion". *Learned publishing*, v. 32, n. 4, pp. 383-393.
<https://doi.org/10.1002/leap.1249>
- Pew Research Center (2010a). *Millennials: A portrait of generation next*.
<https://www.pewsocialtrends.org/2010/02/24/millennials-confident-connected-open-to-change>
- Pew Research Center (2010b). *Millennials will make online sharing in networks a lifelong habit*.
<https://www.pewinternet.org/2010/07/09/millennials-will-make-online-sharing-in-networks-a-lifelong-habit>
- Publishing Research Consortium (PRC) (2016). *Peer review survey 2015*. London, UK: Mark Ware Consulting.
https://www.elsevier.com/_data/assets/pdf_file/0007/655756/PRC-peer-review-survey-report-Final-2016-05-19.pdf
- Rodríguez-Bravo, Blanca; Nicholas, David; Herman, Eti; Boukacem-Zeghmouri, Chérifa; Watkinson, Anthony; Xu, Jie; Abrizah, Abdullah; Świgoń, Marzena** (2017). "Peer review: The experience and views of early career researchers". *Learned Publishing*, v. 30, n. 4, pp. 269-277.
<https://doi.org/10.1002/leap.1111>
- Ross-Hellauer, Tony** (2017). "What is Open Peer Review? A systematic review". *F1000Research*, v. 6, 588.
<https://doi.org/10.12688/f1000research.11369.2>
- Roumbanis, Lambros** (2022). "Disagreement and agonistic chance in peer review". *Science, technology and human values*, v. 47, n. 6, pp. 1302-1333.
<https://doi.org/10.1177/01622439211026>
- Scroope, Chara** (2017). "French culture: Business culture". *Cultural atlas*.
<https://culturalatlas.sbs.com.au/french-culture/french-culture-business-culture>
- Seeber, Marco** (2022). "Efficacy, efficiency, and models of journal peer review: The known and unknown in the social sciences". In T.C.E. Engels; E. Kulczycki (eds.). *Handbook on research assessment in the social sciences*, pp. 67-82. Cheltenham, UK: Edward Elgar Publishing.
<https://10.4337/9781800372559.00011>
- Sevryugina, Yulia V.; Dicks, Andrew J.** (2021). *Publication practices during the COVID-19 pandemic: Biomedical preprints and peer-reviewed literature*. Preprint. BioRxiv, 427563.
<https://doi.org/10.1101/2021.01.21.427563>
- Sevryugina, Yulia V.; Dicks, Andrew J.** (2022). "Publication practices during the COVID-19 pandemic: Expedited publishing or simply an early bird effect?". *Learned publishing*, v. 35, n. 4, pp. 563-573.
<https://doi.org/10.1002/leap.1483>
- Sharma, Kiran** (2021). "Team size and retracted citations reveal the patterns of retractions from 1981 to 2020". *Scientometrics*, n. 126, pp. 8363-8374.
<https://doi.org/10.1007/s11192-021-04125-4>
- Schewe, Charles D.; Debevec, Kathleen; Madden, Thomas J.; Diamond, William D.; Parment, Anders; Murphy, Andrew** (2013). "'If you've seen one, you've seen them all!'" Are young Millennials the same worldwide?". *Journal of international consumer marketing*, v. 25, n. 1, pp. 3-15.
<https://doi.org/10.1080/08961530.2013.751791>
- Shimray, Somipam R.** (2022). "Research done wrong: A comprehensive investigation of retracted publications in COVID-19". *Accountability in research*.
<https://doi.org/10.1080/08989621.2021.2014327>
- Silbiger, Nyssa J.; Stubler, Amber D.** (2019). "Unprofessional peer reviews disproportionately harm underrepresented groups in STEM". *PeerJ*, n. 7, e8247.
<https://doi.org/10.7717/peerj.8247>

- Siler, Kyle; Lee, Kirby; Bero, Lisa** (2015). "Measuring the effectiveness of scientific gatekeeping". *Proceedings of the National Academy of Sciences*, v. 112, n. 2, pp. 360-365.
<https://doi.org/10.1073/pnas.1418218112>
- Sørensen, Lene-Tolstrup; Williams, Idongesit; Khajuria, Samant; Skouby, Knud-Erik** (2017). "Millennial users in a 5G context". In: *Wireless World Research Forum outlook series*, n. 18.
<https://wwrf.ch/wp-content/publications/outlook/Outlook18.pdf>
- Squazzoni, Flaminio; Gandelli, Claudio** (2013). "Opening the black-box of peer review: An agent-based model of scientist behaviour". *Journal of artificial societies and social simulation*, v. 16, n. 2, 3.
<https://doi.org/10.18564/jasss.2128>
- Steen, R. Grant; Casadevall, Arturo; Fang, Ferric C.** (2013). "Why has the number of scientific retractions increased?". *PLoS one*, v. 8, n. 7, e68397.
<https://doi.org/10.1371/journal.pone.0068397>
- Tennant, Jonathan; Dugan, Jonathan M.; Graziotin, Daniel; Jacques, Damien C.; Waldner, François; Mietchen, Daniel; Elkhatib, Yehia; Collister, Lauren B.; Pikas, Christina K.; Crick, Tom; Masuzzo, Paola; Caravaggi, Anthony; Berg, Devin R.; Niemeyer, Kyle E.; Ross-Hellauer, Tony; Mannheimer, Sara; Rigling, Lilian; Katz, Daniel S.; Greshake-Tzovaras, Bastian; Pacheco-Mendoza, Josmel; Fatima, Nazeefa; Poblet, Marta; Isaakidis, Marios; Irawan, Dasapta-Erwin; Renaut, Sébastien; Madan, Christopher R.; Matthias, Lisa; Kjaer, Jesper-Norgaard; O'Donnell, Daniel-Paul; Neylon, Cameron; Kearns, Sarah; Selvaraju, Manojkumar; Colomb, Julien** (2017). "A multi-disciplinary perspective on emergent and future innovations in peer review". *F1000 Research*, n. 6, 1151.
<https://doi.org/10.12688/f1000research.12037.3>
- Tennant, Jonathan P.; Ross-Hellauer, Tony** (2020). "The limitations to our understanding of peer review". *Research integrity and peer review*, v. 5, n. 1, 6.
<https://doi.org/10.1186/s41073-020-00092-1>
- Teixeira-da-Silva, Jaime A.; Bornemann-Cimenti, Helmar; Tsigaris, Panagiotis** (2021). "Optimizing peer review to minimize the risk of retracting COVID-19-related literature". *Medicine, health care and philosophy*, v. 24, pp. 21-26.
<https://doi.org/10.1007/s11019-020-09990-z>
- Teixeira-da-Silva, Jaime A.** (2021). "Challenges that early career researchers face in academic research and publishing: pre-and post-COVID-19 perspectives". *Exchanges: The interdisciplinary research journal*, v. 9, n. 1, pp. 77-106.
<https://doi.org/10.31273/eirj.v9i1.882>
- Tomkins, Andrew; Zhang, Min; Heavlin, William D.** (2017). "Reviewer bias in single-versus double-blind peer review". *Proceedings of the National Academy of Sciences of the United States of America*, v. 114, n. 48, pp. 12708-12713.
<https://doi.org/10.1073/pnas.1707323114>
- Van-Dalen, Hendrik P.; Henkens, Kène** (2012). "Intended and unintended consequences of a publish-or-perish culture: A worldwide survey". *Journal of the American Society for Information Science and Technology*, v. 63, m. 7, pp. 1282-1293.
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1983205
<https://doi.org/10.1002/asi.22636>
- Willis, Jessie V.; Cobey, Kelly D.; Ramos, Janina; Chow, Ryan; Ng, Jeremy Y.; Alayche, Mohsen; Moher, David** (2022a). *Online training in manuscript peer review: A systematic review*. medRxiv, 2022-09.
<https://doi.org/10.1101/2022.09.02.22279345>
- Willis, Jessie V.; Ramos, Janina; Cobey, Kelly D.; Ng, Jeremy Y.; Khan, Hassan; Albert, Marc A.; Alayche, Mohsen; Moher, David** (2022b). *Knowledge and motivations of training in peer review: An international cross-sectional survey*, medRxiv, 2022-09.
<https://doi.org/10.1101/2022.09.03.22279564>
- Woods, Hellen-Buckley; Brumberg, Johanna; Kaltenbrunner, Wolfgang; Pinfield, Stephen; Waltman, Ludo** (2023). "An overview of innovations in the external peer review of journal manuscripts". *Wellcome Open Research*, v. 7, n. 82 [version 2; peer review: 2 approved, 1 not approved].
<https://doi.org/10.12688/wellcomeopenres.17715.2>
- Ziman, John M.** (1968). *Public knowledge: An essay concerning the social dimension of science*. London, England: Cambridge University Press. ISBN: 978 0 521095198

The emerging phenomenon of L2 vlogging on *Bilibili*: characteristics, engagement, and informal language learning

Leticia-Tian Zhang; Boris Vázquez-Calvo; Daniel Cassany

Nota: Este artículo se puede leer en español en:
<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87299>

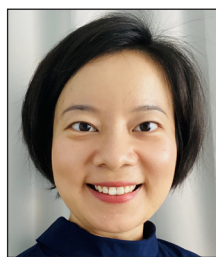
Recommended citation:

Zhang, Leticia-Tian; Vázquez-Calvo, Boris; Cassany, Daniel (2023). "The emerging phenomenon of L2 vlogging on *Bilibili*: characteristics, engagement, and informal language learning". *Profesional de la información*, v. 32, n. 3, e320301.

<https://doi.org/10.3145/epi.2023.may.01>

Manuscript received on 2nd March 2023

Accepted on 31st March 2023



Leticia-Tian Zhang ✉
<https://orcid.org/0000-0003-3905-5933>

Beijing Foreign Studies University
Department of Hispanic and Portuguese Studies
Xisanhuan 2
10089 Beijing, China
zhangtian.leticia@bfsu.edu.cn



Boris Vázquez-Calvo
<https://orcid.org/0000-0001-8574-7848>

Universidad de Málaga
Departamento de Didáctica de las Lenguas, las Artes y el Deporte
Avda. Cervantes, 2
29071 Málaga, Spain
bvazquezcalvo@uma.es



Daniel Cassany
<https://orcid.org/0000-0003-3494-5531>

Universitat Pompeu Fabra
Departament de Traducció i Ciències del Llenguatge
Roc Boronat, 138
08018 Barcelona, Spain
daniel.cassany@upf.edu

Abstract

The rise of digital technology has provided new opportunities for language learning, extending beyond traditional classroom instruction. Video projects have emerged as an effective tool in foreign language education, yet research on self-initiated and regulated video production for language learning is scarce. This study investigates the phenomenon of vlogging in Spanish as a second language on the Chinese video sharing platform *Bilibili*, by analyzing 134 Chinese-produced Spanish-language vlogs. The study aims to understand the vlogs' characteristics, the vloggers' profile, and the ways they utilize the genre for learning Spanish. Through qualitative virtual ethnography, the study uncovers the presence and learning engagement of Spanish L2 vlogs on *Bilibili*. The results reveal a diverse range of vlogs, including daily life experiences and adaptations of popular *YouTube* trends, primarily produced by university students with advanced editing skills. Vloggers incorporate knowledge from both formal education (e.g., the Spanish textbook widely used in China, *Español Moderno*) and informal contexts. In addition to practicing oral Spanish, L2 vloggers use various forms of writing, including Spanish subtitles and Chinese translation, and mobilize multimodal resources, such as *danmu* comments for overlaying corrections. Vloggers also adopt discursive strategies for community interaction, such as self-deprecating metalanguage, feedback solicitation, and metalinguistic reflections. The study highlights the potential of video-sharing platforms like *Bilibili* as tools for language learning, reveals different learning styles in digital environments (self-supervised and interaction-oriented learning), and indicates the direction of integrating daily vlogs and multilingual subtitles into language curricula, emphasizing students' agency, self-directed digital learning, and transmedia literacy development.

Keywords

Spanish as a foreign language; Digital literacy; Video-sharing platform; Multimodal composition; Vlogs; Videovlogs; Participatory culture; Transmedia literacy; *Danmu*; Discourse analysis; Social media; Peer learning; *Bilibili*.



Funding

This research was supported by:

- *Fundamental Research Funds for the Central Universities* (2020QD036, China).
- *Segue: Social Media and Video Games in Language Learning and Teaching of Student Teachers* (B1-2021_33, University of Malaga, Spain).
- *Definers: Digital language learning of preservice and junior language teachers* (TED2021-129984A-I00, Ministry of Science and Innovation, Spain).
- *ForVid: Video as a language learning format in and outside the classroom* (RT2018-100790-B-100, Ministry of Science and Innovation, Spain).

1. Introduction

Vlogs, also known as video blogs, have become an immensely popular visual genre online. With the rise of platforms like *YouTube* and the widespread availability of filming devices such as smartphones and portable cameras, producing and sharing videos has become easier than ever before. As an interactive and engaging format of content creation, vlogging has gained rapid popularity over the past decade (Burgess; Green, 2018). Typically, a vlog focuses on an individual's personal life and features the vlogger speaking directly to the camera or capturing daily activities by holding the device in their hand. Vlogs offer celebrities an effective and intimate channel to connect with their audience, while regular users also use vlogs as a platform for self-presentation and identity expression (Griffith; Papacharissi, 2010; Vizcaíno-Verdú; De-Casas-Moreno; Jaramillo-Dent, 2022).

Furthermore, vlogs are becoming increasingly multilingual and multicultural. Youtubers often self-translate their vlogs to reach a global audience (S. Lee, 2021). L2 learners also intentionally produce vlogs in their second language (L2) to practice their linguistic skills, engage in intercultural interactions, and construct and negotiate alternative L2 identities (Amgott, 2022; Chang; Chang, 2019; Codreanu; Combe, 2018; 2019). Despite the growing popularity of vlogs on social media platforms, little is known about out-of-school L2 vlogging practices or vlogs on platforms other than *YouTube*.

This study seeks to fill this gap by examining vlogs on *Bilibili*, a prominent Chinese video-sharing platform and community renowned for its anime, comic, and game culture. *Bilibili* is particularly notable for its unique video-superimposed commenting feature, which facilitates the practice of *danmu*. Since 2018, *Bilibili* has organized vlog-focused campaigns to promote vlog culture and reward the best vlogs (Wang, 2022). As the vlog genre has gained popularity on Chinese social media, a unique type of production has emerged: L2 vlogging, that is, vlogging using a second language such as English, Spanish, German, or Russian, as the vehicular language. Although *Bilibili* contains some multilingual content, including movies, series, songs, remixes, and fan-made videos, it is still a predominantly Chinese-language platform without an English version, making vlogging in a foreign language an unexpected and intriguing activity within the community.

L2 vlogs, frequently produced by language learners, serve as distinctive examples of language learning in the digital wilds (Sauro; Zourou, 2019) and embody a digital literacy practice (Lankshear; Knobel, 2015). These vlogs involve a variety of skills, such as editing, translation, and the utilization of semiotic resources. Thus, the study of L2 vlogs provides a valuable opportunity to investigate instances of out-of-classroom online learning and interaction, as well as study how L2 learners construct and express their identities and (meta)linguistic practices in digitally mediated spaces. With a focus on Spanish, an increasingly popular L2 among Chinese learners, this research aims to describe the emerging phenomenon of L2 vlogging on *Bilibili* and to identify the potential learning opportunities provided by this practice.

2. Literature review**2.1. Vlogs, identity expression and affective labor**

Vlogging has been extensively researched within the fields of communication and media studies. The pioneering work of Griffith and Papacharissi (2010) uses Goffman's notion of "stage" to theorize vlogs as a means of self-presentation, highlighting its functions as diaries, vehicles for identity expression, and avenues for (quasi-)narcissistic behavior. Subsequent studies have further investigated the potential of vlogs in constructing multifarious transformative identities. For instance, García-Rapp (2016) attributed the commercial success of *YouTube* beauty gurus to two video genres:

- tutorials that strengthen the guru's professional profile, and
- vlogs that foster community and establish affective ties.

Raun (2015) demonstrated how *YouTube* transgender vloggers use the video-sharing platform to motivate others and assert their 'trans' identity. Lovelock (2017) deconstructed the coming-out vlogs of *YouTube* celebrity vloggers, which results in the reinforcement of normative heterosexual identity aligned with the values of neoliberalism.

Another significant area of research investigates vlogging as a form of affective labor, which is commodified and capitalized by interest-driven platforms. For instance, microcelebrities appear to heavily rely on establishing connections with their audience through transparent, open, and authentic communication (Abidin, 2015; Duffy, 2017). This contrasts with mainstream celebrities, who are expected to maintain a public, professional identity while safeguarding their privacy on social media (Marwick, 2013).

Microcelebrities engage in various forms of affective labor, such as accessibility, availability, presence, authenticity, connectedness, and intimacy, effectively commodifying themselves and generating benefits for the platform rather than for themselves (Raun, 2018). As a result, individual creators are influenced by platform affordance, which refers to how digital platforms shape and impact user behavior and interactions. Creators also face challenges due to the instability, intensification, and inequality that emerge within the platform-dependent creative economy (Duffy *et al.*, 2021). For example, although social media workers play a vital role in maintaining a brand's visibility, their contributions are often overlooked and undervalued (Duffy; Sawey, 2022). However, while platforms have become central nodes in virtually every cultural industry, their power is based on mutual dependency and the opportunities provided for negotiation, contestation, and even acts of resistance (Duffy; Poell; Nieborg, 2019).

Vlogging in a second language (L2) has gained popularity on *Bilibili*, a leading Chinese video-sharing platform

While the focus has so far been on *YouTube* and Western social media, there is growing interest to explore vlogging practices on *Bilibili*. In terms of vloggers' identity construction, X. Wang and Picone (2021) revealed that Chinese *toubu* (top) vloggers develop sophisticated strategies to attract traffic and attention. Their identities counter the standardized and institutionalized *wanghong* industry, which is based on perfect self-presentation and blurred commodification. Instead, Chinese *toubu* vloggers embrace niche original personas and openly enact self-commodification. Wang and Picone shift the focus to regular Chinese school vloggers, who are conscious of their precarious position on *Bilibili* due to the platform's algorithm and unclear rules. Still, young school vloggers

“have autonomy in accessible content selection, flexible schedule arrangement, and independent interaction with the audiences” (Wang; Picone, 2022, p. 14).

Beyond the tensions between platforms and content creators, recent studies have begun to highlight the increasing presence of foreign and transnational vloggers on *Bilibili*. Xu and Zhao (2021) provide insights into Chinese international students who exercise digital citizenship to cope with the double blind test during the COVID-19 pandemic. Their vlogging reflects a strong sense of civic engagement, connectivity, and empowerment, contributing to the self-resilience of the affected group, as well as their identity and solidarity building during the crisis. On a related note, Ma (2022), Pérez-González (2022) and Shi, Chang and Gao (2022) illustrate how transnational Western vloggers on *Bilibili* carefully navigate through vernacular scrutinization to maintain their positive images and gain legitimate membership as microcelebrities in the Chinese digitalized third space. Despite the progress of current research on vlogging and *Bilibili*, the emergence of L2 vlogs produced by Chinese speakers remains largely overlooked.

2.2. Video projects and language learning

The advantages and effectiveness of using video recording projects in foreign language teaching have been widely recognized. One useful method is the integration of digital storytelling as an instructional tool in EFL classrooms, which blends traditional storytelling with new technologies such as images, videos, and music. Digital stories can include personal or narrative stories, informative or instructional stories, and stories that retell historical events (Robin, 2008). Students who participated in a digital storytelling project showed significant improvement in English proficiency, particularly in listening, reading, and writing skills, critical thinking, and learning motivation (Yang; Wu, 2012). Collaborative video projects also seem to provide an autonomous environment for digital learners to explore, reflect, and engage in peer teaching and teamwork (Hafner; Miller, 2011). With the increasing use of smartphones, Gromik (2012) implemented a cell phone-based video project with EFL Japanese learners. They would produce a 30-second monologue weekly, with had positive effects on vocabulary knowledge and speaking skills, and was received with positive student feedback, indicating a promising future for mobile-assisted language learning.

The rise of *YouTube* culture has sparked interest in the potential of vlogs for L2 learning. Huang (2021) used a 3-minute smartphone-based collaborative vlog project with college-level EFL learners in China, improving speaking and digital media production skills. Amgott (2021) collected French L2 students' multimodal compositions (vlogs, blogs) from study abroad courses, where they would demonstrate increasing metalinguistic awareness and evolving identities as “multilingual multimodal composers” and self-assumed teachers of imagined audiences.

Regarding learner and teacher perceptions of vlogs, Hung's (2011) study with university ESP students in Taiwan indicated benefits (visual representation, less time constraints, the possibility of self-monitoring, and finding future opportunities) and challenges (technical difficulties, relation to real-time communication) of using vlogs. Similar mixed feedback was reported in Aldukhayel (2021b). Moreover, both students and teachers in this study appreciated the wide accessibility, interesting topics, and appropriate length of vlog resources on *YouTube*. In this line, pictorial content, subtitles, and language proficiency play a key role in L2 vlog comprehension (Aldukhayel, 2021a).

In practice, the consumption of L2 videos remains a predominant practice in class, while a limited number of teachers assign video production tasks (Cassany; Shafirova, 2021). In contrast, in out-of-school settings, students and language learners may have ample opportunities and engagement with viewing and producing L2 videos.

2.3. Language learning, digital wilds and transmedia literacy

The metaphor *digital wilds* refers to the coexistence of two distinct features in language learning in informal learning and socialization contexts:

- the absence of direct curriculum or teacher intervention, and
- the emergence of learning intertwined with digital literacy and digital technologies (Reinhardt, 2022).

A related concept is digital literacy practice or new literacies, which encompasses the essence of vlogging as

“ways in which meaning-making practices are evolving under contemporary conditions that include, but are not limited to, technological changes associated with the rise and proliferation of digital electronics” (Knobel; Lankshear, 2014, p. 97).

Scholars have explored language learning situations in the digital wilds that arise on fandom and fan translation (Sauro, 2017; 2019; Vázquez-Calvo *et al.*, 2019) and online gaming (Chik, 2014). These digital wilds provide ample opportunities for learners to engage in language use, production, and comprehension beyond traditional classroom settings, allowing them to develop their language proficiency in immersive and authentic contexts.

Another crucial concept to comprehend young people’s digital practices and skills is *transmedia literacy*, which encompasses competencies related to the production, sharing, and consumption of digital interactive media. These activities span from problem-solving processes in video games to content production and sharing within web platforms and social networks (Scolari *et al.*, 2018; Guerrero-Pico; Establés; Costa-Sánchez, 2022). This literacy encompasses not only technical proficiencies, but also critical thinking, creativity, and collaboration skills that allow users to effectively interact with diverse media content and participate in online communities.

Proactive fans acquire and utilize fan capital –comprising skills, knowledge, social influence, and dedication to professionalism– accumulated through fan play (Pun, 2021). As a result, they transition from the amateur domain of participatory cultures to the professional realm of cultural industries, particularly in competitive and professional gaming, digital writing, and social networks (Establés; Guerrero-Pico; Contreras-Espinosa, 2018). However, these new environments are still evolving at the legislative and labor levels, and a lack of understanding about economic and power relations can lead to frustration for fans aspiring to become professionals. A case in point is the *Viki* community, a global network of volunteers who collaborate on crowd-sourced translations of television shows. This platform presents an opportunity to expand and normalize practices that professionalize leisure time on a broader scale. Furthermore, this model is reinforced by community members who utilize their experiences to shape work expectations for a global community (Henthorn, 2019).

While video-sharing sites appear to bear great potential for informal language learning, there are relatively few studies investigating this area. Prior studies yield mixed results. For instance, Benson’s (2015) research into *YouTube* comments shows they can be a source of interactional meaning and evidence of language and intercultural learning. However, H. Wang and Chen (2020) claim that social engagement on *YouTube* is limited, because most users participate as ‘lurkers’, ‘likers’ or ‘sharers’: although EFL learners watch English-teaching *YouTube* videos for a variety of learning-related purposes, such as finding resources, increasing attraction, and learning about cultures, few take further actions beyond viewing, liking or sharing the video. *YouTube* content creators are rather exceptional cases. Zhang and Vázquez-Calvo (2022) have demonstrated how, with the translation, singing and video-editing of Korean and Japanese songs into Catalan and Spanish, some youtubers find in vlogging and video production an actual professional pathway driven by their affinities and likes. Other examples include vlogging practices of a French L2 learner, which increased digital literacy skills and multilingual and intercultural awareness, but also highlighted challenges such as content validation and global interpersonal tensions (Codreanu; Combe, 2018). While not directly related to vlogging, studies with younger participants and teenagers reveal that *YouTube* and video production and sharing practices are oftentimes seen as an extension of leisure activity within specific periods of teenage years, where leaving traces of what one does socially is important for group acceptance (Vázquez-Calvo; Elf; Gewerc, 2020). *TikTok* is gaining attention as a platform for mobile-assisted language learning, offering opportunities for multimodal communication and immediate participation (Lee, 2022). Hashtags such as #learnchinese, #learnitalian, and #learnrussian demonstrate the platform’s potential as an active medium for sharing content related to linguistic features, cultural practices, and learner experiences. Language learners can engage in informal learning, intercultural identity negotiation, and emotional support through this platform (Vázquez-Calvo; Shafirova; Zhang, 2022).

Despite the dynamic and novel activities of communities and individuals, the literacy practices of digital learners on Chinese video sharing platforms and social media have received little attention. For example, on *Bilibili*, Spanish learners engage in *danmu* viewing and commenting on the acclaimed Spanish TV series *El Ministerio del Tiempo* (Zhang; Cassany, 2019a). Additionally, advanced learners participate in self-organized fansub groups to translate and subtitle Spanish series and films, utilizing a variety of online Spanish-Chinese resources (Zhang; Cassany,

“L2 vlogging enable learners to practice linguistic skills, engage in intercultural interactions, and construct alternative identities”

2019b). Moreover, language and culture enthusiasts host video channels on *RED* (similar to *Instagram*) and *Bilibili*, sharing language learning content and claiming a new identity as “language influencers” (Han; Reinhardt, 2022). While a case study could provide in-depth observations of learners’ practices, online content analysis on a larger scale is needed to gain insights into emerging phenomena, such as the unique L2 Spanish vlogging on Chinese video sharing platforms and social media.

“ This qualitative study based on virtual ethnography investigates the characteristics and learning engagement of Spanish L2 vlogs on *Bilibili* ”

3. Methods

This exploratory study aims to address the research gap in digital literacy practices of Spanish learners on Chinese video sharing platforms and social media by answering the following research questions:

RQ1. What are the different types of Spanish-language vlogs on *Bilibili*, who produces them, and what are their purposes?

RQ2. How do *Bilibili* users engage with the vlog genre for Spanish language learning?

The present study employs a qualitative research design based on virtual ethnography (Hine, 2015) to investigate Spanish L2 vlogging on *Bilibili*. Virtual ethnography is a research approach that acknowledges cyberspace as a distinct social space and focuses on the social and cultural aspects of online communities. This approach allows for exploration of the ways in which *Bilibili* vloggers influence and are influenced by broader sociocultural forces, such as globalization, interculturality, and language learning trends. Non-participant observation was chosen as the method of data collection because it allows for the analysis of naturally occurring behavior without disrupting the participants or the environment. By collecting and analyzing digital traces and virtual interactions, such as user-generated vlogs and related content, we were able to observe the online behavior of *Bilibili* vloggers and their viewers in a non-intrusive manner and to gain insights into the ways in which language learning and cultural exchange occur in this vernacular out-of-school virtual space.

3.1. Context, observation, and data collection

This study centers on *Bilibili*, a leading Chinese video-sharing platform that is particularly favored by Generation Z individuals (born between the mid-to-late 1990s and early 2010s). *Bilibili* is known for its creative fan community, which is passionate about anime, comics, and games, and is distinguished by a unique real-time commenting system called *danmu* or *danmaku*, which creates a “bullet curtain” of superimposed comments on the video screen. The video creators on *Bilibili* are known as uploaders, up’s, or up-lords [UP主]. As of 2022, *Bilibili* has an average monthly active user base of 332.6 million:

<https://www.statista.com/statistics/1109108/bilibili-average-monthly-active-users>

Over the course of one year, from March 2021 to March 2022, the researcher conducted an ethnographic study of L2 vlogging on *Bilibili*. To collect data, the researcher logged in to the platform at least three times per week to: (1) check platform recommendations for Spanish L2 vlogs, (2) watch recommended videos, and (3) save the videos (57 videos from 34 vloggers) with initial content descriptions. The ethnographer has been a frequent *Bilibili* user for the past decade, with has ample research experience with its functions and community practices. Once the researcher gained a general understanding of Spanish L2 vlogging on *Bilibili*, a sample of vlogs was collected for further analysis.

To sample appropriate videos for this investigation, purposeful sampling was employed, using specific keywords such as ‘vlog’, ‘all Spanish [全西语]’, and ‘Spa/Spanish [西语/西班牙语]’. Videos that were imported from *YouTube* or produced by native Spanish speakers and those that showcased only minimal Spanish competence were excluded. Additionally, only one video per vlogger was included to avoid an unequal representation of cases (Snelson, 2015). The sampling resulted in 134 vlogs from 134 vloggers, which constituted the dataset for analysis.

3.2. Data analysis

Content analysis was adopted to examine the presence and characteristics of L2 Spanish vlogging on *Bilibili*. To conduct this analysis, the field researcher watched each vlog at least twice and took note of several important pieces of information. This included identifying the specific theme and genre of the vlog, as well as noting information such as editing and shooting techniques, the decision to vlog in Spanish, and numerical data such as length, number of likes, comments, and *danmu* (i.e., user-generated comments that scroll across the video screen). Additionally, the date of publication was recorded, allowing for analysis of trends over time. To further contextualize the vlogs, the researcher reviewed the uploader’s homepage to gather information about their profession and location, if available. This helped to provide insight into the background and interests of the vloggers and their motivations for creating content in Spanish.

To protect the privacy of the vloggers, any identifiable information was removed from the figures of the vlogs before they were analyzed. This was done in accordance with the ethical standards of the *International Association of Internet Researchers* (Markham; Buchanan, 2012).

To investigate how learners make use of the vlog genre for L2 Spanish learning on *Bilibili*, qualitative discourse analysis was conducted. Through a third viewing session, the field researcher marked learning-related genres, instances or dis-

courses in the vlog such as accessible daily vlogs and ‘silent’ study vlogs, transcribed the textual or verbal introduction and justification of vlogging in L2 Spanish. The transcription provided insights into vloggers’ discursive engagement with the *Bilibili* community to maximize the learning purpose. It also highlighted lexical and morphosyntactic errors in Spanish, which were not replicated in the English translation provided.

Spanish vlogs on *Bilibili* are multifarious and mainly created by university students in China and Spain, with increasing popularity

In this study, 134 vlogs were carefully examined, resulting in a collection of notes and observations in *Word* and *Excel* documents, comprising 1,258 Chinese characters. The co-authors then engaged in a thorough discussion to identify the most innovative and compelling findings. Several adjustments were made during this process.

Initially, genre categories encompassed themes such as daily routines, food, and travel. However, they were later reorganized based on the characteristics of vlog genres, specifically the “with me” categories. Information that was pertinent to learning and provided insights into students’ digital literacy competence was also incorporated. For example, data on post-editing was included, as numerous videos featured Chinese translations or Spanish transcriptions, as well as multimodal resources like emojis, memes, music, and other annotations (credits, titles). These elements indicate the sophistication of the videos.

Key moments within the vlogs were scrutinized multiple times and analyzed to ensure a consensus on interpreting the vloggers’ learning activities. In doing so, we honed in on the example of *VideoEle*, a significant case that illustrates the relationship between the resources utilized by vloggers and the content of the initial lessons found in official textbooks.

Ultimately, a comprehensive category of analysis emerged, encompassing: 1) audio elements (target audience, presentation, communicative context, presupposed knowledge, explicit references such as deictics, mentions, and place names, and post-editing); 2) visual components (function –whether accompanying or natural setting or intentional action– and gestures); and 3) other inserted written text or resources. Upon reaching a consensus, the study findings were derived, elucidating the nature and scope of Spanish L2 vlogs on *Bilibili* and the ways learners employ the vlog genre for L2 Spanish learning on the platform.

4. Findings

4.1. L2 vlog production and user profile

The production of Spanish-language vlogs on *Bilibili* is diverse, as shown in Table 1. The most common type of video was the ‘with me’ subgenre, where the vlogger invites the audience to ‘participate’ their personal life. Vloggers engaged in various activities while speaking Spanish, such as ‘a day with me (58)’, ‘chat with me (20)’, ‘travel with me (23)’, ‘cook with me (13)’, ‘walk with me (2)’, ‘plan with me (1)’, ‘study with me (1)’, and ‘get ready with me (1)’. Vloggers primarily used two types of shooting techniques: (1) showing-around, where the creator held the camera and explained the surroundings or activities to viewers, such as in ‘a day with me’; and (2) monologue, where the vlogger sat in front of the camera and discussed a specific topic, such as in ‘chat with me’. Additionally, some vloggers adapted popular *YouTube* trends, such as ‘challenge’ (2) and ‘haul’ (1), while others explored less conventional genres like ‘role play’ and ‘short film’, which featured scripted storytelling with artistic dubbing and editing. These findings suggest that *Bilibili* and its users have embraced a broad definition of ‘vlog’, covering a diverse range of topics and subgenres where the vlogger often assumes the dual role of videographer and narrator.

Table 1. Vlog genres

Genres	No. of vlogs	
‘With me’	A day with me	58
	Travel with me	23
	Update/chat with me	20
	Food/cook with me	13
	Study with me	3
	Walk with me	2
	Plan with me	1
	Get ready with me	1
Others	Couple/family	3
	Challenge	2
	Self-introduction	2
	Reading	2
	Short film	2
	Haul	1
	Role play	1
Total	134	

The vlogs, despite being amateur-produced, displayed remarkable sophistication in terms of editing (Table 2). Subtitles were a common feature in most videos, serving as a dual-purpose tool for learning: (1) incorporating written Spanish transcription into oral practice, and (2) using Chinese or English as a support and translation tool for L2 Spanish learning. This blend of advanced multimodal resources with traditional academic practices highlights the ingenuity of the vloggers. As the comments suggested, the deliberate use of subtitles also catered to a specific audience, including regular users and Spanish learners, and was valued by the audience as a learning aid.

Table 2. Post-production

	Degree of post-production	No. of vlogs	%
Subtitles	Chinese translation	90	67
	Spanish transcription	98	73
	English translation	2	1
Inserted multimedia	Emojis, memes, clips/graphics	67	50
	Credits, intros, outros	88	66
	Background music	122	91

The incorporation of multimedia elements was also prevalent in the vlogs, as evident from the use of background music, emojis, and graphics (Figure 1), as well as film clips, scrolling lyrics, and images of famous Spanish quotes. While credits were not commonly seen, many vloggers included video intros and outros to introduce the title or content, thank viewers for watching, and ask for likes, comments, and shares. Some videos also featured outtakes or bloopers, showcasing instances where vloggers stumbled or made comical errors while speaking Spanish.



Figure 1. Inserted emojis and graphics

Regarding the creators, as presented in Table 3, the typical vlogger’s profile is a university student located in China or a Spanish-speaking country, mostly Spain. This aligns with the platform’s targeted user group. Based on the content of the vlogs, many creators were undergraduate students majoring in Spanish, as Spanish Philology is offered in over one hundred universities in China. Others were exchange students or pursuing a master’s or doctoral degree abroad, and some were professionals working in sectors where proficiency in Spanish is required, such as education, interpreting, or tourism. From viewers’ comments, it became evident that three overseas vloggers were second-generation Spanish nationals of Chinese descent.

Table 3. Vloggers’ profile

Location	Employment	Affiliation	No. of vloggers
China	Student	University	69
		High school	2
	Teacher	University / language school	3
	Other professions		2
	Unemployed		1
Overseas	Student	University	25
	Teacher	University / Confucius Institute / others	4
	Other professions		2
Unclear/Undisclosed			26
Total			134

In general, there were three main reasons why L2 vlogging on *Bilibili* occurred (as shown in Table 4): (1) personal interest in practicing Spanish and sharing their lives; (2) creating content for a course assignment (which was then later uploaded to *Bilibili* by the vlogger); and (3) participating in a special activity such as *The First Spanish Vlog Contest*, organized by *Eshelper*, a widely-used Chinese-Spanish online dictionary service. In addition to the types of vlogs mentioned above, there was one vlog produced in collaboration with a local press (bearing the editorial’s watermark).

Table 4. Vlog types

Types	No. of vlogs
Practice and share	96
Course assignment	35
Submission to contest	2
Professional	1
Total	134

Taking a holistic perspective, the number of L2 Spanish vlogs on *Bilibili* has significantly increased from a total of 20 in 2018-2019 to 94 in 2020-2021, likely incentivized by pandemic restrictions, and the trend appears to continue in 2022 with 20 vlogs already uploaded at the time of this research. This demonstrates the growing interest in this genre and the popularization of vlogging culture on *Bilibili*. In the next chapter, we will explore selected vlog subgenres and their potential benefits for learning Spanish as an L2.

4.2. 'A day with me'

Daily vlogs emerged as the most popular genre of Spanish-language vlogs on *Bilibili* due to their linguistic accessibility, opportunities for self-presentation, and genre versatility. They are particularly preferred by learners of Spanish as L2 as they pose relatively low linguistic barriers. Many vloggers described their videos as their *primera experiencia* [first experience], 初尝试 [first try] or 初投稿 [first submission] in Spanish L2 vlogging. The language used in daily vlogs is usually simple, consisting of basic lexicon and syntactic constructions with fundamental grammar. Vloggers commonly use first person singular or plural conjugation or unconjugated verbs in the present or near future tense to describe their daily routines. Describing daily life is a topic covered in the third lesson of *Español Moderno*, the most widely used Spanish textbook in China for university students. This lesson includes reflexive verbs (*cepillarse*, 'to brush one's teeth'), time adverbs (*por la mañana*, 'in the morning'), and the conjugation of the irregular verb *ir*. These linguistic features enable basic communication about daily routines.

Daily vlogs are centered around real-life events, ranging from a vlogger's routine activities to notable moments that occur over a short period of time. For university students (the most common profile), a typical daily vlog might involve getting up in the dormitories (as most Chinese university students live on campus), walking to the canteen, attending classes, and studying in the library. As a result, creating daily vlogs in Spanish creates both a sense of familiarity and novelty: campus activities are relatable to viewers with a similar background on *Bilibili*, while the use of Spanish is original and intriguing for a general audience.

As university-level Spanish learners, vloggers drew inspiration from their class materials and textbooks. V133 creatively imitated the format of *VideoEle*, a popular series of videos for teaching Spanish as a foreign language. Although the production quality of *VideoEle* was lacking (with camera shakes and unstable scene transitions), the videos were available online and used by some Spanish teachers in China. Figure 2a/b shows the similar typography, style, and layout between the title of the vlog ([A day in...]) and *VideoEle* ([This is Barcelona]). The background music used was also taken from *VideoEle*. The vlogger's speech was interactive and conversational (*¿Cómo lo pasaré?* [How will I spend it (today)?]), similar to the register used in *VideoEle*. After an eventful day, the vlogger concluded by quoting a sentence from *Español Moderno* (Figure 2b/c): *¿Estamos contentos con nuestra vida estudiantil? Aunque terminamos la jornada muy cansados, yo digo que sí.* [Are we happy with our student life? Although we finished the day very tired, I say yes.]

Finally, daily vlogs offer a flexible and versatile genre. Depending on the context, such as confinement, festivals, or holidays, daily vlogs can be transformed into various sub-genres, including 'quarantine with me', 'travel with me', or 'cook with me', which require more developed or specialized vocabulary, oral skills, and intercultural awareness. For example, 'quarantine with me' vlogs document daily activities at home during lockdown, detailing coping strategies for confined living, and encourage viewers to stay safe and optimistic. University vloggers have explained in Spanish the extra procedures to enter the campus during the pandemic (such as taking temperature, disinfection, wearing face masks), or showcased a "new normal" landscape (such as Covid-related posters on the campus or making an appointment to enter the library).

The 'travel with me' vlogs recorded the vlogger's experience as a tourist, both within China and in Spanish-speaking countries in South America. The 'food/cook with me' vlogs demonstrated how to cook dishes from China or Spanish-speaking countries, or showed dinners out, providing opportunities for intercultural comparisons. For example, when introducing *callos* (a traditional Spanish stew made with beef tripe), the vlogger referred to *lu zhu* (a traditional Beijing stew made with pork entrails). These intercultural references allow vloggers to act as intercultural brokers, helping the audience to understand the differences and similarities between Spanish and Chinese sociocultural realities (V70).

4.3. 'Chat with me'

'Update/chat with me' vlogs are characterized by interactivity, as vloggers shared recent events or discussed a certain topic with the audience. Compared to other genres, 'chat with me' vlogs are more spontaneous, often involve uninterrupted speaking, and require a certain level of language proficiency. Vloggers in this category adopted two roles:

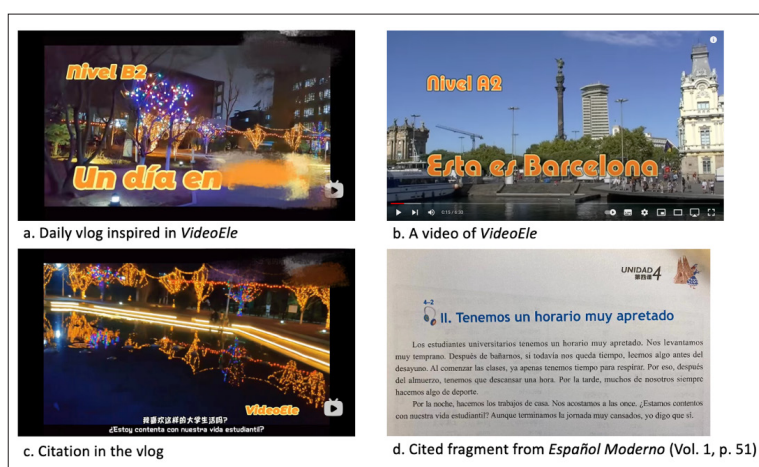


Figure 2. Daily vlogs

“The most popular genres of Spanish vlogs on *Bilibili* are “a day with me,” “chat with me,” and “travel with me.”

- a struggling yet persevering learner who celebrated Spanish learning feats and created an affective link with viewers by sharing emotional moments, such as the eventual pronunciation of the rolled /r/ after two years of practice (V92; Figure 3a), and by sharing anxious moments and mental breakdowns (V110, 128).
- an expert or teacher, who shared and reflected upon their language learning experiences, such as motives to learn Spanish and recounted on study abroad periods (V27, 35; Figure 3b), recommended tools and resources (series, books) for improving Spanish (V36, 110), or work prospects analyses for Spanish graduates (V61).

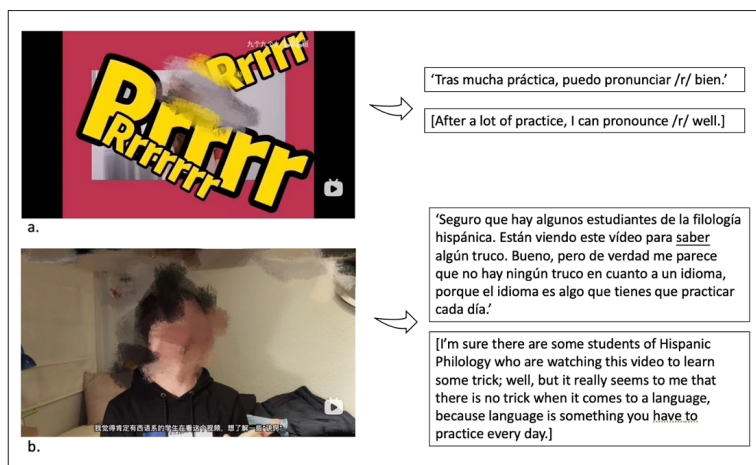


Figure 3. Chat vlogs

4.4. 'Study with me', 'reading' & 'challenge'

Learning-themed genres like 'study with me' and genres to practice and showcase spoken Spanish, such as 'reading' and 'challenge', were also present but to a lesser extent. Unlike daily vlogs or other 'with me' genres, learning-themed genres are less interactive and mainly serve the purpose of self-motivation and self-supervision. In 'study with me' (*estudia conmigo*) vlogs, the learner sits for hours studying Spanish (which is sped up in editing), doing exercises in the textbook *Español Moderno*, taking notes, consulting the Chinese-Spanish dictionary *Eshelper*, or practicing spoken Spanish (including moments of stammering and repetition). Stimulating background music is usually added to the recording. Although there is no interaction with the audience, vloggers can add explanations in the video description. In Figure 4, the vlogger first states the type of vlog (in English), then accentuates the urgency of the task (in Spanish) and provides background information (in Chinese).

Reading aloud is a highly effective and suggestopedic method² for practicing a second language. Vlog creators shared poems read in different languages, including Spanish (by poets such as Pablo Neruda), with synchronous bilingual subtitles set against an edited background. Reading aloud poetry is considered a form of 'close reading' where readers construct interlinguistic and intercultural meaning through careful consideration (Hanauer, 2001). In the 'reading' genre of vlogs, the vlogger's presence was central, and the setting was more curated, with elements such as candles, wool clothing, and ambient music (Figure 5a), creating a space that served both for learning and performance.

The 'challenge' genre provided a fun way to practice Spanish. A popular challenge on *Bilibili* involves creators imitating Chinese and L2 flight announcements, which can be a difficult task due to the speed and pronunciation required. V94 (Figure 5b) is one of two Spanish versions of this type of challenge found in the dataset. However, a recent search for "Spanish flight announcement challenge" revealed over 30 similar challenges available on *Bilibili*. In her video, the creator juxtaposed her image with an edited slide displaying the announcements in Spanish and Chinese. She delivered the speech while reproducing the tone, prosody, and intonation of a flight attendant. By doing so, the learner showcased her language competence and affirmed her identity as a successful learner.

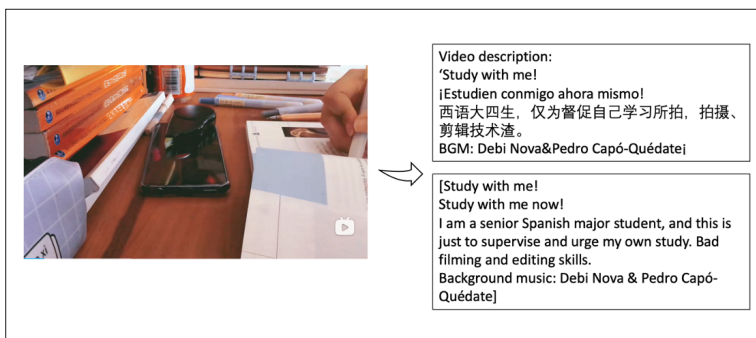


Figure 4. 'Study with me'



Figure 5. 'Reading' and 'challenge'

4.5. Discursive strategies for L2 learning

The analysis revealed various discursive strategies employed by vloggers for L2 learning, including self-deprecating metalinguage, feedback solicitation, and self-improvement and self-correction discourses. Metalinguistic discourses, which involve social media users' self-reflection and comments on their own language use, can be essential to online participation (Lee, 2013). Initially, vlogging in Spanish was perceived as a challenging task:

“Era mucho más difícil que lo que pensaba. No podía dejar de reírme al hablar y, mientras tanto, los peatones me miraron con extrañeza. Me parece muy embarazoso hablarme ante la cámara, sobre todo hablando español.”

[it was much more difficult than I thought. I couldn't stop laughing when I spoke and meanwhile, the pedestrians looked at me strangely. I find it very embarrassing to speak to myself on camera, especially speaking Spanish.] (V103)

In addition, a common discursive strategy observed among vloggers was self-deprecation when it comes to their Spanish proficiency. Many expressed insecurity and expected to make mistakes, often apologizing in advance, or making light of their language skills. Some even explicitly asked their audience for tolerance and forgiveness for any mispronunciations, grammar errors, or limited vocabulary, emphasizing that they were still beginners or 'newbies' in learning Spanish (V71). However, some vloggers also solicited feedback from viewers who are more proficient in Spanish, inviting them to engage in discussions or provide corrective feedback, with the goal of improving their Spanish skills together (V101):

“Estoy segura de que habrá muchos errores cuando hablo español, porque no soy una hispanohablante. Entonces, si también sabes hablar español y si encuentras alguno errores, puedes decírmelo y puedes enviarme mensajes y darme consejos. Y en el futuro voy a hacerlo mejor poco a poco.”

[I am sure there will be many mistakes when I speak Spanish, because I am not a native Spanish speaker. So, if you also know how to speak Spanish and you find any mistakes, you can tell me and you can send me messages and give me advice. In the future I will do better little by little.] (V29)

During quarantine, vloggers frequently made learning-related resolutions. With many students in China and overseas forced to stay confined, opportunities for communication with classmates or locals became scarce. Fearing a decline in their oral language skills, vloggers turned to L2 vlogging in Spanish as a valid way to maintain and eventually improve their proficiency:

“Si no hablar y practicar español durante muchos meses, olvidaré muchos palabras y mi español aprendo empeorará, y creo que hablar español enfrente de una cámara es una buena forma para practicar mi español.”

[If I don't speak and practice Spanish for many months, I will forget many words and my Spanish will get worse, and I think that speaking Spanish in front of a camera is a good way to practice my Spanish.] (V33)

In V36, the vlogger was studying abroad in Mexico and was concerned about the health risks of face-to-face oral interactions with neighbors. By suggesting “practice with me”, he took control and demonstrated the agency and self-efficacy of an L2 digital learner.

La mejor manera para que no se me olvide el español es practicar conmigo, con mí mismo.

[The best way for me to not forget Spanish is to practice with me, with myself.] (V36)

After the video is published, creators have two options to make changes: (1) traditional forum comments, and (2) *danmu* comments. Forum comments are typically found in the comment section, where creators can post self-corrections after publishing the video. The comment can be pinned to the top, making it noticeable for future viewers. For instance, in V111, a food vlog where the vlogger tried street food in Beijing, including *tanghulu*, a northern Chinese snack of candied hawthorn berries, the vlog has subtitles in both Chinese and Spanish. However, the vlogger erroneously translated *espino cocido* [cooked hawthorn berries] into 生山楂 [raw hawthorn berries] in Chinese, and left a comment with the realization of her mistake (Figure 6).



Figure 6. Self-correction through pinned forum comments

In *danmu* comments, which are superimposed moving texts, creators can mark their comment as uploaded, indicated by a blue symbol before the text, to reveal authorship and gain visibility. *Danmu* comments can be used to repair any potential content or linguistic inaccuracies in the video. Figure 7 shows an example where the vlogger used *danmu* to self-repair

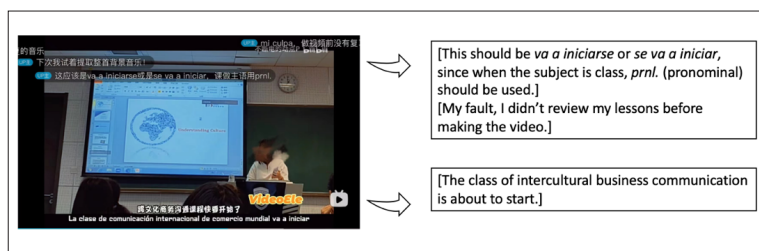


Figure 7. Self-correction through *danmu*

and apologize for a mistake in the original monologue. However, the uploader did not realize that *va a iniciar* is also acceptable in this syntactic construction if the teacher is considered the subject, even though it may not be the most common choice. Despite this, the self-repair *danmu* comment demonstrates learners' metalinguistic awareness through L2 vlogging. However, it's essential to maintain a critical view regarding the quality of such reflections in vernacular productions.

“ L2 vloggers mobilize discursive and multimodal resources for language learning. *Danmu* comments are overlaid on the video to provide corrections ”

5. Discussion

Regarding the types of vlogs, vloggers, and their purposes (RQ1), the analysis revealed a variety of 'with me' genres and adaptations of popular *YouTube* trends. These were primarily created by university students with sophisticated editing skills, demonstrating their competence and aspiration to use Spanish in diverse communicative situations, with different linguistic registers and interactive styles. However, as shown in the examples, vloggers are still in the process of learning and make mistakes in various linguistic domains. The assimilation of *YouTube* vlog genres and trends (intros and outros) suggests that for some users, L2 vlogging is an opportunity to attract attention and create an online persona, such as a language influencer. The motivations for L2 vlogging range from personal and entertainment purposes (recording life, learning L2, participating in vlog campaigns) to formal and even commercial contexts (sharing course assignments, collaborating with the press). While there may be other unidentified reasons, these findings offer insights into the intersection of L2 vlogging on *Bilibili*, where the boundaries between formal and informal learning are blurred, and various contexts (entertainment, learning, working) become interconnected or 'collapsed' on a social media platform (Marwick; Boyd, 2011).

Regarding how users appropriate the vlog genre (RQ2), the analysis identified both content and discursive strategies employed by L2 learners for language learning. In terms of vlog content, learners draw on L2 knowledge and skills gained from both formal and informal contexts. Formal education influences daily and study vlogs, where vloggers discuss topics associated with campus life, record themselves while doing assignments, employ the linguistic repertoire acquired during the first few months of learning, and even the aesthetic and discourse style of teaching materials. Genres such as poems and flight announcements are also closely related to L2 teaching. On the other hand, travel or cooking vlogs are based on the vloggers' personal experiences and pastimes. As *Bilibili* uploaders, vloggers are skilled at leveraging platform affordances such as *danmu* for self-correction and selecting background music for their vlogs.

In terms of discursive strategies, we observed an open space for community-based peer learning, supported by self-deprecating metalanguage, self-improvement discourse, solicitation of feedback, and critical metalinguistic reflection. Vloggers also use writing to facilitate learning more challenging spoken skills. With Spanish subtitles, Chinese translation, and other forms of writing (forum and *danmu* comments), vloggers' learning practices share both traditional (academic) and leisure-like (multimodal) features.

This study specifically focused on informal learning strategies that were identified by *Scolari et al.* (2018). These strategies include learning by doing, which involves the learner engaging in a series of activities (vlogging, interaction with other viewers) that are related to the skill they wish to acquire or improve (oral or written Spanish). Another strategy is imitating or simulating, which requires the learner to self-manage their multimodal resources and time, as well as their own emotions and identity (as a newbie learner). Lastly, the evaluating strategy involves the learner analyzing their own work or the work of others in order to acquire or perfect a skill (asking for feedback).

The vlogs examined in this study demonstrate the complexity of learning styles in digital environments, including interaction-oriented (feedback-seeking) and self-monitored learning. Daily vlogs showcase vloggers sharing their experiences as Spanish learners, connecting with peers, introducing parts of Spanish culture, and expanding their vocabulary and spoken skills. Chat vlogs allow vloggers to leverage intersubjectivity to perform various L2 roles, such as a struggling yet persevering learner or an expert and successful learner teaching an imagined novice audience. While community interaction is a common pursuit, not all vloggers are interested in it. Some 'silent' learners record their practices (e.g., reading, studying) and upload the videos to *Bilibili*, using them as a form of self-supervised learning in the digital age. Although study vloggers may seem disinterested or undisturbed, there is a large audience seeking an effective learning ambiance at a low cost by watching 'study with me' videos (Lee et al., 2021).

6. Conclusions

L2 vlogging can be viewed as a digital literacy practice that involves multiple levels of sophistication. Firstly, L2 vloggers engage in self-translating discourses of various genres and culture-specific terms. Secondly, they add subtitles and use visual texts such as emojis, memes, sound effects, and music. Thirdly, they integrate different multimodal resources into the final audiovisual performance, including the reappropriation of teaching materials such as *VideoEle*. Lastly, they appropriate the platform affordances, such as *danmu* and forum comments, to post-edit the video and interact with the viewers. Learner-vloggers exhibit a combination of agency, autonomy, and self-efficacy, which can lead to cumulative cycles of learning and achievement and create a meaningful learning experience in an immersive environment (Sán-

chez-López; Roig-Vila; Pérez-Rodríguez, 2022). Similar to literacy practices in the digital wilds (fanfic, fansub, and fandub) and short video platforms, vlog making enables language learners to perform various L2 identities and create L2 communities of practices around vlogging.

“Learners draw on knowledge and skills acquired from both formal and informal contexts in their vlogs”

For formal education, L2 vlogging on *Bilibili* can provide an authentic and engaging context for language learning, develop digital literacy skills, offer authentic assessment opportunities, expose students to diverse language use, and promote cultural competence. In particular, integration of daily vlogs, multilingual subtitles, and discursive competence into language curricula and fostering students' agency and self-directed learning styles are some of the directions that this study indicates. Emerging platforms such as *TikTok* facilitate literacy practices like *BookTok*, a subculture within *TikTok* dedicated to reading and discussing young adult literature (Jerasa; Boffone, 2021). By recognizing and appreciating these *TikTok*-mediated literacy practices, educators can bridge the gap between formal and informal reading practices, ultimately promoting increased agency, community, and transformative experiences for students (Espejel; Concheiro; Pujolà, 2022).

This study sheds light on the intricacies of teaching 'minority languages' in China. Although Spanish L2 teaching has experienced exponential growth in the country in the last decade (Yu, 2021), previous research has tended to depict Chinese Spanish learners as uncreative, obedient, and dependent (Castaño Arques, 2019, pp. 50-51), and Spanish teaching methods in China are mostly traditional, focusing on textbooks and written language with limited teacher-student interactions that follow the initiate-response-evaluate model (Li, 2018). However, the L2 vloggers on *Bilibili* showed a great interest in practicing their spoken skills, while writing seemed to take a back seat, with about a third of the vlogs lacking Spanish transcriptions. Educators could consider incorporating L2 vlogging into their teaching to inspire innovation, drawing on experiences from established practices like fanfiction (Sauro; Sundmark, 2018). "A day with me" vlogs could be a starting point for beginners to practice underrepresented oral dialogue, while "chat with me" vlogs could help advanced learners improve their written interactive chatting skills.

This study has limitations due to its exploratory nature. The data analyzed was primarily collected online, and individual vloggers were not interviewed or quantitatively measured for learning outcomes. However, by systematically collecting vloggers' productions, this study provides a broader perspective on L2 vlogging on *Bilibili*, which could give insights into similar practices on global platforms such as *YouTube*. The analysis of learners' comments and self-disclosed information in their profiles highlights the potential of L2 vlogging for language learning. Although not all vlog genres were analyzed, the study provides insight into the most prominent genres and their distinguishing characteristics. Feedback from the audience was not analyzed, but the study identified various discursive strategies used by vloggers to interact with viewers, suggesting the formation of a Spanish-themed learning community on *Bilibili*.

In recent years, the development of L2 Spanish vlogging on *Bilibili* has become increasingly noticeable, and with more active uploaders, the number of L2 vlogs is expected to grow. Future studies could conduct longitudinal research on vloggers or expand the statistical analysis to gain a better understanding of language learners' experiences as vloggers. Another avenue for analysis is to consider vlogs as L2 input and examine their potential benefits for vocabulary learning (Arndt; Woore, 2018). Additionally, analyzing comments on foreign language vlogs on *Bilibili* could be a fruitful area for future research, given the increasing importance of online comments for informal learning. In conclusion, this study offers valuable insights into the emerging phenomenon of L2 vlogging on non-English social media and suggests that further research is needed to fully understand the potential of digital media for language learning.

7. Notes

1. A Chinese term that refers to internet celebrities or influencers who gained their fame and popularity through social media platforms, particularly in China. They often create and share their own content, including vlogs, on various social media platforms such as *Douyin*, *Weibo*, and *Xiaohongshu* (RED).
2. A teaching method that uses relaxation techniques and music to create a comfortable learning environment where students can absorb information more effectively. See Lozanov and Gateva (1988).

8. References

- Abidin, Cystal (2015). "Communicative intimacies: Influencers and perceived interconnectedness". *Ada: A journal of gender, new media, and technology*, n. 8. <https://doi.org/10.7264/N3MW2FFG>
- Aldukhayel, Dukhayel (2021a). "The effects of captions on L2 learners' comprehension of vlogs". *Language learning and technology*, v. 25 n. 2, pp. 178-191. <http://hdl.handle.net/10125/73439>
- Aldukhayel, Dukhayel (2021b). "Vlogs in L2 listening: EFL learners' and teachers' perceptions". *Computer assisted language learning*, v. 34, n. 8, pp. 1085-1104. <https://doi.org/10.1080/09588221.2019.1658608>

- Amgott, Natalie** (2022). “Je suis youtubeur’: Multilingual multimodal composing abroad”. *Journal of language, identity & education*.
<https://doi.org/10.1080/15348458.2022.2057992>
- Arndt, Henriette L.; Woore, Robert** (2018). “Vocabulary learning from watching YouTube videos and reading blog posts”. *Language learning and technology*, v. 22, n. 3, pp. 124-142.
<https://doi.org/10.125/44660>
- Benson, Phil** (2015). “Commenting to learn: Evidence of language and intercultural learning in comments on YouTube videos”. *Language learning & technology*, v. 19, n. 193, pp. 88-105.
<http://llt.msu.edu/issues/october2015/benson.pdf>
- Bilibili** (2022). *Bilibili Inc. announces 2021 fourth quarter and fiscal year financial results*.
<https://ir.bilibili.com/news-releases/news-release-details/bilibili-inc-announces-2021-fourth-quarter-and-fiscal-year>
- Burgess, Jean; Green, Joshua** (2018). *YouTube: online video and participatory culture* (2nd ed.). Cambridge: Polity.
- Cassany, Daniel; Shafirova, Liudmila** (2021). “¡Ya está! Me pongo a filmar: Aprender grabando vídeos en clase”. *Revista signos*, v. 54, n. 107, pp. 893-918.
<https://doi.org/10.4067/s0718-09342021000300893>
- Castaño-Arques, Ana** (2019). *Literacidad crítica en ELE: Una investigación sobre la lectura entre estudiantes sinohablantes* [Doctoral dissertation, Pompeu Fabra University].
<http://hdl.handle.net/10803/668779>
- Chang, Yueh-ching; Chang, Yu-jung** (2019). “Identity negotiation in the third space: An analysis of YouTube channels hosted by expatriates in Taiwan”. *Language and intercultural communication*, v. 19, n. 1, pp. 77-92.
<https://doi.org/10.1080/14708477.2018.1450878>
- Chik, Alice** (2014). “Digital gaming and language learning: Autonomy and community”. *Language learning and technology*, v. 18, n. 2, pp. 85-100.
<https://doi.org/10.125/44658>
- Codreanu, Tatiana; Combe, Christelle** (2018). “Glocal tensions: Exploring the dynamics of intercultural communication through a language learner’s vlog”. In: Kern, Richard; Develotte, Chirstine. *Screens and scenes: Multimodal communication in online intercultural encounters*, pp. 40-61. New York: Routledge. ISBN: 978 0 367589271
- Codreanu, Tatiana; Combe, Christelle** (2019). “Vlogs, video publishing, and informal language learning”. In: Dressman, Mark; Sadler, Randall William. *The handbook of informal language learning*, pp. 153-168. Hoboken, NJ: Wiley.
<https://doi.org/10.1002/9781119472384.ch10>
- Duffy, Brooke Erin; Pinch, Annika; Sannon, Shruti; Sawey, Megan** (2021). “The nested precarities of creative labor on social media”. *Social media + society*, v. 7, n. 2.
<https://doi.org/10.1177/205630512111021368>
- Duffy, Brooke Erin; Poell, Thomas; Nieborg, David B.** (2019). “Platform practices in the cultural industries: Creativity, labor, and citizenship”. *Social media + society*, v. 5, n. 4.
<https://doi.org/10.1177/2056305119879672>
- Duffy, Brooke Erin; Sawey, Megan** (2022). “In/visibility in social media work: The hidden labor behind the brands”. *Media and communication*, v. 10, n. 1, pp. 77-87.
<https://doi.org/10.17645/mac.v10i1.4460>
- Espejel, Olivia; Concheiro, Pilar; Pujolà, Joan-Tomàs** (2022). “TikTok en la enseñanza de español LE/L2: telecolaboración y competencia digital”. *Journal of Spanish language teaching*, v. 9, n. 1, pp. 19-35.
<https://doi.org/10.1080/23247797.2022.2091843>
- Establés, María-José; Guerrero-Pico, Mar; Contreras-Espinosa, Ruth S.** (2018). “Jugadores, escritores e influencers en redes sociales: procesos de profesionalización entre adolescentes”. *Revista latina de comunicación social*, n. 74, pp. 214-236.
<https://doi.org/10.4185/RLCS-2019-1328>
- García-Rapp, Florencia** (2016). “The digital media phenomenon of YouTube beauty gurus: The case of Bubzbeauty”. *International journal of web based communities*, v. 12, n. 4, pp. 360-375.
<https://doi.org/10.1504/IJWBC.2016.080810>
- Godwin-Jones, Robert** (2019). “Riding the digital wilds: Learner autonomy and informal language learning”. *Language learning & technology*, v. 23, n. 1, pp. 8-25.
<https://doi.org/10.125/44667>

- Griffith, Margaret; Papacharissi, Zizi** (2010). "Looking for you: An analysis of video blogs". *First Monday*, v. 15, n. 1. <https://doi.org/10.5210/fm.v15i1.2769>
- Gromik, Nicolas A.** (2012). "Cell phone video recording feature as a language learning tool: A case study". *Computers and education*, v. 58, n. 1, pp. 223-230. <https://doi.org/10.1016/j.compedu.2011.06.013>
- Guerrero-Pico, Mar; Establés, María-José; Costa-Sánchez, Carmen** (2022). "Fan fiction y prácticas de lectoescritura transmedia en Wattpad: una exploración de las competencias narrativas y estéticas de adolescentes". *El profesional de la información*, v. 31, n. 2, e310212. <https://doi.org/10.3145/epi.2022.mar.12>
- Hafner, Christoph A.; Miller, Lindsay** (2011). "Fostering learner autonomy in English for science". *Language learning & technology*, v. 15, n. 3, pp. 68-86. <http://llt.msu.edu/issues/october2011/hafnermiller.pdf>
- Han, Yiting; Reinhardt, Jonathon** (2022). "Autonomy in the digital wilds: Agency, competence, and self-efficacy in the development of L2 digital identities". *Tesol quarterly*, v. 56, n. 3, pp. 985-1015. <https://doi.org/10.1002/tesq.3142>
- Hanauer, David-Ian** (2001). "The task of poetry reading and second language learning". *Applied linguistics*, v. 22, n. 3, pp. 295-323. <https://doi.org/10.1093/applin/22.3.295>
- Hine, Christine** (2015). *Ethnography for the internet: Embedded, embodied and everyday*. London: Taylor & Francis. <https://doi.org/10.4324/9781003085348>
- Huang, Hui-Wen** (2021). "Effects of smartphone-based collaborative vlog projects on EFL learners' speaking performance and learning engagement". *Australasian journal of educational technology*, v. 37, n. 6, pp. 18-40. <https://doi.org/10.14742/ajet.6623>
- Hung, Shao-Ting** (2011). "Pedagogical applications of vlogs: An investigation into ESP learners perceptions". *British journal of educational technology*, v. 42, n. 5, pp. 736-746. <https://doi.org/10.1111/j.1467-8535.2010.01086.x>
- Jerasa, Sarah; Boffone, Trevor** (2021). "BookTok 101: TikTok, digital literacies, and out-of-school reading practices". *Journal of adolescent & adult literacy*, v. 65, n. 3, pp. 219-226. <https://doi.org/10.1002/jaal.1199>
- Knobel, Michele; Lankshear, Colin** (2014). "Studying new literacies". *Journal of adolescent & adult literacy*, v. 58, n. 2, pp. 97-101. <https://doi.org/10.1002/jaal.314>
- Lankshear, Colin; Knobel, Michele** (2015). "Digital literacy and digital literacies: Policy, pedagogy and research considerations for education". *Nordic journal of digital literacy*, v. 10, n. Jubileumsnummer, pp. 8-20. <https://doi.org/10.18261/ISSN1891-943X-2015-Jubileumsnummer-02>
- Lee, Carmen** (2013). "'My English is so poor... so I take photos': metalinguistic discourses about English on Flickr". In: Deborah, Tannen; Trester; Anna Marie. *Discourse 2.0 Language and new media*, pp. 73-84. Washington: Georgetown University Press.
- Lee, Yeong-Ju** (2022). "Language learning affordances of Instagram and TikTok". *Innovation in language learning and teaching*. <https://doi.org/10.1080/17501229.2022.2051517>
- Lee, Yoonjoo; Chung, John-Joon-Young; Song, Jean Y.; Chang, Minsuk; Kim, Juho** (2021). "Personalizing ambience and illusionary presence: How people use study with me videos to create effective studying environments". In: *Conference on human factors in computing systems – Proceedings*. <https://doi.org/10.1145/3411764.3445222>
- Lee, Seryun** (2021). "Translating YouTube vlogs for a global audience: Innovative subtitling and community-building". *International journal of cultural studies*, v. 24, n. 5, pp. 767-790. <https://doi.org/10.1177/1367877920979717>
- Li, Zhiying** (2018). *Creencias de profesores chinos respecto a la enseñanza del español como lengua extranjera en China*. [Doctoral dissertation, Pompeu Fabra University]. <http://hdl.handle.net/10803/664232>
- Lovelock, Michael** (2017). "'Is every YouTuber going to make a coming out video eventually?': YouTube celebrity video bloggers and lesbian and gay identity". *Celebrity studies*, v. 8, n. 1, pp. 87-103. <https://doi.org/10.1080/19392397.2016.1214608>
- Lozanov, Georgi; Gateva, Evalina** (1988). *The foreign language teacher's suggestopedic manual*, ERIC. ISBN: 0677216602 <https://eric.ed.gov/?id=ED363092>

- Ma, Ke** (2022). "Transnational white masculinity on Chinese social media: Western male vloggers' self-representations during the covid-19 pandemic". *Asian anthropology*, v. 21, n. 3, pp. 211-223.
<https://doi.org/10.1080/1683478X.2022.2090054>
- Markham, Annette; Buchanan, Elizabeth** (2012). *Ethical decision-making and Internet research: recommendations from the AoIR ethics working committee (version 2.0)*.
<http://aoir.org/reports/ethics2.pdf>
- Marwick, Alice E.** (2013). *Status update: celebrity, publicity, and branding in the social media age*. New Haven: Yale University Press. ISBN: 978 0 300209389
- Marwick, Alice E.; Boyd, Danah** (2011). "I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience". *New media & society*, v. 13, n. 1, pp. 114-133.
<https://doi.org/10.1177/1461444810365313>
- Henthorn, Jamie** (2019). "International fan professionalization on Viki". *Television & new media*, v. 20, n. 5, pp. 525-538.
<https://doi.org/10.1177/1527476418770742>
- Pérez-González, Luis** (2022). "Subtitling disinformation narratives around COVID-19: 'Foreign'vlogging in the construction of digital nationalism in Chinese social media". In: Baker, Mona. *Unsettling translation*, pp. 232-247. New York: Routledge.
- Pun, Lok-Fai** (2021). *Fans in play and professional in the making: The fan professionalization of comic Doujin in China*. [Doctoral dissertation, The Chinese University of Hong Kong].
<https://repository.lib.cuhk.edu.hk/en/item/cuhk-2876340>
- Raun, Tobias** (2015). "Video blogging as a vehicle of transformation: Exploring the intersection between trans identity and information technology". *International journal of cultural studies*, v. 18, n. 3, pp. 365-378.
<https://doi.org/10.1177/1367877913513696>
- Raun, Tobias** (2018). "Capitalizing intimacy: New subcultural forms of micro-celebrity strategies and affective labour on YouTube". *Convergence*, v. 24, n. 1, pp. 99-113.
<https://doi.org/10.1177/1354856517736983>
- Reinhardt, Jonathon** (2022). "The history of language learning and teaching beyond the classroom". In: Reinders, Hayo; Lai, Chun; Sundqvist, Pia. *The Routledge handbook of language learning and teaching beyond the classroom*, pp. 9-23. New York: Routledge.
<https://doi.org/10.4324/9781003048169-3>
- Robin, Bernard R.** (2008). "Digital storytelling: A powerful technology tool for the 21st century classroom". *Theory into practice*, v. 47, n. 3, pp. 220-228.
<https://doi.org/10.1080/00405840802153916>
- Sauro, Shannon** (2017). "Online fan practices and call". *Calico journal*, v. 34, n. 2, pp. 131-146.
<https://doi.org/10.1558/cj.33077>
- Sauro, Shannon** (2019). "Fan fiction and informal language learning". In: Dressman, Mark; Sadler, Randall William. *The handbook of informal language learning*, pp. 139-151. Hoboken, NJ: Wiley.
<https://doi.org/10.1002/9781119472384.ch9>
- Sauro, Shannon; Zourou, Katerina** (2019). "What are the digital wilds?". *Language learning & technology*, v. 23, n. 1.
<https://doi.org/10125/44666>
- Sauro, Shannon; Sundmark, Björn** (2018). "Critically examining the use of blog-based fanfiction in the advanced language classroom". *Recall*, v. 31, n. 1, pp. 40-55.
<https://doi.org/10.1017/S0958344018000071>
- Sánchez-López, Iván; Roig-Vila, Rosabel; Pérez-Rodríguez, Amor** (2022). "Metaverse and education: The pioneering case of *Minecraft* in immersive digital learning". *Profesional de la información*, v. 31, n. 6, e310610.
<https://doi.org/10.3145/epi.2022.nov.10>
- Scolari, Carlos A.; Masanet, Maria-José; Guerrero-Pico, Mar; Establés, María-José** (2018). "Transmedia literacy in the new media ecology: Teens' transmedia skills and informal learning strategies". *El profesional de la información*, v. 27, n. 4, pp. 801-812.
<https://doi.org/10.3145/epi.2018.jul.09>
- Shi, Xingsong; Chang, Yujie; Gao, Jiawei** (2022). "Investment in transnational identity to become microcelebrities in China: On American uploaders' success in a Chinese video-sharing website". *Language and intercultural communication*, v. 22, n. 6, pp. 642-661.
<https://doi.org/10.1080/14708477.2022.2078343>

- Snelson, Chareen** (2015). "Vlogging about school on YouTube: An exploratory study". *New media and society*, v. 17, n. 3, pp. 321-339.
<https://doi.org/10.1177/1461444813504271>
- Vázquez-Calvo, Boris; Elf, Nikolaj; Gewerc, Adriana** (2020). "Catalan Teenagers' identity, literacy and language practices on YouTube". In: Freiermuth, Mark R.; Zarrinabadi, Nourollah. *Technology and the psychology of second language learners and users*, pp. 251-278. Cham: Springer International Publishing.
https://doi.org/10.1007/978-3-030-34212-8_10
- Vázquez-Calvo, Boris; Zhang, Leticia-Tian; Pascual, Mariona; Cassany, Daniel** (2019). "Fan translation of games, anime, and fanfiction". *Language learning and technology*, v. 23, n. 1, pp. 49-71.
<https://doi.org/10.125/44672>
- Vázquez-Calvo, Boris; Shafirova, Liudmila; Zhang, Leticia-Tian** (2022). "Language learning hashtags on TikTok in Chinese, Italian, and Russian". In: Klimanova, Liudmila. *Identity, multilingualism & call*, pp. 104-134. Sheffield: Equinox.
<https://doi.org/10.1558/equinox.43411>
- Vizcaíno-Verdú, Arantxa; De-Casas-Moreno, Patricia; Jaramillo-Dent, Daniela** (2022). "Thanks for joining our life: Intimacy as performativity on YouTube parenting vlogs". *Profesional de la información*, v. 31, n. 4, e310407.
<https://doi.org/10.3145/epi.2022.jul.07>
- Wang, Hung-Chun; Chen, Cheryl-Wei-Yu** (2020). "Learning English from youtubers: English L2 learners' self-regulated language learning on YouTube". *Innovation in language learning and teaching*, v. 14, n. 4, pp. 333-346.
<https://doi.org/10.1080/17501229.2019.1607356>
- Wang, Xiaoxian** (2022). "Popularising vlogging in China: Bilibili's institutional promotion of vlogging culture". *Global media and China*, v. 7, n. 4, pp. 441-462.
<https://doi.org/10.1177/20594364221133536>
- Wang, Xiaoxian; Picone, Ike** (2021). "The art of attracting attention: A process model of chinese toubu vloggers' strategies to create online identities and self-brands". *Celebrity studies*.
<https://doi.org/10.1080/19392397.2021.1991408>
- Wang, Xiaoxian; Picone, Ike** (2022). "Mobilized, negotiated and balanced: Chinese school vloggers' platform engagements and layered identity construction on Bilibili". *Journal of youth studies*.
<https://doi.org/10.1080/13676261.2022.2098705>
- Xu, Jian; Zhao, Xinyu** (2022). "Coping with the 'double bind' through vlogging: Pandemic digital citizenship of Chinese international students". *Continuum*, v. 36, n. 2, pp. 260-273.
<https://doi.org/10.1080/10304312.2021.2008319>
- Yang, Ya-Ting C.; Wu, Wan-Chi I.** (2012). "Digital storytelling for enhancing student academic achievement, critical thinking.; Learning motivation: A year-long experimental study". *Computers and education*, v. 59, n. 2, pp. 339-352.
<https://doi.org/10.1016/j.compedu.2011.12.012>
- Yu, Man** (2021). El español en China. In: *El español en el mundo: Anuario del Instituto Cervantes 2021*. Instituto Cervantes.
https://cvc.cervantes.es/lengua/anuario/anuario_21/asia_oceania/china.htm
- Zhang, Leticia-Tian; Cassany, Daniel** (2019a). "The 'danmu' phenomenon and media participation: Intercultural understanding and language learning through 'The Ministry of Time'". *Comunicar*, v. 27, n. 58, pp. 19-29.
<https://doi.org/10.3916/C58-2019-02>
- Zhang, Leticia-Tian; Cassany, Daniel** (2019b). "Practices on audiovisual comprehension and Spanish to Chinese translation in a fansub community". *Revista española de lingüística aplicada/Spanish journal of applied linguistics*, v. 32, n. 2, pp. 620-649.
<https://doi.org/10.1075/resla.17013.zha>
- Zhang, Leticia-Tian; Vázquez-Calvo, Boris** (2022). "'¿Triste estás? I don't know nan molla': Multilingual pop song fandubs by @miree_music, ITL" - *International journal of applied linguistics*, v. 173, n. 2, pp. 210-211.
<https://doi.org/10.1075/itl.21007.zha>

Voice search optimization in digital media: challenges, use and training

Carlos Lopezosa; Lluís Codina; Javier Guallar; Mario Pérez-Montoro

Nota: Este artículo se puede leer en español en:
<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87214>

Recommended citation:

Lopezosa, Carlos; Codina, Lluís; Guallar, Javier; Pérez-Montoro, Mario (2023). "Voice search optimization in digital media: challenges, use and training". *Profesional de la información*, v. 32, n. 3, e320307.
<https://doi.org/10.3145/epi.2023.may.07>

Manuscript received on 15th December 2022
Accepted on 13th April 2023



Carlos Lopezosa ✉
<https://orcid.org/0000-0001-8619-2194>
Universitat de Barcelona
Centro de Investigación en Información,
Comunicación y Cultura (CRICC)
Melcior de Palau, 140
08014 Barcelona, Spain
lopezosa@ub.edu



Lluís Codina
<https://orcid.org/0000-0001-7020-1631>
Universitat Pompeu Fabra
Facultad de Comunicación
Roc Boronat, 138
08018 Barcelona, Spain
lluis.codina@upf.edu



Javier Guallar
<https://orcid.org/0000-0002-8601-3990>
Universitat de Barcelona
Centro de Investigación en Información,
Comunicación y Cultura (CRICC)
Melcior de Palau, 140
08014 Barcelona, Spain
jguallar@ub.edu



Mario Pérez-Montoro
<https://orcid.org/0000-0003-2426-8119>
Universitat de Barcelona
Centro de Investigación en Información,
Comunicación y Cultura (CRICC)
Melcior de Palau, 140
08014 Barcelona, Spain
perez-montoro@ub.edu

Abstract

In view of the widespread use of virtual voice assistants and/or voice searches on smartphones to find all kinds of information, this article explores voice search optimisation (VSO) and its application in the journalistic sector. To this end, 32 semi-structured interviews were conducted with experts representing different professional profiles in the fields of journalism, search engine optimisation (SEO) and academic research. On the basis of the data, eight semantic categories were created and the experts' perceptions were correlated to identify response patterns. The results confirm the existence of various degrees of convergence and divergence between these three professional profiles in relation to different dimensions of VSO, such as its definition, its techniques, its current and future strategic role in digital media, and its application in journalistic writing. This study confirms that although the use of VSO in digital news media is still in its embryonic stages, it will be useful in the medium and long term to train journalists in basic aspects of voice searches. In addition, internal SEO departments should be prepared to optimise the visibility of news for virtual voice assistants when they become widespread and when the technology companies that develop these assistants define a viable business model.

Keywords

Digital news media; Voice search optimization; VSO; Digital journalism; Web visibility; Search engine optimization; SEO; Voice search; Training.

Funding

This work is part of the Project "Parameters and strategies to increase the relevance of media and digital communication in society: curation, visualisation and visibility (Cuvicom)". PID2021-123579OB-I00 (Micin). Spanish Ministry of Science and Innovation (Spain)



1. Introduction

Search engine optimisation (SEO), understood as the set of techniques that help position web content in the top results on *Google* or *Bing* (among other search engines), has become an important business strategy for digital news media to increase their audience (Carlson, 2007). Such search engines help users select the most relevant content (Machill; Beiler; Zenke, 2008) in response to their searches, and it is for this reason, as well as the increase in hyper-competitive internet news, that digital media are highly dependent on SEO (Lopezosa et al., 2020).

In fact, there is a complex relationship between the media and *Google* and other technology companies (Lee; Chyi, 2015), characterized by great ups and downs (Marcos-Recio; Sánchez-Vigil; Olivera-Zaldua, 2015) that have caused conflicts ranging from audience issues and web traffic to financial aspects through accusations of oligopoly among many others (Guallar, 2015). All of this has resulted in the dynamics between both parties conditioning and affecting the journalistic industry and its readers (Barr, 2014). However, the relationship between *Google* and the media can be characterized by the term *frenemies*, since *Google* ensures that its products are also platforms for the media, e.g., *Google News* and *Google Discover*, among others.

This situation has led to a boom in research on web visibility in journalism, with very interesting studies on the application of SEO in digital media, including those by

- Giomelakis and Veglis (2015) and Charlton (2016), which demonstrate that there is considerable room for improvement in the alignment between journalistic routines and SEO routines;
- Dick (2011) and Smyrniaios and Sire (2014), who recognise the need to constantly pursue optimal convergence between the best journalism and the best SEO; and
- Lopezosa et al. (2021) and Pedrosa and De-Morais (2021), which identify practical SEO techniques applied in journalistic writing.

All these studies analyse digital media visibility on digital platforms. However, visibility is constantly evolving, and nowadays users also consume information via other channels, such as voice assistants (also called virtual assistants), physical assistants such as *Alexa* and *Google Home* and others integrated into mobiles, such as *Siri* or *Google Assistant*.

In the case of voice searches, i.e., queries spoken to a virtual assistant that then returns a result, there are a number of academic studies published in the social sciences (Sa, 2016; Shokouhi; Ozertem; Craswell, 2016; Mairesse; Raccuglia; Vitaladevuni, 2016; Hurwitz et al., 2017; Guy, 2018; Lovato; Piper, 2019), several of which include a consideration of journalism (Lochrie et al., 2018; Kischinhevsky, 2019; Jung et al., 2019; Fagundes-Pase et al., 2020; Turow, 2021) but not of visibility, that is, not of voice search optimisation (hereafter VSO). There have also been numerous industry reports and news stories published on the use of voice search technology, both at a general level (Huffman, 2014; Cachón, 2019; *Isidigitaldata*, 2020; *Rabit and Pork*, 2021) and in the context of digital media (Newman, 2018). It is also possible to find technical documentation (Checa, 2020; Makhyan, 2022) with a broad and generalist application, although not directly focused on news. However, to date no academic study has focused on voice search optimisation and its application to digital media.

Likewise, the dynamics of voice-based SEO have not yet been studied by the academy. However, we can find sectoral reports and user guides by companies and specialized media (Bonelli, 2017; Gareth, 2019). In this sense, voice-based SEO is understood as the discipline in charge of optimizing content so that voice assistants such as *Alexa*, *Google Assistant*, *Siri*, and *Cortana*, among others, offer the most accurate results for conversational queries (Codedesign, 2023). Among the recommendations to optimise websites for voice searches are using interrogative keywords, marking the content with the *Schema.org Speakable* scheme (Schema.org, 2018), having a frequently asked questions section, and having a website that loads fast, is compatible with mobile devices, and works for local inquiries, among others (Bonelli, 2017; Gareth, 2019; Codedesign, 2023).

On this premise, and based on the assumptions (1) that the use of voice searches is recognised as growing in line with the consolidation of the market for virtual assistants (Sterling, 2016), and (2) that digital media need to know how to programme their news for voice assistants (Newman, 2018), this study offers an analysis of VSO in digital news media. To this end, we use semi-structured interviews, a methodology that has already been effectively applied to the study of search engine optimisation (Schultheiß; Lewandowski, 2020; Lopezosa et al., 2020; 2021).

The general objective of this study is to carry out a diagnosis of VSO in the journalistic sector to identify the extent of its application in digital media and the challenges involved in implementing voice search optimisation strategies in journalistic writing, both on a strategic business level and on the level of training for journalists and SEO teams integrated into digital media.

On the basis of this main objective, the specific objectives are:

- 1) To explore the opinions of search engine optimisation experts working in digital media, consultancy firms and universities in relation to VSO applied to digital media.
- 2) To determine whether there is a convergence of perceptions among these different professional profiles with respect to the dimensions of VSO applied to digital media.
- 3) To identify the knowledge journalists and internal media web visibility teams should have about VSO and to propose training recommendations in this regard, whether for editorial staff or in university programmes.

2. Material and methods

For this study, 32 semi-structured interviews (Coller, 2000; Valles, 2002; Alves; Díaz-Noci, 2019) were carried out during the first quarter of 2022 with journalists and SEO professionals working for digital media, web visibility experts in the field of consultancy and entrepreneurship, and teachers and researchers whose field of study or teaching is search engine optimisation.

The participants were selected based on a previous design that covered the different professional profiles mentioned above, for which specific data are provided in tables 1 and 2). Secondly, all of them had to have competencies in SEO and work in media, business consultancy, teaching, or as a researcher in the field. Thirdly, in accordance with a series of criteria questions (Valles, 2002), identifying who has the relevant information, is more physically and socially accessible, is more willing to provide the information and is more capable of communicating it accurately. After requesting interviews with the experts identified as indicated above, 32 interviews were carried out (Table 1).

Table 1. Participants in the semi-structured interviews

Journalists/professionals responsible for SEO at media outlets	
Participant	Digital media
Natalia Alaminos	<i>apuntmedia.co.uk</i>
Alexis Apablaza	<i>player8.org</i>
José Luis Cases	<i>rankia.com</i>
Nacho Delgado	<i>brand.com</i>
Edgard Matsuki	<i>boatos.org</i>
Richard Nazarewicz	<i>wsj.com</i>
Víctor Pérez	<i>prensaiberica.es</i>
Alexandra Ptachick	<i>usatoday.com</i>
Ángeles Sánchez	<i>grupojoly.com</i>
José Soto	<i>eleconomista.com.mx</i>
SEO experts in the field of consultancy	
Participant	Company/website
David Ayala	<i>davidayala.com</i>
Emilio Berenguer	<i>interamplify.com</i>
María José Cachón	<i>laikateam.com</i>
John Campbell	<i>wearerabbitandpork.com</i>
Esteve Castells	<i>estevecastells.com</i>
Esther Checa	<i>t2o.com</i>
Emilio García	<i>campoweb.com</i>
Fernando Maciá	<i>Humanlevel.com</i>
Fernando Muñoz	<i>fernando.senormunoz.es</i>
Jordi Ordóñez	<i>jordiob.com</i>
Juan Pérez	<i>agenciaseo.eu</i>
Bruno Ramos	<i>brunoramos.es</i>
Álvaro Rondón	<i>alvarorondon.com</i>
Natzir Turrado	<i>analistaseo.es</i>
Teachers and researchers	
Participant	Institution
Raquel Escandell	<i>University of Alicante</i>
Dimitrios Giomelakis	<i>Aristotle University of Thessaloniki</i>
Carlos Gonzalo Penela	<i>Pompeu Fabra University</i>
Noelia Herrero	<i>Barcelona School of Management</i>
Nic Newman	<i>Reuters Institute</i>
Leyberson Pedrosa	<i>São Paulo State University</i>
María Pérez	<i>Pompeu Fabra University</i>
Andreu Sulé	<i>University of Barcelona</i>

The interviews were conducted by email and/or via audio recording and then transcribed. The results were then coded and analysed using the qualitative data analysis software *Atlas.ti* (version 22), a highly prestigious tool that has been

used for over 30 years (Muhr, 1991) and has become one of the most widely used in the sector (Paulus; Lester, 2015). Specifically, this tool has helped to create categories that have been applied to all dimensions of interview analysis.

Table 2 outlines the design of the professional profiles for the selection of the experts interviewed according to the criteria indicated above (Valles, 2002).

Table 2. Professional profiles

Journalists/professionals responsible for SEO at media outlets	SEO experts in the field of consultancy	University professors teaching SEO subjects
Journalist responsible for audiences and SEO at a leading international online digital news media outlet	External SEO advising digital news media	Lecturer in Digital Information and Documentation at undergraduate level
Technical SEO for an internationally important online business media outlet	Voice Search Optimisation Consultant at a national SEO agency	Lecturer in SEO and Big Data at master's level
Product, engineering and SEO manager for a specialised news website	Founder of an international voice experience agency	Senior Research Associate of a prestigious international research group dedicated to digital news media studies
SEO Manager of a national corporate digital media group	SEO expert for a multinational company	Lecturer in search engines at master's level
SEO for a regional business digital media group	SEO expert for a medium-sized company	Lecturer in advertising and public relations at undergraduate level
SEO manager at a leading online sports media company	SEO consultant with own agency	PhD and researcher in web visibility and digital news media
Head of Digital Content Service and SEO for a regional digital news media outlet.	Independent consultant offering SEO services to companies.	PhD student researching web quality and SEO
Owner of a small digital media outlet who performs basic SEO tasks	Owner of an international SEO agency	Researcher on semantic SEO and structured data
Editor of an international fact-checking digital news media outlet with expertise in web visibility.	SEO entrepreneur of digital projects	Researcher on web visibility and big data
Journalist with extensive knowledge of SEO working for a public media outlet	Independent SEO consultant who also provides training and lectures	PhD and researcher on professional profiling and SEO

The interview questions and their associated categories configured for *Atlas.ti* with the intention of coding the interviews and obtaining response patterns are listed in Table 3.

Table 3. Coding of interviews

Questions/Categories	Objectives/categories
1. Do you know the term 'voice search optimisation'?	(1) Identify whether respondents are familiar with VSO and (2) develop a consensual definition of VSO
2. Do you know or would you recommend any positioning strategies for voice searches?	(1) Identify whether respondents are aware of VSO strategies, and (2) attempt to standardise VSO strategies expressed by respondents to obtain a practical guide for use
3. Do you think voice searches will play a major role for digital news media in the short to medium term?	(1) Determine the degree of importance of VSO in digital media according to the opinions and arguments of the experts interviewed
4. Do you think digital news media adequately apply voice search optimisation strategies?	(1) Reflect expert opinion on the effective implementation of VSO in digital news media
5. Do you think journalists should be aware of aspects of voice search optimisation?	(1) Ascertain whether journalists and digital media SEO teams should have VSO skills in their work, and if so, (2) what VSO skills they should learn

Finally, to develop the training models for efficient VSO competencies focusing on the journalist/editor and media outlet SEO manager profiles, the following process was adopted:

- 1) VSO strategies (question 2) and VSO knowledge areas (question 5) were identified based on the results of the responses.
- 2) The answers were semantically compared across the professional profiles (journalists/professionals responsible for SEO in digital media, SEO experts in the consultancy field and university professors and researchers) to determine the degree of convergence or divergence between them.
- 3) Levels of perceived agreement on VSO knowledge areas resulting from the experts' statements were identified. A high level of perceived agreement was when all three parties (journalism professionals, SEO consultants/experts and university professors) expressed the same ideas, a medium level was when two of the three parties appeared to agree, and a low level was when all three parties expressed different responses.
- 4) Training models were developed by selecting high and medium levels of perceived agreement on VSO strategies, skills, and abilities. Low levels of perceived agreement were discarded.

3. Results

This section outlines the key ideas emerging from the interviews. It begins with an analysis of the degree of consensus among the different professional profiles on each of the survey questions. This is followed by the establishment of categories to summarise the most important ideas, and finally, the proposition of possible training models for VSO competencies focusing on two professional profiles (the journalist/editor profile and the SEO expert profile) based on the level of perceived agreement among the professionals in the scenarios studied.

In general, the term VSO (question 1) is widely known by the experts: 29 of them were aware of the concept, and although for three of them it was a new term, they were aware of the existence of techniques capable of positioning content for voice assistants. In this respect, 29 experts recommended specific VSO techniques (question 2); conversely, two recommended traditional SEO techniques (those applied to rankings in *Google*, *Bing*, etc.) as a process to position content for virtual assistants, and one expert said he did not know of any specific strategy.

Regarding the strategic role of VSO for digital news media (question 3), 23 interviewees considered that it did or would play a significant role in the medium or long term; in contrast, seven experts considered that its role was irrelevant or did not directly affect digital media, while two were unsure of its role or could not give a clear answer.

There was a low level of consensus on the effective application of VSO in digital media (question 4). Specifically, 12 of the interviewees stated that digital media outlets in general were not applying VSO techniques, although ten respondents considered that some digital media outlets were engaging in good VSO practices, five reported that, in general, the number or quality of VSO techniques being developed was low, and the other five did not know whether or not the digital media were currently applying voice search optimisation techniques.

However, 23 interviewees stated that there was a need for VSO in journalistic practices (question 5), although these needs vary, as will be discussed below, depending on the professional profile (journalist/editor or in-house SEO) and on changes in the industry and in news consumption in the short, medium and long term.

Below is an integrative synthesis (connecting the different answers to the same questions) developed using the categories with *Atlas.ti*.

3.1. Voice search and VSO

Just as users type their search query into *Google* in a web browser or mobile app, voice searches involve the same process but using the spoken word instead. It is the search engine's task to select the results using information retrieval techniques, applying relevance algorithms, and of course, the whole process of understanding the users' voice, using natural language processing techniques and translating voice into data. Voice searches could therefore be defined as entities that use the spoken word to query a search engine so that the auditory interface returns a unique, unambiguous answer provided by either a virtual voice assistant (*Alexa* or *Google Home*) or an integrated assistant (*Siri*, *Google Assistant*, *Cortana*, etc.).

“Voice searches could therefore be defined as entities that use the spoken word to query a search engine so that the auditory interface returns a unique, unambiguous answer provided by either a virtual voice assistant or an integrated assistant”

On the other hand, voice search optimisation is a variant of SEO that is defined as the process of semantic and technical optimisation of content from the web so that it can be processed, understood and ranked when users ask virtual voice assistants.

The VSO must ensure that the website's text content is correctly configured to appear as a suggested result when a user searches with voice signals. This incorporates the application of a set of technical and content practices aimed at aligning a website's pages with the search intentions of its potential audience when the audience uses voice rather than a keyboard as an input interface.

VSO also incorporates the process used by publishers to make programmes or audio items easy to find via voice interfaces such as virtual assistants. In short, voice search optimisation is the process of gaining control of voice search results.

3.2. VSO strategies

Some of the interviewees consider it unnecessary to carry out specific actions to position themselves in voice searches and that it is only necessary to apply traditional SEO strategies, as they believe that there is no specific optimisation process for the voice environment for web-based content that has a presence in virtual assistants.

However, most of the responses of the experts interviewed refer to certain strategies additional to or adapted from traditional SEO that can help rank content for voice searches.

General voice search positioning strategies applicable to all types of virtual assistants are outlined below, followed by specific strategies for *Google Assistant* (a virtual assistant mainly active on *Google Home* and *Android* mobile devices).

At a general level, three scenarios are observable in descriptions of techniques for optimising voice searches:

- Strategic scenario: The content needs to be adapted to the way users search by voice, i.e., it is necessary to know what morphology and syntax tend to predominate when queries are made by means of speech. Next, it is necessary to carry out keyword research on these queries and then apply them to the website, as according to the experts consulted, there are notable differences between voice searches and written queries (keywords), since the former use natural language phrases in the form of a question while the latter use single keywords. Finally, tools such as *alsoasked.com* or *answerthepublic.com* can be used to identify key phrases in the form of questions.
- Technical scenario: It is essential to ensure that the metadata (HTML elements containing information about each web page) is adequate, that it is complete, and that the content can be easily read by voice navigators and robots. In addition, a high-quality website needs to be developed in terms of both its mobile version and its loading times and, finally, page results should have excellent user experience and satisfaction values.
- Content scenario: On the website to be positioned, it is advisable to use expressions that are easy to understand with the use of connectors that give rise to concrete explanations. It is also advisable to avoid metaphors that may confuse the interpretation of the text. Additionally, the content must be correctly structured, while answering specific questions since users make queries in the form of questions. To position the questions, it is necessary to use long-tail keywords (long query phrases) in question format, which describe the answer briefly, concisely, and directly in the first paragraph. On the other hand, local content should also be considered, as many voice searches are local (e.g., film premieres in my city; traffic today, etc.). To this end, it is recommended to include specific names of places in prominent spaces on the web page, such as in the title or subtitle.

At a general level, three scenarios are observable in descriptions of techniques for optimising voice searches: Strategic scenario, technical scenario, and content scenario

Other aspects to consider have to do specifically with *Google Assistant*, which tends to read the answer to informational questions that are positioned at position zero (Spencer, 2017). In addition, it can read up to a maximum of 45 - 50 words. To appear in position zero, experts recommend developing content like lists, rankings, or comparisons. Another element that plays a fundamental role in appearing in position zero is the authority of a website's domain and the factors related to what *Google* calls E-A-T, which stands for Expertise, Authoritativeness, and Trustworthiness (Google, 2021).

Moreover, *Google* is working on *Actions*, automatic voice apps derived from web content; *Amazon's* version of this is called *Alexa Skills Kit* (Google, 2022; Amazon, 2022). The types of compatible content it has started with are podcasts, recipes, news, how-to guides, and FAQs. In order to be able to position this type of content for virtual assistants, it is technically advisable to include JSON-LD (data encoding method) structured data fragments, i.e., to mark them semantically with the specific entities of *Schema.org* (an initiative of *Google*, *Bing*, *Yahoo* and *Yandex* that allows web content to be marked semantically so that search engines can interpret it more easily).

It is also recommended to use the schema (metadata) called *Speakable* from the *Schema.org* language for news. This schema, launched in 2018, was created for news publishers to mark the most relevant sections of an article, so that *Google Assistant* can read them aloud (Google, 2018; *Schema.org*, 2018).

3.3. The role of VSO in digital media

Those experts who believe that VSO does or will in the short/medium term play a significant role in digital media say that this is due to the following circumstances: (1) because traffic coming from voice searches is growing; (2) because we are moving towards a reality where we seek to satisfy our needs almost instantaneously, and voice assistants offer that immediacy; (3) because in the medium term, voice assistants will become mainstream; and (4) because new forms of information consumption always tend to add to what already exists.

Conversely, those interviewed who consider that VSO does not play an important role in the digital news media, at least in the short and medium term, agree that this is mainly because for at present there is no clear business opportunity that can be monetised (either for digital media outlets or for the companies that develop this technology), and that the search results offered by voice assistants are concise, making it difficult at this stage to include reproducible in-depth news stories or reports. Added to this is the existence of an obstacle that is slowing down the development of VSO within the journalistic sector: the infrastructure itself, which is in the hands of technology companies. For example, if you ask *Siri* a question, *Apple's Siri* will look for the result in one of its trusted sources. *Apple* compiles this list of trusted sources, so if a media outlet is not on this list, it is not possible for it to appear as a voice search result.

3.4. Implementation of VSO in digital media

It is generally acknowledged that there is currently no clear application of VSO by digital media, but that this will most likely change in the medium/long term.

Optimisation for voice searches is very similar to the usual content writing strategy (question-answer) established prior to introduction of voice search technology, which is why digital media usually have optimised content for voice searches

even when such optimisation is not explicit. The major digital news media outlets are making great efforts to optimise their content from a semantic point of view by marking their content with *Schema.org* language schemas, which can contribute the most to voice search optimisation.

On the other hand, many respondents observe that digital media outlets are working towards developing the necessary skills to become integrated as sources into existing ecosystems (*Apple, Alexa, Google, etc.*). They also identify that digital media companies are currently focusing on work with podcasts or streaming, which is not specifically VSO, or some kind of app in assistants that can recite the news. This is very clear in the strategies of radio stations and their websites, as most radio broadcasters are very interested in voice search technology because they recognise it as key to their future. For example, the *British Broadcasting Corporation (BBC)* and the *American Broadcasting Company (ABC)* are all aware of this, as demonstrated by their strategies to expand their visibility through voice assistants.

3.5. Expertise with VSO among journalists and SEO teams

Some of the experts believe that journalists should create journalistic content and should not focus on voice search optimisation. However, they recognise that SEO teams should be engaged in creating specific content for VSO.

Regarding those who consider that journalists and SEO teams should have specific knowledge of VSO, the levels of perceived agreement among the professionals for the knowledge areas identified are shown in Table 4. The responses suggest that there is an opportunity for digital media outlets to take advantage of the voice search niche as part of their differentiation strategy.

Table 4. Level of perceived agreement among the professionals of the scenarios studied

Knowledge of VSO	Level of perceived agreement		
	High	Medium	Low
Journalists should have a basic understanding of voice search technology.		●	
Journalists should consider how a user searches via voice assistants before creating their news story.	●		
Journalists will have to adapt the wording of news stories to how people make voice searches.		●	
Journalists should be aware of the importance of SEO for today's media outlets.	●		
Journalists should learn how to write headlines for digital publishing that will have an impact on voice searches.	●		
Journalists need to know web technologies and understand the culture of online news consumption.		●	
Journalists should better structure their interviews to take into account the possibility of being listened to via virtual assistants.			●
SEOs will advise journalists on how existing news metadata should be adapted and changed for voice searches.	●		
SEOs will adapt content and headlines to voice searches.		●	
SEOs will adapt the website on a technical level through microformats or new types of content to improve VSO.		●	
SEO departments will help and support editorial staff in order to achieve better results in these searches.		●	

3.6. Training recommendations

Finally, based on the analysis of the interviews and the results obtained, a possible training model has been drawn up in the form of a table (Table 5) that includes the profile of a journalist/editor with SEO skills and the profile of a digital media SEO expert (but not a journalist).

4. Discussion and conclusions

On a general level, this research offers a diagnosis of VSO in the journalistic sector in order to identify the extent of its application in digital media and the challenges involved in its implementation. This represents a new contribution to the research on search engine optimisation in the journalistic sector that has been pursued intensively since 2007 (**Carlson, 2007; Smyrnaio; Rebillard, 2009; Giomelakis; Veglis, 2015; Charlton, 2016; García-Carretero et al., 2016; Pedrosa; De-Morais, 2021**). This area may offer an opportunity for business initiatives with a low opportunity cost, as it does not require special investments.

Research on voice search technology and digital media published to date has highlighted the need to

- work on adapting content to voice interfaces (**Fagundes-Pase et al., 2020**);
- adopt agreements between digital news media outlets and voice assistant developers to work towards updating the provision of these new ways of summarising news (**Kischinhevsky, 2019**);
- be aware of the types of news searches performed by virtual assistants (**Jung et al., 2019**);
- consider how news consumption could shape the news agenda (**Turow, 2021**); and
- promote the use of technologies that allow news to be read aloud for playback by virtual assistants (**Lochrie et al., 2018**).

All these needs have been explored in this research from the point of view of VSO, resulting in a new research approach.

Table 5. Journalist/editor training model and in-house SEO team in digital media

Profile	Subjects	Training in
Journalist	Theoretical orientation to voice searches	Introduction to voice searches
		Types of virtual assistants
		How users are using voice assistants to get information
	User-orientation of voice searches	Identifying news topics based on long-tail keywords
		Adaptation of content to the user's search intention by applying a question-answer structure
	Content-orientation of voice searches	Learning to headline for digital publishing
		Keyword optimisation in headlines
		Use of uncomplicated semantics giving clear and concise answers within the content
		Creation of audio content (podcasting) or summary content tailored to the requirements of voice assistants
Organisation of content by blocks that provide answers to users' questions		
SEO team	Productive routines	Voice-adapted morphology and syntax
		Keyword research tailored to voice searches (usually in question form)
		Ways to optimise metadata
		Ways to optimise responsive websites and loading speed
		Ways to optimise content for local/geographical searches
		Ways to optimise content to appear in position zero
		Ways of optimising EAT factors
		Semantic content tagging taking into account <i>Actions for Google Assistant</i>
		<i>Schema.org Speakable</i> content markup
		JSON-LD

At a specific level, this research offers a multidisciplinary expert view of VSO applied to online media outlets through interviews with Spanish and international professionals, in some cases including sectoral reports (Cachón, 2019; Newman, 2018; *Rabit & Pork*, 2021) and academic studies (Sulé-Duesa, 2015; Giomelakis; Veglis, 2015; Gonzalo-Penela; Codina; Rovira, 2015; Pedrosa; De-Morais, 2021). The expert view expressed here generally tend to have a high degree of convergence, except in relation to the application of VSO in journalistic writing, which has a low level of consensus, possibly because the implementation of VSO in digital media is still in its embryonic stages.

The definition offered by experts of the main dimensions of VSO studied here is in line with the main sources of industry information, such as *Search Engine Journal* (Makhyan, 2022) and *Search Engine Land* (Jones, 2019), in which they explain that optimizing for voice search involves, among other things, answering questions in featured search results, paying attention to local content, and improving adaptability to mobile devices. However, this research broadens that definition by linking it to the process adopted by publishers to make audio elements easier to find via voice interface. On the other hand, in relation to VSO strategies in digital news media and *Schema.org*, there is a previous study of SEO in digital media (Lopezosa *et al.*, 2020) that identifies the importance of VSO, which includes the optimization of technical aspects of semantic markup with *Schema.org*, but does not deal with it in detail. The present study is in line with studies asserting that voice search technology will play an important role in information consumption in general (Cachón, 2019; *Isidigitaldata*, 2020), and news consumption in particular (Newman, 2018), but it offers a new level of precision to the research. Additionally, this study confirms that there is currently no consolidated view on VSO in digital media, although current trends suggest that this will change in the medium term, especially as the use of voice assistants increases.

Limitations

First, there are other interview methods that could have been used, such as ethnographic methods or case studies, which could certainly be utilized in future research. Along these lines, a more immersive body of work to analyze, contextualize, and enrich these findings could be obtained in the future. The idea of future research for our team or others who wish to explore this area is to build a critical perspective on VSO, for which our research could be a good base. The present research could also be a good starting point for future research opportunities regarding issues such as the risk of large media groups entering into commercial agreements with search engines to appear at the top of searches, or even actors who promote misinformation appropriating these techniques to circulate negative or hate speech with political and/or economic motivations.

The second limitation refers to the very technological context of web visibility and the consumption of information through search engines. Although voice assistants have been working for years to become essential devices, for now they have not succeeded. This has allowed for the emergence of new technologies, among them artificial intelligence; more specifically, chatbots such as *ChatGPT* (Edell, 2023), for which both *Google* and *Bing* have opted (Pichai, 2023; Mehdi, 2023). However, there is no evidence at the time of carrying out this study that AI has made fundamental con-

tributions linked to these new recent technological waves, and we believe that in the future it will be necessary to study web visibility from the point of view of conversational AI.

Finally, this article has presented a proposal on what journalists and in-house media web visibility teams should know about VSO, along with training recommendations. This proposal could be useful in several scenarios: for example, as part of the ongoing strategies for training editorial staff, especially in the case of digital media with a perspective focusing more on entrepreneurship and new product creation, particularly given the low opportunity cost. This research complements previous studies on SEO and the possibilities of SEO training both for editorial staff and for students in journalism degree or postgraduate programmes. At the same time, it offers insights for digital media SEO experts who may be considering voice searches for their news content. Overall, this research and other studies published on digital media visibility optimisation strategies show that scholars and practitioners need to be vigilant as the information retrieval algorithms used by search engines such as *Google*, *Bing*, etc., and their platforms are constantly changing. Rather than lamenting these unstable conditions, new entrepreneurs need to be mindful of the opportunities they offer. Media editors need to be constantly monitoring trends and incorporating them at a reasonable pace in order to provide the public with the best services possible. This means naturally embracing the kind of user-centred attitude that more and more media outlets are adopting. This article is a modest attempt to support that endeavour.

5. References

- Alves, Kellyanne; Díaz-Noci, Javier** (2019). "Informatius de televisió i audiència activa: Estudi de cas de TV3-Televisió de Catalunya". *Comunicació*, n. 36, n. 1, pp. 9-29.
<https://doi.org/10.2436/20.3008.01.176>
- Amazon (2022). *Alexa skills kit*.
<https://amzn.to/3KsLmMS>
- Barr, Alistair** (2014). "Google to shut Google News in Spain in response to new law". *The Wall Street Journal*, 10 December.
<https://www.wsj.com/articles/google-shutting-google-news-in-spain-1418265199>
- Bonelli, Sherry** (2017). "How to optimize for voice search". *Search engine land*, May 1.
<https://searchengineland.com/optimize-voice-search-273849>
- Cachón, María-José** (2019). "¡Esto es La Voz!". *Sistrix toolbox*, 27 abril.
<https://es.slideshare.net/sistrixES/esto-es-la-voz-mjcatchon-en-ensaladaseo>
- Carlson, Matt** (2007). "Order versus access: News search engines and the challenge to traditional journalistic roles". *Media, culture & society*, v. 29, n. 6, pp. 1014-1030.
<https://doi.org/10.1177/0163443707084346>
- Charlton, Graham** (2016). "How much do journalists and editors need to know about SEO?". *Search engine watch*, March 15.
<https://bit.ly/1UIOyp5>
- Checa, Esther** (2020). "Qué oportunidad tienen las marcas si quieren optimizar para voz". *Vamos Talegon*, 28 abril.
<https://www.slideshare.net/EstherCheca/que-oportunidad-tienen-las-marcas-si-quieren-optimizar-para-voz-234779052>
- Codedesign (2023). *Voice search optimization. A complete guide in 2023*.
<https://codedesign.org/voice-search-optimization-complete-guide>
- Coller, Xavier** (2000). *Estudio de caso. Colección cuadernos metodológicos*. Madrid: Centro de Investigaciones Sociológicas.
<https://bit.ly/2NjqLPZ>
- Dick, Murray** (2011). "Search engine optimization in UK news production". *Journalism practice*, v. 5, n. 4, pp. 462-477.
<https://doi.org/10.1080/17512786.2010.551020>
- Edell, Patrick** (2023). "How Siri, Alexa and Google Assistant lost the A.I. race". *The New York Times*, 16 March.
<https://www.nytimes.com/2023/03/15/technology/siri-alexa-google-assistant-artificial-intelligence.html>
- Fagundes-Pase, André; Noll, Giselle; Gomes-Da Fontoura, Mariana; Dallegrave, Leticia** (2020). "Who controls the voice? The journalistic use and the informational domain in vocal transactors". *Brazilian journalism research*, v. 16, n. 3.
<https://doi.org/10.25200/BJR.v16n3.2021.1316>
- García-Carretero, Lucía; Codina, Lluís; Díaz-Noci, Javier; Iglesias-García, Mar** (2016). "Herramientas e indicadores SEO: características y aplicación para análisis de cybermedios". *El profesional de la información*, v. 25, n. 3, pp. 497-504.
<https://doi.org/10.3145/epi.2016.may.19>
- Gareth, Brian** (2022). "Voice search: 10 ways to optimize your site for future SEO success". *Search engine journal*, May 16.
<https://www.searchenginejournal.com/voice-search-optimization-seo-webceo/446158>
- Giomelakis, Dimitrios; Veglis, Andreas** (2015). "Investigating search engine optimization factors in media websites. The case of Greece". *Digital journalism*, v. 4, n. 3, pp. 379-400.
<https://doi.org/10.1080/21670811.2015.1046992>

- Gonzalo-Penela, Carlos; Codina, Lluís; Rovira, Cristòfol** (2015). "Recuperación de información centrada en el usuario y SEO: Categorización y determinación de las intenciones de búsqueda en la Web". *Index.Comunicación*, v. 5, n. 3, pp. 19-27. <https://indexcomunicacion.es//index.php/indexcomunicacion/article/view/197>
- Google (2018). *Speakable beta*. <https://bit.ly/3LLOvrg>
- Google (2021). *Search quality rating guidelines*. <https://bit.ly/3LlmdOc>
- Google (2022). *Actions on Google*. <https://bit.ly/3v1Rrcw>
- Guallar, Javier** (2015). "Prensa digital en 2013-2014". *Anuario ThinkEPI*, v. 9, pp. 153-160. <https://doi.org/10.3145/thinkepi.2015.37>
- Guy, Ido** (2018). "The characteristics of voice search: Comparing spoken with typed-in mobile web search queries". *Transactions on information systems*, v. 36, n. 3. <https://doi.org/10.1145/3182163>
- Huffman, Scott** (2014). "OMG! Mobile voice survey reveals teens love to talk". *Google*, October 14. <https://blog.google/products/search/omg-mobile-voice-survey-reveals-teens>
- Isidigitaldata* (2020). *El uso de la tecnología de voz en España*. <https://bit.ly/3jkSjd6>
- Jones, Kristopher** (2019). "How SEOs can master voice search now". *Search engine land*, 22 April. <https://bit.ly/3umAWsw>
- Jung, Hyunhoon; Hwang, Changhoon; Hwang, Gilhwan; Yoonjung, Cindy; Lee, Joonhwan; Suh, Bongwon** (2019). "Tell me more: Understanding user interaction of smart speaker news powered by conversational search". In: *CHI Conference on human factors in computing systems extended abstracts (CHI'19 Extended Abstracts)*, May 4-9, Glasgow, Scotland Uk. ACM, New York, NY, USA. <https://doi.org/10.1145/3290607.3312979>
- Kischinhevsky, Marcelo** (2019). "De las síntesis informativas a los resúmenes para altavoces inteligentes, desafíos al periodismo radiofónico de carácter local". *Estudios sobre el mensaje periodístico*, v. 26, n. 1, pp. 167-175. <https://doi.org/10.5209/esmp.67296>
- Lee, Angela M.; Chyi, Hsiang-Iris** (2015). "The rise of online news aggregators: Consumption and competition". *International journal on media management*, v. 17, n. 1, pp. 3-24. <https://doi.org/10.1080/14241277.2014.997383>
- Lochrie, Mark; De-Neef, Robin; Mills, John; Davenport, Jack** (2018). "Designing immersive audio experiences for news and information in the Internet of things using text-to-speech objects". In: *Proceedings of the 32nd International BCS Human computer interaction conference (HCI)*. <https://www.scienceopen.com/hosted-document?doi=10.14236/ewic/HCI2018.90>
- Lopezosa, Carlos; Codina, Lluís; Díaz-Noci, Javier; Ontalba, José-Antonio** (2020). "SEO and the digital news media: From the workplace to the classroom. [SEO y cibermedios: De la empresa a las aulas]". *Comunicar*, v. 63, pp. 65-75. <https://doi.org/10.3916/C63-2020-06>
- Lopezosa, Carlos; Trillo-Domínguez, Magdalena; Codina, Lluís; Cabrera-Méndez, Marga** (2021). "SEO in the journalistic company: perceptions and key elements for its adoption in writing". *Revista latina de comunicación social*, n. 79, pp. 27-45. <https://www.doi.org/10.4185/RLCS-2020-1487>
- Lovato, Silvia B.; Piper, Anne-Marie** (2019). "Young children and voice search: What we know from human-computer interaction". *Frontiers in psychology*, v. 10, n. 8. <https://doi.org/10.3389/fpsyg.2019.00008>
- Machill, Marcel; Beiler, Markus; Zenke, Martin** (2008). "Search-engine research: A European-American overview and systematization of an interdisciplinary and international research field". *Media, culture & society*, v. 30, n. 5, pp. 591-608. <https://doi.org/10.1177/0163443708094010>
- Mairesse, François; Raccuglia, Paul; Vitaladevuni, Shiv** (2016). "Search-based evaluation from truth transcripts for voice search applications". In: *Proceedings of the 39th International ACM SIGIR conference on research and development in information retrieval*, pp. 985-988. <https://doi.org/10.1145/2911451.2914735>
- Makhyan, Ludwig** (2022). "Voice search optimization at scale: A guide for enterprise marketers". *Search engine journal*, March 24. <https://www.searchenginejournal.com/voice-search-optimization-guide/437467>

- Marcos-Recio, Juan-Carlos; Sánchez-Vigil, Juan-Miguel; Olivera-Zaldua, María** (2015). "Google News y el impacto de la Ley de propiedad intelectual en la prensa: un nuevo amanecer para la información". *Documentación de las ciencias de la información*, v. 38, pp. 67-81.
https://doi.org/10.5209/rev_DCIN.2015.v38.50809
- Mehdi, Yusuf** (2023). "Reinventing search with a new AI-powered *Microsoft Bing* and *Edge*, your copilot for the web". *Official Microsoft blog*, February 7.
<https://blogs.microsoft.com/blog/2023/02/07/reinventing-search-with-a-new-ai-powered-microsoft-bing-and-edge-your-copilot-for-the-web>
- Muhr, Thomas** (1991). "Atlas/ti. A prototype for the support of text interpretation". *Qualitative sociology*, v. 14, n. 4, pp. 349-371.
<https://link.springer.com/article/10.1007/BF00989645>
- Newman, Nic** (2018). *The future of voice and the implications for news*. Digital News Project. Reuters Institute; University of Oxford.
<https://bit.ly/3LOJ1fj>
- Paulus, Trena M.; Lester, Jessica-Nina** (2015). "Atlas.ti for conversation and discourse analysis studies". *International journal of social research methodology*, v. 19, n. 4, pp. 405-428.
<https://doi.org/10.1080/13645579.2015.1021949>
- Pedrosa, Leyberson; De-Morais, Osvando-José** (2021). "Visibilidad web en buscadores. Factores algorítmicos de SEO on page (FAOP) como técnica e práctica periodística". *Estudios sobre el mensaje periodístico*, v. 27, n. 2, pp. 579-591.
<https://doi.org/10.5209/esmp.71291>
- Pichai, Sundar** (2023). "An important next step on our AI journey". *Google. The keyword*, February 6.
<https://blog.google/technology/ai/bard-google-ai-search-updates>
- Rabit & Pork** (2021). *Voice search ranking report*.
<https://bit.ly/38BRynl>
- Sa, Ning** (2016). "Improving query reformulation in voice search system". *CHIIR '16: Proceedings of the 2016 ACM on Conference on human information interaction and retrieval*, pp. 365-367. ISBN: 978 1 4503 3751
<https://doi.org/10.1145/2854946.2854951>
- Schema.org** (2018). *Schema Speakable*.
<https://schema.org/speakable>
- Schultheiß, Sebastian; Lewandowski, Dirk** (2020). "'Outside the industry, nobody knows what we do' SEO as seen by search engine optimizers and content providers". *Journal of documentation*, v. 77, n. 2, pp. 542-557.
<https://doi.org/10.1108/JD-07-2020-0127>
- Shokouhi, Milad; Ozertem, Umut; Craswell, Nick** (2016). "Did you say U2 or *YouTube*?: Inferring implicit transcripts from voice search logs". In: *WWW'16: Proceedings of the 25th international conference on World Wide Web*, pp. 1215-1224. ISBN: 978 1 45034143
<https://doi.org/10.1145/2872427.2882994>
- Smyrnaio, Nikos; Rebillard, Franck** (2009). "L'actualité selon *Google*. L'emprise du principal moteur de recherche sur l'information en ligne". *Communication et langages*, n. 160, pp. 95-109.
<https://doi.org/10.4074/S0336150009002087>
- Smyrnaio, Nikos; Sire, Guillaume** (2014). "The news according to *Google* how does algorithmic infomedia frame the work of French journalists?". In: *JSS-Ecrea 2014 Conference*, Thessaloniki.
<https://bit.ly/33bA5tN>
- Spencer, Stephan** (2017). "How to rank for 'position 0' in 3 simple steps: A featured snippets primer". *Search engine land*, 30 April.
<https://bit.ly/3rwzdiD>
- Sterling, Greg** (2016). "Google says 20 percent of mobile queries are voice searches. Voice search growing as virtual assistant market heats up". *Search engine land*, May 18.
<https://bit.ly/3rbBwY2>
- Sulé-Duesa, Andreu** (2015). "Schema.org, la millora de la visualització dels resultats en els cercadors i molt més". *BiD: textos universitaris de biblioteconomia i documentació*, v. 34.
<https://doi.org/10.1344/BiD2015.34.23>
- Turow, Joseph** (2021). "Journalism and the voice intelligence industry". *Digital journalism*, v. 9, n. 7, pp. 1000-1006.
<https://doi.org/10.1080/21670811.2020.1829979>
- Valles, Miguel S.** (2002). *Entrevistas cualitativas*. Colección cuadernos metodológicos. Madrid: Centro de Investigaciones Sociológicas. ISBN: 978 84 74763423

Proposal for an index measuring the reputation of open data portals: The *Odapri*

Marta Ortiz-de-Urbina-Criado; Alberto Abella; Carmen De-Pablos-Heredero

Nota: Este artículo se puede leer en español en:
<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87151>

Recommended citation:

Ortiz-de-Urbina-Criado, Marta; Abella, Alberto; De-Pablos-Heredero, Carmen (2023). "Proposal for an index measuring the reputation of open data portals: The *Odapri*". *Profesional de la información*, v. 32, n. 3, e320312.
<https://doi.org/10.3145/epi.2023.may.12>

Manuscript received on 25th October 2022
Accepted on 13th April 2023



Marta Ortiz-de-Urbina-Criado ✉

<https://orcid.org/0000-0001-7527-6798>

Universidad Rey Juan Carlos
Facultad de Ciencias de la Economía y de la Empresa
Paseo de los Artilleros, s/n
28032 Madrid, Spain
marta.ortizdeurbina@urjc.es



Alberto Abella

<https://orcid.org/0000-0003-0243-4606>

Fiware Foundation
Franklinstrasse 13A
10587 Berlin, Germany
alberto.abella@fiware.org



Carmen De-Pablos-Heredero

<https://orcid.org/0000-0003-0457-3730>

Universidad Rey Juan Carlos
Facultad de Ciencias de la Economía y de la Empresa
Paseo de los Artilleros, s/n
28032 Madrid, Spain
carmen.depablos@urjc.es

Abstract

The demand for open data has led to the creation and the availability of numerous portals releasing data. However, a large percentage of them are not properly designed for professional use. One way to analyze the value that a portal delivers is through its reputation, but this is a concept that has not yet been well defined or measured. This work will focus on the reputation of open data portals with two objectives. The first objective is to delimit the concept of reputation for open data portals, therefore leading to the first proposed research question: How can the reputation of open data portals be defined? The second objective is to propose criteria for measuring reputation and to create a reputation index for open data portals, with the second research question: How can the reputation of open data portals be assessed? This work is conceptual and descriptive and proposes a multidimensional definition of this concept, which includes whether it is known, being known for something, and its generalized favorability. In addition, a proposal is made for a reputation index identifying its dimensions and measurements, and finally, an analysis of its usefulness is presented. To this end, a figure has been created that summarizes the dimensions and benefits for creators and developers of open data portals and for the different application sectors (public, private, academic, and third sector). This research can help public administrations and other sectors to have a reference of good practices when offering open data to citizens and organizations oriented towards creating value in society.

Keywords

Open data; Open data portals; Reputation; Indexes; Indicators; Metric; Public administration; Rankings; *Odapri*; Value creation; *Meloda*.



Funding

This work was supported by

- *Ministry of Science, Innovation and Universities* (Spain) [grant number RTI2018-097447-B-I00];
- Bridge project [grant number V-1066] of *Rey Juan Carlos University (URJC)*, Madrid;
- “Strategor” and “OpenInnova” excellent research groups of *URJC*.

1. Introduction

“Data matters” as proclaimed by **Zhang, Sun and Zhang** (2022). In that context, public interest in accessing open information on topics of general interest, such as the use of public budgets, economic and social development projects, and demographic information has led to the popularization and development of open government initiatives in recent years (**Oviedo; Mazón; Zubcoff**, 2015; **Reggi; Dawes**, 2022; **Zhao; Fan**, 2021). As an aspect of open government, the publication of open government data is a rapidly expanding phenomenon, motivated by the desire to democratize data access and knowledge production (**Aguilera et al.**, 2017; **Zhenbin et al.**, 2020), increase transparency (**Bisogno; Cuadrado-Balasteros; Santis**, 2022), and improve the social and economic dimensions of cities (**Hu; Zheng**, 2021; **Jetzek; Avital; Bjorn-Andersen**, 2019).

The most widely used tools providing access to open data are open data portals. These portals allow data to be published from a wide range of information sources, both formal and structured as well as informal. Sabri, Emran and Harum explain that

“open data portals serve as big data sources that support data-based discoveries and software applications development. Thus, open data are characterized by the properties of volume, variety, speed, and veracity of big data” (**Sabri; Emran; Harum**, 2019).

In this regard, it has been observed that the quality of veracity is difficult to measure (**Sabri; Emran; Harum**, 2019). This technical difficulty, together with the fact that most open data portals useful today come from public administration initiatives (**Quarati; De-Martino; Rosim**, 2021), has led to an increased interest in the definition of aspects such as the reputation of the published data together with their publisher. This can help users and re-users to be able to trust available data and appreciate more public administration services.

For data to be consulted and reused, they should be of good quality. Several studies focus on assessing the quality of open data (**Abella; Ortiz-de-Urbina-Criado; De-Pablos-Heredero**, 2014; 2019a; 2022; **Gil-García; Dawes; Pardo**, 2018; **Lnenicka; Nikiforova**, 2021; **Neumaier; Umbrich; Polleres**, 2016; **Oviedo; Mazón; Zubcoff**, 2015; **Sadiq; Indulska**, 2017) and propose criteria for evaluating the quality of open data and the portals that publish them (**Abella; Ortiz-de-Urbina-Criado; De-Pablos-Heredero**, 2018; **Fernández-Ardèvol; Rosales**, 2022; **Kubler et al.**, 2018; **Moghadami; Malekol-kalami**, 2022; **Zuiderwijk; Pirannejad; Susha**, 2021). **Kubler et al.** (2018) present a model for monitoring and evaluating the quality of open data portals on the basis of a set of open data quality indicators. The authors emphasize that the quality of the data impacts the publisher organization’s reputation, but in their model, they only analyze the status and quality of the metadata. To refer to metadata quality is a typical approach when dealing with both the quality of open government data (OGD) and the quality of OGD portals. However, it is limited in nature and does not provide insights regarding the actual quality. One of the limitations that **Kubler et al.** (2018) highlight is that the quality indicators considered are not enough to demonstrate the use of a set of data (for example, a data publisher or consumer might be interested in knowing to what extent a dataset is used by third parties), which relates to “reputation” or “participation and collaboration” metrics.

Some metrics already consider open data reputation among their criteria. **Oviedo, Mazón and Zubcoff** (2015) include the criterion “reputation of the published data” in their model, which is defined as the degree of credibility of the portal that publishes the data. Moreover, **Abella, Ortiz-de-Urbina-Criado and De-Pablos-Heredero** (2019a) have developed version 5 of *Meloda*, a metric that analyzes the degree of reusability of open data, and for which they consulted a panel of experts who suggested including the reputation criteria. *Meloda* has already been applied to open data portals constructed by public administrations (**Abella; Ortiz-de-Urbina-Criado; De-Pablos-Heredero**, 2019a). A way of measuring reputation is proposed in the first study in which this metric is applied, **Abella et al.** (2022), which presents a reputation ranking of open data publishers wherein a set of portals is evaluated through the opinions of the managers of other open data portals. The authors highlight the interest in having a reputation ranking for open data portals, and propose the development of more robust, complete, and objective measures of the reputation of open data portals as a future line of research.

Thus, in the context of open data, it is necessary to consider how we might define the reputation of open data portals. However, the previous literature has not yet defined this concept, and it is therefore necessary to make an approximation to the concept of reputation applied in the case of open data portals. This is very important for public administrations, as they dedicate public resources to the creation and maintenance of open data portals (**Alzamil; Vasarhelyi**, 2019), and they are aware of the impact that citizen evaluations of their services have on their own reputations (**Grossi; Meijer; Sargiacomo**, 2020).

It is necessary to further analyze how the reputation of open data portals might be measured. Although work has been carried out regarding open government reputation (Veljković; Bogdanović-Dinić; Stoimenov, 2014) and on the intensity of open data publication in smart cities (Prieto; Mazón; Lozano-Tello, 2019), we have not found previous research that studies the creation of reputational assessments while taking into account the multidisciplinary nature of this concept as well as of the agents of the open data re-users' ecosystem. Consequently, it is useful to propose a reputation index for open data and assess its usefulness.

The goal of this paper is to study the reputation of open data portals to define the concept and then develop an index to measure it.

- The first objective is to delimit the concept of reputation for open data portals (*Odapre*), and therefore the first research question is proposed: How can the reputation of open data portals be defined?
- The second objective is to recommend criteria for measuring reputation and to create a reputation index for open data portals (*Odapri*), leading to the second proposed research question: How can the reputation of open data portals be assessed?

Following the introduction, the second section presents open data portals. The third section provides a definition of the concept of reputation in terms of open data portals (*Odapre*) and their dimensions. In the fourth section, a proposal of a reputation index (*Odapri*) and its benefits are explained. Finally, the conclusions section presents the theoretical and practical contributions of this work, as well as its limitations, and proposes future lines of research.

This work is conceptual and descriptive, and is supported by the discussion of literature regarding the need to create a reputation index, as well as its proposed usefulness. A literature search on the concept of "open data portal reputation" has been carried out in the *Web of Science* and *Scopus* databases. No clear definition or delimitation of this concept was found, and no specific metrics have been developed to evaluate it. For this reason, the authors have carried out a process of brainstorming and reflection on the basis of their experience in the evaluation of open data portals. In a previous study, a first proposal was made to measure reputation and a questionnaire was sent to the heads of open data portals in Spain. On the basis of the results obtained and their comments, the need to develop a multidimensional concept and a more complete metric able to analyze the evaluation of the reputation of open data portals from a more complete and comprehensive perspective was observed, which is the main contribution made in this paper. The results of this study are of interest for public administrations and other sectors (private, academic, and third party), as it is imperative they have a metric that can help to identify potential issues and develop active policies and practices oriented toward improving the quality of their open data portals.

2. Open data portals

Open data portals are the main tools used to publish open data. Open data portals often hold a significant and valuable amount of data with the potential to directly impact citizens (Barcellos; Bernardini; Viterbo, 2022).

The utility of open data portals lies not only in their ability to provide reliable data, but also in the fact that these data can be reused to create value. Cetina (2021) calls them "data with purpose" and points out the importance of identifying, before designing open data portals, the most urgent challenges that open data can help solve. When data are open, they can be reused to create value. Ferrer-Sapena *et al.* point out that, in an open government, information must be well structured enough

"so that the effort of the public, through appropriate re-use, can make it possible to improve services for citizens and at the same time, be a source of wealth creation and modernization of the public sector" (Ferrer-Sapena *et al.*, 2020, p. 6).

The lack of quality in information published by organizations is a barrier to the feasibility of the reuse process (Cadena-Vela, 2019). According to Cadena-Vela, data are created, stored, processed, exchanged, shared, added to, and reused as long as they are useful and the value obtained from them meets expectations. In addition, to ensure data quality in an organization, Cadena-Vela (2019) proposes five steps:

- consider the needs of users;
- define a clear data generation process;
- identify the data life cycle;
- appoint a data manager; and
- evaluate the quality of the data product.

Creators and publishers face several challenges, such as making published information useful for developing new products and services and allowing for beneficial interaction between users and citizens. They also face the need to make public/private service delivery more efficient, effective, and democratic, and must also take into consideration the restoration of trust and satisfaction with government and data openness policies (Barcellos; Bernardini; Viterbo, 2022). Therefore, one of the issues that has been given more attention in the literature is the quality of open data (Vetrò *et al.*, 2016) and the development of models and metrics to assess it (Park; Gil-García, 2022).

“ A definition of open data portal reputation is proposed ”

Best governance practices enable citizens, businesses, and entrepreneurs to participate as key stakeholders in open data projects (Zuiderwijk; Pirannejad; Sussha, 2021), which makes having reliable metrics to assess the quality of the data published in open data portals very important for reuse.

3. Reputation of open data portals: concept and dimensions

There has been little exploration about the use of the reputation of open data portals to improve the satisfaction of users and citizens. For Taylor *et al.* (2018), trust and reputation allow agents to make informed decisions about possible interactions. Trust in an agent derives from direct experience with that agent, while reputation is determined by experiences related by other witnesses with potentially different points of view. Reputation systems often include the collection, dissemination, and aggregation of actors' experiences (Hoelz; Ralha, 2015).

The quality of the data and their description have an impact not only on the reputation of the organization that publishes the data, but also on the decision-making and business income that can be generated from open data (Kubler *et al.*, 2018). Kubler *et al.* (2018) present the *Open Data Portal Watch* model, which allows the monitoring and evaluation of quality in open data portals on the basis of a set of open data quality indicators. This model only analyzes the status and quality of metadata, providing quality indicators for the applications that use them, but the proposed set of indicators are not enough to capture the idea of "reputation" or "participation and collaboration" (Kubler *et al.*, 2018).

Other metrics have considered the reputation of data among their criteria. Oviedo, Mazón and Zubcoff (2015) define the "reputation of published data" as the degree of credibility of the portal where the published data are located. In their analysis, they raise three questions:

- 1) Is it clear from which sources the information being published on the portal was obtained, so that these sources can be verified?
- 2) Do the sources from which the information published on the portal is derived carry prestige within your environment?
- 3) Is the original source respected, or is there any manipulation that might call into question the degree of credibility of the published data?

Abella, Ortiz-de-Urbina-Criado and De-Pablos-Heredero (2019a) include reputation in the latest version of the *Meloda* metric (version 5), in which they consulted a panel of experts who suggested the inclusion of reputation. Reputation, in this metric, has three levels:

- there is no information about the reputation of the data portal;
- there are statistics or reports published about the opinions of users; and
- there are indicators or classifications about the reputation of the data portal.

The first study in which this metric is applied is conducted by Abella *et al.* (2022) and presents a reputation ranking of open data publishers in which the portals analyzed in the study are valued through the opinion of the managers of open data portals. To accomplish this task, a survey is filled out by the managers of open data portals regarding the reputation of the other open data portals. The sample portals were divided into groups of ten, and each group was assigned a set of portals, ensuring that no portal could vote for itself. The survey asked about the knowledge and prestige of other open data portals, and the final reputation level was obtained from the voters' most frequent value over the reputation of other open data portals, weighted by the knowledge of the portal by respondent as claimed in the survey.

To shed light on the concept of reputation of open data portals, we will consider not only the idea of reputation of the data, but also that of the entity that publishes open data, thus alluding to the concept of organizational reputation. The reputation of an organization, together with any changes in its reputation, influences an organization's relationships with its stakeholders. Multiple multidimensional definitions of this concept have been proposed in the literature. Gwebu, Wang and Wang (2018) define a company's reputation as an intangible asset based on the collective recognition of the company's demonstrated ability to offer stakeholders quality and value in relation to their peers (Schultz; Mouritsen; Gabrielsen, 2001). Organizations build a good reputation by systematically demonstrating quality and participating in quality behaviors (Gardberg; Fombrun, 2002). Reputation diagnosis is a stable and credible signal that is difficult to discredit (Roberts; Dowling, 2002).

Lange, Lee and Dai (2011) review the literature on organizational reputation, and from the definitions of reputation propose three fundamental dimensions to understand organizational reputation:

- being known: widespread knowledge or visibility of the company; prominence of the company in collective perception (Barnett; Jermier; Lafferty, 2006; Rindova *et al.*, 2005);
- being known for something: perceived predictability of organizational outcomes and behavior relevant to a specific audience's interests (Fischer; Reuber, 2007; Love; Kraatz, 2009; Rindova *et al.*, 2005); and
- generalized favorability: perceptions or judgments of the organization in general as good, attractive, and appropriate (Barnett; Jermier; Lafferty, 2006).

The dimension of being known refers to the degree to which the group of beneficiaries maintains a strong, lasting, and non-evaluative image of the organization. This is the extent to which all agents share a deep and common knowledge of

the organization (Lange; Lee; Dai, 2011). The dimension of being known for something is the result of the assessments of the group of people who perceive it in terms of

How can the reputation of open data portals be assessed?

the probability that the company meets, or does not meet, its specific needs. Within a given set of recipients, there may be different competing, and perhaps conflicting, needs. Therefore, it is possible that within a defined set of beneficiaries, a company may have more than one significant reputation in the dimension of being known for something. Rindova *et al.* refer to the dimension of being known for something of organizational reputation as “perceived quality,” that is,

“the extent to which stakeholders positively assess an organization on a specific attribute, such as the ability to produce quality products” (Rindova *et al.*, 2005, p. 1035).

Love and Kraatz (2009, p. 317) label this dimension of reputation as “technical effectiveness” and describe it as the public’s assessment of the company’s ability to meet the material needs of the public, which means that the organizational reputation is “closely linked to the consequences and tangible results of the organization.” Pfarrer, Pollock and Rindova (2010) refer to this dimension of organizational reputation as the result of judgments regarding “the demonstrated ability of the company to create value.”

The dimension of generalized favorability, although it also entails judgments of the beneficiary, refers to the evaluation of the organization as an aggregate whole, rather than through assessments of its ability to deliver certain outcomes that meet the specific needs of recipients:

“the esteem, the consideration of the enterprise and how attractive the enterprise is” or “the public assessment of a company in relation to other companies” (Lange; Lee; Dai, 2011).

All dimensions depend largely on the audience that researchers or professionals decide to investigate (Lange; Lee; Dai, 2011).

Additionally, to define and measure the reputation of open data portals, the perspectives of the actors involved in the open data ecosystem must be considered. Following the proposal of Abella, Ortiz-de-Urbina-Criado and De-Pablos-Heredero (2019b), we find the following agents.

- One group consists of the data publishers, which are the organizations that release data: they are mostly public organizations, and the organizations related to publishers are, in the majority, other public organizations that depend on data publishers.
- The second group consists of direct re-users of open data, who provide products/services for others, whether they be professional re-users such as for-profit entities that use data to create innovative products/services offered to others, or social re-users, such as not-for-profit organizations that provide services to others.
- The third group consists of end users, including professional users who are for-profit entities consuming products/services based on open data and/or academic users, who directly access published open data or use direct reuse services.

On the basis of these ideas, we contend that the reputation of open data portals is the collective recognition of the ability demonstrated by the portal to systematically offer reusable open data, allowing the creation of value from them. This concept can be understood multidimensionally in the three dimensions proposed by Lange, Lee and Dai (2011):

- the degree to which it is “known” (dissemination and knowledge of the data portal),
- the means in which it is “known for something” (for example, by their degree of maturity, by their datasets, by the services developed by their data, or by their innovation), and
- the generalized favorability (opinion of the re-users of the ecosystem).

Therefore, reputation is determined by the experiences reported by other ecosystem actors with potentially different views, and their analysis requires the collection, dissemination, and aggregation of the knowledge and experiences of agents. Considering this precedent, we will define the reputation concept of open data portals as the degree to which a portal is known, is known for something, and is valued in the opinion of ecosystem agents who reuse their data. Table 1 presents the dimensions of recognition, its definition, and its inspiration from academic literature.

Table 1. Dimensions of the reputation of open data portals

Dimension	Definition	Sources of origin of the dimension
Known portal	It is known to exist. It is distributed by agents with potential interest	Rindova <i>et al.</i> (2005)
Portal known for something	It has characteristics that make it known for some specific attribute	Deutsch & Ross (2003); Carter (2006); Dimov, Shepherd & Sutcliffe (2007)
Valued portal	It presents favorable opinions of those who use the portal data	Barnett, Jermier & Lafferty (2006); Fischer & Reuber (2007); Highhouse <i>et al.</i> (2009)

4. Open data portal’s reputation assessment

Here, a proposal for building a reputation index for open data portals is presented on the basis of the definition previously given for the reputation of open data portals.

4.1. The Open data portal reputation index: Odapri

To establish criteria to assess the reputation of open data, we will consider the two elements of analysis of an open data portal: the creators/developers and the dataset. We will also consider the three dimensions of reputation proposed in the definition, that is, the degree to which “it is known,” how it is “known for something,” and the general assessment (generalized favorability).

Therefore, our measurement will consider the assessment of the following players in the open data re-user ecosystem: the data publishing organization itself, other public and private organizations publishing data on open data portals, professional data re-users, and academic data users.

Considering these preliminary assumptions, the analysis of the reputation of an open data portal requires not only that portal managers analyze knowledge about it, but also that experts in the reuse of open data should objectively assess the characteristics of the portal (degree of maturity), its datasets (degree of reuse), and the applications and services created from them (innovation and features of services and business models). The level of value of the open data portal, as perceived by both creators of other portals and the re-users of the published data, should also be taken into consideration.

Therefore, the following criteria are to be analyzed:

- degree of knowledge of the portal by the agents of the open data ecosystem;
- degree of maturity of the portal;
- degree of reuse of their datasets;
- products and services created from their datasets: if they have a section in which to leave comments, or if they are known by the creators of the portal;
- degree of innovation;
- opinion on the reputation and prestige of the creators and their portal developers (institution, company, etc.).

Table 2 presents dimensions, measurement indicators, and proposed items to measure the reputation of open data portals.

Table 2. Reputation dimensions and measurements of open data portals

Reputation dimension and weight	Open data ecosystem agent and weighting	Measurement indicator	Items
Being known (20%)	Open data portal managers (20%)	Degree of knowledge of the portal by agents of the open data ecosystem	<ol style="list-style-type: none"> 1. Indicate to what extent you know any entity that reuses the data published in its open data portal on a regular basis (Likert scale, 0-10) 2. Indicate to what extent different types of entities reuse your data on a regular basis (e.g., individual citizens, professional re-users, researchers, and academic staff including students; the data of the publishing organization itself, and other public entities) (Likert scale, 0-10) 3. Indicate to what extent your entity has carried out some type of activity to promote the use of your data in the last year (e.g., application competitions, own events, external events, meetings with re-users and others) (Likert scale, 0-10) 4. Indicate the extent to which you use the access statistics for your open data portal (Likert scale, 0-10)
Being known for something (20%)	Academic experts and researchers (20%)	Degree of maturity of the portal	Degree of maturity of the portal according to the metric of Abella; Ortiz-de-Urbina-Criado; De-Pablos-Heredero (2017) (Likert scale, 0-10)
		Degree of reusability of datasets	Degree of reusability according to the <i>Meloda 5</i> metric of Abella; Ortiz-de-Urbina-Criado; De-Pablos-Heredero (2019a) (Likert scale, 0-10)
		Products and services created from datasets	<ol style="list-style-type: none"> 1. To what extent your portal has a section identifying services and/or applications based on portal data (Likert scale, 0-10) 2. How many applications you have identified in your open data portal
		Degree of innovation	Indicate to what extent these types of innovations are produced by data reuse; product innovation; process innovation (Likert scale, 0-10)
General favorability (60%)	Managers of other portals (20%) Professional re-users (40%)	Opinion on the reputation and prestige of the creators and their portal developers	<ol style="list-style-type: none"> 1. Indicate the level of knowledge of the entity: I) high: you know the portal well and its features; you use it frequently and/or have used them; II) medium: you know the portal and its features; you have briefly used it and/or have ever used your data; III) low: did not know of its existence or visited it without using its data 2. Indicate the level of reputation of the entity that develops the portal: I) reference portal of the sector: data of the portal are a reference in the sector and/or appear as a recognized reference in rankings of the sector or in reputation reports; 2) mature portal: the data are known and quality is valued and/or user opinions are known; 3) incipient portal: no information on the reputation of the portal data

Once the items have been identified for each dimension, a quantitative assessment will be made for each one on a scale of 0-10. This way, it will be possible to give a quantitative assessment for each portal regarding the three aspects analyzed: whether it is known, whether it is known for something, and whether it is valued by the re-users. A weighted average of the combined three will be calculated after considering the measurement for each dimension. Each dimension has been weighted in accordance with the importance of each ecosystem agent. As the value of the data lies in its reuse, the most important agents are professional re-users, who have been assigned double importance (40%). The average value obtained will be the reputation index of each portal (scale 0-10). This value will allow for the development of a global reputation ranking of portals and the analysis of their positioning.

Results claim for offering open data to citizens and organizations oriented to create value in society

4.2. Potential benefits of the *Open data portal reputation index*

The main functions of the reputation indicator for open data portals are listed below.

First, it serves to evaluate the concept of an open data portal and its usefulness as an open data provider that can be reused.

Second, the indicator assists with the identification and dissemination of best practices in the development of open data portals.

Third, the reputation indicator helps to discourage the creation of only apparently open data portals and helps to improve the degree of maturity of the portals as well as the degree of reusability of the datasets they publish.

Fourth, the reputation indicator for open data portals helps to promote better open government development policies, favoring the fulfillment of the objectives proposed by **Bello-García (2017)**:

- more transparent government operations;
- creation of a better policy through greater participation;
- provision of better public services through greater collaboration; and
- unlocking the economic potential of public resources.

Fifth, the reputation indicator for open data portals can help to make policymakers aware of the impact of open data portals. When publishing and sharing open data, there are numerous factors and barriers (**Alcaide-Muñoz; Rodríguez-Bolívar; Villamayor-Arellano, 2022; Huang et al., 2020; Janssen; Charalabidis; Zuiderwijk, 2012; Nikiforova; Lnenicka, 2021; Zuiderwijk; De-Reuver, 2021**) that discourage public authorities from adopting more innovative and collaborative approaches, and that, among other consequences, lead to the creation of pretender open data portals (PODP) (**Abella; Ortiz-de-Urbina-Criado; De-Pablos-Heredero, 2022**). **Bello-García (2017)** identifies 15 barriers:

- lack of leadership and political commitment;
- inertia of the status quo;
- lack of financial means;
- lack of capacity and competencies at the individual and institutional levels;
- legal restrictions;
- lack of representativeness;
- multilingualism;
- lack of common standards and specifications (interoperability);
- perception of loss of control;
- uncertainty about sustainability issues and business models;
- legal gaps in knowledge regarding responsibilities;
- poor quality data;
- difficulties in identifying and creating demand from citizens and businesses;
- lack of confidence; and
- unrealistic or false expectations.

The last four in particular are closely related to issues with the reputation of publishers and data quality. Therefore, having a reputation indicator for portals can help to reduce these barriers, making both public institutions and other sectors aware of the importance of having useful open data portals.

Sixth, encouraging other sectors to open their data by having a measure of reputation is a good initiative, and is an example of good practices from public and private organizations in the publication and reuse of open data. In this sense, although open data have been more focused on the public sector (open government data, OGD), other sectors that are currently considered private sector, such as the academic sector or the so-called third sector, which are alternative or complementary sources, might

How useful is it to have a ranking based on the application of a reputation index for open data portals?

be encouraged to open their data. Each sector has different motivations, but their collaboration is essential to enriching information and achieving a more efficient production and reuse of data, resulting in greater social, economic, and scientific impact (Iglesias, 2020).

The private sector is one of the largest generators of data, but most of these data are not available as open data. However, some companies are already considering the advantages that some of their data can offer in this modality, for example, for research and/or social purposes. This can serve to improve their corporate reputation and attract new users, customers, employees, and investors (Iglesias, 2020). Data spaces provide the mechanisms for these data transactions to take place (Autolitano; Pawlowska, 2021).

IT User (2019) observes in its work that the exchange of data can benefit the public sector as a re-user, and the private sector as a publisher, equally. Open data can help improve decision-making and policymaking, helping to optimize public services and increase efficiency in both internal administrative processes and public service delivery. Private companies can also benefit from the openness of their data, as it will aid in achieving improvements in reputation, since open data will be considered as a corporate social responsibility action and an improvement in the perception of the brand as an employer. This can serve to create an open and transparent business image: one that shares information and helps to attract and retain talent. This can also serve to better the knowledge of the company, helping to solve questions that were not already answered internally, thus encouraging innovation. However, to be able to share their data, companies must overcome organizational, technical, and legal challenges.

The academic sector is another producer of data through scientific activity. The motivations of this sector to open their data are to facilitate open access to scientific results; to provide public domain data for the benefit of the community; to develop data laboratories (or data labs) as experimental environments to provide innovative and practical solutions to social problems; and to create entities such as Data Innovation Hubs to share data and be able to work collaboratively (Iglesias, 2020).

The third sector is a sector that is not directly covered either by the private or the public sector, and generally includes not-for-profit organizations that carry out concrete actions in social areas. The openness of data in this sector can help to improve transparency in its management as a best practice in the social and solidarity economy, as well as to facilitate multilateral collaboration (Iglesias, 2020).

Figure 1 shows the dimensions of *Odapri* and its benefits.

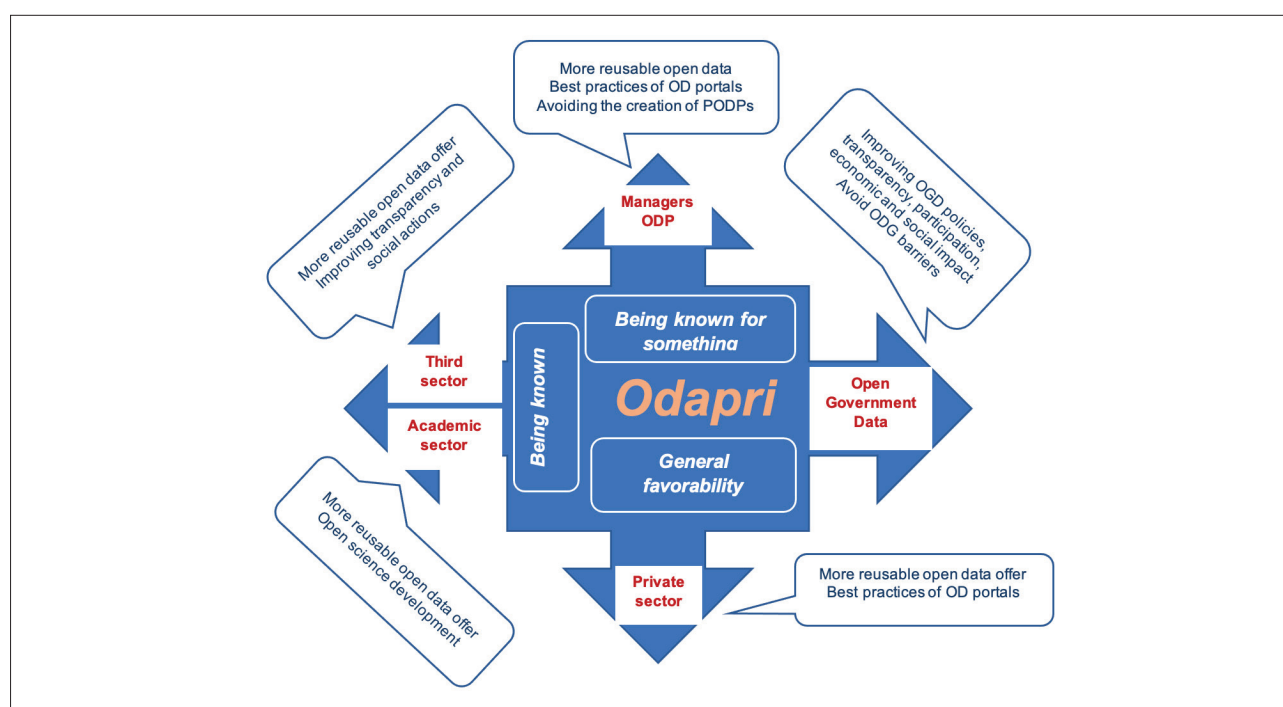


Figure 1. *Odapri*: Dimensions and benefits

5. Conclusions

This work has developed advances in the definition and assessment of the reputation of open data portals, providing answers to the three research questions posed. A definition of the reputation of open data has been proposed, and based on this definition, criteria have been proposed for the measurement of reputation and the creation of a reputation index. Finally, we have reflected on the usefulness that this reputation index can offer to each of the agents of the data re-user ecosystem, the different sectors that publish data, and for society in general.

Public Administrations can have a reference of quality and best practices

Open data have been useful in improving many areas, including job creation, public services, transparency between government and citizens, accountability, and citizen participation in government decision-making. In addition, by increasing government transparency, open data can serve to improve the level of trust of citizens, enhance the participation and collaboration of the public and private sectors, and encourage more innovative developments (Sabri; Emran; Harum, 2019).

Public administrations can analyze themselves according to this reputation index. This will help them to redesign the open data portals they have created according to the perceptions and expectations citizens and organizations have of them.

This work is conceptual and only presents a proposal for measuring open data reputation. Although some of the items proposed in the indicator have already been used in previous studies (Abella; Ortiz-de-Urbina-Criado; De-Pablos-Heredero, 2017; Abella *et al.*, 2019; 2022), further research is needed to refine these indicators and apply them to representative samples of actual data. Future research can validate the dimensions of *Odapri* as well as the weights and involved agents of *Odapri*. The validation of this index can be done through qualitative methodologies with the Delphi method. Quantitative methodologies can also be applied by sampling open data portals to measure their reputation. All of this can help in the creation of a reputation ranking of open data portals.

6. References

- Abella, Alberto; Ortiz-de-Urbina-Criado, Marta; De-Pablos-Heredero, Carmen** (2014). "Meloda, métrica para evaluar la reutilización de datos abiertos". *El profesional de la información*, v. 23, n. 6, pp. 582-588.
<https://doi.org/10.3145/epi.2014.nov.04>
- Abella, Alberto; Ortiz-de-Urbina-Criado, Marta; De-Pablos-Heredero, Carmen** (2017). *La reutilización de datos abiertos: una oportunidad para España*. Cotec. Report.
<https://cotec.es/proyecto/reutilizacion-de-datos-abiertos-una-oportunidad/ee68d249-51dd-4e53-8a3c-b6e32442ffc3>
- Abella, Alberto; Ortiz-de-Urbina-Criado, Marta; De-Pablos-Heredero, Carmen** (2018). "Indicadores de calidad de datos abiertos: el caso del portal de datos abiertos de Barcelona". *El profesional de la información*, v. 27, n. 2, pp. 375-383.
<https://doi.org/10.3145/epi.2018.mar.16>
- Abella, Alberto; Ortiz-de-Urbina-Criado, Marta; De-Pablos-Heredero, Carmen** (2019a). "Meloda 5: A metric to assess open data reusability". *El profesional de la información*, v. 28, n. 6, e280620.
<https://doi.org/10.3145/epi.2019.nov.20>
- Abella, Alberto; Ortiz-de-Urbina-Criado, Marta; De-Pablos-Heredero, Carmen** (2019b). "The process of open data publication and reuse". *Journal of the Association for Information Science and Technology*, v. 70, n. 3, pp. 296-300.
<https://doi.org/10.1002/asi.24116>
- Abella, Alberto; Ortiz-de-Urbina-Criado, Marta; De-Pablos-Heredero, Carmen** (2022). "Criteria for the identification of ineffective open data portals: pretender open data portals". *Profesional de la información*, v. 31, n. 1, e310111.
<https://doi.org/10.3145/epi.2022.ene.11>
- Abella, Alberto; Ortiz-de-Urbina-Criado, Marta; De-Pablos-Heredero, Carmen; García-Luna, Diego** (2022). *Reusing open data in Spain III*. Madrid: ESIC Editorial.
https://www.meloda.org/wp-content/uploads/2022/03/open_data_in_Spain_III_2021.pdf
- Abella, Alberto; Ortiz-de-Urbina-Criado, Marta; De-Pablos-Heredero, Carmen; Vidal-Cabo, Chistian; Ferrer-Sapena, Antonia** (2019). *La reutilización de datos abiertos en España II*. Barcelona: DesideDatum.
https://www.meloda.org/wp-content/uploads/2022/12/La_reutilizacion_datos_abiertos_en_espana_2019.pdf
- Aguilera, Unai; Peña, Óscar; Belmonte, Óscar; López-de-Ipiña, Diego** (2017). "Citizen-centric data services for smarter cities". *Future generation computer systems*, v. 76, pp. 234-247.
<https://doi.org/10.1016/j.future.2016.10.031>
- Alcaide-Muñoz, Laura; Rodríguez-Bolívar, Manuel-Pedro; Villamayor-Arellano, Cinthia L.** (2022). "Factors in the adoption of open government initiatives in Spanish local governments". *Government information quarterly*, v. 39, n. 4, 101743.
<https://doi.org/10.1016/j.giq.2022.101743>
- Alzamil, Zamil S.; Vasarhelyi, Micklos A.** (2019). "A new model for effective and efficient open government data". *International journal of disclosure and governance*, v. 16, n. 4, pp. 174-187.
<https://doi.org/10.1057/s41310-019-00066-w>
- Autolitano, Simona; Pawlowska, Agnieszka** (2021). *Europe's quest for digital sovereignty: Gaia-x as a case study*. IAI Papers, 21/14. ISBN: 978 88 9368 188 9
<https://www.iai.it/sites/default/files/iaip2114.pdf>

- Barcellos, Raisa; Bernardini, Flavia; Viterbo, José** (2022). "Towards defining data interpretability in open data portals: Challenges and research opportunities". *Information systems*, v. 106, 101961.
<https://doi.org/10.1016/j.is.2021.101961>
- Barnett, Michael L.; Jermier, John M.; Lafferty, Barbara A.** (2006). "Corporate reputation: the definitional landscape". *Corporate reputation review*, v. 9, pp. 26-38.
<https://doi.org/10.1057/palgrave.crr.1550012>
- Bello-García, Antonio** (2017). "Datos abiertos y participación en el gobierno social". *Economía industrial*, n. 405, pp. 99-111.
<https://www.mincotur.gob.es/Publicaciones/Publicacionesperiodicas/EconomiaIndustrial/RevistaEconomiaIndustrial/405/BELLO%20GARC%C3%8DA.pdf>
- Bisogno, Marco; Cuadrado-Ballesteros, Beatriz; Santis, Serena** (2022). "Do e-government initiatives and e-participation affect the level of budget transparency?". *International public management journal*, v. 25, n. 3, pp. 365-391.
<https://doi.org/10.1080/10967494.2022.2042437>
- Cadena-Vela, Susana** (2019). *Marco de referencia para la publicación de datos abiertos comprensibles basados en estándares de calidad*. Tesis doctoral. Universidad de Alicante.
<http://rua.ua.es/dspace/handle/10045/102507>
- Carter, Suzanne M.** (2006). "The interaction of top management group, stakeholder, and situational factors on certain corporate reputation management activities". *Journal of management studies*, v. 43, n. 5, pp. 1145-1176.
<https://doi.org/10.1111/j.1467-6486.2006.00632.x>
- Cetina, Camilo** (2021). *Gobernanza de datos y capacidades estatales para la pos-pandemia*. Policy Brief #28, Corporación Andina de Fomento, 2021. Caracas: CAF.
<https://scioteca.caf.com/handle/123456789/1765>
- Deutsch, Yuval; Ross, Thomas W.** (2003). "You are known by the directors you keep: Reputable directors as a signaling mechanism for young firms". *Management science*, v. 49, n. 8, pp. 1003-1017.
<http://www.jstor.org/stable/4133933>
- Dimov, Dimo; Shepherd, Dean A.; Sutcliffe, Kathleen M.** (2007). "Requisite expertise, firm reputation, and status in venture capital investment allocation decisions". *Journal of business venturing*, v. 22, n. 4, pp. 481-502.
<https://doi.org/10.1016/j.jbusvent.2006.05.001>
- Fernández-Ardèvol, Mireia; Rosales, Andrea** (2022). "Quality assessment and biases in reused data". *American behavioral scientist*, First online.
<https://doi.org/10.1177/00027642221144855>
- Ferrer-Sapena, Antonia; Calabuig, José M.; Sánchez-Pérez, Enrique A.; Vidal-Cabo, Christian** (2020). "Gobierno abierto y acceso a la información: un estudio de caso sobre el impacto en la economía local". *BiD: textos universitaris de biblioteconomia i documentació*, 45.
<https://doi.org/10.1344/BiD2020.45.17>
- Fischer, Eileen; Reuber, Rebecca** (2007). "The good, the bad, and the unfamiliar: the challenges of reputation formation facing new firms". *Entrepreneurship theory and practice*, v. 31, n. 1, pp. 53-75.
<https://doi.org/10.1111/j.1540-6520.2007.00163.x>
- Gardberg, Naomi A.; Fombrun, Charles J.** (2002). "The global reputation quotient project: First steps towards a cross-nationally valid measure of corporate reputation". *Corporate reputation review*, v. 4, pp. 303-307.
<https://doi.org/10.1057/palgrave.crr.1540151>
- Gil-García, José-Ramón; Dawes, Sharon S.; Pardo, Theresa A.** (2018). "Digital government and public management research: finding the crossroads". *Public management review*, v. 20, n. 5, pp. 633-646.
<https://doi.org/10.1080/14719037.2017.1327181>
- Grossi, Giuseppe; Meijer, Albert; Sargiacomo, Massimo** (2020). "A public management perspective on smart cities: 'Urban auditing' for management, governance and accountability". *Public management review*, v. 22, n. 5, pp. 633-647.
<https://doi.org/10.1080/14719037.2020.1733056>
- Gwebu, Kholekile L.; Wang, Jing; Wang, Li** (2018). "The role of corporate reputation and crisis response strategies in data breach management". *Journal of management information systems*, v. 35, n. 2, pp. 683-714.
<https://doi.org/10.1080/07421222.2018.1451962>
- Highhouse, Scott; Broadfoot, Alison; Yugo, Jennifer E.; Devendorf, Shelba A.** (2009). "Examining corporate reputation judgments with generalizability theory". *Journal of applied psychology*, v. 94, n. 3, pp. 782-789.
<https://doi.org/10.1037/a0013934>

- Hoelz, Bruno W. P.; Ralha, Célia G.** (2015). "Towards a cognitive meta-model for adaptive trust and reputation in open multi-agent systems". *Autonomous agents and multi-agent systems*, v. 29, n. 6, pp. 1125-1156.
<https://doi.org/10.1007/s10458-014-9278-9>
- Hu, Qian; Zheng, Yueping** (2021). "Smart city initiatives: a comparative study of American and Chinese cities". *Journal of urban affairs*, v. 43, n. 4, pp. 504-525.
<https://doi.org/10.1080/07352166.2019.1694413>
- Huang, Hsini; Zhou-Peng-Liao, Calvin; Liao, Hsin-Chung; Chen, Don-Yun** (2020). "Resisting by workarounds: unraveling the barriers of implementing open government data policy". *Government information quarterly*, v. 37, n. 4, 101495.
<https://doi.org/10.1016/j.giq.2020.101495>
- Iglesias, Carlos** (2020). *Datos abiertos más allá del sector público. Publicadores, motivaciones y modelos de colaboración*. Iniciativa Aporta del Ministerio de Asuntos Económicos y Transformación Digital, a través de la Entidad Pública Empresarial Red.es.
<https://datos.gob.es/es/documentacion/datos-abiertos-mas-alla-del-sector-publico>
- IT User* (2019). *Datos abiertos: base de nuevos servicios y beneficios para el ciudadano*. GMV Innovation Solutions.
<https://www.ituser.es/whitepapers/content-download/8a926896-a959-4955-a5c0-986ae3d717fa/especial-gmv-datos-abiertos-it-user-45.pdf>
- Janssen, Marijn; Charalabidis, Yannis; Zuiderwijk, Anneke** (2012). Benefits, adoption barriers and myths of open data and open government. *Information systems management*, v. 29, n. 4, pp. 258-268.
<https://doi.org/10.1080/10580530.2012.716740>
- Jetzek, Thorhildur; Avital, Michel; Bjorn-Andersen, Niels** (2019). "The sustainable value of open government data". *Journal of the Association for Information Systems*, v. 20, n. 6.
<https://doi.org/10.17705/1jais.00549>
- Kubler, Sylvain; Robert, Jérémie; Neumaier, Sebastian; Umbrich, Jürgen; Le-Traon, Yves** (2018). "Comparison of metadata quality in open data portals using the analytic hierarchy process". *Government information quarterly*, v. 35, n. 1, pp. 13-29.
<https://doi.org/10.1016/j.giq.2017.11.003>
- Lange, Donald; Lee, Peggy M.; Dai, Ye** (2011). "Organizational reputation: a review". *Journal of management*, v. 37, n. 1, pp. 153-184.
<https://doi.org/10.1177/0149206310390963>
- Lnenicka, Martin; Nikiforova, Anastasija** (2021). "Transparency-by-design: what is the role of open data portals?". *Telematics and informatics*, v. 61, 101605.
<https://doi.org/10.1016/j.tele.2021.101605>
- Love, E. Geoffrey; Kraatz, Matthew** (2009). "Character, conformity, or the bottom line? How and why downsizing affected corporate reputation". *Academy of management journal*, v. 52, n. 2, pp. 314-335.
<https://doi.org/10.5465/AMJ.2009.37308247>
- Moghadami, Maryam; Malekolkalami, Mila** (2022). "Evaluating the quality of open data portals in Iran". *International journal of knowledge processing studies*, v. 2, n. 1, pp. 43-53.
<https://doi.org/10.22034/kps.2022.145407>
- Neumaier, Sebastian; Umbrich, Jürgen; Polleres, Axel** (2016). "Automated quality assessment of metadata across open data portals". *Journal of data and information quality*, v. 8, n. 1.
<https://doi.org/10.1145/2964909>
- Nikiforova, Anastasija; Lnenicka, Martin** (2021). "A multi-perspective knowledge-driven approach for analysis of the demand side of the open government data portal". *Government information quarterly*, v. 38, n. 4, 101622.
<https://doi.org/10.1016/j.giq.2021.101622>
- Oviedo, Edgar; Mazón, José-Norberto; Zubcoff, José-Jacobo** (2015). "Quality and maturity model for open data portals". *2015 XLI Latin American computing conference (CLEI)*.
<https://doi.org/10.1109/CLEI.2015.7360011>
- Park, Sora; Gil-García, José-Ramón** (2022). "Open data innovation: visualizations and process redesign as a way to bridge the transparency-accountability gap". *Government information quarterly*, v. 39, n. 1, 101456.
<https://doi.org/10.1016/j.giq.2020.101456>
- Pfarrer, Michael D.; Pollock, Timothy G.; Rindova, Violina P.** (2010). "A tale of two assets: the effects of firm reputation and celebrity on earnings surprises and investors' reactions". *Academy of management journal*, v. 53, n. 5, pp. 1131-1152.
<https://doi.org/10.5465/AMJ.2010.54533222>

- Prieto, Álvaro E.; Mazón, José-Norberto; Lozano-Tello, Adolfo** (2019). "Framework for prioritization of open data publication: an application to smart cities". *IEEE transactions on emerging topics in computing*, v. 9, n. 1, pp. 131-143.
<https://doi.org/10.1109/tetc.2019.2893016>
- Quarati, Alfonso; De-Martino, Mónica; Rosim, Sergio** (2021). "Geospatial open data usage and metadata quality". *ISPRS international journal of geo-information*, v. 10, n. 1, 30.
<https://doi.org/10.3390/ijgi10010030>
- Reggi, Luigi; Dawes, Sharon S.** (2022). "Creating open government data ecosystems: network relations among governments, user communities, NGOs and the media". *Government information quarterly*, v. 39, n. 2, 101675.
<https://doi.org/10.1016/j.giq.2022.101675>
- Rindova, Violina P.; Williamson, Ian O.; Petkova, Antoaneta P.; Sever, Joy-Marie** (2005). "Being good or being known: an empirical examination of the dimensions, antecedents, and consequences of organizational reputation". *Academy of management journal*, v. 48, n. 6, pp. 1033-1049.
<https://doi.org/10.5465/amj.2005.19573108>
- Roberts, Peter W.; Dowling, Grahame R.** (2002). "Corporate reputation and sustained superior financial performance". *Strategic management journal*, v. 23, n. 12, pp. 1077-1093.
<https://doi.org/10.1002/smj.274>
- Sabri, Nor A. M.; Emran, Nurul A.; Harum, Norharyati** (2019). "Government open data portals: a measurement of data veracity coverage". *International journal of innovative technology and exploring engineering*, v. 8, n. 12, pp. 1975-1983.
<https://doi.org/10.35940/ijitee.L2908.1081219>
- Sadiq, Shazia; Indulska, Marta** (2017). "Open data: quality over quantity". *International journal of information management*, v. 37, n. 3, pp. 150-154.
<https://doi.org/10.1016/j.ijinfomgt.2017.01.003>
- Schultz, Majken; Mouritsen, Jan; Gabrielsen, Gorm** (2001). "Sticky reputation: analyzing a ranking system". *Corporate reputation review*, v. 4, n. 1, pp. 24-41.
<https://doi.org/10.1057/palgrave.crr.1540130>
- Taylor, Phillip; Barakat, Lina; Miles, Simon; Griffiths, Nathan** (2018). "Reputation assessment: a review and unifying abstraction". *The knowledge engineering review*, v. 33, E6.
<https://doi.org/10.1017/S0269888918000097>
- Veljković, Nataša; Bogdanović-Dinić, Sanja; Stoimenov, Leonid** (2014). "Benchmarking open government: an open data perspective". *Government information quarterly*, v. 3, n. 2, pp. 278-290.
<https://doi.org/10.1016/j.giq.2013.10.011>
- Vetrò, Antonio; Canova, Lorenzo; Torchiano, Marco; Orozco-Minotas, Camilo; Iemma, Raimondo; Morando, Federico** (2016). "Open data quality measurement framework: definition and application to Open Government Data". *Government information quarterly*, v. 33, n. 2, pp. 325-337.
<https://doi.org/10.1016/j.giq.2016.02.001>
- Zhang, Qingqiang; Sun, Xinbo; Zhang, Mingchao** (2022). "Data matters: a strategic action framework for data governance". *Information & management*, v. 59, n. 4, 103642.
<https://doi.org/10.1016/j.im.2022.103642>
- Zhao, Yupan; Fan, Bo** (2021). "Effect of an agency's resources on the implementation of open government data". *Information & management*, v. 58, n. 4, 103465.
<https://doi.org/10.1016/j.im.2021.103465>
- Zhenbin, Yang; Kankanhalli, Atreyi; Ha, Sangwook; Tayi, Giri-Kumar** (2020). "What drives public agencies to participate in open government data initiatives? An innovation resource perspective". *Information & management*, v. 57, n. 3, 103179.
<https://doi.org/10.1016/j.im.2019.103179>
- Zuiderwijk, Anneke; De-Reuver, Mark** (2021). "Why open government data initiatives fail to achieve their objectives: categorizing and prioritizing barriers through a global survey". *Transforming government: people, process and policy*, v. 15, n. 4, pp. 377-395.
<https://doi.org/10.1108/TG-09-2020-0271>
- Zuiderwijk, Anneke; Pirannejad, Ali; Susha, Iryna** (2021). "Comparing open data benchmarks: which metrics and methodologies determine countries' positions in the ranking lists?". *Telematics and informatics*, v. 62, 101634.
<https://doi.org/10.1016/j.tele.2021.101634>

Thirty years of research on high-growth entrepreneurship: bibliometric overview of its H-Classics

María-Paula Lechuga-Sancho; Salustiano Martínez-Fierro; Antonio R. Ramos-Rodríguez

Nota: Este artículo se puede leer en español en:
<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87121>

Recommended citation:

Lechuga-Sancho, María-Paula; Martínez-Fierro, Salustiano; Ramos-Rodríguez, Antonio R. (2023). “Thirty years of research on high-growth entrepreneurship: bibliometric overview of its H-Classics”. *Profesional de la información*, v. 32, n. 3, e320320.
<https://doi.org/10.3145/epi.2023.may.20>

Manuscript received on 20th September 2022
Accepted on 18th April 2023



María-Paula Lechuga-Sancho

<https://orcid.org/0000-0003-2340-7615>

Universidad de Cádiz
Instituto de Investigación para el
Desarrollo Social Sostenible (Indess)
Avda. de la Universidad, 4
11408 Jerez de la Frontera (Cádiz), Spain
paula.lechuga@uca.es



Salustiano Martínez-Fierro ✉

<https://orcid.org/0000-0003-0848-8007>

Universidad de Cádiz
Instituto de Investigación para el
Desarrollo Social Sostenible (Indess)
Avda. de la Universidad, 4
11408 Jerez de la Frontera (Cádiz), Spain
salustiano.martinez@gm.uca.es



Antonio R. Ramos-Rodríguez

<https://orcid.org/0000-0002-1799-8474>

Universidad de Cádiz
Instituto de Investigación para el
Desarrollo Social Sostenible (Indess)
Avda. de la Universidad, 4
11408 Jerez de la Frontera (Cádiz), Spain
rafael.ramos@gm.uca.es

Abstract

In recent years, high-growth entrepreneurship (HGE) research has gained increasing importance. For this reason, it is considered necessary to analyze papers that have had the most significant impact on the development of the discipline and that should be familiar to all researchers. Building new knowledge on these works is important because it provides legitimacy and coherence to the future development of this research field. Thus, this paper aims to identify and characterize the classic articles in the field of HGE, a line of research that has seen significant growth in the last 30 years. The H-Classics method is used to identify these papers. Subsequently, several bibliometric aspects of this collection are analyzed, such as the forums of journals where they have been published, the most productive authors, the patterns of collaboration, and an analysis of the conceptual structure through co-word analysis. Exhaustive content analysis is carried out to complement this vision, identifying the proposed objectives, methodologies, types of data, analysis techniques used, and their main contributions in three consecutive periods. The results are of value to researchers interested in high-growth firms because they allow us to understand the foundations on which this discipline has been built through its classics and to determine its main challenges for the future.

Keywords

Entrepreneurship; High growth; Rapid-growth firms; Gazelle companies; H-Classics; Highly cited papers; Citation classics; Bibliometrics; Biblioshiny; Co-word analysis; WoS; Web of Science; Social Sciences Citation Index; SSCI.

JEL codes: L26; M13; M21.



Funding

This work has been supported by:

- The Spanish *State Research Agency (AEI)* through the project PID2019-105381GA-I00/10.13039/501100011033 (iScience);
- The *Program of Promotion and Development of Research Activity of the University of Cadiz (Programa de Fomento e Impulso de la Actividad Investigadora)*, with reference PR2020-011.

1. Introduction

The relationship between entrepreneurship and economic development is widely recognized, and has been analyzed in-depth in previous literature (**Audretsch; Belitski; Desai**, 2015; **Platzek; Pretorius**, 2020; **Urbano; Aparicio; Audretsch**, 2019). As part of this relationship, high-growth enterprises (HGEs) are key to economic growth in developed countries (**Acs; Parsons; Tracy**, 2008; **BERR**, 2008; **Henrekson; Johansson**, 2010). Moreover, there is empirical evidence on the impact of HGEs on economic development (**Kemp; Nieuwenhuijsen; Bruins**, 2000; **Mason**, 1985; **Wong; Ho; Autio**, 2005). HGEs have a greater impact on the business environment than other firms (**Martínez-Fierro; Biedma-Ferrer; Ruiz-Navarro**, 2020). They are seen as important drivers of employment and economic dynamics (**Friesenbichler; Hölzl**, 2020).

The concept of a high-growth firm can be ambiguous, as a standard definition has not yet been established (**Martínez-Fierro; Biedma-Ferrer; Ruiz-Navarro**, 2020). The *Organization for Economic Cooperation and Development (OECD)* considers HGEs to be firms with an annual growth rate of > 20% over three years and more than 10 employees at the beginning of this three-year period (*Eurostat-OCDE*, 2007). On the basis of this conceptualization, 2.7% of the entire firm population can be classified as HGEs, out of which only 3.9% can repeat their HGE status in two successive periods. Hence, only 0.11% of the entire population are repeatedly classified as HGEs, being firms with persistent high potential.

The importance of HGE for employment and, more generally, for economic development has led to a considerable increase in the number of papers published in high-impact journals within this discipline. Although there is a vast number of bibliometric papers in the field of entrepreneurship (**Baier-Fuentes et al.**, 2019; **Vallaster et al.**, 2019; **Xu et al.** 2021), hardly any scientific research has been carried out in the form of a bibliometric analysis of the HGE literature produced. Such analysis would provide a solid basis for discussion to estimate its degree of consolidation and development (**Ramos-Rodríguez; Ruiz-Navarro**, 2004). Therefore, a literature review with a quantitative approach, based on bibliometric methods and that is not conditioned by the subjectivity of the authors and allows for an inventory of the work carried out, would form a valuable contribution to this field of research.

In the field of bibliometrics, citation analyses are aimed at assessing highly cited articles in a given scientific field, on the basis of the assumption that articles with a high number of citations correspond to a more influential impact in the area (**De-la-Flor-Martínez et al.**, 2016). Recently, several specialties have analyzed their so-called citation classics by setting the threshold value at the number of citations received or according to the number of highly cited papers (**Cascón-Katchadourian**, 2020). However, these studies of “citation classics” suffer from certain shortcomings. For example, the articles with the maximum number of citations have been shown without calibrating factors such as time (in terms of scientific evolution of research areas), and citation patterns are not considered (**Chiang et al.**, 2018). Furthermore, such analysis is only quantitative and does not provide detailed information on the qualitative aspects of citations. Additionally, these approaches are not based on an objective criterion. Unanswered questions may arise, such as:

- “Why do we use a threshold of 100 or 500 citations and not another similar number?” (**Cascón-Katchadourian**, 2020);
- “Why do we use the 100 most cited articles and instead of 300?” (**Corbella et al.**, 2017; **Feijoo et al.**, 2014).

To overcome the drawbacks and problems related to these traditional approaches to identifying highly cited articles, **Martínez-Sánchez et al.** (2014) presented the concept of H-Classics. The concept is based on the popular H-index by **Hirsch** (2005), which allows the classical citation search procedure to be systematized for any field of research (**Martínez-Sánchez et al.**, 2014). According to **Martínez-Sánchez et al.** (2014) and **Chiang et al.** (2018), the main advantages of this index are that it provides an objective, transparent, and impartial criterion to identify highly cited articles in the scientific literature as it identifies—in a single procedure—the number of papers published in the field and the impact of those publications. The criterion it applies is sensitive to the characteristics and the evolution of all research disciplines, as well as to the citation pattern of all research areas. However, the H-index has some limitations, namely that H-index values are not comparable across different academic fields, considers self-citations, and rewards academic seniority over more recent works (**Alonso et al.**, 2009).

This methodology, which is becoming increasingly relevant in the literature on bibliometrics (**Martínez-Sánchez et al.**, 2015), has recently been used (**Cabeza-Ramírez; Sánchez-Cañizares; Fuentes-García**, 2018) in the field of entrepreneurship to identify the main scientific articles (using the term “knowledge base” for the most cited articles) in the discipline. In this paper, we use the H-Classics methodology as a way to retrospectively determine the most cited articles within *Web of Science (WoS)* HGE research, that are essential reference points (the most influential) in this field of study. This paper aims to provide an overview of the scientific structure of high-potential entrepreneurship publications, using the

H-Classics methodology to identify classic articles. Their bibliometric characteristics are then assessed, such as the research performance of the main authors in the area, the journals, the countries, the research trends, the nature of the articles in terms of the main topics addressed, the type of articles, the research methods used, the methodological procedures, the statistical techniques, and the research samples, as well as the knowledge base on which the published works are based.

Highly cited papers are an essential reference point in a research field

This paper contributes to the literature on entrepreneurship by providing a comprehensive overview of the scientific structure of the high-potential entrepreneurship literature through a historical perspective of the evolution of this topic over time. The insights gained into this field illuminate the potential directions future research will follow. The article is structured as follows. Section 2 presents the basis of the methods used in the study and the procedure used to identify the primary papers for the analysis. Section 3 reports the results of our bibliometric analysis, and Section 4 presents the conclusions, limitations, and future lines of research.

2. Methodology

We followed the steps proposed by **Martínez-Sánchez et al.** (2014) to identify HGE H-Classics. First, we selected the bibliographic database to identify the scientific output in this field and citations received. In this study, we used the *Social Sciences Citation Index (SSCI)* of the *Web of Science (WoS)*. Although other databases, such as *Scopus* or *Google Scholar*, report the number of citations received, many studies consider *WoS* to be the most reliable database for applying the bibliometric methodology to the field of management and organization (for a review, see **Zupic; Čater**, 2015).

Second, to establish the research area under study and to obtain a collection of papers representative of HGE research, the following *WoS* advanced search query was performed:

TS = ("high-growth entrepreneur*" OR "high-growth firm*" OR "high-growth new firm*" OR "high-growth start-up" OR "high-growth SME*" OR "high-growth business" OR "high-growth venture*" OR "high-growth new venture*") OR TS = ("high-potential entrepreneur*" OR "high-potential firm*" OR "high-potential new firm*" OR "high-potential start-up" OR "high-potential SME*" OR "high-potential business" OR "high-potential venture*" OR "high-potential new venture*") OR TS = ("rapid-growth entrepreneur*" OR "rapid-growth firm*" OR "rapid-growth new firm*" OR "rapid-growth start-up" OR "rapid-growth SME*" OR "rapid-growth business" OR "rapid-growth venture*" OR "rapid-growth new venture*" OR "gazelle*."

The TS prefix performs a search of the title, abstract, and author keywords fields, and in *Keywords Plus*. Subsequently, we established a filter in the document type field to select only articles, discarding other types of documents such as reviews, editorial material, book chapters, proceedings papers, book reviews, or corrections. The search terms were consensually agreed upon by a group of experts belonging to the *Global Entrepreneurship Monitor (GEM)*, which is the most important global network on entrepreneurship worldwide. We carried out the search in mid-March 2021.

As a result of this search, 343 articles were downloaded from *WoS*. These articles were reviewed, one by one, by entrepreneurship experts. By consensus, 107 articles were eliminated because they did not contain evidence of research on the object of study. Of the remaining 236 articles, four that were published in 2021 were removed because the year was not over yet. Among the remaining 232 papers, 20 early-access articles were identified. These documents have an empty year of publication (PY) field, as there was no printed edition available. Given that the bibliometric software used (*Biblioshiny*) automatically discards this type of record because this field is empty, it was decided to rescue these articles for use in the analysis. Thus, the year of publication as early access was considered as the year of final publication, and these data were manually entered in the database. Of these 20 early-access articles, two appeared in 2019 and two in 2021. The latter two were removed. With these changes, the final sample consisted of 230 articles. Next, analysis of the collaboration and co-citation networks was performed with the Louvain cluster logarithm, considering the interface default options –star layout and normalisation by association– using the *InfoMap* clustering algorithm with 50 nodes. Amongst the method parameters, we used a repulsion force of 0,1 and a minimum number of 2 edges. The isolated nodes were removed. Each circular node represents a keyword, and the thickness of the connection between the nodes indicates the frequency of the co-occurrence between them.

Third, it was necessary to calculate the H-index associated with the research area. Accordingly, the articles were sorted in decreasing order based on the number of citations received. By definition, the H-Classics of the field are those articles that are located above the H-index, i.e., those whose position in the ranking is lower than or equal to the number of citations received. In this case, the H-index was equal to 50. Therefore, following the recommendations of **Martínez-Sánchez et al.** (2014), the first 50 articles were considered the H-Classics in HGE, and are listed in Appendix A.

To characterize the H-Classics in HGE, we used the *Bibliometrix R-package* for calculation of citation metrics and analysis (**Aria; Cuccurullo**, 2017). In particular, the distribution of documents per year of publication, the most productive authors, journals, institutions, countries, and some collaboration indicators were calculated. In addition, a co-word analysis was performed using the *Biblioshiny* interface to identify the conceptual structure of this field of research. More specifically, to identify the thematic areas that can be found in the conceptual structure of this discipline, we used the co-occurrence network based on the authors' keywords.

Finally, following the recommendations of **Moral-Muñoz et al. (2015)** and **Köseoglu et al. (2016)**, an exhaustive content analysis of H-Classics in HGE in three consecutive periods was carried out to identify the approaches adopted, the types of data and statistical techniques most used, and their evolution over the last thirty years.

3. Results

3.1. Distribution of high-growth entrepreneurship (HGE) H-Classics by year of publication

The scientific literature on HGE comprises 50 papers classified as H-Classics. The earliest two papers classified as H-Classics were published in 1992. **Larson (1992)** published one of these papers in the *Administrative science quarterly* journal, with the highest citation rate (1,278). The other was published by **Willard, Krueger and Feeser (1992)** in the *Journal of business venturing*, with 93 citations. In contrast, the most recent article within these H-Classics was published in the *Strategic entrepreneurship journal* by **Spigel and Harrison (2018)** with a reasonably high citation rate (111). Therefore, despite not being very common in the analysis of highly cited literature –because newly published articles require more time to accumulate citations– we find a highly cited paper from a recent year in this field of study. Furthermore, it is generally the case that highly cited papers tend to be published a longer time ago due to the citation window. However, in the field of HGE, we found that although the largest number of highly cited articles was published in the first decade of the 21st century (46% corresponding to 23 articles), the decrease in the second decade is not very significant (44% corresponding to 22 articles). This shows that HGE is currently a popular and contemporary topic.

3.2. Most productive authors, institutions, and countries with most published HGE H-Classics

One of the most important factors contributing to the structure and growth of any field (**Cuccurullo; Aria; Sarto, 2016; Ramos-Rodríguez; Ruiz-Navarro, 2004**) is its key authors. Authors such as **Furrer, Thomas and Goussevskaia (2008)** argue that author characteristics have the most explanatory power over the impact of an article. Although 106 different authors have written the 50 highly cited articles identified here, with an average of 2.12 authors per paper, Table 1 lists only the authors with two or more highly cited articles in HGE research. As can be seen, the authors who contributed most to the field are affiliated with European universities. The author with the highest number of H-Classics articles is Hölzl, from the *Austrian Institute of Economic Research (WIFO)*. The other most productive, highly cited authors have published a total of two articles published in the area, and four of these authors belong to universities or research centers in Sweden. These results are not surprising given that HGE has been widely defined as the main driver of innovation (see, e.g., **Brown; Manson, 2014**), and Sweden, in turn, is considered one of the most highly innovative countries in the world (**Kander et al., 2019**). Furthermore, in the results, we observe collaborations between the authors of the most cited papers. The author with the most collaborations is Johansson, from *The Ratio Institute* of Sweden, who collaborates with five other authors between the two papers. In the first of these (**Henrekson; Johansson, 2010**), together with Henrekson of the *Research Institute of Industrial Economics* of Sweden, Johansson carried out a meta-analysis of the empirical evidence regarding whether net employment growth is generated by a few fast-growing –not necessarily small and young– firms. The second study (**Coad et al., 2014**) was carried out with Coad, Daunfeldt, Hölzl, and Nightingale, from the *University of Sussex* (UK), *Dalarna University* (Sweden), *Dalarna University* (Sweden), and the *Austrian Institute of Economic Research* (Austria). Specifically, they analyze the reasons for the academic community's increased interest in high-growth firms, summarize the existing literature, and highlight the methodological considerations that constrain and bias research. Both papers review previous results, applying meta-analysis and literature review as methodologies. Conceptual papers using these techniques tend to be widely cited.

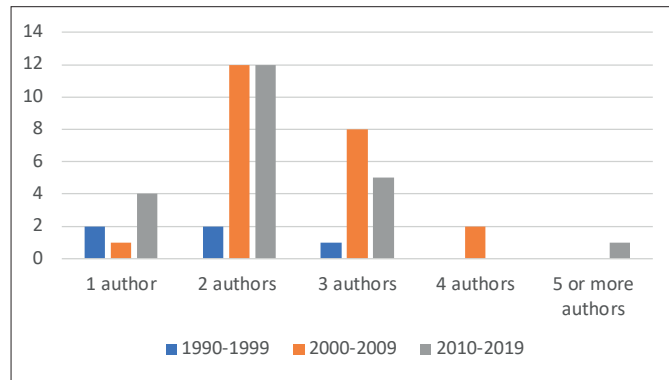
Table 1. Most productive highly cited authors in the field of HGE

Author	Institution	H-Classics	
		n	%
Hölzl, W.	<i>Austrian Institute of Economic Research (WIFO)</i> ; Austria.	3	6
Acs, Z. J.	<i>School of Policy and Government. George Mason University</i> ; Virginia, USA.	2	4
Marron, R.	<i>School of Management, University of St Andrews</i> ; Scotland, UK.	2	4
Daunfeldt, S. O.	<i>HUI Research</i> ; Stockholm, Sweden.	2	4
Davidsson, P.	<i>Jönköping International Business School</i> ; Sweden.	2	4
Henrekson, M.	<i>Research Institute of Industrial Economics</i> ; Stockholm, Sweden.	2	4
Johansson, D.	<i>The Ratio Institute</i> ; Stockholm, Sweden.	2	4
Mason, C.	<i>Hunter Centre for Entrepreneurship. University of Strathclyde</i> ; Glasgow, Scotland.	2	4
Stam, E.	<i>Utrecht University School of Economics</i> ; Utrecht, The Netherlands.	2	4

Following **Köseoglu et al. (2016)**, articles were classified into four groups according to the number of authors: single author, two authors, three authors, and four or more authors. Based on the premise that co-authorship in the publication of research results is a clear exponent of scientific collaboration, we can say that single authorship does not seem to be common in HGE H-Classics, with collaboration between two authors (26 papers) and three authors (14 papers) being the most frequent among all research output (Graph 1). This consolidates the perception of **Ramos-Rodríguez,**

Lechuga-Sancho, and Martínez-Fierro (2021) and Ramos-Rodríguez and Lechuga-Sancho (2020) that in current times, scientific research needs to be undertaken collaboratively.

According to Kalantari *et al.* (2017), each author makes an individual and unique contribution to a paper. Therefore, the institution and country to which the author belongs could be considered important contributors when assessing research. Accordingly, the number of publications from each country was used to assess the research contribution of a given region/country in a related field. Table 2 lists the countries from which highly cited papers in the field of HGE originated. In the analysis of highly cited literature (Baltussen; Kindler, 2004; Paladugu *et al.*, 2002), it is usual to find a small number of countries. However, in this particular discipline, we identified a high number of countries publishing highly cited articles. As can be seen, the USA (68%) was the most productive country with the highest number of publications. The USA was followed at a distance by the UK (34%); Sweden (28%); Netherlands (16%); Belgium and Germany (12% each); Spain (10%); Austria, Canada, and Finland (6% each); and Denmark, and New Zealand (2% each). The USA's dominance in this list of highly cited articles is not surprising given the top set of journals is biased in favor of journals of US origin, as is the case in other scientific fields and in the WoS itself (Martínez-Sánchez *et al.*, 2014; 2015).



Graph 1. Collaboration/co-authors per article in HGE H-Classics research

There are collaborations between authors from different countries, the most frequent collaborations being between authors from the Netherlands and Germany, the UK and Belgium, or the UK and Finland. For example, the works of Sternberg and Wennekers (2005) and Grimm, Knorringa, and Lay (2012) correspond to collaborations between the Netherlands and Germany. Collaboration is also carried out between the following institutions: the University of Cologne (Germany), EIM Business and Policy Research (Netherlands), Erasmus University Rotterdam (Netherlands), the Institute for the Study of Labor (Germany), the German Institute of Global and Area Studies (Germany), and the University of Göttingen (Germany).

In-depth analysis focusing on the institutions at which the research is conducted offers additional interesting clues about the history and evolution of HGE H-Classics over the last 30 years. Table 3 shows the institutions that had at least two productions or more. The USA is the most productive country, and therefore it is not surprising that five of the 16 most productive universities are American. First and second places in the list are held by Syracuse University, with four H-Classics, and the University of North Carolina, with three H-Classics. Five institutions from Sweden also stand out, namely The Ratio Institute, Dalarna University, Hui Research, Jönköping International Business School and the Stockholm School of Economics.

Table 2. Most productive countries in terms of publication of HGE H-Classics

Country	H-Classics (%)	
	n	%
USA	34	68
UK	17	34
Sweden	14	28
Netherlands	8	16
Belgium	6	12
Germany	6	12
Spain	5	10
Austria	3	6
Canada	3	6
Finland	3	6
Denmark	1	2
New Zealand	1	2

Table 3. Most productive institutions in terms of publication of HGE H-Classics

Institution	Country	H-Classics	
		n	%
Syracuse University	USA	4	8
University of North Carolina	USA	3	6
The Ratio Institute	Sweden	3	6
Imperial College of Science, Technology and Medicine	UK	3	6
University of Utrecht	Netherlands	3	6
Aalto University	Finland	2	4
Austrian Institute of Economic Research (WIFO)	Austria	2	4
Dalarna University	Sweden	2	4
HUI Research	Sweden	2	4
Jönköping International Business School (JIBS)	Sweden	2	4
Stockholm School of Economics	Sweden	2	4
University of Central Florida	USA	2	4
University of Colorado Denver	USA	2	4
University of Ghent	Belgium	2	4
University of Sussex	UK	2	4
University of Utah	USA	2	4

3.3. Journals

Table 4 shows all the scientific journals that make up our HGE H-Classics database. With 19 papers out of 50, *Small business economics* is the journal with the most highly cited publications. The *Journal of business venturing* also stands out (with seven publications). The *Journal of small business management* and *Research policy*, with three papers each, have also contributed significantly to the development of this scientific area. However, in our area of study, there is a wide variety of journals (29.31% of the total output) containing only one H-Classic.

Table 4. Sources in which the highest number of HGE H-Classics are published

Sources	Papers	SSCI citations	Impact factor (JCR 2019)
<i>Small business economics</i>	19	975	4.803
<i>Journal of business venturing</i>	7	630	7.590
<i>Journal of small business management</i>	3	308	3.461
<i>Research policy</i>	3	1,525	5.351
<i>Industrial and corporate change</i>	2	208	1.981
<i>Journal of finance</i>	2	872	6.813
<i>Strategic entrepreneurship journal</i>	2	279	6.200
<i>Administrative science quarterly</i>	1	382	8.304
<i>Entrepreneurship and regional development</i>	1	251	2.885
<i>Human resource management</i>	1	3,910	2.476
<i>Journal of applied psychology</i>	1	1,053	5.818
<i>Journal of banking & finance</i>	1	912	2.269
<i>Journal of business research</i>	1	3,573	4.874
<i>Journal of financial and quantitative analysis</i>	1	471	2.707
<i>Journal of financial economics</i>	1	1,341	5.731
<i>Organization science</i>	1	345	2.782
<i>Public administration review</i>	1	520	4.063
<i>Technovation</i>	1	338	5.729
<i>World development</i>	1	2,472	3.869

3.4. Content of reviewed articles

3.4.1. Research themes

To establish the knowledge structure of these most cited works in the specific discipline of HGE, we developed clusters of the different scientific aspects. The co-word analysis provided sets of groups of textual information representing the conceptual basis of various topics in the field. Graph 2 shows the main research topics of the most cited articles within the HGE discipline. As the image shows, there are four main themes:

- determinants (26 articles, 52%),
- strategic management (10 articles, 20%),
- finance (nine articles, 18%), and
- innovation (five articles, 10%).

Determinants

Many studies have addressed the determinants of high-potential firm growth (Delmar; Davidsson; Gartner, 2003; Baerlinger; Jones; Neubaum, 2005; Moreno; Casillas, 2007). However, most studies report the same key characteristic of HGE firms within the determinants cluster, namely their ability to create employment and to be important drivers of economic and structural change (Sternberg; Wennekers, 2005; Acs; Mueller, 2008; Henrekson; Johansson, 2010; Mason; Brown, 2013; Coad *et al.*, 2014). Other important aspects of the articles within this central theme are aimed at

- studying how policy (initiative design) can foster HGE;
- determining appropriate and inappropriate policies for the promotion of venture capital and high-potential entrepreneurs;
- determining public policies to foster regional development through HGE.

The impact on job creation appears to have put HGE on the public policy agenda. Other papers within this cluster focus on identifying the attributes of entrepreneurs who tend to generate a significant number of HGEs or those factors that increase or decrease the entrepreneur's willingness to grow the firm (Siegel; Siegel; Macmillan, 1993; Schindehutte; Morris; Allen, 2006; Baum; Bird, 2010; Goedhuys; Sleuwaegen, 2010). Such factors include:

- the strategic origin of the firm (i.e., the methods and paths through which the firm was founded);
- the previous experience of the founder/owner; and
- the entrepreneur's ability to set realistic and measurable goals and to manage conflict effectively.

In general, the samples vary according to the sector, country, region of study analysis, firm size, use of high or low technology, and the stage of the firm’s life cycle, i.e., start-ups versus more mature and consolidated firms.

High-growth enterprises (HGEs) are key to economic growth in developed countries

Strategic management

The next cluster comprises papers that mainly analyze the various strategies and environmental variables that have a predictable influence on the results of HGEs. In this regard, many studies argue that HGEs manage resources according to the specific requirements of the competitive environment. In particular, in the context of the companies in this group, we find gender-based research and other works that focus on the strategic management of human resources. More specifically, issues such as recruitment, training, and remuneration are addressed. Regarding remuneration, some papers link remuneration to top management, analyzing the negative consequences of pay disparity in family and non-family top management teams. As might be expected, the governance and management of HGEs are also issues that have been analyzed empirically. In this area, fast-growing start-ups quickly outgrow the management capacity of the founder. Therefore, the research literature finds that unless the founder is replaced or supplemented by professional management, firm performance is likely to stagnate or decline rapidly.

Financial perspective

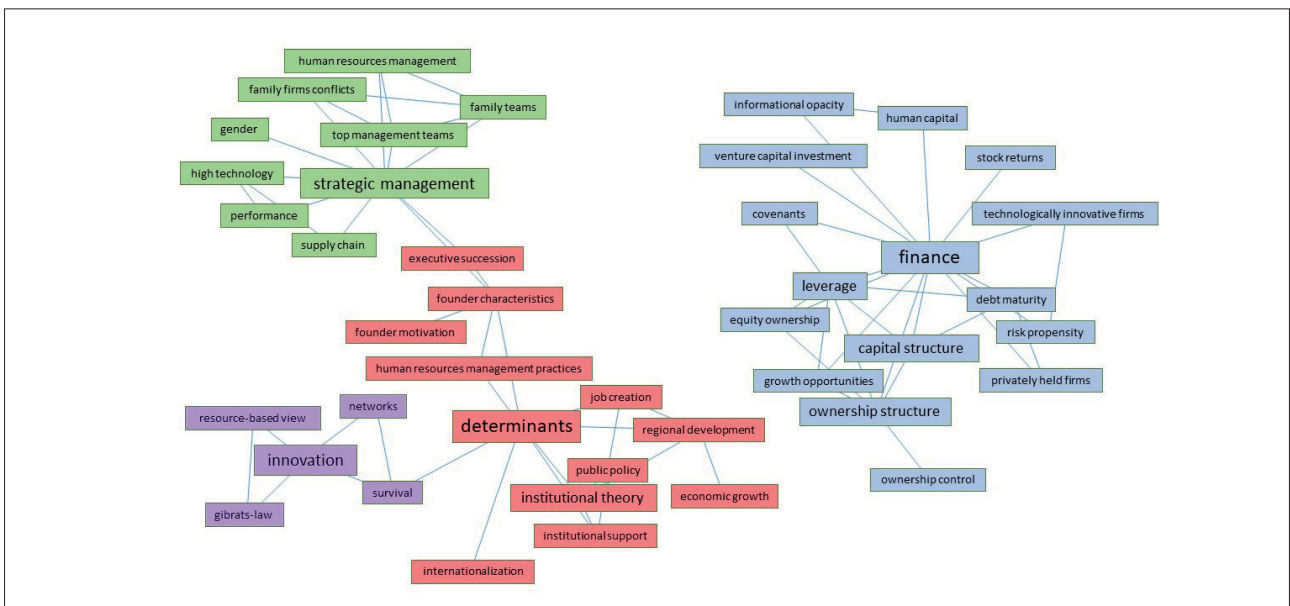
In this cluster, we find papers of different types, namely those focusing on:

- finance sources used by high-growth entrepreneurs;
- empirical analysis of the relationship between firm value; leverage, and ownership;
- debt, debt maturity, and covenants;
- corporate governance/capital structure;
- venture capital analysis.

Venture capitalists (VCs) are considered experts in identifying high-potential start-ups. Additional studies analyze the differences in risk propensity between managers and entrepreneurs and between low- and high-growth entrepreneurs.

Innovation

Finally, in the innovation group, most papers provide empirical support for the recommendations of policymakers and business leaders to maintain R&D investments, especially in high-tech sectors, even when faced with a recession. Several papers show that high-potential firms are increasingly crucial to achieving innovation and formulating entrepreneurship policies. R&D is one of the innovation activities that is most often measured in the most cited papers in the discipline. Thus, among the main analyses carried out in these papers are those addressing the following questions: Are high-growth SMEs more active in R&D than comparable firms that are not growing as fast? In other words, what is the role of innovation in the probability of being a high-growth firm? To what extent do R&D activities affect the growth of new firms, directly or indirectly, through the development of new products and/or alliances with other firms? Is R&D more important for high-growth SMEs in countries that are closer to the technological frontier?



Graph 2. Main topics cited in keywords of HGE H-Classics documents

3.4.2. Nature of high-growth entrepreneurship (HGE) papers

The nature of the papers was examined depending on their approach, the research methods, the type of data, statistical techniques, and countries they report. The nature of the articles included in this study is presented by type in Table 5.

Table 5. Nature of H-Classic articles in HGE

Category	1992-1999		2000-2009		2010-2018		Total	
	n	%	n	%	n	%	n	%
Article type								
Empirical	5	100	21	91.3	16	72.7	42	84.0
Conceptual	0	0	1	4.4	5	22.7	6	12.0
Review	0	0	1	4.4	1	4.5	2	4.0
Research methods								
Quantitative	3	60.0	17	73.9	11	50.0	31	62.0
Qualitative	2	40.0	4	17.4	11	50.0	17	34.0
Hybrid	0	0	2	8.7	0	0	2	4.0
Primary or secondary data								
Primary- questionnaire	1	20.0	4	17.4	2	12.5	7	16.7
Primary-interview	0	0	3	13.0	3	18.8	6	14.3
Primary-case study	0	0	0	0	1	6.3	1	2.4
Primary hybrid	1	20.0	2	8.7	0	0	3	7.1
Secondary	3	60.0	11	47.8	7	43.8	21	50.0
Primary-Secondary	0	0	1	4.3	3	18.8	4	9.5
Statistical techniques								
Variance analysis [AN(C)OVA, MANOVA]	0	0	2	8.7	0	0	2	4.0
Descriptive/content	0	0	1	4.3	0	0	1	2.0
t-test, χ^2 , correlation, Mann-Whitney U	1	25.0	0	0	1	7.1	2	4.0
Regression	2	50.0	9	39.1	10	71.4	21	42.0
Factor, cluster, discriminant	1	25.0	4	17.4	0	0	5	10.0
Structural equation model	0	0	1	4.3	2	14.3	3	6.0
Others	0	0	3	13.0	1	7.1	4	8.0
Samples								
Firms	4	100	8	34.8	9	81.8	21	42.0
Entrepreneurs	0	0	3	13.0	0	0	3	6.0
Territorial areas	0	0	1	4.3	0	0	1	2.0
Papers	0	0	1	4.3	1	9.1	2	4.0
Others	0	0	3	13.0	1	9.1	4	8.0
Country								
USA	3	100	3	13.0	0	0	6	12.0
Sweden	0	0	2	8.7	1	9.1	3	6.0
Finland	0	0	1	4.3	1	9.1	2	4.0
UK	0	0	0	0	2	18.2	2	4.0
Spain	0	0	0	0	2	18.2	2	4.0
European countries	0	0	1	4.3	1	9.1	2	4.0
African countries	0	0	0	0	2	18.2	2	4.0
Netherlands	0	0	1	4.3	0	0	1	2.0
Austria	0	0	0	0	1	9.1	1	2.0
Scotland	0	0	0	0	1	9.1	1	2.0

Primary methods

The primary methods used in the articles published were also identified. During the 27-year period from 1992 to 2018, 84% of the articles were empirical studies, 12% were conceptual studies, and 4% were reviews. Empirical studies account for 100% of the works in the first sub-period analyzed (1992-1999), decreasing to 72.7% for 2010-2018. In the latter period, the proportion of conceptual studies increased to 22.7%.

The evolution of research methods

The majority (62%) of articles published between 1992 and 2018 used a quantitative research method, and 34% used a qualitative research method. Approximately 4% used a hybrid research method. In the period before 2010, quantitative research was more frequent than qualitative research. Both forms of research were equally represented during the 2010-2018 period.

Data-gathering methods

During the 27-year period investigated, half of the empirical articles used secondary data, while 40.5% of the empirical articles used primary data, with almost 10% of the articles using both types of information for their research. It can be observed that over this period, the use of secondary data decreased from 60% in the first sub-period to 43.8% in the years 2010 to 2018.

Primary data-gathering methods

For the sourcing of primary data, questionnaires were the most commonly used method (16.7%), followed by interviews at 14.3%. Some papers used other techniques such as hybrid or multiple primary methods (7.1%) and case studies (2.4%). When the three sub-periods are considered separately, questionnaires and interviews were also the most frequently used data-collection tools. However, the use of questionnaires decreased in favor of interviews, and the emergence of case studies was observed in the last sub-period analyzed. The combination of primary and secondary information followed an increasing trend over the years, reaching 18.8% in 2010-2018.

Most cited articles in HGE provide signals of the evolution and knowledge structure of the classic papers on the topic

Statistical techniques

Several different statistical techniques were represented in the articles examined in this study. The most common category of statistical methods included regression analysis (42.0%). The second most common category included factor, cluster, and discriminant analysis (10.0%). The third most common category included structural equation models (8.0%), and the fourth and fifth categories used variance analysis [AN(C)OVA, MANOVA] and t-tests, chi-square, correlation, and the Mann-Whitney U test, at 4.0%. It should be noted that 8.0% of the articles used other techniques that are not mentioned above. In the period analyzed, there was an increasing trend for the use of regression models from 50.0% in 1992-2000 to 71.4% in 2010-2018. In particular, there was an increase in the use of structural equation models. In the first sub-period, not a single article used structural equation models, whereas in 2010-2018, 14.3% of the articles used this statistical technique. There was a decrease in t-tests, chi-square, correlation, and Mann-Whitney U tests, from 25.0% to 7.1%. It is also noteworthy that in the sub-period 2010-2018, no articles employing factor, cluster, and discriminant analysis techniques were published.

Standard group sampling

The most common group sampled in the articles examined during the 27-year period of this study was that focused on firms, representing 42.0% of all articles. The second most common group was on entrepreneurs (6.0%), followed by researchers using other papers as units of analysis (4.0%) and those focused on territorial areas (2.0%). During the first sub-period, 1992-1999, all the articles used a sample of companies. From 2000 onwards, researchers began to produce empirical research based on other types of samples, such as entrepreneurs, territorial areas, and academic papers.

Countries in the samples

The country of origin of the sample is only provided in 22 of the papers analyzed. Articles using samples from the USA account for 12% of the total. Sweden accounts for 6%. Other countries follow, with smaller percentages of papers. In the sub-period 1992-1999, only samples from the USA were used. In the following sub-period, articles with samples from Sweden, Finland, Netherlands, and other European countries were added. From 2010-2018, papers were published with samples from African countries, Spain, UK, Austria, and Scotland.

4. Conclusions, limitations, and directions for future research

The academic field of entrepreneurship research has evolved from isolated groups of scholars researching small businesses, to an international community of departments, institutes, and foundations promoting research on start-ups and high-growth ventures (Aldrich, 2012).

This article provides a retrospective of the most cited articles published in *WoS* journals on high-potential entrepreneurship. To this end, the H-Classics methodology, based on the H-index, was applied, identifying the 50 most cited articles on the subject out of the total of 230 articles in the sample. Bibliometric characteristics, authors, journals, institutions, collaboration patterns, research trends, and the nature of the most cited articles from 1989 to 2020 are analyzed. Using H-Classics, new findings were discovered and more precise results were obtained on high-growth entrepreneurship (HGE) than previous work that has carried out bibliometric analyses without taking into account the most-cited papers.

Conclusions

- Research on high-potential firms is a current and hot topic.
- Of the most cited articles, 90% were published in the 21st century, and 44% were published in the second decade of this century.
- Collaboration between researchers is the most common way of researching this sub-area of entrepreneurship. Two or three authors have produced 80% of the 50 most cited articles. The most productive author is Hölzl from the *Austrian Institute of Economic Research (WIFO)*. Four of the most productive authors are from Swedish universities or research centers.
- Several countries are represented in the most-cited articles on HGE, with the USA being the most productive, followed by the UK and Sweden.
- The institutions with three or more of the top 50 most-cited publications are located in these countries and the Netherlands.
- Analysis of the knowledge structure of these 50 most cited articles on HGEs results in three clusters. The key themes in these clusters are the determinants of HGEs, strategic management, and finance. These research themes highlight that while all forms of entrepreneurship play a key role in the economy, there is growing evidence that HGEs are the main drivers of innovation and job creation.
- The vast majority (84%) of the papers were empirical in nature, although in recent years, there has been an increase in conceptual papers. The research method most commonly used was the quantitative method. The empirical studies mainly used primary data from questionnaires and interviews, with the combination of primary and secondary data becoming more and more frequent.
- The most commonly used statistical techniques were regression analysis, factor analysis, cluster analysis, and discriminant analysis. Over the years, there has been an increase in the use of regression analysis and in papers using structural equation modeling.

Contributions

This paper makes several contributions to the literature:

- First, an analysis of the most cited articles in HGE provides valuable insights into the evolution and knowledge structure of the classic articles on the topic.
- Second, this paper establishes HGE research as an important and current sub-area within entrepreneurship, to be explored even further in future research.
- Third, HGE is an area of knowledge whose development has been fostered by European countries, especially Sweden, given its high link with innovation, and by the United States. The present study confirms that there is a body of literature focused on HGE research. This finding supports the feasibility of creating a specific category for HGE within the main lines of research in the field of entrepreneurship.

Limitations

- First, the articles analyzed are derived from a single source, the *WoS*. Although *WoS* is one of the most relevant databases, other papers published in relevant journals not indexed in this database have been omitted.
- Second, more recent publications had less time to be cited, so citation counts as a measure of influence could be biased toward older publications (Zupic; Čater, 2014). Furthermore, both old and new articles receive citations and the citation scores are constantly changing (Piwowar-Sulej, 2021). Therefore, as established by Aksnes, Langfeldt, and Wouters, (2019), further analyses conducted in different periods, with the same methodology, are warranted.

Future research

Future research should include other data sources to validate the results obtained. In addition, this analysis could be complemented by another study exploring the knowledge structure of HGE research in recent decades through other bibliometric methodologies such as science mapping (Cobo *et al.*, 2012), to obtain a complete picture of this area of entrepreneurship. An analysis of the latest HGE research could explain future trends in the area. Finally, new studies with specific geographical foci (Europe, USA, etc.) would be interesting.

“ HGE research is an important and current sub-area within entrepreneurship to be explored further in future research ”

Finally, the information provided in this study is invaluable to understanding the scientific structure within the field of research on high-growth firms, and could therefore serve as a reference for future research development in this field.

5. References

Acs, Zoltan J.; Mueller, Pamela (2008). “Employment effects of business dynamics: mice, gazelles and elephants”. *Small business economics*, v. 30, n. 1, pp. 85-100.
<https://doi.org/10.1007/s11187-007-9052-3>

- Acs, Zoltan J.; Parsons, William; Tracy, Spencer** (2008). *High impact firms: gazelles revisited*. Washington DC: Office of Advocacy, US Small Business Administrations.
- Aksnes, Dag W.; Langfeldt, Liv; Wouters, Paul** (2019). "Citations, citation indicators, and research quality: an overview of basic concepts and theories". *SAGE open*, v. 9, n. 1.
<https://doi.org/10.1177/2158244019829575>
- Aldrich, Howard E.** (2012). "The emergence of entrepreneurship as an academic field: a personal essay on institutional entrepreneurship". *Research policy*, v. 41, n. 7, pp. 1240-1248.
<https://doi.org/10.1016/j.respol.2012.03.013>
- Alonso, Sergio; Cabrerizo, Francisco-Javier; Herrera-Viedma, Enrique; Herrera, Francisco** (2009). "H-index: a review focused in its variants, computation and standardization for different scientific fields". *Journal of informetrics*, v. 3, n. 4, pp. 273-289.
<https://doi.org/10.1016/j.joi.2009.04.001>
- Aria, Massimo; Cuccurullo, Corrado** (2017). "Bibliometrix: an R-tool for comprehensive science mapping analysis". *Journal of informetrics*, v. 11, n. 4, pp. 959-975.
<https://doi.org/10.1016/j.joi.2017.08.007>
- Audretsch, David B.; Belitski, Maksim; Desai, Sameeksha** (2015). "Entrepreneurship and economic development in cities". *The annals of regional science*, v. 55, n. 1, pp. 33-60.
<https://doi.org/10.1007/s00168-015-0685-x>
- Baier-Fuentes, Hugo; Merigó, José M.; Amorós, José-Ernesto; Gaviria-Marín, Magaly** (2019). "International entrepreneurship: a bibliometric overview". *International entrepreneurship and management journal*, v. 15, pp. 385-429.
<https://doi.org/10.1007/s11365-017-0487-y>
- Baltussen, Anja; Kindler, Christoph H.** (2004). "Citation classics in critical care medicine". *Intensive care medicine*, v. 30, n. 5, pp. 902-910.
<https://doi.org/10.1007/s00134-004-2195-7>
- Barringer, Bruce R.; Jones, Foard F.; Neubaum, Donald O.** (2005). "A quantitative content analysis of the characteristics of rapid-growth firms and their founders". *Journal of business venturing*, v. 20, n. 5, pp. 663-687.
<https://doi.org/10.1016/j.jbusvent.2004.03.004>
- Baum, J. Robert; Bird, Barbara J.** (2010). "The successful intelligence of high-growth entrepreneurs: links to new venture growth". *Organization science*, v. 21, n. 2, pp. 397-412.
<https://doi.org/10.1287/orsc.1090.0445>
- BERR** (2008). *High growth firms in the UK: lessons from an analysis of comparative UK performance*. BERR economics paper n. 3. London: Department for Business Enterprise & Regulatory Reform.
<https://www.econbiz.de/Record/high-growth-firms-in-the-uk-lessons-from-an-analysis-of-comparative-uk-performance/10003880888>
- Brown, Ross; Mason, Colin** (2014). "Inside the high-tech black box: a critique of technology entrepreneurship policy". *Technovation*, v. 34, n. 12, pp. 773-784.
<https://doi.org/10.1016/j.technovation.2014.07.013>
- Cabeza-Ramírez, L. Javier; Sánchez-Cañizares, Sandra-María; Fuentes-García, Fernando J.** (2018). "Caracterización de los clásicos del emprendimiento (1968-2016). Un análisis basado en la *Web of Science*". *Revista española de documentación científica*, v. 41, n. 2.
<https://doi.org/10.3989/redc.2018.2.1488>
- Cascón-Katchadourian, Jesús** (2020). "Identificación y análisis de los H-Classics de las universidades públicas andaluzas en la categoría economics del *Journal Citation Reports (JCR)*". *Revista de estudios empresariales. Segunda época*, v. 1, pp. 75-90.
<https://doi.org/10.17561/ree.v2020n1.6>
- Chiang, Ho-Sheng; Huang, Ren-Yeong; Weng, Pei-Wei; Mau, Lian-Ping; Tsai, Yi-Wen-Cathy; Chung, Ming-Pang; Chung, Chi-Hsiang; Yeh, Hsiao-Wen; Shieh, Yi-Shing; Cheng, Wang-Chien** (2018). "Prominence of scientific publications towards peri-implant complications in implantology: a bibliometric analysis using the H-classics method". *Journal of oral rehabilitation*, v. 45, n. 3, pp. 240-249.
<https://doi.org/10.1111/joor.12606>
- Coad, Alex; Daunfeldt, Sven-Olov; Hözl, Werner; Johansson, Dan; Nightingale, Paul** (2014). "High-growth firms: introduction to the special section". *Industrial and corporate change*, v. 23, n. 1, pp. 91-112.
<https://doi.org/10.1093/icc/dtt052>

- Cobo, Manuel J.; López-Herrera, Antonio-Gabriel; Herrera-Viedma, Enrique; Herrera, Francisco** (2012). "SciMAT: a new science mapping analysis software tool". *Journal of the American Society for Information Science and Technology*, v. 63, n. 8, pp. 1609-1630.
<https://doi.org/10.1002/asi.22688>
- Corbella, Stefano; Francetti, Luca; Taschieri, Silvio; Weinstein, Roberto; Del-Fabbro, Massimo** (2017). "Analysis of the 100 most-cited articles in periodontology". *Journal of investigative and clinical dentistry*, v. 8, n. 3, e12222.
<https://doi.org/10.1111/jicd.12222>
- Cuccurullo, Corrado; Aria, Massimo; Sarto, Fabrizia** (2016). "Foundations and trends in performance management. A twenty-five years bibliometric analysis in business and public administration domains". *Scientometrics*, v. 108, n. 2, pp. 595-611.
<https://doi.org/10.1007/s11192-016-1948-8>
- De-la-Flor-Martínez, María; Galindo-Moreno, Pablo; Sánchez-Fernández, Elena; Piattelli, Adriano; Cobo, Manuel J.; Herrera-Viedma, Enrique** (2016). "H-classic: a new method to identify classic articles in implant dentistry, periodontics, and oral surgery". *Clinical oral implants research*, v. 27, n. 10, pp. 1317-1330.
<https://doi.org/10.1111/clr.12749>
- Delmar, Frédéric; Davidsson, Per; Gartner, William B.** (2003). "Arriving at the high-growth firm". *Journal of business venturing*, v. 18, n. 2, pp. 189-216.
[https://doi.org/10.1016/S0883-9026\(02\)00080-0](https://doi.org/10.1016/S0883-9026(02)00080-0)
- Eurostat–OECD** (2007). *Eurostat–OECD manual on business demography statistics*. Office for Publications of the European Communities.
<https://www.oecd.org/sdd/business-stats/eurostat-oecdmanualonbusinessdemographystatistics.htm>
- Feijoo, Javier F.; Limeres, Jacobo; Fernández-Varela, Marta; Ramos, Isabel; Diz, Pedro** (2014). "The 100 most cited articles in dentistry". *Clinical oral investigations*, v. 18, pp. 699-706.
<https://doi.org/10.1007/s00784-013-1017-0>
- Friesenbichler, Klaus; Hölzl, Werner** (2020). "High-growth firm shares in Austrian regions: the role of economic structures". *Regional studies*, v. 54, n. 11, pp. 1585-1595.
<https://doi.org/10.1080/00343404.2020.1742316>
- Furrer, Olivier; Thomas, Howard; Goussevskaia, Anna** (2008). "The structure and evolution of the strategic management field: A content analysis of 26 years of strategic management research". *International journal of management reviews*, v. 10, n. 1.
<https://doi.org/10.1111/j.1468-2370.2007.00217.x>
- Goedhuys, Micheline; Sleuwaegen, Leo** (2010). "High-growth entrepreneurial firms in Africa: a quantile regression approach". *Small business economics*, v. 34, n. 1, pp. 31-51.
<https://doi.org/10.1007/s11187-009-9193-7>
- Grimm, Michael; Knorringa, Peter; Lay, Jann** (2012). "Constrained gazelles: high potentials in West Africa's informal economy". *World development*, v. 40, n. 7, pp. 1352-1368.
<https://doi.org/10.1016/j.worlddev.2012.03.009>
- Henrekson, Magnus; Johansson, Dan** (2010). "Gazelles as job creators: a survey and interpretation of the evidence". *Small business economics*, v. 35, pp. 227-244.
<https://doi.org/10.1007/s11187-009-9172-z>
- Hirsch, Jorge E.** (2005). "An index to quantify an individual's scientific research output". *Proceedings of the National Academy of Sciences of the United States of America*, v. 102, n. 46, pp. 16569-16572.
<https://doi.org/10.1073/pnas.0507655102>
- Kalantari, Ali; Kamsin, Amirrudin; Kamaruddin, Halim-Shukri; Ebrahim, Nader-Ale; Gani, Abdullah; Ebrahimi, Ali; Shams Shirband, Shahaboddin** (2017). "A bibliometric approach to tracking big data research trends". *Journal of big data*, v. 4.
<https://doi.org/10.1186/s40537-017-0088-1>
- Kander, Astrid; Taalbi, Josef; Oksanen, Juha; Sjöö, Karolin; Rilla, Nina** (2019). "Innovation trends and industrial renewal in Finland and Sweden 1970-2013". *The Scandinavian economic history review*, v. 67, n. 1, pp. 47-70.
<https://doi.org/10.1080/03585522.2018.1516697>
- Kemp, Ron G. M.; Nieuwenhuijsen, Henk R.; Bruins, Andries** (2000). *Bedrijvendynamiek, snelgroeiende bedrijven en regionaal-economische ontwikkeling*. Zoetermeer: EIM Business and Policy Research.
<https://ondernemerschap.panteia.nl/pdf-ez/a200006.pdf>
- Köseoglu, Mehmet-Ali; Sehitoglu, Yasin; Ross, Gary; Parnell, John A.** (2016). "The evolution of business ethics research in the realm of tourism and hospitality". *International journal of contemporary hospitality management*, v. 28, n. 8, pp. 1598-1621.
<https://doi.org/10.1108/IJCHM-04-2015-0188>

- Larson, Andrea** (1992). "Network dyads in entrepreneurial settings: a study of the governance of exchange relationships". *Administrative science quarterly*, v. 37, n. 1, pp. 76-104.
<https://doi.org/10.2307/2393534>
- Martínez-Fierro, Salustiano; Biedma-Ferrer, José-María; Ruiz-Navarro, José** (2020). "Impact of high-growth start-ups on entrepreneurial environment based on the level of national economic development". *Business strategy and the environment*, v. 29, n. 3, pp. 1007-1020.
<https://doi.org/10.1002/bse.2413>
- Martínez-Sánchez, María-Ángeles; Herrera, Manuel; Jiménez-Contreras, Evaristo; Ruiz, Aldrick; Herrera-Viedma, Enrique** (2015). "Characterizing highly cited papers in social work through H-Classics". *Scientometrics*, v. 102, n. 2, pp. 1713-1729.
<https://doi.org/10.1007/s11192-014-1460-y>
- Martínez-Sánchez, María-Ángeles; Herrera, Manuel; López-Gijón, Javier; Herrera-Viedma, Enrique** (2014). "H-Classics: characterizing the concept of citation classics through h-index". *Scientometrics*, v. 98, n. 3, pp. 1971-1983.
<https://doi.org/10.1007/s11192-013-1155-9>
- Mason, Colin M.** (1985). "The geography of 'successful' small firms in the United Kingdom". *Environment and planning A: Economy and space*, v. 17, n. 11, pp. 1499-1513.
<https://doi.org/10.1068/a171499>
- Mason, Colin M.; Brown, Ross** (2013). "Creating good public policy to support high-growth firms". *Small business economics*, v. 40, n. 2, pp. 211-225.
<https://doi.org/10.1007/s11187-011-9369-9>
- Moral-Muñoz, José A.; Cobo, Manuel J.; Chiclana, Francisco; Collop, Andrew; Herrera-Viedma, Enrique** (2015). "Analyzing highly cited papers in intelligent transportation systems". *IEEE transactions on intelligent transportation systems*, v. 17, n. 4, pp. 993-1001.
<https://doi.org/10.1109/TITS.2015.2494533>
- Moreno, Ana M.; Casillas, José C.** (2007). "High-growth SMEs versus non-high-growth SMEs: a discriminant analysis". *Entrepreneurship and regional development*, v. 19, n. 1, pp. 69-88.
<https://doi.org/10.1080/08985620601002162>
- Paladugu, Ramesh; Schein, Moshe; Gardezi, Syed; Wise, Leslie** (2002). "One hundred citation classics in general surgical journals". *World journal of surgery*, v. 26, n. 9, pp. 1099-1105.
<https://doi.org/10.1007/s00268-002-6376-7>
- Piwowar-Sulej, Katarzyna** (2021). "Core functions of sustainable human resource management. A hybrid literature review with the use of H-classics methodology". *Sustainable development*, v. 29, n. 4, pp. 671-693.
<https://doi.org/10.1002/sd.2166>
- Platzek, Bernd P.; Pretorius, Leon** (2020). "Regional cooperation in a thriving entrepreneurial economy: a holistic view on innovation, entrepreneurship and economic development". *International journal of innovation and technology management*, v. 17, n. 3, 2050014.
<https://doi.org/10.1142/S0219877020500145>
- Ramos-Rodríguez, Antonio-Rafael; Lechuga-Sancho, María-Paula** (2020). "Patrones de coautoría en las publicaciones españolas en Ciencias Económicas y Empresariales indexadas en el SSC". *Revista de estudios empresariales*, n. 1, pp. 113-136.
<https://doi.org/10.17561/ree.v2020n1.8>
- Ramos-Rodríguez, Antonio-Rafael; Lechuga-Sancho, María-Paula; Martínez-Fierro, Salustiano** (2021). "Authorship trends and collaboration patterns in hospitality and tourism research". *International journal of contemporary hospitality management*, v. 33, n. 4, pp. 1344-1367.
<https://doi.org/10.1108/IJCHM-09-2020-0981>
- Ramos-Rodríguez, Antonio-Rafael; Ruiz-Navarro, José** (2004). "Changes in the intellectual structure of strategic management research: a bibliometric study of the *Strategic management journal*, 1980-2000". *Strategic management journal*, v. 25, n. 10, pp. 981-1004.
<https://doi.org/10.1002/smj.397>
- Schindehutte, Minet; Morris, Michael; Allen, Jeffrey** (2006). "Beyond achievement: entrepreneurship as extreme experience". *Small business economics*, v. 27, n. 4-5, pp. 349-368.
<https://doi.org/10.1007/s11187-005-0643-6>

- Siegel, Robin; Siegel, Eric; Macmillan, Ian C.** (1993). "Characteristics distinguishing high-growth ventures". *Journal of business venturing*, v. 8, n. 2, pp. 169-180.
[https://doi.org/10.1016/0883-9026\(93\)90018-Z](https://doi.org/10.1016/0883-9026(93)90018-Z)
- Spigel, Ben; Harrison, Richard** (2018). "Toward a process theory of entrepreneurial ecosystems". *Strategic entrepreneurship journal*, v. 12, n. 1, pp. 151-168.
<https://doi.org/10.1002/sej.1268>
- Sternberg, Rolf; Wennekers, Sander** (2005). "Determinants and effects of new business creation using global entrepreneurship monitor data". *Small business economics*, v. 24, n. 3, pp. 193-203.
<https://doi.org/10.1007/s11187-005-1974-z>
- Urbano, David; Aparicio, Sebastian; Audretsch, David** (2019). "Twenty-five years of research on institutions, entrepreneurship, and economic growth: what has been learned?". *Small business economics*, v. 53, n. 1, pp. 21-49.
<https://doi.org/10.1007/s11187-018-0038-0>
- Vallaster, Christine; Kraus, Sascha; Merigó-Lindahl, José M.; Nielsen, Annika** (2019). "Ethics and entrepreneurship: a bibliometric study and literature review". *Journal of business research*, v. 99, pp. 226-237.
<https://doi.org/10.1016/j.jbusres.2019.02.050>
- Willard, Gary E.; Krueger, David A.; Feeser, Henry R.** (1992). "In order to grow, must the founder go: a comparison of performance between founder and non-founder managed high-growth manufacturing firms". *Journal of business venturing*, v. 7, n. 3, pp. 181-194.
[https://doi.org/10.1016/0883-9026\(92\)90025-M](https://doi.org/10.1016/0883-9026(92)90025-M)
- Wong, Poh-Kam; Ho, Yuen-Ping; Autio, Erkki** (2005). "Entrepreneurship, innovation and economic growth: evidence from GEM data". *Small business economics*, v. 24, n. 3, pp. 335-350.
<https://doi.org/10.1007/s11187-005-2000-1>
- Xu, Zeshui; Wang, Xindi; Wang, Xinxin; Skare, Marinko** (2021). "A comprehensive bibliometric analysis of entrepreneurship and crisis literature published from 1984 to 2020". *Journal of business research*, v. 135, pp. 304-318.
<https://doi.org/10.1016/j.jbusres.2021.06.051>
- Zupic, Ivan; Čater, Tomaz** (2015). "Bibliometric methods in management and organization". *Organizational research methods*, v. 18, n. 3, pp. 429-472.
<https://doi.org/10.1177/1094428114562629>

Appendix A. High-growth entrepreneurship (HGE) core research documents list

Article	Cites
Larson, Andrea (1992). "Network dyads in entrepreneurial settings: a study of the governance of exchange relationships". <i>Administrative science quarterly</i> , v. 37, n. 1, pp. 76-104. https://doi.org/10.2307/2393534	1.278
Delmar, Frédéric; Davidsson, Per; Gartner, William B. (2003). "Arriving at the high-growth firm". <i>Journal of business venturing</i> , v. 18, n. 2, pp. 189-216. https://doi.org/10.1016/S0883-9026(02)00080-0	608
Loughran, Tim; Ritter, Jay R. (1997). "The operating performance of firms conducting seasoned equity offerings". <i>The journal of finance</i> , v. 52, n. 5, p. 1823-1850. https://doi.org/10.1111/j.1540-6261.1997.tb02743.x	390
Henrekson, Magnus; Johansson, Dan (2010). "Gazelles as job creators: a survey and interpretation of the evidence". <i>Small business economics</i> , v. 35, pp. 227-244. https://doi.org/10.1007/s11187-009-9172-z	338
Stenholm, Pekka; Acs, Zoltan J.; Wuebker, Robert (2013). "Exploring country-level institutional arrangements on the rate and type of entrepreneurial activity". <i>Journal of business venturing</i> , v. 28, n. 1, p. 176-193. https://doi.org/10.1016/j.jbusvent.2011.11.002	274
McConnell, John J.; Servaes, Henri (1995). "Equity ownership and the two faces of debt". <i>Journal of financial economics</i> , v. 39, n. 1, p. 131-157. https://doi.org/10.1016/0304-405X(95)00824-X	271
Barringer, Bruce R.; Jones, Foard F.; Neubaum, Donald O. (2005). "A quantitative content analysis of the characteristics of rapid-growth firms and their founders". <i>Journal of business venturing</i> , v. 20, n. 5, pp. 663-687. https://doi.org/10.1016/j.jbusvent.2004.03.004	256
Billett, Matthew T.; King, Tao-Hsien-Dolly; Mauer, David C. (2007). "Growth opportunities and the choice of leverage, debt maturity, and covenants". <i>The journal of finance</i> , v. 62, n. 2, p. 697-730. https://doi.org/10.1111/j.1540-6261.2007.01221.x	243
Morris, Michael H.; Miyasaki, Nola N.; Watters, Craig E.; Coombes, Susan M. (2006). "The dilemma of growth: understanding venture size choices of women entrepreneurs". <i>Journal of small business management</i> , v. 44, n. 2, p. 221-244.	225

Gundry, Lisa K.; Welsch, Harold P. (2001). "The ambitious entrepreneur: high growth strategies of women-owned enterprises". <i>Journal of business venturing</i> , v. 16, n. 5, p. 453-470. https://doi.org/10.1016/S0883-9026(99)00059-2	214
Sternberg, Rolf; Wennekers, Sander (2005). "Determinants and effects of new business creation using global entrepreneurship monitor data". <i>Small business economics</i> , v. 24, n. 3, pp. 193-203. https://doi.org/10.1007/s11187-005-1974-z	207
Mason, Colin M.; Brown, Ross (2013). "Creating good public policy to support high-growth firms". <i>Small business economics</i> , v. 40, n. 2, pp. 211-225. https://doi.org/10.1007/s11187-011-9369-9	200
Acs, Zoltan J.; Mueller, Pamela (2008). "Employment effects of business dynamics: mice, gazelles and elephants". <i>Small business economics</i> , v. 30, n. 1, pp. 85-100. https://doi.org/10.1007/s11187-007-9052-3	192
Zacharakis, Andrew L.; Meyer, G. Dale (2000). "The potential of actuarial decision models: can they improve the venture capital investment decision?". <i>Journal of business venturing</i> , v. 15, n. 4, p. 323-346. https://doi.org/10.1016/S0883-9026(98)00016-0	172
Stam, Erik; Wennberg, Karl (2009). "The roles of R&D in new firm growth". <i>Small business economics</i> , v. 33, n. 1, p. 77-89. https://doi.org/10.1007/s11187-009-9183-9	169
Lemmon, Michael L.; Zender, Jaime F. (2010). "Debt capacity and tests of capital structure theories". <i>Journal of financial and quantitative analysis</i> , v. 45, n. 5, p. 1161-1187.	144
Davidsson, Per; Henrekson, Magnus (2002). "Determinants of the prevalence of start-ups and high-growth firms". <i>Small business economics</i> , v. 19, n. 2, p. 81-104. https://doi.org/10.1023/A:1016264116508	132
Goedhuys, Micheline; Sleuwaegen, Leo (2010). "High-growth entrepreneurial firms in Africa: a quantile regression approach". <i>Small business economics</i> , v. 34, n. 1, pp. 31-51. https://doi.org/10.1007/s11187-009-9193-7	126
Siegel, Robin; Siegel, Eric; Macmillan, Ian C. (1993). "Characteristics distinguishing high-growth ventures". <i>Journal of business venturing</i> , v. 8, n. 2, pp. 169-180. https://doi.org/10.1016/0883-9026(93)90018-Z	125
Coad, Alex; Daunfeldt, Sven-Olov; Hözl, Werner; Johansson, Dan; Nightingale, Paul (2014). "High-growth firms: introduction to the special section". <i>Industrial and corporate change</i> , v. 23, n. 1, pp. 91-112. https://doi.org/10.1093/icc/dtt052	124
Miner, John B.; Raju, Nambury S. (2004). "Risk propensity differences between managers and entrepreneurs and between low-and high-growth entrepreneurs: a reply in a more conservative vein". <i>Journal of applied psychology</i> , v. 89, n. 1, p. 3-13. https://doi.org/10.1037/0021-9010.89.1.3	122
Hözl, Werner (2009). "Is the R&D behaviour of fast-growing SMEs different? Evidence from CIS III data for 16 countries". <i>Small business economics</i> , v. 33, n. 1, p. 59-75. https://doi.org/10.1007/s11187-009-9182-x	115
Moreno, Ana M.; Casillas, José C. (2007). "High-growth SMEs versus non-high-growth SMEs: a discriminant analysis". <i>Entrepreneurship and regional development</i> , v. 19, n. 1, pp. 69-88. https://doi.org/10.1080/08985620601002162	113
Vos, Ed; Yeh, Andy-Jia-Yug; Carter, Sara; Tagg, Stephen (2007). "The happy story of small business financing". <i>Journal of banking & finance</i> , v. 31, n. 9, p. 2648-2672. https://doi.org/10.1016/j.jbankfin.2006.09.011	111
Spigel, Ben; Harrison, Richard (2018). "Toward a process theory of entrepreneurial ecosystems". <i>Strategic entrepreneurship journal</i> , v. 12, n. 1, pp. 151-168. https://doi.org/10.1002/sej.1268	111
Parker, Simon C.; Storey, David J.; Van-Witteloostuijn, Arjen (2010). "What happens to gazelles? The importance of dynamic management strategy". <i>Small business economics</i> , v. 35, n. 2, p. 203-226. https://doi.org/10.1007/s11187-009-9250-2	108
Fischer, Eileen; Reuber, A. Rebecca (2003). "Support for rapid-growth firms: a comparison of the views of founders, government policymakers, and private sector resource providers". <i>Journal of small business management</i> , v. 41, n. 4, p. 346-365. https://doi.org/10.1111/1540-627X.00087	105
Rutherford, Matthew W.; Buller, Paul F.; McMullen, Patrick R. (2003). "Human resource management problems over the life cycle of small to medium-sized firms". <i>Human resource management</i> , v. 42, n. 4, p. 321-335. https://doi.org/10.1002/hrm.10093	99
Lerner, Josh (2010). "The future of public efforts to boost entrepreneurship and venture capital". <i>Small business economics</i> , v. 35, n. 3, p. 255-264. https://doi.org/10.1007/s11187-010-9298-z	98
Willard, Gary E.; Krueger, David A.; Feeser, Henry R. (1992). "In order to grow, must the founder go: a comparison of performance between founder and non-founder managed high-growth manufacturing firms". <i>Journal of business venturing</i> , v. 7, n. 3, pp. 181-194. https://doi.org/10.1016/0883-9026(92)90025-M	93
García-Manjón, Juan V.; Romero-Merino, M. Elena (2012). "Research, development, and firm growth. Empirical evidence from European top R&D spending firms". <i>Research policy</i> , v. 41, n. 6, p. 1084-1092. https://doi.org/10.1016/j.respol.2012.03.017	90

Schindehutte, Minet; Morris, Michael; Allen, Jeffrey (2006). "Beyond achievement: entrepreneurship as extreme experience". <i>Small business economics</i> , v. 27, n. 4-5, pp. 349-368. https://doi.org/10.1007/s11187-005-0643-6	79
Baum, J. Robert; Bird, Barbara J. (2010). "The successful intelligence of high-growth entrepreneurs: links to new venture growth". <i>Organization science</i> , v. 21, n. 2, pp. 397-412. https://doi.org/10.1287/orsc.1090.0445	77
Clarysse, Bart; Bruneel, Johan; Wright, Mike (2011). "Explaining growth paths of young technology-based firms: structuring resource portfolios in different competitive environments". <i>Strategic entrepreneurship journal</i> , v. 5, n. 2, p. 137-157. https://doi.org/10.1002/sej.111	72
Heyman, Dries; Deloof, Marc; Ooghe, Hubert (2008). "The financial structure of private held Belgian firms". <i>Small business economics</i> , v. 30, n. 3, p. 301-313. https://doi.org/10.1007/s11187-006-9031-0	66
Autio, Erkkö; Rannikko, Heikki (2016). "Retaining winners: can policy boost high-growth entrepreneurship?". <i>Research policy</i> , v. 45, n. 1, p. 42-55. https://doi.org/10.1016/j.respol.2015.06.002	65
Daunfeldt, Sven-Olov; Halvarsson, Daniel (2015). "Are high-growth firms one-hit wonders? Evidence from Sweden". <i>Small business economics</i> , v. 44, n. 2, p. 361-383. https://doi.org/10.1007/s11187-014-9599-8	64
Littunen, Hannu; Tohmo, Timo (2003). "The high growth in new metal-based manufacturing and business service firms in Finland". <i>Small business economics</i> , v. 21, n. 2, p. 187-200. https://doi.org/10.1023/A:1025014427294	64
Terjesen, Siri; Bosma, Niels; Stam, Eri (2016). "Advancing public policy for high-growth, female, and social entrepreneurs". <i>Public administration review</i> , v. 76, n. 2, p. 230-239. https://doi.org/10.1111/puar.12472	63
Hözl, Werner (2014). "Persistence, survival, and growth: a closer look at 20 years of fast-growing firms in Austria". <i>Industrial and corporate change</i> , v. 23, n. 1, p. 199-231. https://doi.org/10.1093/icc/dtt054	62
Ensley, Michael D.; Pearson, Allison W.; Sardeshmukh, Shruti R. (2007). "The negative consequences of pay dispersion in family and non-family top management teams: an exploratory analysis of new venture, high-growth firms". <i>Journal of business research</i> , v. 60, n. 10, p. 1039-1047. https://doi.org/10.1016/j.jbusres.2006.12.012	62
Feindt, Sylvie; Jeffcoate, Judith; Chappell, Caroline (2002). "Identifying success factors for rapid growth in SME e-commerce". <i>Small business economics</i> , v. 19, n. 1, p. 51-62. https://doi.org/10.1023/A:1016165825476	61
Segarra, Agustí; Teruel, Mercedes (2014). "High-growth firms and innovation: an empirical analysis for Spanish firms". <i>Small business economics</i> , v. 43, n. 4, p. 805-821. https://doi.org/10.1007/s11187-014-9563-7	57
Grimm, Michael; Knorringer, Peter; Lay, Jann (2012). "Constrained gazelles: high potentials in West Africa's informal economy". <i>World development</i> , v. 40, n. 7, pp. 1352-1368. https://doi.org/10.1016/j.worlddev.2012.03.009	56
Lee, Neil (2014). "What holds back high-growth firms? Evidence from UK SMEs". <i>Small business economics</i> , v. 43, n. 1, p. 183-195. https://doi.org/10.1007/s11187-013-9525-5	54
Brown, Ross; Mason, Colin (2014). "Inside the high-tech black box: a critique of technology entrepreneurship policy". <i>Technovation</i> , v. 34, n. 12, pp. 773-784. https://doi.org/10.1016/j.technovation.2014.07.013	54
Aldrich, Howard E. (2012). "The emergence of entrepreneurship as an academic field: a personal essay on institutional entrepreneurship". <i>Research policy</i> , v. 41, n. 7, pp. 1240-1248. https://doi.org/10.1016/j.respol.2012.03.013	52
Friar, John H.; Meyer, Marc H. (2003). "Entrepreneurship and start-ups in the Boston region: factors differentiating high-growth ventures from micro-ventures". <i>Small business economics</i> , v. 21, n. 2, p. 145-152. https://doi.org/10.1023/A:1025045828202	52
López-García, Paloma; Puente, Sergio (2012). "What makes a high-growth firm? A dynamic probit analysis using Spanish firm-level data". <i>Small business economics</i> , v. 39, n. 4, p. 1029-1041. https://doi.org/10.1007/s11187-011-9321-z	52
Beekman, Amy V.; Robinson, Richard B. (2004). "Supplier partnerships and the small, high-growth firm: selecting for success". <i>Journal of small business management</i> , v. 42, n. 1, p. 59-77. https://doi.org/10.1111/j.1540-627X.2004.00097.x	50

Body perception and frequency of exposure to advertising on social networks among adolescents

Beatriz Feijoo; Belén Cambronero-Saiz; Begoña Miguel-San-Emeterio

Nota: Este artículo se puede leer en español en:
<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87224>

Recommended citation:

Feijoo, Beatriz; Cambronero-Saiz, Belén; Miguel-San-Emeterio, Begoña (2023). "Body perception and frequency of exposure to advertising on social networks among adolescents". *Profesional de la información*, v. 32, n. 3, e320318.

<https://doi.org/10.3145/epi.2023.may.18>

Manuscript received on 21th December 2022

Accepted on 10th May 2023



Beatriz Feijoo ✉

<https://orcid.org/0000-0001-5287-3813>

Universidad Internacional de La Rioja
Facultad de Empresa y Comunicación
Avenida de La Paz, 137
26006 Logroño, Spain
beatriz.feijoo@unir.net



Belén Cambronero-Saiz

<https://orcid.org/0000-0001-5273-9534>

Universidad Internacional de La Rioja
Facultad de Empresa y Comunicación
Avenida de La Paz, 137
26006 Logroño, Spain
belen.cambronero@unir.net



Begoña Miguel-San-Emeterio

<https://orcid.org/0000-0003-0272-7476>

Universidad Internacional de La Rioja
Facultad de Empresa y Comunicación
Avenida de La Paz, 137
26006 Logroño, Spain
begona.miguel@unir.net

Abstract

The marketing actions of influencers promoting unhealthy targeted products can aggravate the problem of childhood obesity and thus generate self-esteem problems. The influence of exposure to this type of sponsored content can also have an effect on the emotional well-being of adolescents. This is the context of this study, which seeks to analyze the relationship between exposure to food or body care advertising on social networks (*YouTube*, *Instagram*, and *TikTok*) of adolescents and their perception of their physical appearance. Online surveys were distributed between April and June 2022 to 1,055 adolescents aged 11-17 years and residing in Spain. The main results show that 16- and 17-year-olds have a heavier figure as an ideal reference, but they are also the group that presents the thinnest evaluation of their current figure. Their level of satisfaction is the lowest compared with the rest of the age groups, and they attach greater importance to the social role of a good physical appearance. Thus, exposure to advertising by influencers on social networks is directly related to lower satisfaction with their bodies. This assessment is based not so much on individual reasons related to health or personal well-being, but rather on fundamentally social reasons, and considers that physical appearance is a determining factor for social success. Therefore, it is advisable to study in more detail the beliefs that directly affect adolescents' self-esteem to improve their critical competence in the face of this idealized content.

Keywords

Adolescents; Teenagers; Young people; Childhood; Social media; Social networks; *YouTube*; *Instagram*; *TikTok*; Perception; Personal satisfaction; Self-esteem; Well-being; Physical appearance; Body; Gender; Advertising; Influencers; Diet; Health.



Funding

This study is conducted within the *Digital_Fit* research project: “Body and diet as sales arguments on social networks: Impact of brand content published by influencers on the beliefs of minors on physical appearance and healthy eating habits,” funded by the *Mapfre Foundation*.

1. Introduction

The increased time spent by adolescents in the digital environment, may involve a risk that has an impact on health, both directly and indirectly (*WHO*, 2022). Thus, on the one hand, the increased consumption of content through social networks such as *TikTok*, *YouTube*, and *Instagram*, means less time available for sports practice. On the other hand, the greater number of hours of use also means being impacted to a greater extent by advertising (**González-Díaz**, 2014; *WHO*, 2022). Food products with high levels of saturated fats, salt, and/or sugar are some of the most recurrent advertising categories (**Boyland et al.**, 2016; **Coates et al.**, 2019; **Feijoo; Sádaba; Bugueño**, 2020), which can lead to increased consumption of unhealthy foods by children (**Sadeghirad et al.**, 2016; **Murphy et al.**, 2020; **Reyes-Pedraza; García-González; Téllez-Castilla**, 2020).

The high levels of prevalence of both childhood obesity and being overweight have forced governments around the world to develop surveillance tools to assess and design corrective measures (*WHO*, 2021; *Gobierno de España*, 2022). Although this is an international problem, data indicate higher levels of obesity and being overweight in Mediterranean and Eastern European countries (*WHO*, 2021).

In Spain specifically, the prevalence of childhood obesity stands at 17.3%, while being overweight is at 23.3%; that is, 4 out of 10 children suffer from one of these disorders (*Aesan*, 2019). The causes that have led to the increase in the prevalence of childhood obesity are diverse and multifactorial (**González-Díaz**, 2014; **De-Jongh-González; Escalante-Izeta; Ojeda-García**, 2023) and cover very different areas, including health, education, or the economy (*WHO*, 2022).

Furthermore, there is also the influence of media exposure on the emotional well-being of children. A private research leak from the *Facebook* team has highlighted that simply consuming certain content on social networks influences the mood of younger people: more than 40% of *Instagram* users acknowledged that the feeling of being unattractive started while using the app (**Milmo; Skopeliti**, 2021).

A growing number of studies (**Coates et al.**, 2019; **Tiggemann; Anderberg**, 2020; **De-Jans et al.**, 2021; **Lowe-Calverley; Grieve**, 2021) have shown that publications on social networks and by influencers impact users' choices (**Zozaya; Feijoo; Sádaba**, 2023), both in the perception of their body image and in their dietary patterns. Experts insist on the need to control and regulate algorithms that choose and adapt the photos and videos a user sees according to their usage, potentially creating a spiral of harmful content; others reinforce the idea of educating young people on how to navigate a world dominated by social networks, equipping them with tools to make healthy choices for themselves (**Milmo; Skopeliti**, 2021).

Following *Unicef Spain's* recommendations to address the concerns of children and adolescents and to promote the practice of healthy habits (*Unicef*, 2021), this study seeks to analyze the influence that exposure to food and body care advertising on social networks has on their perceptions of their physical appearance.

The novelty of this research lies in studying the effect that this type of advertising generates on the body perception of adolescents. The survey utilized, distributed to a sample of more than 1000 adolescents, enriches previous findings on children, adolescents, marketing, eating habits, and body image, and is especially focused on content analysis (**Fernández-Gómez; Díaz-del-Campo**, 2014; **González-Oñate; Martínez-Sánchez**, 2020; **Castelló-Martínez; Tur-Viñes**, 2021; **Tur-Viñes; Castelló-Martínez**, 2021; **Feijoo; Fernández-Gómez**, 2021).

1.1. Advertising in the digital environment

The decline in the effectiveness of online advertising in its most classic formats (banners, pop-ups, or skyscrapers) and the growing trend toward the use of tools to block it (ad blockers) prompted the emergence of so-called native advertising, which developed mainly on social networks and is characterized by respecting the editorial style of the platform on which it is added (**Wojdynski; Golan**, 2016; **Gómez-Nieto**, 2018; **Eyada; Milla**, 2020).

Therefore, the delimitation of advertising spaces and formats is not as clear as in advertising disseminated through traditional mass media, which requires that this content be accompanied by a legend mentioning their advertising nature to facilitate their identification as commercial messages (**Wojdynski; Golan**, 2016).

It is in this context that influencers appear, who, according to the definition developed by *Autocontrol*, are

“people with a high level of influence on the public due to their high number of followers on social networks and/or in digital communications media and who interact through tweets, videos, and posts as well as through messages on blogs or other websites” (*Autocontrol*, 2021).

In addition, their main arena of action is social networks.

In the specific case of their advertising, the dividing line between advertising and content becomes even blurrier (*Autocontrol*, 2021; **Tur-Viñes; Núñez-Gómez; González-Río**, 2018; **Feijoo; Pavez**, 2019; **López-Villafranca; Olmedo-Salar**, 2019; **Van-Dam; Van-Reijmersdal**, 2019; **De-Jans; Hudders**, 2020). One of the most common advertising actions among influencers is the use of product placement strategies, in which products are seemingly casually displayed within the content of a post or video (**Alruwaily et al.**, 2020; **Du; Rui; Yu**, 2023).

“ The publications of the influencers in social networks affect the mood of the youngest, both in the perception of their body image and in the choice of their dietary guidelines ”

This implies that the audience may have greater difficulty in perceiving that it is content for commercial purposes (**Tur-Viñes; Núñez-Gómez; González-Río**, 2018; **Feijoo; Pavez**, 2019; **López-Villafranca; Olmedo-Salar**, 2019; **Van-Dam; Van-Reijmersdal**, 2019; **De-Jans; Hudders**, 2020; *Autocontrol*, 2021), with adolescents generally being the most vulnerable to advertising messages owing to their lower critical capacity for identifying marketing strategies and greater propensity to being influenced (*WHO*, 2016; **Coates et al.**, 2019; **Balaban; Mucundorfeanu; Muresan**, 2022).

Marketing by influencers also benefits from the parasocial relationships that are established between the sender and the receiver, which result in influencers' being perceived not as content creators with an economic objective but rather as “close and intimate friends” (**Lim et al.**, 2017; **Meyers et al.**, 2017; **Conde; Casais**, 2023). The stronger such parasocial interaction with influencers, the more likely it is that the audience will perceive authenticity in the messages they convey and interest in the recommendations they make (**Djafarova; Rushworth**, 2017; **Audrezet; De-Kerviler; Moulard**, 2020; **Silva et al.**, 2021).

In a context in which physical appearance is increasingly important and in which influencers consolidate ideal physical traits that then turn them into archetypes of beauty and influence the care of physical health (**Rosara; Luthfia**, 2020), more caution should be taken with these types of self-interested recommendations. Scientific evidence has shown, particularly in the case of adolescents, that they are influenced in terms of not only their eating patterns, but also their body image perception (**Coates et al.**, 2019; **Tiggemann; Anderberg**, 2020; **De-Jans et al.**, 2021; **Lowe-Calverley; Grieve**, 2021). This means that, by promoting unhealthy food products, influencers may be contributing to the aggravation of the childhood obesity problem, as well as related emotional and self-esteem problems (**De-Jans et al.**, 2021).

The aim of this research is to analyze the relationship between exposure to advertising on social networks of adolescents (*YouTube*, *Instagram*, and *TikTok*) and the perception they have regarding their body, as well as the impact on the level of satisfaction with their physical appearance and its social relevance. To this end, the following research questions are posed:

RQ 1. Is there an association between the level of exposure that adolescents have with food or body care advertising on social networks and:

- the perception they have of their bodies and
- their level of satisfaction with their physical appearance and its social relevance?

RQ 2. Do the age and gender of adolescents introduce significant differences around:

- the perception they have of their bodies and
- their level of satisfaction with their physical appearance and its social relevance?

2. Methodology

To collect information, a questionnaire was developed on the basis of previous work (**Lou; Yuan**, 2019; **Smit et al.**, 2020; **Cambroner-Saiz; Segarra-Saavedra; Cristófol-Rodríguez**, 2021; **De-Jans et al.**, 2021), and was given to adolescents between 11 and 17 years of age residing in Spain. A total of 1055 individuals participated, with a 95% confidence level and a $\pm 3\%$ margin of error.

The sampling procedure was multistage, stratified with proportional allocation. It used as first stratum four ad hoc aggregated geographical areas (following the classification of NUT areas used by the EU), and a second level of stratification according to the socioeconomic level of the families (low, medium, and high). The final selection of the individuals to be surveyed was made according to cross quotas of gender and age.

The distribution of the sample according to the adolescents' age is 28.3% aged 11-12 years; 44% aged 13-15 years, and 27.7% aged 16-17 years; while by gender, 53.6% are male, 46.3% female, and 0.1% other. In terms of socioeconomic level, 30.2% were low, 50.4% medium, and 19.3% high. The questionnaire was distributed online through the survio.com platform, which, with the support of a social studies company, enabled the participation of adolescents from all the autonomous communities of Spain. Field work was conducted between April and June 2022.

To safeguard the integrity of the study participants and the researchers, authorization for the collection of information was requested from the minor's guardian by means of a signed informed consent form, previously validated by the *Ethics Committee* of the university to which this research is attached (*Universidad Internacional de la Rioja*), which also reviewed and approved the methodological design of the project.

2.1. Measurement

This study included the measurement of three groups of variables: perception of the adolescent's current and ideal figures; level of satisfaction with their physical appearance and its social relevance; and frequency of exposure to food or body care advertising on social networks (*YouTube*, *Instagram*, and *TikTok*).

1) For the variables relating to the adolescent's perception of their ideal (FIG_ID) and current (FIG_AC) figures, the **Thompson and Gray** (1995) Silhouette Test was used as a reference to establish the body image that they consider ideal and the body image they have of themselves. This scale is composed of nine drawings that identify physical shape, from value 1 as very thin to value 9 as very fat, for both men and women. The calculation of the means indicates that the adolescents surveyed tend to perceive their current figure as fatter than their ideal figure (FIG_ID M = 5.76; FIG_AC M = 5.82).

2) Level of satisfaction with their physical appearance and its social relevance: in this case, the aim was to assess the degree of acceptance adolescents have of their physical appearance and the influence they give to body and appearance in social relations. Thus, the following four questions were formulated, previously tested in similar studies (**Fanjul-Peyró; López-Font; González-Oñate**, 2019):

- "Do you feel satisfied with your body and physical appearance in general?" (ASP_FIS1);
- "Do you think that others' perception of you would improve if your physical appearance changed?" (ASP_FIS2);
- "Do you think that having a good physique helps you to be more successful (more friends, be liked more, be more accepted, etc.)?" (ASP_FIS3);
- "How important do you think your friends think your physical appearance is?" (ASP_FIS4),

We used a five-point Likert scale in which 1 = not at all and 5 = very much. The averages were as follows: ASP_FIS1 M = 3.68; ASP_FIS2 M = 2.76; ASP_FIS3 M = 3.32; and ASP_FIS4 M = 3.38, indicating that in general their degree of satisfaction with their body is higher than the impact that physical appearance can have in society.

3) Frequency of exposure to food or body care advertising on social networks (EXP): in this case, the frequency with which adolescents receive advertising about food or body care on *YouTube* (EXP_YT), *Instagram* (EXP_IG), and *TikTok* (EXP_TK) was assessed. The following question was asked: "How often do you receive food or body care advertising on the following platforms?" The frequency of exposure to advertising was measured using a six-point Likert scale, where 1 = never (because I do not use this platform); 2 = never; 3 = infrequent; 4 = somewhat frequent; 5 = frequent; and 6 = very frequent. The calculation of averages reflects that adolescents perceive more advertising on this topic on *YouTube*, followed by *TikTok*, and in third place, *Instagram* (EXP_YT M = 4.04; EXP_TK M = 3.51; EXP_IG M = 3.27).

Statistical analysis was performed using the *SPSS* software package version 25.0. To answer question 1, bivariate analyses were performed using Pearson's tests to analyze the relationship between exposure to advertising, perception of their bodies, and level of satisfaction with their bodies. regarding RQ2, the Anova test was applied to determine whether there is a relationship of dependence between the variables analyzed and the filter variables (gender and age).

3. Results

3.1. Relationship between exposure to social media advertising and body perception and satisfaction

3.1.1. Body perception and exposure to advertising on social networks

Bivariate analyses using the overall sample revealed no significant correlations between adolescents' perception of ideal/actual figure and their exposure to food and body care advertising on *YouTube*, *Instagram*, and *TikTok* ($p > 0.05$), as reported in Table 1.

However, what was observed is that there is a positive and relevant association between the respondents' ideal and real perceptions of their body ($p < 0.01$). Although exposure to advertising on social networks does not affect the perception of the ideal figure, the variables defined to measure the social relevance given to physical appearance do; thus, the ideal perception correlates negatively with the importance of physique in the perception of others ($p < 0.01$), with the perception that success goes hand in hand with a good physique ($p < 0.01$), and with the importance given to physical appearance by their friends ($p < 0.01$). However, the current perception variable correlates negatively with the level of satisfaction with their body ($p < 0.05$).

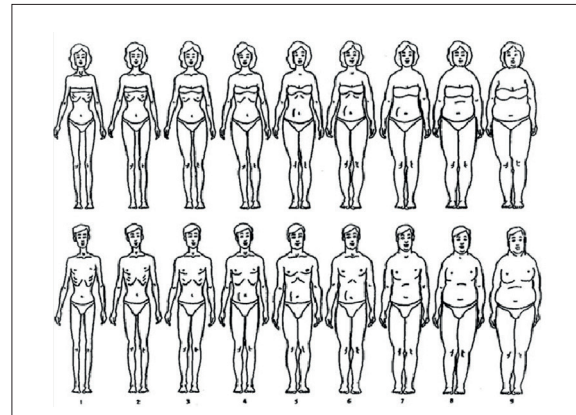


Figure 1. **Thompson and Gray** Silhouette Test (1995)

Exposure to advertising on social networks influences the level of satisfaction that minors express with their own bodies, more pronouncedly among those between 16 and 17 years of age

Table 1. Association between body perception, level of satisfaction with physical appearance, social relevance, and exposure to advertising on social networks

		FIG_ID	FIG_AC	ASP_FIS1	ASP_FIS2	ASP_FIS3	ASP_FIS4	EXP_YT	EXP_IG	EXP_TK
FIG_ID	Pearson's correlation	1	0.420**	0.004	-0.084**	-0.108**	-0.121**	-0.056	-0.044	-0.016
	Sig. (bilateral)		0.000	0.902	0.006	0.000	0.000	0.068	0.149	0.601
FIG_AC	Pearson's correlation	0.420**	1	-0.075*	0.036	-0.045	-0.017	-0.020	-0.047	-0.030
	Sig. (bilateral)	0.000		0.014	0.238	0.144	0.583	0.524	0.129	0.330
ASP_FIS1	Pearson's correlation	0.004	-0.075*	1	-0.335**	-0.201**	-0.111**	-0.086**	-0.100**	-0.139**
	Sig. (bilateral)	0.902	0.014		0.000	0.000	0.000	0.005	0.001	0.000
ASP_FIS2	Pearson's correlation	-0.084**	0.036	-0.335**	1	0.546**	0.427**	0.159**	0.209**	0.205**
	Sig. (bilateral)	0.006	0.238	0.000		0.000	0.000	0.000	0.000	0.000
ASP_FIS3	Pearson's correlation	-0.108**	-0.045	-0.201**	0.546**	1	0.529**	0.125**	0.206**	0.220**
	Sig. (bilateral)	0.000	0.144	0.000	0.000		0.000	0.000	0.000	0.000
ASP_FIS4	Pearson's correlation	-0.121**	-0.017	-0.111**	0.427**	0.529**	1	0.150**	0.182**	0.187**
	Sig. (bilateral)	0.000	0.583	0.000	0.000	0.000		0.000	0.000	0.000
EXP_YT	Pearson's correlation	-0.056	-0.020	-0.086**	0.159**	0.125**	0.150**	1	0.235**	0.357**
	Sig. (bilateral)	0.068	0.524	0.005	0.000	0.000	0.000		0.000	0.000
EXP_IG	Pearson's correlation	-0.044	-0.047	-0.100**	0.209**	0.206**	0.182**	0.235**	1	0.489**
	Sig. (bilateral)	0.149	0.129	0.001	0.000	0.000	0.000	0.000		0.000
EXP_TK	Pearson's correlation	-0.016	-0.030	-0.139**	0.205**	0.220**	0.187**	0.357**	0.489**	1
	Sig. (bilateral)	0.601	0.330	0.000	0.000	0.000	0.000	0.000	0.000	

* Correlation significant at the 0.05 level (bilateral).

** Correlation significant at the 0.01 level (bilateral).

3.1.2. Level of satisfaction with physical appearance, social relevance, and exposure to advertising on social networks

In this case, bivariate analyses show significant correlations between the level of satisfaction expressed by adolescents regarding their bodies and exposure to advertising on social networks (Table 1). A negative association was observed between the satisfaction variable and exposure to *YouTube* ($p < 0.01$), *Instagram* ($p < 0.01$), and *TikTok* ($p < 0.01$) variables, with the strongest association being with that which is seen on *TikTok*.

Likewise, all the variables designed to measure the social relevance that adolescents give to their bodies—the importance of physique in the perception of others (ASP_FIS2); the level of social success according to physique (ASP_FIS3); the relevance that friends give to a good physique (ASP_FIS4)—correlate positively ($p < 0.01$) with the variables of advertising exposure on *YouTube*, *Instagram*, and *TikTok*. Across all variables, the correlation is strongest on *Instagram* and *TikTok*.

Furthermore, it is also interesting to note that the level of satisfaction shows a negative association ($p < 0.01$) with the social relevance of the body variables (ASP_FIS2; ASP_FIS3; ASP_FIS4).

3.2. Differences by gender and age

If the gender-differentiated means are compared, it can be seen that girls tend to choose a fatter ideal figure as a reference than boys, something that is also registered in the choice of figure that represents their current state (FIG_AC), as presented in Table 2. Likewise, boys reported a higher level of satisfaction with their bodies than girls and also perceive to a greater extent than girls that physique influences the perception that others have of us and that the level of social success is commensurate with the physique we have. Girls, however, are more likely than boys to believe that body care is important to their friends.

Table 2. Body perception, level of satisfaction with physical appearance, and social relevance according to gender

	Male (N = 566)		Female (N = 488)		Total (N = 1,054)	
	Average	Deviation	Average	Deviation	Average	Deviation
FIG_ID	5.54	8.930	6.02	14.846	5.76	12.032
FIG_AC	5.52	9.020	6.16	14.224	5.82	11.719
ASP_FIS1	3.73	0.883	3.63	1.003	3.69	0.942
ASP_FIS2	2.78	1.083	2.73	1.217	2.76	1.147
ASP_FIS3	3.36	1.121	3.28	1.199	3.32	1.158
ASP_FIS4	3.33	1.113	3.44	1.076	3.38	1.097

However, the Anova test did not show significant differences according to gender, as presented in Table 3.

Table 3. Anova test to test the level of significance between gender and the variables of body perception, level of satisfaction with physical appearance, and social relevance

			Sum of squares	gl	Root mean square	F	Sig.
FIG_ID * Gender	Between groups	(Combined)	58.437	1	58.437	0.403	0.525
FIG_AC * Gender	Between groups	(Combined)	108.758	1	108.758	0.792	0.374
ASP_FIS1 * Gender	Between groups	(Combined)	3.053	1	3.053	3.452	0.063
ASP_FIS2 * Gender	Between groups	(Combined)	0.509	1	0.509	0.387	0.534
ASP_FIS3 * Gender	Between groups	(Combined)	1.531	1	1.531	1.143	0.285
ASP_FIS4 * Gender	Between groups	(Combined)	3.149	1	3.149	2.620	0.106

When the sample is differentiated by age, it can be observed that the oldest participants (16-17 years old) have a fatter figure as an ideal reference, however, this is also the group that presents a thinner evaluation of their current figure. In addition, their level of satisfaction is the lowest compared with the rest of the age groups. Along the same lines, the oldest group is the one that gives more importance to the social role of a good physical appearance, while the youngest group (11-12 years old) is the one that gives less value to these variables (Table 4).

Table 4. Body perception, level of satisfaction with physical appearance, and social relevance according to age

	11-12 years (N = 299)		13-15 years (N = 464)		16-17 years (N = 292)		Total (N = 1,055)	
	Average	Deviation	Average	Deviation	Average	Deviation	Average	Deviation
FIG_ID	5.86	12.266	5.47	10.814	6.12	13.546	5.76	12.027
FIG_AC	6.00	12.310	6.23	13.184	4.98	8.013	5.82	11.714
ASP_FIS1	3.77	0.978	3.66	0.948	3.63	0.901	3.68	0.945
ASP_FIS2	2.62	1.151	2.78	1.185	2.87	1.072	2.76	1.148
ASP_FIS3	3.04	1.198	3.34	1.169	3.58	1.034	3.32	1.158
ASP_FIS4	3.16	1.202	3.43	1.072	3.52	0.993	3.38	1.098

As a function of age, certain significant differences are observed in the variables defined to assess social relevance of the body (ASP_FIS2; ASP_FIS3; ASP_FIS4), as presented in Table 5.

Table 5. ANOVA test to test the level of significance between age and the variables of body perception, level of satisfaction with their physical appearance, and its social relevance

			Sum of squares	gl	Root mean square	F	Sig.
FIG_ID * Age	Between groups	(Combined)	80.373	2	40.186	0.277	0.758
FIG_AC * Age	Between groups	(Combined)	293.889	2	146.945	1.071	0.343
ASP_FIS1 * Age	Between groups	(Combined)	3.296	2	1.648	1.849	0.158
ASP_FIS2 * Age	Between groups	(Combined)	10.110	2	5.055	3.854	0.021
ASP_FIS3 * Age	Between groups	(Combined)	42.694	2	21.347	16.376	0.000
ASP_FIS4 * Age	Between groups	(Combined)	21.204	2	10.602	8.929	0.000

4. Discussion

Although the body positive movement, which promotes the acceptance and normalization of body diversity (Cohen; Newton-John; Slater, 2021), is currently gaining more and more strength on social networks, there are still many representations on social networks that continue to be based on traditional models in which the archetype of the ideal body has attributes such as youth and thinness.

Furthermore, continuing to utilize these types of unattainable ideals has led to a high number of people reporting that they do not feel attractive after having started to use social networks (Markey; Daniels, 2022; Fioravanti *et al.*, 2022).

In this sense, the results of our work have confirmed that there are significant correlations between the level of exposure of adolescents to food or body care advertising on social networks and the level of satisfaction with their physical appearance (RQ 1). Specifically, the data show that exposure to social networks increases the dissatisfaction that adolescents have with their own bodies, which, as previous studies have shown, can lead to the development of mental health problems such as depressive disorders (Murray *et al.*, 2023), self-esteem problems due to feeling that they do not fit in with the canons of beauty established by society (De-Jans *et al.*, 2021) and agreed upon by social networks, and eating disorders (Cordero *et al.*, 2022; Pink *et al.*, 2022; Sanzari *et al.*, 2023).

“ When comparing ideal and real figure, girls consider that they are above the ideal weight, while boys consider that they are below ”

In terms of level of exposure and social relevance, we also observed that, the greater the exposure to advertising on social networks such as *YouTube*, *Instagram*, and *TikTok*, the more important they believe physical appearance is to others, both to friends and to people with whom they have a more distant or no relationship.

There is no association between exposure to advertising and perception of one's own body

In addition, they also consider that people who are closer to the established ideal of beauty will have a higher level of social success. Although this relationship was found in both genders, boys reported having this perception more frequently than girls.

In contrast to previous studies, no significant differences were found between the beauty ideals of boys and girls when analyzing the results by gender (RQ 2). However, in the case of girls, the ideal figure corresponds to one that is comparatively thinner than in the case of the figure chosen by boys. Furthermore, when it comes to comparing ideal and actual figures, girls consider themselves to be above the ideal weight, while boys consider themselves to be below.

This once again highlights the need to incorporate the gender perspective both in the analysis of health problems and in the proposal of interventions (Álvarez-Díaz, 2020), since gender is a key structural determinant (WHO, 2018) that seeks the identification and consideration of differences in the socialization of women and men. In addition, it is also considered a pillar on which the development of values, attitudes, and behaviors is based (Borrell; Artazcoz, 2008).

Significant differences in the level of satisfaction with physical appearance and its social relevance (RQ 2) are seen when it comes to age: it was found that the level of satisfaction is lower in the group between 16 and 17 years of age, and that in turn, this group is the one that attaches the greatest importance to the social role of having a good physical appearance. Therefore, although they have greater cognitive development to critically process advertising content, they are the most vulnerable audience to this type of advertising exposure on social networks.

5. Conclusions

The results of our study show that exposure to advertising by influencers on social networks such as *YouTube*, *Instagram*, and *TikTok* does not influence adolescents' perception of their own bodies, but it is directly related to lower body satisfaction.

This assessment is based not so much on individual reasons related to health or personal well-being, but mainly on social reasons, which increase in importance with age as the participants consider that physical appearance can affect the consideration that others have of them and be a determining factor for social success.

Among the main limitations of the study are that the methodological tool used does not allow us to delve into the impact that advertising by influencers on social networks has on the choice of products, nor does it evaluate the advertised products' nutritional quality or the usefulness of care, so it would be interesting to complement the results of this work with other qualitative methodologies or with a content analysis of the products advertised.

In addition, given these findings, we believe that it would be interesting for future lines of research to conduct a study that allows us to delve into the impact that the physical appearance of the influencer may have on adolescents' intention to purchase food and body care products, as well as to delve into the beliefs that condition their body satisfaction.

Finally, one of the main contributions of the study is the establishment of a tested and updated quantitative methodology that allows for its replication in the future in other countries and regions, with the aim of evaluating the effectiveness of possible actions to improve critical competencies in the face of idealized content and that which has a persuasive intentionality, as recommended in the *European regional report on obesity* (WHO, 2022).

6. References

Aesan (2019). *Estudio Aladino*.

https://www.aesan.gob.es/AECOSAN/web/nutricion/detalle/aladino_2019.htm

Alruwaily, Amaal; Mangold, Chelsea; Greene, Tenay; Arshonsky, Josh; Cassidy, Omni; Pomeranz, Jennifer L.; Bragg, Marie (2020). "Child social media influencers and unhealthy food product placement". *Pediatrics*, v. 146, n. 5, pp. 2-10. <https://doi.org/10.1542/peds.2019-4057>

Álvarez-Díaz, Jorge A. (2020). "La necesaria perspectiva de género para el análisis de problemas de salud". *Cirugía y cirujanos*, v. 88, n. 3, pp. 383-388. <http://doi.org/10.24875/ciru.19000865>

Audrezet, Alice; De-Kerviler, Gwarlann; Moulard, Julie-Guidry (2020). "Authenticity under threat: when social media influencers need to go beyond self-presentation". *Journal of business research*, v. 117, pp. 557-569. <http://doi.org/10.1016/j.jbusres.2018.07.008>

Autocontrol (2021). *Código de conducta sobre el uso de influencers en la publicidad*.

<https://www.autocontrol.es/wp-content/uploads/2020/10/codigo-de-conducta-publicidad-influencers.pdf>

Balaban, Delia-Cristina; Mucundorfeanu, Meda; Mureșan, Larisa-Ioana (2022). "Adolescents' understanding of the model of sponsored content of social media influencer *Instagram* stories". *Media and communication*, v. 10, n. 1, pp. 305-316.

<https://doi.org/10.17645/mac.v10i1.4652>

Borrell, Carme; Artazcoz, Lucía (2008). "Las desigualdades de género en salud: retos para el futuro". *Revista española de salud pública*, v. 82, n. 3, pp. 241-249.

<https://doi.org/10.1590/s1135-57272008000300001>

Boyland, Emma J.; Nolan, Sarah; Kelly, Bridget; Tudur-Smith, Catrin; Jones, Andrew; Halford, Jason C.; Robinson, Eric (2016). "Advertising as a cue to consume: a systematic review and meta-analysis of the effects of acute exposure to unhealthy food and nonalcoholic beverage advertising on intake in children and adults". *The American journal of clinical nutrition*, v. 103, n. 2, pp. 519-533.

<https://doi.org/10.3945/ajcn.115.120022>

Cambronero-Saiz, Belén; Segarra-Saavedra, Jesús; Cristófol-Rodríguez, Carmen (2021). "Análisis desde la perspectiva de género del engagement de los principales youtubers de divulgación científica". *Cuestiones de género: de la igualdad y la diferencia*, n. 16, pp. 511-525.

<https://doi.org/10.18002/cg.v0i16.6914>

Castelló-Martínez, Araceli; Tur-Viñes, Victoria (2021). "Una combinación de alto riesgo: obesidad, marcas de alimentación, menores y retos en *YouTube*". *Gaceta sanitaria*, v. 35, n. 4, pp. 352-354.

<https://doi.org/10.1016/j.gaceta.2020.06.018>

Coates, Anna-Elizabeth; Hardman, Charlotte-Alice; Grovenor-Halford, Jason-Christian; Christiansen, Paul; Boyland, Emma-Jane (2019). "The effect of influencer marketing of food and a 'protective' advertising disclosure on children's food intake". *Pediatric obesity*, v. 14, n. 10, e12540.

<https://doi.org/10.1111/ijpo.12540>

Cohen, Rachel; Newton-John, Toby; Slater, Amy (2021). "The case for body positivity on social media: perspectives on current advances and future directions". *Journal of health psychology*, v. 26, n. 13, pp. 2365-2373.

<http://doi.org/10.1177/1359105320912450>

Conde, Rita; Casais, Beatriz (2023). "Micro, macro and mega-influencers on Instagram: the power of persuasion via the parasocial relationship". *Journal of business research*, v. 158, 113708.

<https://doi.org/10.1016/j.jbusres.2023.113708>

Cordero, Christina; Pulgaron, Elizabeth; Marchante-Hoffman, Ashley; Llabre, María; Perreira, Krista, Sotres-Álvarez, Daniela; Isasi, Carmen; Elder, John; Delamater, Alan (2022). "Body image and disordered eating behaviors in Hispanic/Latino youth: findings from the Hispanic community health study/study of Latino youth". *Appetite*, v. 175, 106079.

<https://doi.org/10.1016/j.appet.2022.106079>

De-Jans, Steffi; Hudders, Liselot (2020). "Disclosure of vlog advertising targeted to children". *Journal of interactive marketing*, v. 52.

<https://doi.org/10.1016/j.intmar.2020.03.003>

De-Jans, Steffi; Spielvogel, Ines; Naderer, Brigitte; Hudders, Liselot (2021). "Digital food marketing to children: How an influencer's lifestyle can stimulate healthy food choices among children". *Appetite*, v. 162, 105182.

<https://doi.org/10.1016/j.appet.2021.105182>

De-Jongh-González, Olivia; Escalante-Izeta, Ericka-Ileana; Ojeda-García, Angélica (2023). "Comprendiendo la obesidad infantil como una condición bio-psico-social". *Revista cubana de psicología*, v. 4, n. 6.

<https://revistas.uh.cu/psicocuba/article/view/311>

Djafarova, Elmira; Rushworth, Chloe (2017). "Exploring the credibility of online celebrities' *Instagram* profiles in influencing the purchase decisions of young female users". *Computers in human behavior*, v. 68.

<https://doi.org/10.1016/j.chb.2016.11.009>

Du, Yuetong; Rui, J. Raymond; Yu, Nan (2023). "How parasocial relationship and influencer-product congruence shape audience's attitude towards product placement in online videos: the mediation role of reactance". *Psychology research and behavior management*, v. 16, pp. 1315-1329.

<https://doi.org/10.2147/PRBM.S406558>

Eyada, Bassant; Milla, Asli (2020). "Native advertising: challenges and perspectives". *Journal of design sciences and applied arts*, v. 1, n. 1, pp. 67-77.

<https://doi.org/10.21608/JDSAA.2020.70451>

- Fanjul-Peyró, Carlos; López-Font, Lorena; González-Oñate, Cristina** (2019). "Adolescentes y culto al cuerpo: influencia de la publicidad y de internet en la búsqueda del cuerpo masculino idealizado". *Doxa comunicación*, v. 29, pp. 61-74.
<https://doi.org/10.31921/doxacom.n29a3>
- Feijoo, Beatriz; Fernández-Gómez, Erika** (2021). "Niños y niñas influyentes en *YouTube* e *Instagram*: contenidos y presencia de marcas durante el confinamiento". *Cuadernos.info*, n. 49, pp. 302-330.
<https://doi.org/10.7764/cdi.49.27309>
- Feijoo, Beatriz; Pavez, Isabel** (2019). "Audiovisual content with advertising intention in children's videos on *YouTube*: the case of the *Soy Luna* series". *Communication & society*, v. 32, n. 1, pp. 313-331.
<https://doi.org/10.15581/003.32.1.313-331>
- Feijoo, Beatriz; Sádaba, Charo; Bugueño, Simón** (2020). "Anuncios entre vídeos, juegos y fotos. Impacto publicitario que recibe el menor a través del teléfono móvil". *Profesional de la información*, v. 29, n. 6, e290630.
<https://doi.org/10.3145/epi.2020.nov.30>
- Fernández-Gómez, Erika; Díaz-Campo, Jesús** (2014). "La publicidad de alimentos en la televisión infantil en España: promoción de hábitos de vida saludables". *Observatorio*, v. 8, n. 4, pp. 133-150.
<https://www.doi.org/10.15847/obsOBS842014802>
- Fioravanti, Giulia; Bocci-Benucci, Sara; Ceragioli, Giulia; Casale, Silvia** (2022). "How the exposure to beauty ideals on social networking sites influences body image: A systematic review of experimental studies". *Adolescent research review*, v. 7, n. 3, pp. 419-458.
<https://doi.org/10.1007/s40894-022-00179-4>
- Gobierno de España (2022). *Plan Estratégico Nacional para la reducción de la obesidad infantil (2022-2030)*. En *Plan bien. Resumen ejecutivo*.
https://www.lamoncloa.gob.es/presidente/actividades/Documents/2022/100622-plan-estrategico-nacional-reduccion-obesidad-infantil_en-plan-bien.pdf
- Gómez-Nieto, Begoña** (2018). "El influencer: herramienta clave en el contexto digital de la publicidad engañosa". *Me-thaodos revista de ciencias sociales*, v. 6, n. 1.
<https://doi.org/10.17502/m.rcs.v6i1.212>
- González-Díaz, Cristina** (2014). "La publicidad dirigida a niños en el sector de la alimentación: un estudio atendiendo al tipo de producto". *Historia y comunicación social*, v. 18, pp. 175-187.
https://doi.org/10.5209/rev_hics.2013.v18.44235
- González-Oñate, Cristina; Martínez-Sánchez, Adela** (2020). "Estrategia y comunicación en redes sociales: Un estudio sobre la influencia del movimiento *RealFooding*". *Ámbitos. Revista internacional de comunicación*, n. 48, pp. 79-101.
<https://doi.org/10.12795/Ambitos.2020.i48.05>
- Lim, Xin-Jean; Radzol, Aifa-Rozaini; Cheah, Jun-Hwa; Wong, Mun-Wai** (2017). "The impact of social media influencers on purchase intention and the mediation effect of customer attitude". *Asian journal of business research*, v. 7, n. 2, pp. 19-36.
<https://doi.org/10.14707/ajbr.170035>
- López-Villafranca, Paloma; Olmedo-Salar, Silvia** (2019). "Menores en *YouTube*, ¿ocio o negocio? Análisis de casos en España y EUA". *Profesional de la información*, v. 28, n. 5, e280520.
<https://doi.org/10.3145/epi.2019.sep.20>
- Lou, Chen; Yuan, Shupe** (2019). "Influencer marketing: how message value and credibility affect consumer trust of branded content on social media". *Journal of interactive advertising*, v. 19, n. 1, pp. 58-73.
<http://doi.org/10.1080/15252019.2018.1533501>
- Lowe-Calverley, Emily; Grieve, Rachel** (2021). "Do the metrics matter? An experimental investigation of *Instagram* influencer effects on mood and body dissatisfaction". *Body image*, v. 36.
<https://doi.org/10.1016/j.bodyim.2020.10.003>
- Markey, Charlotte H.; Daniels, Elizabeth A.** (2022). "An examination of preadolescent girls' social media use and body image: Type of engagement may matter most". *Body image*, v. 42, pp. 145-149.
<https://doi.org/10.1016/j.bodyim.2022.05.005>
- Meyers, Cynthia B.** (2017). "Social media influencers: a lesson plan for teaching digital advertising media literacy". *Advertising & society quarterly*, v. 18, n. 2, pp.
<https://doi.org/10.1353/asr.2017.0018>

- Milmo, Dan; Skopeliti, Clea** (2021). "Teenage girls, body image and *Instagram's* 'perfect storm'". *The Guardian*, 18 September.
<https://www.theguardian.com/technology/2021/sep/18/teenage-girls-body-image-and-instagrams-perfect-storm>
- Murphy, Gráinne; Corcoran, Ciara; Tatlow-Golden, Mimi; Boyland, Emma; Rooney, Brendan** (2020). "See, like, share, remember: adolescents' responses to unhealthy-, healthy- and non-food advertising in social media". *International journal of environmental research and public health*, v. 17, n. 7, pp. 21-81.
<https://doi.org/10.3390/ijerph17072181>
- Murray, Marisa A.; Obeid, Nicole; Gunnell, Katie E.; Buchholz, Annick, Flament; Martine F.; Goldfield, Gary S.** (2023). "Appearance satisfaction mediates the relationship between recreational screen time and depressive symptoms in adolescents". *Child and adolescent mental health*, v. 28, n. 1, pp. 12-21.
<https://doi.org/10.1111/camh.12576>
- OMS (2018). *Género y salud*. Organización Mundial de la Salud.
<https://www.who.int/es/news-room/fact-sheets/detail/gender>
- Pink, Aimee E.; Lim, Phoebe X. H.; Sim, Aaron Y.; Cheon, Bobby K.** (2022). "The effects of acute social media exposure on body dissatisfaction and eating behavior of male and female students". *Journal of social and clinical psychology*, v. 41, n. 4, pp. 365-397.
<https://doi.org/10.1521/jscp.2022.41.4.365>
- Reyes-Pedraza, María-Eugenia; García-González, Janet; Téllez-Castilla, María-Delia** (2018). "Impacto de la publicidad en los hábitos alimenticios en los niños". *Revista española de comunicación en salud*, v. 9, n. 2, pp. 116-126.
<https://doi.org/10.20318/recs.2018.4490>
- Rosara, Nadia-Annisa; Luthfia, Amia** (2020). "Factors influencing consumer's purchase intention on beauty products in YouTube". *Journal of distribution science*, v. 18, n. 6, pp. 37-46.
<https://doi.org/10.15722/jds.18.6.202006.37>
- Sadeghirad, Behnam; Duhaney, Tara; Motaghipisheh, Shahrzad; Campbell, Norm R. C.; Johnston, Bradley C.** (2016). "Influence of unhealthy food and beverage marketing on children's dietary intake and preference: a systematic review and meta-analysis of randomized trials". *Obesity reviews*, v. 17, n. 10, pp. 945-959.
<https://doi.org/10.1111/obr.12445>
- Sanzari, Christina M.; Gorrell, Sasha; Anderson, Lisa M.; Reilly, Erin E.; Niemiec, Martha A.; Orloff, Natalia C.; Anderson, Drew A.; Hormes, Julia M.** (2023). "The impact of social media use on body image and disordered eating behaviors: Content matters more than duration of exposure". *Eating behaviors*, 101722.
<https://doi.org/10.1016/j.eatbeh.2023.101722>
- Silva, Marianny-Jessica-De Brito; De-Farias, Salomão-Alencar; Grigg, Michelle-Helena-Kovacs; Barbosa, María-de-Lourdes-de-Azevedo** (2021). "The body as a brand in social media: analyzing digital fitness influencers as product endorsers". *Athenea digital*, v. 21, n. 1.
<https://doi.org/10.5565/rev/athenea.2614>
- Smit, Crystal R.; Buijs, Laura; Van-Woudenberg, Thabo J.; Bevelander, Kirsten E.; Buijzen, Moniek** (2019). "The impact of social media influencers on children's dietary behaviors". *Frontiers in psychology*, v. 10, 2975.
<https://doi.org/10.3389/fpsyg.2019.02975>
- Thompson, Marjorie A.; Gray, James J.** (1995). "Development and validation of a new body-image assessment scale". *Journal of personality assessment*, v. 64, n. 2, pp. 258-269.
https://doi.org/10.1207/s15327752jpa6402_6
- Tiggemann, Marika; Anderberg, Isabella** (2020). "Muscles and bare chests on *Instagram*: the effect of influencers' fashion and fitspiration images on men's body image". *Body image*, v. 35, pp. 237-244.
<https://doi.org/10.1016/j.bodyim.2020.10.001>
- Tur-Viñes, Victoria; Castelló-Martínez, Araceli** (2021). "Food brands, *YouTube* and children: media practices in the context of the PAOS selfregulation code". *Communication & society*, v. 34, n. 2, pp. 87-105.
<https://doi.org/10.15581/003.34.2.87-105>
- Tur-Viñes, Victoria; Núñez-Gómez, Patricia; González-Río, María-José** (2018). "Menores influyentes en *YouTube*. Un espacio para la responsabilidad". *Revista latina de comunicación social*, n. 73, pp. 1211-1230.
<https://doi.org/10.4185/RLCS-2018-1303>
- Unicef (2021). *¿Qué opinan los niños, niñas y adolescentes? Resultados de la segunda edición del Barómetro de opinión de infancia y adolescencia 2020-2021*.
<https://www.unicef.es/sites/unicef.es/files/comunicacion/Barometro%20infancia%202021.pdf>

Van-Dam, Sophia; Van-Reijmersdal, Eva (2019). "Insights in adolescents' advertising literacy, perceptions and responses regarding sponsored influencer videos and disclosures". *Cyberpsychology: journal of psychosocial research on cyberspace*, v. 13, n. 2, article 2.

<http://doi.org/10.5817/CP2019-2-2>

WHO (2016). *Tackling food marketing to children in a digital world: trans-disciplinary perspectives: children's rights, evidence of impact, methodological challenges, regulatory options and policy implications for the WHO European Region*. World Health Organization. Regional Office for Europe.

<https://apps.who.int/iris/handle/10665/344003>

WHO (2021). *European childhood obesity surveillance initiative (COSI). Report on the fourth round of data collection, 2015-2017*. World Health Organization. Regional Office for Europe. ISBN: 978 92 890 5773 8

<https://apps.who.int/iris/bitstream/handle/10665/341189/WHO-EURO-2021-2495-42251-58349-eng.pdf>

WHO (2022). *European regional obesity report*. World Health Organization. Regional Office for Europe.

<https://apps.who.int/iris/bitstream/handle/10665/353747/9789289057738-eng.pdf>

Wojdyski, Bartosz W.; Golan, Guy J. (2016). "Native advertising and the future of mass communication". *American behavioral scientist*, v. 60, n. 12, pp. 1403-1407.

<http://doi.org/10.1177/0002764216660134>

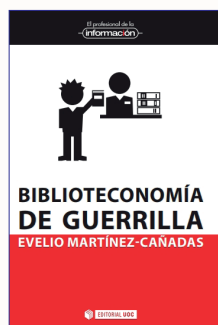
Zozaya, Luisa; Feijoo, Beatriz; Sádaba, Charo (2023). "El papel de los influencers en las decisiones de consumo de los menores españoles". *Doxa comunicación*, n. 36, pp. 401-413.

<https://doi.org/10.31921/doxacom.n36a1685>

Colección de libros de bolsillo

El profesional de la información (Editorial UOC)

Últimos títulos publicados



Media

Publishing

Editorial UOC



Más información:

<https://www.profesionaldelainformacion.com/producto/libros-epi-uoc/>

Political communication on social media in Latin America: unequal use of *Twitter* by members of parliament

Michael Haman

Nota: Este artículo se puede leer en español en:
<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87242>

Recommended citation:

Haman, Michael (2023). "Political communication on social media in Latin America: unequal use of *Twitter* by members of parliament". *Profesional de la información*, v. 32, n. 3, e320316.

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

Article received on December 31st 2022

Approved on May 17th 2023



Michael Haman

<https://orcid.org/0000-0001-5772-2045>

University of Hradec Králové
Philosophical Faculty
Department of Political Science
Rokitanského 62/26
500 03 Hradec Králové, Czech Republic
michael.haman@uhk.cz

Abstract

This article focuses on the under-researched topic of the use of *Twitter* by members of parliament (MPs) in Latin America. There have not yet been any thorough comparative studies on this topic, and the majority of publications on the region focus on single case studies. Previous studies have primarily concentrated on presidents, or in the case of MPs, on specific case studies. A total of 2,353,138 tweets were accessed via the *Twitter* API, and 3,215 MPs were examined. Regression models and correlations were used to answer research questions, and the main variables examined concerned individual characteristics of MPs (gender and age) and socioeconomic indicators of the country (number of people on *Twitter*, internet access, Human Development Index - HDI). As a result, this paper offers a report on how MPs in Latin America are currently utilizing *Twitter*. While this social network is used by more than 90% of MPs in some countries (Argentina, Colombia, Ecuador, El Salvador, Peru, and Uruguay), there are countries where fewer than half of MPs use it (Bolivia, Honduras, and Nicaragua). The results show that female MPs are more likely than male MPs to use *Twitter*. In addition, *Twitter* is being adopted more by younger MPs. Other results show that country characteristics such as internet penetration, *Twitter* population, and the HDI are significant predictors regarding the adoption and use of *Twitter* by MPs. These results are consistent with assumptions based on cost-benefit calculus. Thus, it does not make as much sense for politicians to adopt *Twitter* in countries where there are fewer people on *Twitter* and low internet penetration. In particular, if politicians want to be elected or inform citizens about their activities, they have an opportunity to reach voters through *Twitter*. However, this is only true if *Twitter* is used in their countries.

Keywords

Twitter; Social media; Legislators; *Twitter* adoption; Social media adoption; Gender; Members of parliaments; Parliamentarians; Politicians; Political communication; Latin America; Social networks.

Funding

This article is a result of the project of specific research "Using survey experiments in political science" supported by the *Philosophical Faculty of the University of Hradec Králové* in 2021.



1. Introduction

In recent years, there has been an increasing amount of research looking at social media and politics globally (**Casero-Ripollés**, 2018; **Jungherr**, 2016; **Matassi**; **Boczkowski**, 2020). Among social media publications, articles focusing on the United States and Europe have consistently dominated, and more recently, research on Latin America has increased. (**Matassi**; **Boczkowski**, 2020). This article looks at parliamentarians and their adoption of *Twitter* in Latin America as a currently unexplored topic. Most publications on Latin America are single case studies, and to date, there is no comprehensive comparative study of the adoption of social media by members of parliament (MPs) in Latin America.

Political actors are increasingly using *Twitter* for various purposes, including disseminating political proposals, engaging with the public, broadcasting their messages, and shaping public debate (**Casero-Ripollés**; **Alonso-Muñoz**; **Marcos-García**, 2022; **López-Meri et al.**, 2017). In the 2016 US presidential election, *Twitter* played a significant role in both Hillary Clinton's and Donald Trump's communication strategies (**Buccoliero et al.**, 2020; **Enli**, 2017). During electoral campaigns, political actors may use *Twitter* to spread political messages, reference candidate activities, emphasize personal aspects, and mobilize supporters (**López-García**, 2016). Emerging parties tend to focus on mobilization and generic announcements, while traditional parties are more likely to publish policy proposals (**López-García**, 2016). However, the use of *Twitter* for dialogue with citizens remains limited (**Alonso-Muñoz et al.**, 2016).

In the European context, **Fazekas et al.** (2021) found that most political actors did not engage with the public about EU issues, resulting in less public interaction. Meanwhile, **Vergeer**, **Hermans** and **Cunha** (2013) noted that candidates in the 2009 *European Parliament* elections primarily used *Twitter* for electoral campaigning. Members of the *European Parliament* have been found to use *Facebook* for national audiences and *Twitter* for international audiences, with language preferences varying across countries (**Haman**; **Školník**; **Čopík**, 2022). In the Spanish context, **Casero-Ripollés**, **Alonso-Muñoz** and **Marcos-García** (2022) identified ideology, political initiative, and political career as factors influencing the authority and digital influence of political actors on *Twitter*. **López-Meri**, **Marcos-García** and **Casero-Ripollés** (2017) observed a trend towards hybridization between new and conventional digital media in the platform's usage. **Jivkova-Semova**, **Requeijo-Rey** and **Padilla-Castillo** (2017) found that virality on *Twitter* did not guarantee electoral victory, but certain communication trends could provide valuable insights. The use of *Twitter* by Latin American political leaders has also been examined, with **Segado-Boj**, **Díaz-Campo** and **Lloves-Sobrado** (2015) finding no common strategy for using the platform during times of crisis.

Overall, political actors use *Twitter* for various purposes, including disseminating messages, engaging with the public, and shaping political discourse. However, the extent of their interaction with citizens and the effectiveness of their communication strategies on the platform vary across different contexts and political actors.

Most of the research on political communication on social media in Latin America concerns presidential election campaigns, with dozens of studies. Most studies examining presidential campaigns have been produced in :

- Brazil: **Alves et al.**, 2019; **Brito et al.**, 2019; **Calais-Guerra et al.**, 2011; **Canavilhas**; **Bittencourt**; **De-Andrade**, 2019; **De-Carvalho**; **Massuchin**; **Mitoto**, 2018; **Cremonese**, 2012; **Ferreira et al.**, 2021; **Ribeiro-Ferreira**, 2018; **Gilmore**, 2012; **Hargreaves et al.**, 2020; **Ituassu et al.**, 2018; **Kobellarz et al.**, 2019; **Levy**; **Sarmiento**, 2020; **Machado et al.**, 2019; **Massuchin**; **Campos-Domínguez**, 2016; **Mitoto**; **Massuchin**; **De-Carvalho**, 2017; **Novais**; **De-Araújo**, 2012; **D. J. S. Oliveira et al.**, 2017; **Passos et al.**, 2019; **Recuero**; **Zago**; **Bastos**, 2015; **Recuero**; **Bonow-Soares**; **Gruzd**, 2020; **Reis-Longhi**; **Santos-Oliveira**, 2020; **Santana**; **Vanin**, 2020; **Santos**, 2020; **Soares**; **Recuero**; **Zago**, 2019; **Teixeira et al.**, 2019; **Tomaz**; **Tomaz**, 2020.
- Argentina: **Filer**; **Fredheim**, 2017; **Gulías**; **López-López**; **Boubeta**, 2020; **López-López**; **Oñate**; **Chavero-Ramírez**, 2018; **López-López**; **Oñate**; **Rocha**, 2020; **López-López**; **Oñate**, 2019; **López-López**; **Vásquez-González**, 2018; **Mussi-Reyero et al.**, 2021.
- Chile: **Castillo et al.**, 2019; **Olivares et al.**, 2019; **Santana**; **Huerta-Cánepa**, 2019; **Santander**; **Elórtegui**; **Buzzo**, 2020.
- Mexico: **Andrade-del-Cid**; **Flores-González**; **Pablo-Contreras**, 2020; **Beltrán**, 2020; **Bernábe-Loranca**; **González-Velázquez**; **Cerón-Garnica**, 2020; **Camp**, 2013; **Montes-de-Oca-López**; **Sandoval-Almazán**, 2019; **González-Tule**; **Restrepo-Echavarría**, 2020; **Green**, 2021a; 2021b; **Kavanaugh et al.**, 2016; **De-León**; **Vermeer**; **Trilling**, 2021; **López-Chau**; **Valle-Cruz**; **Sandoval-Almazán**, 2019; **Ortiz-Espinoza**; **Espejel-Trujillo**, 2021; **Rodríguez-Fidalgo**; **Ruiz-Paz**; **Paíno-Ambrosio**, 2019; **Pérez-Salazar**, 2019; **Sandoval**; **Matus**; **Rogel**, 2012; **Jiménez-Zarate**, 2018.
- Colombia: **Acosta-Valencia et al.**, 2021; **Alvarado-Vivas**; **López-López**; **Pedro-Carañana**, 2020; **Cerón-Guzmán**; **León-Guzmán**, 2016; **Dajer**, 2021; **Pedro-Carañana**; **Alvarado-Vivas**; **López-López**, 2020; **Ruiz-Rojas**; **Boguslavskaya**, 2018; **Ruano**; **López**; **Mosquera**, 2018.
- Ecuador: **Rofrío et al.**, 2019; **Vélez-Loor**; **Córdova**, 2021; **Zumárraga-Espinosa**; **Reyes-Valenzuela**; **Carofilis-Cedeño**, 2017.
- Costa Rica: **Cruz-Romero**, 2015.
- Peru: **Cabrera-Méndez et al.**, 2021.

But at the same time, there are comparative studies on presidents' communications on *Twitter* (**Puertas-Hidalgo**; **Carpio-Jiménez**; **Suing**, 2019; **Waisbord**; **Amado**, 2017).

Several studies have been conducted on the topic of parliamentarians and the adoption of social media or Internet technologies in Latin America, and these are primarily single case studies, for example, case studies on Brazil (**Amaral; Pinho, 2017; Brandt; Vidotti, 2020; García-Sánchez et al., 2021; Marques; De-Aquino; Miola, 2014a; 2014b; Oliveira et al., 2018**); Chile (**Fuente-Alba-Cariola; Parada-Gavilán, 2019; Henríquez et al., 2022**); or studies comparing Argentina, Paraguay, and Uruguay (**Welp; Marzuca, 2014; 2016**).

Latin America has presidential regimes, so the main focus of the research is on the political communication of presidents or during presidential elections and less on members of parliament. However, MPs are very important players in the political system that deserve attention.

The contribution of this paper is that it analyses the unexplored topic of the use of *Twitter* by Latin American MPs. This study, therefore, provides up-to-date data on the *Twitter* activity of MPs in Latin America. Previous research has focused almost exclusively on presidents or, in the case of MPs, single case studies. This paper is divided into four parts. The first section presents the research questions and hypotheses. The second section presents the methodology, and the third section presents the results. The last section is the conclusion of the paper.

2. Research questions and hypotheses

This paper aims to answer two research questions. These are:

1. What are the differences between the adoption and use of *Twitter* by parliamentarians in Latin America across countries?
2. What factors influence the adoption and use of *Twitter* in Latin America?

To answer the first research question, it will first be necessary to examine the current state of *Twitter* adoption, i.e. to identify the individual legislators who use *Twitter*. This is not an easy task, as no databases exist yet where the usernames (*Twitter* handles) of legislators on *Twitter* are present.

Multiple methods and data are used to answer the second research question. The dataset will consist of variables for each legislator and information on the number of citizens on *Twitter*, the Internet, and the HDI in each country.

In this research, *Twitter* was chosen for several reasons. Firstly, at the time of the research, *Twitter* offered the *Twitter* API for academic research to researchers, which provides access to almost all data on *Twitter*. In contrast, *Facebook* provided an API that was considerably limited, and researchers often did not have access to it or only had temporary access through various tools. Overall, *Facebook* has been criticized by academics for providing limited access to a restricted amount of posts (**Ho, 2020**). Among social networks, *Twitter* is most frequently the subject of research concerning digital communication technologies for legislators (**Neihouser; Tremblay-Antoine, 2021**), which demonstrates that *Twitter* is used by legislators as a common communication tool. Another issue associated with *Facebook* is that Latin American lawmakers often have a *Facebook* profile rather than *Facebook* pages, from which it is even more challenging to obtain data, and sometimes they use both. *Instagram*, on the other hand, suffers from similar problems as *Facebook* and is predominantly used for publishing images. Additionally, there is limited space in a single article, so for all the reasons mentioned above, only *Twitter* was analyzed in this research.

2.1. Gender

Although some of the earliest research on social media use suggested that women generally use social media more actively (**Hargittai, 2007**) while at the same time considering that gender may influence how online campaigning is approached (**Druckman; Kifer; Parkin, 2007**), most studies focusing on politics on social media later showed that there are no differences between women and men in the use of social media (**Chi; Yang, 2010; Grant; Moon; Grant, 2010; Lappas et al., 2016; Lappas; Triantafillidou; Yannas, 2019; Macková; Štětka, 2016; Metag; Marcinkowski, 2012; Neihouser, 2021; Obholzer; Daniel, 2016; Rauchfleisch; Metag, 2016; 2020; Sandberg; Öhberg, 2017; Scherpereel; Wohlgemuth; Schmelzinger, 2017; Strandberg, 2009; 2013; Vergeer; Hermans, 2013**). But that does not mean that no studies are showing such a difference. While some studies show that men have adopted social media more or used it more actively (**Ausserhofer; Maireder, 2013; Lappas; Triantafillidou; Yannas, 2018; Vergeer; Hermans; Sams, 2011**), some studies say that women are more active (**Cook, 2017; Evans; Cordova; Sipole, 2014; Sullivan, 2021**). Differences between studies about politicians may be related to the adoption of the technology at a given time within a given country and among citizens. Since most studies have not found a difference between women and men, the hypothesis is:

H1: There will be no relationship between gender and adoption or *Twitter* activity among MPs.

2.2. Age

The adoption of new technology is often related to age according to the diffusion of innovation theory (**Rogers, 1962**). So, it should be young politicians as early adopters and active users of social media. Early studies showed that young candidates in elections are more familiar with the internet and make more use of its potential. (**Gibson; McAllister, 2006**). Subsequent studies have confirmed this assumption and indeed found a relationship between age and social media use (**Gulati; Williams, 2013; Larsson, 2015; 2015; Larsson; Kalsnes, 2014; Larsson; Moe, 2012; Lassen; Brown, 2011; Metag; Marcinkowski, 2012; Obholzer; Daniel, 2016; Peterson, 2012; Scherpereel; Wohlgemuth; Schmelzinger, 2017; Strand-**

berg, 2009; 2013; Straus *et al.*, 2013; Vergeer; Hermans, 2013). But there are also studies where age is not a significant predictor (Macková; Štětka, 2016; Rauchfleisch; Metag, 2020). Given the majority of studies confirming the relationship between age and social media use, the hypothesis is:

H2: Younger MPs will adopt *Twitter* more and also use it more actively.

2.3. Technological development of the country

From a classical microeconomic approach, demand should influence supply. It is the adoption of new technologies by the population that should increase the demand for political information in the online environment and politicians should respond to this fact. Parliamentarians should use *Twitter* primarily when they believe that the benefits outweigh the costs, at least according to a simple calculation based on rational choice theory. If Internet penetration is very low, then a politician should have less incentive to use the Internet for his or her purposes. Studies of political communication on the Internet and social media tend to include a variable in the form of the popularity of the social media network in the politician's constituency (Haman; Školník, 2021; Scherpereel; Wohlgemuth; Schmelzinger, 2017), alternatively proxy variable as internet penetration (Obholzer; Daniel, 2016; Sudulich; Wall, 2009) or socio-economic variables of the constituency. For example, this could be the number of people with a university degree, the number of people of retirement age, and the level of urbanization or wealth (Carlson; Djupsund; Strandberg, 2013; Chi; Yang, 2010; Cook, 2016; 2017; Gulati; Williams, 2007; 2010; 2013; Herrnson; Stokes-Brown; Hindman, 2007; Lappas *et al.*, 2016; Lassen; Brown, 2011; Metag; Marcinkowski, 2012; Peterson, 2012; Southern, 2015; Southern; Lee, 2019; Strandberg, 2009, p. 20; 2013). Sometimes these variables are used as independent variables, sometimes they have a primary control function. Either way, they are expected to have an impact on the adoption and use of social networking by politicians. Indeed, in many of them, variables giving information about potential demand from voters are significant predictors (Cook, 2016; 2017; Gulati; Williams, 2010; Haman; Školník, 2021; Herrnson; Stokes-Brown; Hindman, 2007; Peterson, 2012; Scherpereel; Wohlgemuth; Schmelzinger, 2017; Straus *et al.*, 2013). However, only rarely (Haman; Školník, 2021) has research directly used data on the percentage of residents using a given social media. This is primarily because this data is not publicly available.

H3: The higher the number of citizens on *Twitter* (as well as Internet penetration and the Human Development Index - HDI), the higher the adoption of *Twitter* by MPs.

3. Methodology

First, it was necessary to collect data. Unfortunately, there is no up-to-date list of Latin American MPs and their *Twitter* accounts. Some parliaments allow an MP to put a link to their official profile on the parliamentary website. However, even in this case, not all MPs make use of this. In any case, parliamentary websites were the primary source of data. If a *Twitter* account was not found, I proceeded to use the *Google* search engine that offers the *Google Knowledge Graph*, which often lists the social media accounts of politicians and other known persons:

<https://cloud.google.com/enterprise-knowledge-graph/docs/search-api>

However, not all MPs are searchable in this way. Thus, there was also a direct search on *Twitter*, using various combinations of names. On parliamentary websites, MPs often have their full names listed, which makes direct *Twitter* searches problematic as some MPs do not use all their names and are therefore not easily traceable. Hispanic names can include five or more names, and a politician may only use two of them. At the same time, some parliaments use *Twitter* accounts and have lists of MPs on them. In this case, this source was also used.

There are a total of three independent variables in the paper. The first two independent variables are individual and based on data from parliamentary websites. These are age and gender. In the previous section, reference was made to studies that utilized variables such as gender and age. This study employed variables constructed in a similar manner. Gender was coded as 1 for male and 0 for female. Age was measured in years. Not all parliaments reported information on the date of birth and therefore the age of MPs, so some countries have this variable missing. Also, not all parliaments reported the sex of the MP, but gender could be inferred from the name of the MP or other characteristics. *Twitter* population data comes from the latest *Latinobarómetro* survey in 2020 (*Latinobarómetro*, 2022). Information on the number of individuals using the Internet in a given country is provided by the *International Telecommunication Union* (*International Telecommunication Union*, 2022), while not offering data for Venezuela. The HDI is then offered by the *United Nations Development Programme* (2022).

I considered *Twitter* adoption if MPs tweeted at least once in the period between 1 January 2021 and 31 March 2022. For some countries, this period was different and followed elections in those states. Thus, in states that held elections during this period, the observation period was only from the newly elected legislature after the election. In the case of subsequent states, the period began on a different date:

- Argentina - from December 10, 2021.
- Ecuador - from May 14, 2021.
- El Salvador - from May 1, 2021.
- Mexico - from September 1, 2021.
- Peru - from July 28, 2021.
- Venezuela - from January 5, 2021.

The year 2021 and the first three months of 2022 were chosen arbitrarily. However, it is not possible to choose a long period during which there would be no elections in all countries. Elections make the analysis problematic because one would have to examine two or more legislative periods in one country, which could cause problems for cross-country comparisons if other countries have only one legislative period. By selecting this time frame, even countries like Argentina had more than three months' worth of tweets, which should be sufficient for analysis.

At the same time, as will be shown below, a variable of an average number of posts per week and a dichotomous variable that takes the value of 1 if an MP sent at least one tweet per week on average were created for each MP. Indeed, it is not enough to include mere adoption, it is also necessary to include how much *Twitter* usage the MP has. The threshold of one tweet per week was chosen. The models also use the number of tweets per week to check the actual activity among the MPs present on *Twitter*. A total of 2,353,138 tweets were collected via the *Twitter* API using the *rtweet* package (Kearney, 2019) and subsequently analyzed in the R programming language, and a total of 3,215 MPs were examined.

4. Results

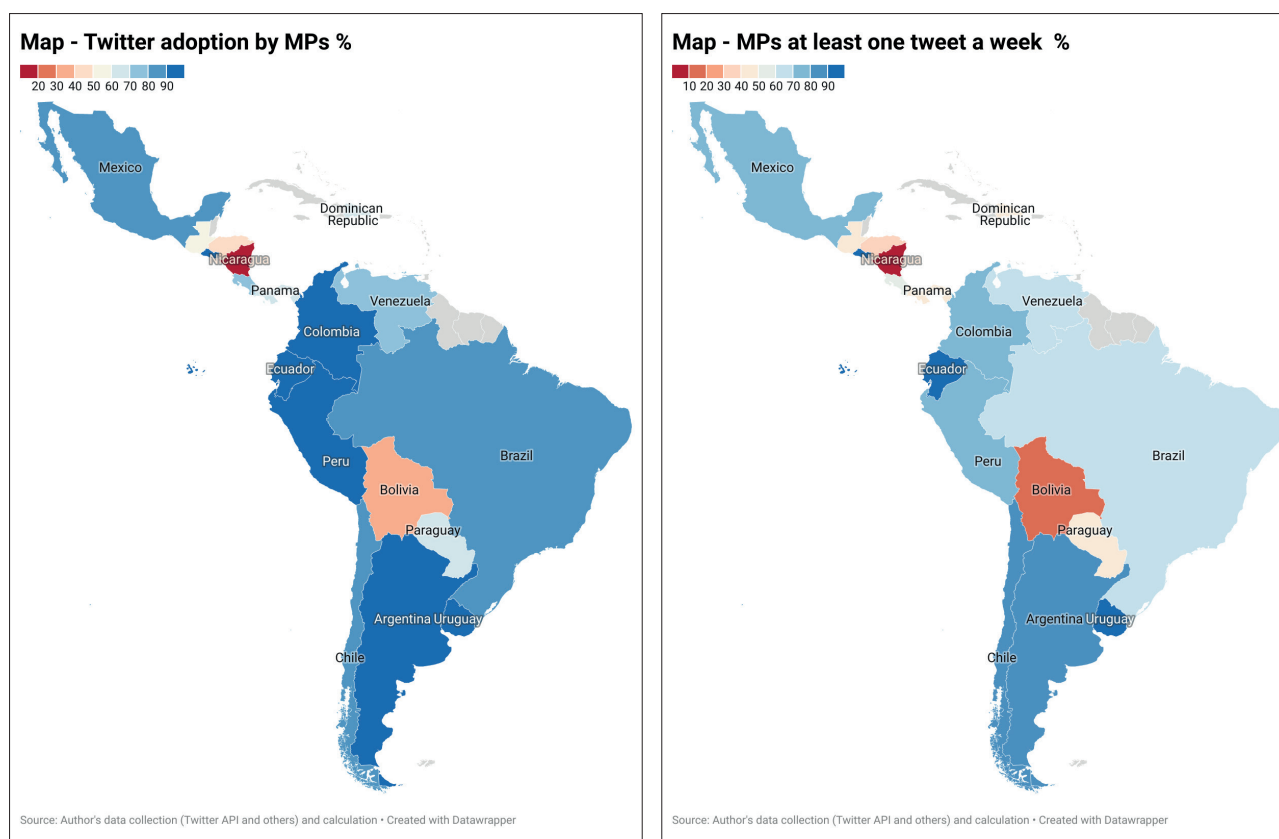
4.1. Use of *Twitter* by MPs

Table 1 provides some data on the adoption and activity of parliamentarians. The second and third columns give information on *Twitter* adoption. As already mentioned, in the second column I mean adoption in the form of posting at least one tweet during the period under study, The third column then requires the activity of at least one tweet per week during the period under study. These data are also displayed on two maps. The mean value of presence on *Twitter* in Latin American countries is 73.6%, while the percentage of the mean decreases to 61.5% if the activity of at least one tweet per week is required. However, as the table shows there are significant differences between countries. Six countries have more than 90% of MPs present on *Twitter* (Argentina, Ecuador, Colombia, Peru, El Salvador, Uruguay) and three countries have more than 80% of MPs (Brazil, Chile, and Mexico). Conversely, three countries in Latin America do not have even half of their MPs on *Twitter* (Bolivia, Honduras, Nicaragua). Then six states have between 50% and 80% (Dominican Republic, Guatemala, Costa Rica, Panama, Paraguay, and Venezuela).

Table 1. MPs in Latin American countries on *Twitter*

Country	Twitter adoption %		Number of tweets per week		Distribution of tweet		Analyzed MPs	
	Min. 1 tweet	Min. 1 tweet per week	Mean	Median	Hashtags	Retweets	Male	Female
Argentina	90.7%	87.9%	17.5	9.4	24.3%	41.0%	142	115
Bolivia	36.9%	17.7%	2.4	0.9	51.6%	10.4%	69	61
Brazil	84.2%	66.5%	17.4	5.3	35.3%	13.0%	437	76
Dominican Republic	62.1%	43.2%	8.5	2.3	14.4%	36.7%	144	46
Ecuador	97.1%	95.6%	33.3	23.3	65.1%	58.5%	84	53
Guatemala	58.8%	43.1%	8.3	3.6	24.2%	43.6%	129	31
Honduras	48.8%	35.8%	7.6	2.2	29.2%	47.0%	95	28
Chile	89.0%	81.3%	20.0	10.0	49.4%	50.8%	119	36
Colombia	90.7%	79.6%	26.1	9.4	52.9%	40.9%	131	31
Costa Rica	73.7%	54.4%	11.0	4.2	14.2%	15.4%	35	22
Mexico	83.8%	75.2%	18.7	9.2	56.0%	38.6%	250	250
Nicaragua	12.1%	8.8%	5.0	1.9	60.9%	35.8%	47	44
Panama	64.8%	45.1%	6.2	2.6	18.4%	38.7%	55	16
Paraguay	65.0%	42.5%	12.1	3.8	15.0%	38.1%	66	14
Peru	96.2%	79.2%	13.1	7.1	31.9%	37.8%	80	50
El Salvador	98.8%	97.6%	64.9	50.1	31.0%	41.2%	58	24
Uruguay	94.9%	90.9%	16.6	10.5	13.7%	53.2%	75	24
Venezuela	77.3%	62.2%	20.7	8.1	56.0%	78.8%	183	95
Median	80.6%	64.4%	14.9	6.2	35.7%	40.0%		
Mean	73.6%	61.5%	17.2	9.1	31.5%	39.8%		

It is crucial to compare the results from the table with those obtained in previous research to analyze trends. Several prior studies have examined the adoption of *Twitter* among Brazilian parliamentarians. In 2013, 64.3% of Brazilian MPs actively used a *Twitter* account (Amaral; Pinho, 2017). Another study revealed that in 2019, 84.9% of MPs had a *Twitter* account (García-Sánchez et al., 2021). The current results indicate that 84.2% of MPs have sent at least one tweet, representing an increase of approximately 20% compared to 2013. According to a study on communication and adoption of social media by Chilean MPs in 2018, 82.5% of MPs had a *Twitter* account (Fuente-Alba-Cariola; Parada-Gavilán, 2019). In this case, this signifies a 7% increase in *Twitter* usage by Chilean MPs.

Map 1. *Twitter* adoption by MPs %

Map 2. MPs that post at least one tweet a week %

Two studies examining the adoption of *Twitter* by MPs in Argentina, Paraguay, and Uruguay showed that in Argentina in 2012, 53% of MPs were present on *Twitter*, while in Paraguay, only 11% of MPs were on *Twitter*, and in Uruguay, 46% of MPs used *Twitter* (Welp; Marzucca, 2014; 2016). Thus, since 2012, all three countries have experienced an increase of several tens of percent in MPs' presence on *Twitter*. Prior studies indicate that there has been an increase of several tens of percent in the number of active MPs on *Twitter*, with a noticeable rise in the number of MPs using *Twitter* in each country.

However, the percentages of MPs on *Twitter* decrease, often significantly, if we look at the MPs who sent on average at least one post per week. Subsequently, only Ecuador, El Salvador, and Uruguay reach values above 90%. The third and fourth columns give more detailed activity information. It can be seen that the MPs from El Salvador have developed the absolute highest activity, with an average of 64.9 and a median of 50.1 tweets per week, followed by Ecuador with an average of 33.3 and a median of 23.3 tweets per week. Several countries have an average that corresponds to sending at least one tweet per week. However, it is important to note that the average and median include only MPs who are present on *Twitter*. Therefore, these two columns need to be combined with the information from the second column to see the overall state of political communication on *Twitter* in a given country.

The sixth and seventh columns show the distribution of retweets in terms of the number of hashtags and retweets. The sixth and seventh columns show additional data on the use of *Twitter* by MPs. The second column provides information on how many tweets had a hashtag in the period under study. The mean for Latin American countries is 31.5%. However, there is considerable variation between countries. While 65% of tweets from Ecuador used a hashtag, only 13.7% of tweets in Uruguay did. These differences suggest that it is hardly possible to speak of a pattern of Latin American MP communication. This fact is confirmed by the third column, which shows what percentage of the tweets were retweets. The average for Latin American countries is 39.8%, but the figures range from 10.4% in Bolivia to 78.8% in Venezuela.

4.2. Adoption of *Twitter* –gender and age– individual level

In this part of the paper, only data at the first level, which is the level of the parliamentarian, will be used. The following section will then use hierarchical models (multilevel models). This is primarily a robustness testing of the results using other models as well. At the same time, for hierarchical models, the lack of cases at the second level, that is, the state level can be problematic. This paper works with a total of 18 Latin American countries, which may be considered methodologically insufficient for hierarchical models. For example, Krefth and Bokhee (1996), Hox (2010, p. 235), or Snijders and Bosker (1999, p. 154) propose a "30/30 rule" whereby there should be at least 30 cases at each level.

There are four models in the following Table 2. The first two models have the dependent variable as a dichotomous variable of adoption (1) of the MP on *Twitter*, i.e., posting at least one tweet during the observation period. The third

and fourth models have the dependent dichotomous variable of whether the MP sent on average at least one tweet per week (1) during the observation period. Because of the dichotomous variable, logistic regression was chosen. There are two variables in the model, namely gender, with the male taking the value of 1 and this variable being present for all MPs. On the other hand, age is not present for all MPs and therefore has separate models. Dummy variables have also been used for each country but are not presented here to save space.

Table 2. Logistic regressions

	Dependent variable:			
	Twitter adoption		At least one tweet per week	
	(1)	(2)	(3)	(4)
Gender (Male = 1)	-0.252*	-0.162	-0.273**	-0.301*
	(0.109)	(0.151)	(0.096)	(0.135)
Age		-0.032***		-0.020***
		(0.006)		(0.005)
Constant	1.891***	3.523***	0.918***	2.008***
	(0.154)	(0.357)	(0.125)	(0.297)
Observations	3,215	1,615	3,215	1,615
Log Likelihood	-1,408.712	-734.983	-1,722.229	-900.350

Note: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Table 2 shows that in both the first and third models, where there are the most cases, gender is statically significant and with a negative value, hence implying that women used *Twitter* more, both in the form of adoption and when including requiring a minimum activity of one tweet per week. In the second model, however, gender loses statistical significance, which is due to the fact that a number of countries where women were more prevalent in *Twitter* use than men dropped out of the model because there is no information on age in these countries. Models 1 and 3, therefore, reject the first hypothesis that there is no difference between the genders. The age variable is statistically significant in both models 2 and 4 and takes negative values, i.e. adoption and activity (in the form of one tweet per week) increase with lower age of the MPs. Models 2 and 4, therefore, confirm the second hypothesis that younger MPs adopt *Twitter* more and are more active.

4.3. Activity on *Twitter* –gender and age– individual level

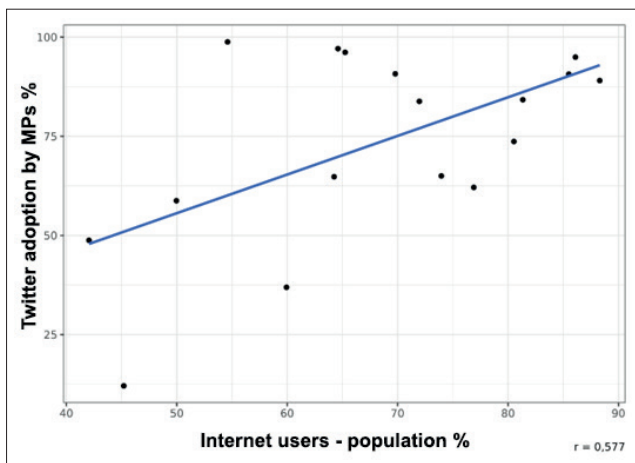
Whereas previous models looked at adoption and general activity in the form of at least one tweet per week. The following table shows four models where the subject of interest is MPs' *Twitter* activity. The dependent variable in the first and second models is the logarithmic average number of tweets per week, logarithmic because of the skewed distribution of tweets. Thus, the first two models are linear regressions (OLS). However, in political communication research, negative binomial regression is often used when examining the number of tweets, as these are the counts (Jacobs; Spierings, 2019; Peterson, 2012; Sandberg; Öhberg, 2017; Scherpereel; Wohlgemuth; Schmelzinger, 2017; Scherpereel; Wohlgemuth; Lievens, 2018). For example, one study rounded up the average number of tweets per week to deal with whole numbers and capped the upper value at 250 to avoid distorting the general pattern with outliers that could lead to incorrect conclusions (Jacobs; Spierings, 2019).

Table 3. OLS and negative binomial regressions

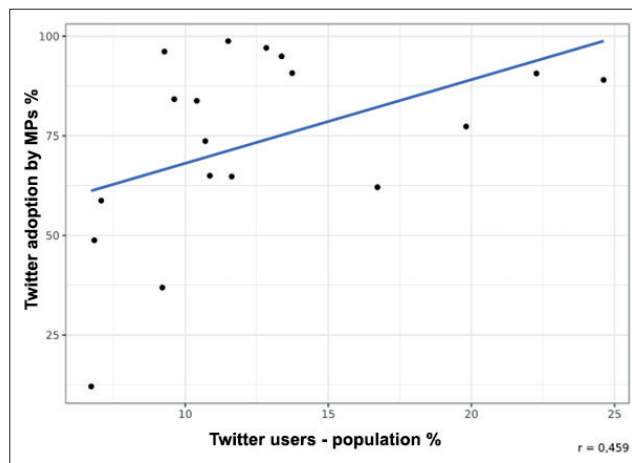
	Dependent variable:			
	Tweets per week (log)		Tweets per week (rounded)	
	OLS		negative	binomial
	(1)	(2)	(3)	(4)
Gender (Male = 1)	-0.191*	-0.353***	-0.105*	-0.167*
	(0.076)	(0.107)	(0.051)	(0.072)
Age		-0.002		0.004
		(0.004)		(0.003)
Constant	1.671***	1.926***	2.971***	2.839***
	(0.104)	(0.237)	(0.070)	(0.161)
Observations	2,488	1,266	2,488	1,266
R2	0.145	0.106		
Adjusted R2	0.139	0.101		
Log likelihood			-9,598.981	-4,782.554

Note: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

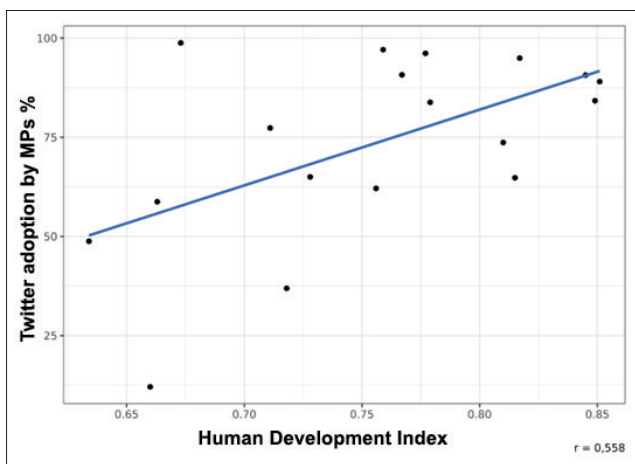
In these models, only MPs with a *Twitter* account are included. In the second and fourth models, age is then added. All four models show that the gender variable is negatively statistically significant, i.e. women use *Twitter* more actively. At the same time, no statistical significance is found for the age variable. Thus, while the models in the previous table implied that women are more likely to be on *Twitter*, these models confirm that women are also more likely to be more active among MPs who are on *Twitter*. In contrast, the models showed that younger MPs are more likely to adopt *Twitter*. However, when the analysis includes MPs who have a *Twitter* account and their activity, it loses statistical significance, so there is no way to confirm the second hypothesis in the activity part. Younger MPs adopt *Twitter* more, but this is no longer the case for activity in terms of the number of tweets.



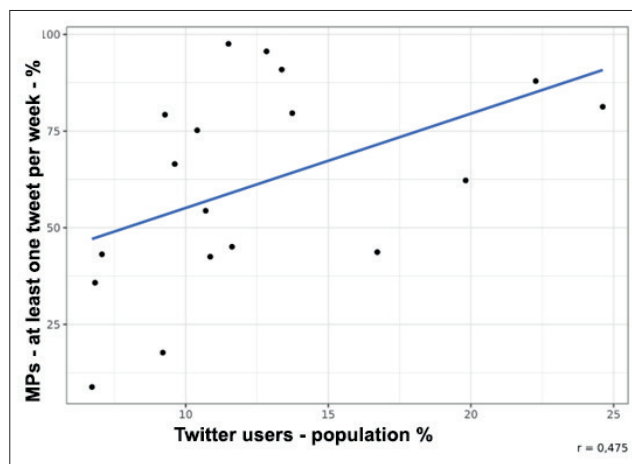
Graph 1. Correlation between Internet users - population in countries and the percentage of MPs who have adopted *Twitter*



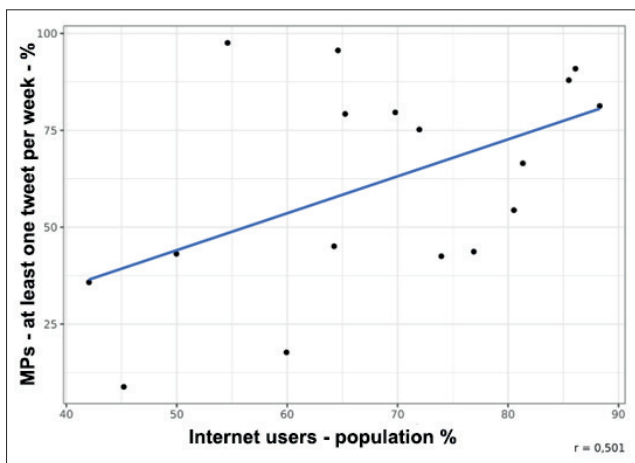
Graph 2. Correlation between *Twitter* users - population in countries and the percentage of MPs who have adopted *Twitter*



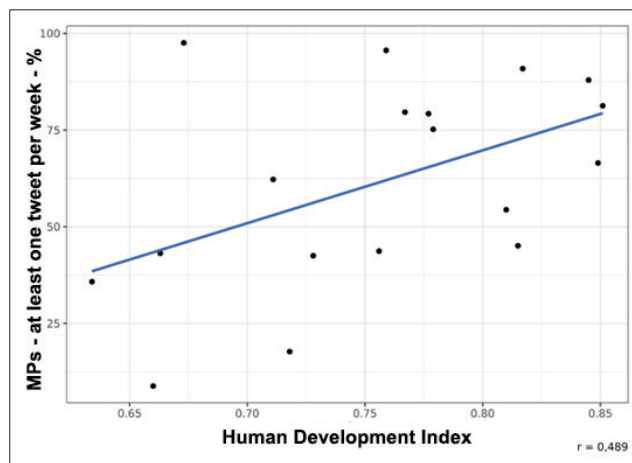
Graph 3. Correlation between countries' HDI and the percentage of MPs who have adopted *Twitter*



Graph 4. Correlation between *Twitter* users - population in countries and the percentage of MPs who sent at least one tweet per week



Graph 5. Correlation between Internet users - population in countries and the percentage of MPs who sent at least one tweet per week



Graph 6. Correlation between countries' HDI - population in countries and the percentage of MPs who sent at least one tweet per week

4.4. Technological development of countries

The following Graphs 1 to 6 show the relationship between the presence of MPs in each country on *Twitter* and the number of people on the internet, the number of people on *Twitter*, and the Human Development Index (HDI) for that country. The data on the *Twitter* population comes from *Latinobarómetro's* latest survey for 2020 (*Latinobarómetro*, 2022). Information on the number of individuals using the Internet in a given country is provided by the *International Telecommunication Union* (2022), while not offering data for Venezuela. The HDI is then offered by the *United Nations Development Programme* (2022). The relationship between the percentage of MPs who, on average, sent at least one tweet during the period under review is then plotted against these three indicators. These three variables are moderately to strongly correlated. The highest correlation is between the HDI and the percentage of people on the internet ($r = 0.905$), while the relationship between the percentage of people on the internet and on *Twitter* ($r = 0.717$) also shows a very strong correlation. The least correlation, but still significant, is between the HDI and the percentage of people on *Twitter* ($r = 0.534$). Strong correlations are not surprising, as they give similar information, namely about the socio-economic and technological progress of a given country.

All the graphs show at least a moderate correlation between the variables. The figures and the Pearson correlation coefficient values at the bottom right of each figure suggest that the MP's adoption of *Twitter* increases with the percentage of people on the Internet, on *Twitter*, and the value of human development. The relationships between the percentage of MPs on *Twitter* with at least one tweet per week and these three indicators are very similar in Pearson correlation coefficient values. These data are further used in the multilevel models.

Table 4 shows the values for each country. It can be seen that there are a few exceptions where the Internet is not very prevalent, yet it is very much used by MPs. The most notable exception is El Salvador, where just over half of the people have access to the internet, but almost all of its MPs use *Twitter*. Other countries with internet penetration of around 50%, while having a lower HDI, are among the countries with the lowest adoption of *Twitter* by MPs.

Table 4. Countries - the relationship between the number of MPs on *Twitter* and state characteristics

Country	<i>Twitter</i> adoption %	Min. 1 tweet per week %	Population on <i>Twitter</i> %	Population - Penetration of Internet %	HDI
Argentina	90.7	87.9	22.3	85.5	0.845
Bolivia	36.9	17.7	9.2	59.9	0.718
Brazil	84.2	66.5	9.6	81.3	0.849
Chile	89.0	81.3	24.6	88.3	0.851
Colombia	90.7	79.6	13.7	69.8	0.767
Costa Rica	73.7	54.4	10.7	80.5	0.81
Dominican Republic	62.1	43.2	16.7	76.9	0.756
Ecuador	97.1	95.6	12.8	64.6	0.759
Guatemala	58.8	43.1	7.1	50.0	0.663
Honduras	48.8	35.8	6.8	42.1	0.634
Mexico	83.8	75.2	10.4	72.0	0.779
Nicaragua	12.1	8.8	6.7	45.2	0.66
Panama	64.8	45.1	11.6	64.3	0.815
Paraguay	65.0	42.5	10.9	74.0	0.728
Peru	96.2	79.2	9.3	65.3	0.777
El Salvador	98.8	97.6	11.5	54.6	0.673
Uruguay	94.9	90.9	13.4	86.1	0.817
Venezuela	77.3	62.2	19.8		0.711

4.4. Multilevel models

In Table 5, there are six multilevel models, where the variables gender or age are at the first level. Then at the second level are the variables percentage of citizens on the Internet, HDI, and percentage of people on the Internet. The data source is the same as mentioned above. The dependent variable is whether the MP was on *Twitter* (1) with at least one tweet sent during the period under study. The six models are because the three second-level indicators are always in the model separately due to the strong correlation but also the fact that they give similar information to some extent and the age variable was not available for some countries. The models confirm what the correlations and models in the previous part of the paper already showed. MPs adopt *Twitter* more when the country's technological development is greater. Younger MPs also adopt *Twitter* more. At the same time, all three variables are statistically significant. Thus, *Twitter* is used more in countries where there is a higher demand, i.e. where citizens have adopted new technologies more,

in the form of the Internet or *Twitter* directly, to find political information. This result is in line with the assumption that in these countries MPs will benefit more from using it as it reaches a larger percentage of the electorate.

Table 5. Multilevel models - the adoption of *Twitter* by MPs

	Dependent variable:					
	Adoption of <i>Twitter</i>					
	(1)	(2)	(3)	(4)	(5)	(6)
Gender (male)	-0.291*	-0.152	-0.248*	-0.153	-0.248*	-0.156
	(0.117)	(0.150)	(0.109)	(0.150)	(0.109)	(0.150)
Age		-0.031***		-0.031***		-0.031***
		(0.006)		(0.006)		(0.006)
Population - penetration of Internet	0.048*	0.065**				
	(0.024)	(0.020)				
Population on <i>Twitter</i>			0.107+	0.133*		
			(0.064)	(0.067)		
HDI					9.703*	12.196**
					(4.532)	(3.728)
Constant	-1.682	-1.880	0.212	1.215	-5.780	-6.714*
	(1.645)	(1.452)	(0.876)	(0.883)	(3.443)	(2.901)
N	2,937	1,615	3,215	1,615	3,215	1,615
Log Likelihood	-1,296.891	-745.550	-1,449.047	-747.065	-1,448.418	-745.529

Note: +p <0.1; *p<0.05; **p<0.01; ***p<0.001

Therefore, it can be concluded that the first hypothesis was not confirmed. Existing global research suggests that there is generally no difference in the use of social media between male and female politicians (Chi; Yang, 2010; Macková; Štětka, 2016; Metag; Marcinkowski, 2012; Strandberg, 2009; 2013; Vergeer *et al.*, 2013). However, in this case, a difference was found, as there is a gender gap in Latin America, with female parliamentarians using *Twitter* more. This finding represents an original contribution to the field of study, as it significantly contradicts previous findings from other countries, and is based on a large dataset. One potential reason for female MPs being more active on *Twitter* could be their desire to bypass traditional media, which has been known to portray them less favorably (Heith, 2003; Kahn, 1996).

By using *Twitter*, female politicians can present a more open, personal, and interactive image (Carlson; Djupsund; Strandberg, 2013) and communicate directly with voters, mobilizing them and targeting specific groups like young women. This direct communication allows female politicians to promote both themselves and their parties more effectively while simultaneously sidestepping gendered coverage (Lawless, 2012). Another possible explanation for the observed results could be the different communication styles adopted by women and men. Studies have indicated that women tend to use technology in a more sociable manner (Walton; Rice, 2013; Lasorsa, 2012), which may contribute to their more active engagement on *Twitter*. This communication style can help female politicians forge stronger connections with their audience, enabling them to better address their constituents' concerns and needs.

The second hypothesis is confirmed as younger MPs adopt *Twitter* more frequently. Existing research suggests that younger representatives are more likely to embrace social media (Gulati; Williams, 2013; Larsson, 2015; 2015; Scherpereel; Wohlgemuth; Schmelzinger, 2017; Strandberg, 2009, 2013; Straus *et al.*, 2013). In this regard, the results align with current research. Younger age cohorts in developed nations are often labeled as "digital natives" due to their lifelong exposure to computers, demonstrating a more intuitive command of online technologies compared to "digital immigrants" from older cohorts (Scherpereel; Wohlgemuth; Schmelzinger, 2017). Individuals from younger generations typically exhibit greater proficiency in using online platforms for political engagement (Bakker; De-Vreese, 2011). This expertise is also likely to manifest among younger politicians (Larsson, 2015). Younger and freshman politicians might be more motivated to leverage all available communication channels to consolidate political support and cultivate their brand (Peterson, 2012), unlike older politicians who might have already built their brand via traditional media.

At the same time, the hypothesis regarding country characteristics was also confirmed when variables such as the number of people on the Internet, on *Twitter*, and the HDI were statistically significant in the models. Based on assumptions from rational choice theory and cost-benefit analysis, if politicians aim to be elected or re-elected, they should consider the number of citizens on a given social network. If no one uses the social network, no one will read the politician's *Twitter* posts, making it illogical for the politician to be present on *Twitter* in a country where almost no one utilizes the platform. In this case, correlations initially revealed a strong relationship between the number of people on the internet in a given country, the quality of life as a socioeconomic indicator, the number of *Twitter* users, and the adoption of *Twitter* by MPs. Subsequently, these variables were statistically significant in regression models, confirming the hypothesis. Of course, the results do not imply that there are no exceptions with a smaller percentage of internet users where MPs

simultaneously use *Twitter*. These findings corroborate a previous study comparing the relationship between *Twitter* usage by citizens and MPs in Europe (Haman; Školník, 2021). The most significant exception is El Salvador, which has a low standard of living and fewer people on both the internet and *Twitter*; nevertheless, nearly all its MPs are present on *Twitter*. This intriguing phenomenon is worth investigating further. Consequently, it does not solely depend on the number of people on the internet in a given country, but other factors are also involved.

5. Conclusion

This paper has provided a unique analysis of the use of *Twitter* by Latin American MPs. More than 3,000 MPs were examined and two million tweets were collected. Research on political communication in Latin America was dominated by research on elections, especially presidential elections, and presidential political communication. The paper, therefore, provides an update on the current state of the MPs' adoption in Latin America. The use of *Twitter* by MPs in Latin America varies considerably. While in several countries more than 90% of MPs use *Twitter* (Argentina, Colombia, Ecuador, El Salvador, Peru, and Uruguay). Then there are countries where less than half of MPs use *Twitter* (Bolivia, Honduras, Nicaragua). At the same time, in several countries there is very low activity, so we cannot speak of much active use, even if a higher percentage of MPs are on *Twitter*. Differences were also found between the use of hashtags and the proportion of retweets among tweets. While MPs in some countries actively use hashtags, in others they do not use this form of communication at all. A high proportion of retweets means that MPs in a given country are spreading ideas already formulated by someone rather than creating their content.

In terms of factors that may help explain *Twitter* use, the first hypothesis was rejected when a significant relationship was found in that female MPs were more likely to use *Twitter* than male MPs. When analyzing MPs on *Twitter*, it was also found that women are more active. The second hypothesis was confirmed and younger MPs adopt *Twitter* more. However, when comparing MPs already present on *Twitter*, the models did not show younger MPs to be more active at a statistically significant level. The third hypothesis was also confirmed when MPs in states with higher numbers of *Twitter* users adopted *Twitter* more. Results were similar when using Internet penetration and human development variables. Similar results are consistent with assumptions based on rational choice theory and cost-benefit calculus. Thus, in countries where *Twitter* or the Internet is less used, it does not make as much sense to use *Twitter* as in a country where many people are connected to the Internet. Especially if a politician wants to get elected, he or she has a unique opportunity to reach voters through social media. But this is only true if social media are used in the country.

Thus, these results have contributed to the scholarly debate on the adoption and use of social media by MPs in different countries (Amaral; Pinho, 2017; Fuente-Alba-Cariola; Parada-Gavilán, 2019; García-Sánchez *et al.*, 2021; Marques; De-Aquino; Miola, 2014a; Welp; Marzuca, 2014; 2016) with new and updated data. As has been mentioned several times, for many countries in Latin America, no study has yet been conducted on MPs' use of *Twitter*, and therefore it was not clear how many MPs in those countries use *Twitter*. So far, comparative research on a larger number of countries has primarily focused on Europe (Castanho-Silva; Proksch, 2022; Haman; Školník, 2021; Van-Vliet; Törnberg; Uitermark, 2020). Just as studies focusing on European countries have noted differences between countries, the Latin American area is similar in this regard. At the same time, the results show an increasing tendency among MPs to use *Twitter* compared to previous studies. For example, in Argentina in 2012, 53% of MPs used *Twitter* (Welp; Marzuca, 2014; 2016), while the data of this paper show that more than 90% of MPs are currently active on *Twitter*.

One limitation of the article is that it primarily concentrates on the adoption and activity of Latin American parliamentarians on *Twitter*, without delving deeply into the content of their tweets. This focus may overlook important nuances or trends in the messages being communicated by these politicians. Consequently, the main objective of the article was to map the current adoption of Latin American parliamentarians on *Twitter*, rather than to examine the precise purposes for which they use *Twitter*. A more in-depth analysis of the content could have revealed patterns and differences among the parliamentarians, as well as their stance on critical matters in the region. Future research could build upon the findings of this article to explore these aspects more thoroughly, thus providing a more holistic perspective on the role of *Twitter* in Latin American politics.

Of course, there are several other limitations. For example, a shorter period was observed, but this was primarily to offer the most recent data. However, research looking at longer periods would certainly be useful in the future. At the same time, there is also a problem in identifying MPs on *Twitter*. As there are no official complete lists, and thus multiple techniques had to be used to collect the data. Therefore, it is not possible to ensure that an MP could have not missed through data collection, especially if they did not use their official name and were not registered on any lists or found through *Google* search. At the same time, of course, only two individual variables are used in the research, which is primarily due to comparisons across states. Important variables such as an MP's ideology or their position in the government were not examined. While several other variables can be collected on a single country, data on multiple parliaments does not exist in a comprehensive and uniform form. Even in this study, it was not possible to find the age of all MPs. However, the above reasons are relevant in interpreting the paper's results.

6. References

- Acosta-Valencia, Gladys-Lucía; Maya-Franco, Claudia-María; Acevedo-Merlano, Álvaro-Alfonso; Crawford-Visbal, J. Livingston** (2021). “Campaña presidencial colombiana en *Twitter* (2018): Apropiación, comunicación y subjetividades”. *Observatorio (OBS*)*, v. 15, n. 4, pp. 116-143.
<https://doi.org/10.15847/obsOBS15420211809>
- Alonso-Muñoz, Laura; Marcos-García, Silvia; Casero-Ripollés, Andreu** (2016). “Political leaders in (inter)action. *Twitter* as a strategic communication tool in electoral campaigns”. *Trípodos*, v. 39, pp. 71-90.
http://www.tripodos.com/index.php/Facultat_Comunicacio_Blanquerna/article/view/381
- Alvarado-Vivas, Sergio; López-López, Juan; Pedro-Carañana, Joan** (2020). “Los debates electorales en *Twitter* y su correspondencia con las preocupaciones ciudadanas en la contienda presidencial en Colombia 2018”. *Signo y pensamiento*, v. 39, n. 77.
<https://doi.org/10.11144/Javeriana.syp39-76.detc>
- Alves, Jairo L.; Weitzel, Leila; Quesma, Paulo; Cardoso, Carlos E.; Cunha, Luan** (2019). “Brazilian presidential elections in the era of misinformation: A machine learning approach to analyse fake news”. In: Nyström, Ingela; Hernández-Hereidia, Yanio; Milián-Núñez, Vladimir (eds.). *Progress in pattern recognition, image analysis, computer vision, and applications*. Springer International Publishing, pp. 72-84.
https://doi.org/10.1007/978-3-030-33904-3_7
- Amaral, Marcelo-Santos; De-Pinho, José-Antonio-Gomes** (2017). “Ideologias partidárias em 140 caracteres: Uso do *Twitter* pelos parlamentares brasileiros”. *Revista de administração pública*, v. 51, n. 6, pp. 1041-1057.
<https://doi.org/10.1590/0034-7612155837>
- Andrade-del-Cid, Patricia; Flores-González, Rubén; Pablo-Contreras, Mariangel** (2020). “Comportamiento de las comunidades digitales en *Twitter* durante las elecciones México 2018”. *Revista de comunicación*, v. 19, n. 1, pp. 19-36.
<https://doi.org/10.26441/RC19.1-2020-A2>
- Ausserhofer, Julian; Maireder, Axel** (2013). “National politics on Twitter: Structures and topics of a networked public sphere”. *Information, communication & society*, v. 16, n. 3, pp. 291-314.
<https://doi.org/10.1080/1369118X.2012.756050>
- Bakker, Tom P.; De-Vreese, Claes H.** (2011). “Good news for the future? Young people, internet use, and political participation”. *Communication research*, v. 38, n. 4, pp. 451-470.
<https://doi.org/10.1177/0093650210381738>
- Beltrán, Ulises** (2020). “Traditional media and social networks in the 2018 presidential election”. *Política y gobierno*, v. 27, n. 2.
https://www.scielo.org.mx/scielo.php?pid=S1665-20372020000200012&script=sci_arttext
- Bernábe-Loranca, M. Beatriz; González-Velázquez, Enrique-Espinoza; Cerón-Garnica, Carmen** (2020). “Algorithm for collecting and sorting data from Twitter through the use of dictionaries in Python”. *Computación y sistemas*, v. 24, n. 2.
<https://doi.org/10.13053/cys-24-2-3408>
- Brandt, Mariana-Baptista; Vidotti, Silvana-Aparecida-Borsetti-Gregório** (2020). “Dados de pesquisa em informação legislativa”. *Encontros Bibli: Revista eletrônica de biblioteconomia e ciência da informação*, n. 25.
<https://doi.org/10.5007/1518-2924.2020.e72208>
- Brito, Kellyton; Paula, Natalia; Fernandes, Manoel; Meira, Silvio** (2019). “Social media and presidential campaigns - Preliminary results of the 2018 Brazilian presidential election”. *Proceedings of the 20th Annual international conference on digital government research*, pp. 332-341.
<https://doi.org/10.1145/3325112.3325252>
- Buccoliero, Luca; Bellio, Elena; Crestini, Giulia; Arkoudas, Alessandra** (2020). “Twitter and politics: Evidence from the US presidential elections 2016”. *Journal of marketing communications*, v. 26, n. 1, pp. 88-114.
<https://doi.org/10.1080/13527266.2018.1504228>
- Cabrera-Méndez, Marga; Cisternas-Osorio, Rodrigo; López-Navarrete, Alberto J.; Díez-Somavilla, Rebeca** (2021). “Misleading discourse on Instagram: A multimodal study of Latin American presidential candidates in the face of Covid-19”. *Anàlisi*, n. 64, pp. 27-47.
<https://doi.org/10.5565/rev/analisi.3404>
- Calais-Guerra, Pedro-Henrique; Veloso, Adriano; Meira, Wagner; Almeida, Virgílio** (2011). “From bias to opinion: A transfer-learning approach to real-time sentiment analysis”. *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '11*, 150.
<https://doi.org/10.1145/2020408.2020438>

- Camp, Roderic A. I.** (2013). "The 2012 presidential election and what it reveals about Mexican voters". *Journal of Latin American studies*, v. 45, n. 3, pp. 451-481.
<https://doi.org/10.1017/S0022216X1300076X>
- Canavilhas, João; Bittencourt, Maíra; De-Andrade, Marco** (2019). "Viral content on Facebook: A case study on the run-up to the Brazilian 2018 presidential elections". *Brazilian journalism research*, v. 15, n. 3, pp. 562-589.
<https://doi.org/10.25200/BJR.v15n3.2019.1171>
- Carlson, Tom; Djupsund, Göran; Strandberg, Kim** (2013). "Taking risks in social media campaigning: The early adoption of blogging by candidates". *Scandinavian political studies*, v. 37, n. 1, pp. 21-40.
<https://doi.org/10.1111/1467-9477.12011>
- Casero-Ripollés, Andreu** (2018). "Research on political information and social media: Key points and challenges for the future". *El profesional de la información*, v. 27, n. 5, pp. 964-974.
<https://doi.org/10.3145/epi.2018.sep.01>
- Casero-Ripollés, Andreu; Alonso-Muñoz, Laura; Marcos-García, Silvia** (2022). "The influence of political actors in the digital public debate on twitter about the negotiations for the formation of the government in Spain". *American behavioral scientist*, v. 66, n. 3, pp. 307-322.
<https://doi.org/10.1177/00027642211003159>
- Castanho-Silva, Bruno; Proksch, Sven-Oliver** (2022). "Politicians unleashed? Political communication on Twitter and in parliament in Western Europe". *Political science research and methods*, v. 10, n. 4, pp. 776-792.
<https://doi.org/10.1017/psrm.2021.36>
- Castillo, Samara; Allende-Cid, Héctor; Palma, Wenceslao; Alfaro, Rodrigo; Ramos, Heitor S.; González, Cristian; Elortegui, Claudio; Santander, Pedro** (2019). "Detection of bots and cyborgs in Twitter: A study on the Chilean presidential election in 2017". In: G. Meiselwitz (ed.). *Social computing and social media. Design, human behavior and analytics*, pp. 311-323. Springer International Publishing.
https://doi.org/10.1007/978-3-030-21902-4_22
- Cerón-Guzmán, Jhon-Adrian; León-Guzmán, Elizabeth** (2016). "A sentiment analysis system of Spanish tweets and its application in Colombia 2014 presidential election". *2016 IEEE International conferences on big data and cloud computing (BDCloud), Social computing and networking (SocialCom), Sustainable computing and communications (SustainCom) (BDCloud-SocialCom-SustainCom)*, pp. 250-257.
<https://doi.org/10.1109/BDCloud-SocialCom-SustainCom.2016.47>
- Chi, Feng; Yang, Nathan** (2010). "Twitter adoption in congress". *Review of network economics*, v. 30, n. 1.
<https://doi.org/10.2202/1446-9022.1255>
- Cook, James M.** (2016). "Twitter adoption in U.S. legislatures: A fifty-state study", article 10.
<https://doi.org/10.1145/2930971.2930982>
- Cook, James M.** (2017). "Twitter adoption and activity in U.S. legislatures: A 50-state study". *American behavioral scientist*, v. 61, n. 7, pp. 724-740.
<https://doi.org/10.1177/0002764217717564>
- Cremonese, Dejalma** (2012). "Política on-line: A utilização do Twitter como ferramenta de capital social nas eleições presidenciais de 2010". *Sociedade e cultura*, v. 15, n. 1.
<https://doi.org/10.5216/sec.v15i1.20679>
- Cruz-Romero, Roberto** (2015). "Politics 2.0 ICTs in Costa Rica's 2014 Presidential campaign". *International conference on information society, i-Society*, pp. 115-120.
<https://doi.org/10.1109/i-Society.2015.7366870>
- Dajer, Diana** (2021). "Colombia. The use of social media in democratic spaces: a double-edged sword. In: L. Schirch (ed.). *Social media impacts on conflict and democracy*, 1st ed., pp. 59-80. Routledge.
<https://doi.org/10.4324/9781003087649-4>
- De-Carvalho, Fernanda-Cavassana; Massuchin, Michele-Goulart; Mitozo, Isabelle-Batista** (2018). "Radicalização nas redes sociais: Comentários no Facebook durante a disputa presidencial em 2014 no Brasil". *Análise social*, v. 53, n. 229, pp. 2182-2999.
<https://doi.org/10.31447/AS00032573.2018229.03>
- De-León, Ernesto; Vermeer, Susan; Trilling, Damian** (2021). "Electoral news sharing: A study of changes in news coverage and Facebook sharing behaviour during the 2018 Mexican elections". *Information, communication & society*, v. 26, n. 6, pp. 1193-1209.
<https://doi.org/10.1080/1369118X.2021.1994629>

- Druckman, James N.; Kifer, Martin J.; Parkin, Michael** (2007). "The technological development of congressional candidate web sites: How and why candidates use web innovations". *Social science computer review*, v. 25, n. 4, pp. 425-442. <https://doi.org/10.1177/0894439307305623>
- Enli, Gunn** (2017). "Twitter as arena for the authentic outsider: Exploring the social media campaigns of Trump and Clinton in the 2016 US presidential election". *European journal of communication*, v. 32, n. 1, pp. 50-61. <https://doi.org/10.1177/0267323116682802>
- Evans, Heather K.; Cordova, Victoria; Sipole, Savannah** (2014). "Twitter style: An analysis of how house candidates used Twitter in their 2012 campaigns". *PS: Political science & politics*, v. 47, n. 2, pp. 454-462. <https://doi.org/10.1017/S1049096514000389>
- Fazekas, Zoltan; Popa, Sebastian-Adrian; Schmitt, Hermann; Barberá, Pablo; Theocharis, Yannis** (2021). "Elite-public interaction on Twitter: EU issue expansion in the campaign". *European journal of political research*, v. 60, n. 2, pp. 376-396. <https://doi.org/10.1111/1475-6765.12402>
- Ferreira, Gabriel-Estavaringo; Santos, Bianca-Lima; Torres-do-Ó, Marcelo; Braz, Rafael-Rodrigues; Digiampietri, Luciano-Antonio** (2021). "Social bots detection in Brazilian presidential elections using natural language processing". *XVII Brazilian symposium on information systems*. <https://doi.org/10.1145/3466933.3466991>
- Filer, Tanya; Fredheim, Rolf** (2017). "Popular with the robots: Accusation and automation in the Argentine presidential elections, 2015". *International journal of politics, culture, and society*, v. 30, n. 3, pp. 259-274. <https://doi.org/10.1007/s10767-016-9233-7>
- Fuente-Alba-Cariola, Fernando; Parada-Gavilán, Carolina** (2019). "Eficacia de los sitios webs como herramienta comunicacional de los diputados chilenos". *Revista de comunicación*, v. 18, n. 2, pp. 139-154. <https://doi.org/10.26441/RC18.2-2019-A7>
- García-Sánchez, Efraín; Benetti, Pedro-Rolo; Higa, Gustavo-Lucas; Álvarez, Marcos-César; Gómez-Nieto, Erick** (2021). "Political discourses, ideologies, and online coalitions in the Brazilian Congress on Twitter during 2019". *New media & society*, v. 25, n. 5. <https://doi.org/10.1177/14614448211017920>
- Gibson, Rachel K.; McAllister, Ian** (2006). "Does cyber-campaigning win votes? Online communication in the 2004 Australian election". *Journal of elections, public opinion and parties*, v. 16, n. 3, pp. 243-263. <https://doi.org/10.1080/13689880600950527>
- Gilmore, Jason** (2012). "Ditching the pack: Digital media in the 2010 Brazilian congressional campaigns". *New media & society*, v. 14, n. 4, pp. 617-633. <https://doi.org/10.1177/1461444811422429>
- González-Tule, Luis; Restrepo-Echavarría, Néstor-Julián** (2020). "Campañas presidenciales de 2018 en México: Nivel de profesionalización, equipo de campaña y estrategias políticas". *América Latina hoy*, v. 84, pp. 137-161. <https://doi.org/10.14201/alh.21009>
- Grant, Will J.; Moon, Brenda; Grant, Janie-Busby** (2010). "Digital dialogue? Australian politicians' use of the social network tool Twitter". *Australian journal of political science*, v. 45, n. 4, pp. 579-604. <https://doi.org/10.1080/10361146.2010.517176>
- Green, Andrew** (2021a). "'Yo te AMLO': Sentimentalism, threat and affective flows in political campaign song videos in Mexico". *Media, culture & society*, v. 43, n. 3, pp. 411-427. <https://doi.org/10.1177/0163443720974248>
- Green, Andrew** (2021b). "'Demo' and 'cracy': Music, trust, and authentication in Mexico's 2018 elections". *Cultural studies*, v. 37, n. 3, pp. 418-443. <https://doi.org/10.1080/09502386.2021.1978518>
- Gulati, Girish J.; Williams, Christine B.** (2007). "Closing the gap, raising the bar: Candidate web site communication in the 2006 campaigns for Congress". *Social science computer review*, v. 25, n. 4, pp. 443-465. <https://doi.org/10.1177/0894439307305624>
- Gulati, Girish J.; Williams, Christine B.** (2010). "Congressional candidates' use of YouTube in 2008: Its frequency and rationale". *Journal of information technology & politics*, v. 7, n. 2, pp. 93-109. <https://doi.org/10.1080/19331681003748958>
- Gulati, Girish J.; Williams, Christine B.** (2013). "Social media and campaign 2012: Developments and trends for Facebook adoption". *Social science computer review*, v. 31, n. 5, pp. 577-588. <https://doi.org/10.1177/0894439313489258>

- Gulías, Erika-Jaráiz; López-López, Paulo-Carlos; Boubeta, Miguel-Anxo-Bastos** (2020). "Economía, política social y Twitter: Análisis de las emociones negativas en cuatro elecciones presidenciales latinoamericanas a través del LIWC". *Risti. Revista ibérica de sistemas e tecnologías de informação*, n. E26, pp. 270-282.
<https://shorturl.at/rGNO8>
- Haman, Michael; Školník, Milan** (2021). "Politicians on social media. The online database of members of national parliaments on Twitter". *Profesional de la información*, v. 30, n. 2, e300217.
<https://doi.org/10.3145/epi.2021.mar.17>
- Haman, Michael; Školník, Milan; Čopík, Jan** (2022). "Members of the European Parliament on social media: Why do they use Facebook and Twitter?". *Political studies review*, v. 21, n. 2.
<https://doi.org/10.1177/14789299221082444>
- Hargittai, Eszter** (2007). "Whose space? Differences among users and non-users of social network sites". *Journal of computer-mediated communication*, v. 13, n. 1, pp. 276-297.
<https://doi.org/10.1111/j.1083-6101.2007.00396.x>
- Hargreaves, Eduardo; Mangabeira, Eduardo F.; Oliveira, Jonice; Franca, Tiago C.; Mcnasché, Daniel S.** (2020). "Facebook news feed personalization filter: A case study during the Brazilian elections". *IEEE/ACM International conference on advances in social networks analysis and mining (ASONAM)*, pp. 615-618.
<https://doi.org/10.1109/ASONAM49781.2020.9381301>
- Heith, Diane J.** (2003). "11 - The lipstick watch: Media coverage, gender, and presidential campaigns". In: R. P. Watson; A. Gordon (eds.). *Anticipating Madam President*, pp. 121-130. Lynne Rienner Publishers.
<https://doi.org/10.1515/9781685855185-013>
- Henríquez, Pablo A.; Sabat, Jorge; Sullivan, José-Patricio** (2022). "Politicians' willingness to agree: Evidence from the interactions in Twitter of Chilean deputies". *Journal of information technology & politics*, v. 20, n. 1, pp. 92-111.
<https://doi.org/10.1080/19331681.2022.2056278>
- Herrnson, Paul S.; Stokes-Brown, Atiya-Kai; Hindman, Matthew** (2007). "Campaign politics and the digital divide: Constituency characteristics, strategic considerations, and candidate internet use in state legislative elections". *Political research quarterly*, v. 60, n. 1, pp. 31-42.
<https://doi.org/10.1177/1065912906298527>
- Ho, Justin-Chun-Ting** (2020). "How biased is the sample? Reverse engineering the ranking algorithm of Facebook's graph application programming interface". *Big data & society*, v. 7, n. 1.
<https://doi.org/10.1177/2053951720905874>
- Hox, Joop J.** (2010). *Multilevel analysis: Techniques and applications*. Routledge. ISBN: 978 1 138121362
- International Telecommunication Union* (2022). *Percentage of individuals using the Internet*.
<https://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx>
- Ituassu, Artur; Lifschitz, Sergio; Capone, Leticia; Vaz, Maria-Beatriz; Mannheimer, Vivian** (2018). "Compartilhamento de mídia e preferência eleitoral no Twitter: Uma análise de opinião pública durante as eleições de 2014 no Brasil". *Palavra chave*, v. 21, n. 3, pp. 860-884.
<https://doi.org/10.5294/pacla.2018.21.3.9>
- Jacobs, Kristof; Spierings, Niels** (2019). "A populist paradise? Examining populists' Twitter adoption and use". *Information, communication & society*, v. 22, n. 12, pp. 1681-1696.
<https://doi.org/10.1080/1369118x.2018.1449883>
- Jiménez-Zarate, Carlos-Augusto** (2018). "Social neointeraction on Facebook, presidential campaign Mexico 2018". In: A. J. Morales; C. Gershenson; D. Braha; A. A. Minai; Y. Bar-Yam (eds.). *Unifying themes in complex systems IX*, pp. 406-413. Springer International Publishing. ISBN: 978 3 319 96661 8
https://doi.org/10.1007/978-3-319-96661-8_42
- Jivkova-Semova, Dimitrina; Requeijo-Rey, Paula; Padilla-Castillo, Graciela** (2017). "Usos y tendencias de Twitter en la campaña a elecciones generales españolas del 20D de 2015: hashtags que fueron trending topic". *El profesional de la información*, v. 26, n. 5, pp. 824-837.
<https://doi.org/10.3145/epi.2017.sep.05>
- Jungherr, Andreas** (2016). "Twitter use in election campaigns: A systematic literature review". *Journal of information technology & politics*, v. 13, n. 1, pp. 72-91.
<https://doi.org/10.1080/19331681.2015.1132401>
- Kahn, Kim-Fridkin** (1996). *The political consequences of being a woman: How stereotypes influence the conduct and consequences of political campaigns*. Columbia University Press. ISBN: 978 0 231103039

- Kavanaugh, Andrea L.; Sheetz, Steven D.; Sandoval-Almazan, Rodrigo; Tedesco, John C.; Fox, Edward A.** (2016). "Media use during conflicts: Information seeking and political efficacy during the 2012 Mexican elections". *Government information quarterly*, v. 33, n. 3, pp. 595-602.
<https://doi.org/10.1016/j.giq.2016.01.004>
- Kearney, Michael W.** (2019). "rtweet: Collecting and analyzing Twitter data". *Journal of open source software*, v. 4, n. 42, pp. 1829.
<https://doi.org/10.21105/joss.01829>
- Kobellarz, Jordan K.; Graeml, Alexandre R.; Reddy, Michelle; Silva, Thiago H.** (2019). "Parrot talk: Retweeting among Twitter users during the 2018 Brazilian presidential election". *Proceedings of the 25th Brazillian symposium on multimedia and the web*, pp. 221-228.
<https://doi.org/10.1145/3323503.3349559>
- Kreft, Ita G. G.; Bokhee, Yon** (1996). *Are multilevel techniques necessary? An overview, including simulation studies*. California State University.
<https://eric.ed.gov/?id=ED371033>
- Lappas, Georgios; Triantafillidou, Amalia; Yannas, Prodromos** (2018). "Social media adoption and popularity of members of the European Parliament". In: *Proceedings of the 5th Multidisciplinary international social networks conference (Misc)*, article 22.
<https://doi.org/10.1145/3227696.3227722>
- Lappas, Georgios; Triantafillidou, Amalia; Yannas, Prodromos** (2019). "Members of European Parliament (MEPs) on social media: Understanding the underlying mechanisms of social media adoption and popularity". *The review of socio-network strategies*, v. 13, n. 1, pp. 55-77.
<https://doi.org/10.1007/s12626-019-00033-5>
- Lappas, Georgios; Triantafillidou, Amalia; Yannas, Prodromos; Kavada, Anastasia; Kleftodimos, Alexandros; Vasileiadou, Olga** (2016). "Social media battles: Their impact during the 2014 Greek municipal elections". *Journal of universal computer science*, n. 22, pp. 375-393.
<https://doi.org/10.3217/jucs-022-03-0375>
- Larsson, Anders-Olof** (2015). "The EU Parliament on Twitter - Assessing the permanent online practices of parliamentarians". *Journal of information technology & politics*, v. 12, n. 2, pp. 149-166.
<https://doi.org/10.1080/19331681.2014.994158>
- Larsson, Anders-Olof; Kalsnes, Bente** (2014). "'Of course we are on Facebook': Use and non-use of social media among Swedish and Norwegian politicians". *European journal of communication*, v. 29, n. 6, pp. 653-667.
<https://doi.org/10.1177/0267323114531383>
- Larsson, Anders-Olof; Moe, Hallvard** (2012). "Studying political microblogging: Twitter users in the 2010 Swedish election campaign". *New media & society*, v. 14, n. 5, pp. 729-747.
<https://doi.org/10.1177/1461444811422894>
- Lasorsa, Dominic** (2012). "Transparency and other journalistic norms on Twitter: The role of gender". *Journalism studies*, v. 13, n. 3, pp. 402-417.
<https://doi.org/10.1080/1461670X.2012.657909>
- Lassen, David S.; Brown, Adam R.** (2011). "Twitter: The electoral connection?". *Social science computer review*, v. 29, n. 4.
<https://doi.org/10.1177/0894439310382749>
- Latinobarómetro* (2022). *Latinobarómetro 2020*.
<https://www.latinobarometro.org/lat.jsp>
- Lawless, Jennifer L.** (2012). "Twitter and Facebook: New ways for members of Congress to send the same old messages?". In: R. L. Fox; J. M. Ramos (eds.). *iPolitics*, 1st ed., pp. 206-232. Cambridge University Press.
<https://doi.org/10.1017/CBO9781139059893.013>
- Levy, Helton; Sarmiento, Claudia** (2020). "Understanding viral Communism: A thematic analysis of Twitter during Brazil's 2018 elections". *Westminster papers in communication and culture*, v. 15, n. 1, pp. 19-36.
<https://doi.org/10.16997/wpcc.322>
- López-Chau, Asdrúbal; Valle-Cruz, David; Sandoval-Almazán, Rodrigo** (2019). "Analyzing polarization through social media with artificial intelligence: The Mexican presidential election in 2018". In: *Proceedings of the 20th Annual international conference on digital government research*, pp. 502-503.
<https://doi.org/10.1145/3325112.3328214>

- López-García, Guillermo** (2016). "New' vs 'old' leaderships: The campaign of Spanish general elections 2015 on Twitter". *Communication & society*, v. 29, n. 3, pp. 149-168.
<https://doi.org/10.15581/003.29.3.149-168>
- López-López, Paulo-Carlos; Oñate, Pablo** (2019). "De la videopolítica a la ciberpolítica: Debate entre candidatas y tele visiones en cinco elecciones presidenciales". *El profesional de la información*, v. 28, n. 5, e280512.
<https://doi.org/10.3145/epi.2019.sep.12>
- López-López, Paulo-Carlos; Oñate, Pablo; Chavero-Ramírez, Palmira** (2018). "The horse-race campaign in front of the issue-campaign in the digital agenda: Presidential elections in Latin America in the period 2015-2017". *Risti. Revista ibérica de sistemas e tecnologias de informação*, (E16), pp. 270-282.
- López-López, Paulo-Carlos; Oñate, Pablo; Rocha, Álvaro** (2020). "Social media mining, debate and feelings: Digital public opinion's reaction in five presidential elections in Latin America". *Cluster computing*, v. 23, n. 3, pp. 1875-1886.
<https://doi.org/10.1007/s10586-020-03072-8>
- López-López, Paulo-Carlos; Vásquez-González, Javier** (2018). "Agenda temática y *Twitter*: Elecciones presidenciales en América Latina durante el período 2015-2017". *El profesional de la información*, v. 27, n. 6, pp. 1204-1214.
<https://doi.org/10.3145/epi.2018.nov.04>
- López-Meri, Amparo; Marcos-García, Silvia; Casero-Ripollés, Andreu** (2017). "What do politicians do on *Twitter*? Functions and communication strategies in the Spanish electoral campaign of 2016". *El profesional de la información*, v. 26, n. 5, pp. 795-804.
<https://doi.org/10.3145/epi.2017.sep.02>
- Machado, Caio; Kira, Beatriz; Narayanan, Vidya; Kollanyi, Bence; Howard, Philip** (2019). "A study of misinformation in WhatsApp groups with a focus on the Brazilian presidential elections". *Companion proceedings of The 2019 World Wide Web conference*, pp. 1013-1019.
<https://doi.org/10.1145/3308560.3316738>
- Macková, Alena; Štětka, Václav** (2016). "Walking the party line? The determinants of Facebook's adoption and use by Czech parliamentarians". *Medijske studije*, v. 7, n. 14, pp. 157-175.
<https://doi.org/10.20901/ms.7.14.11>
- Marques, Francisco-Paulo-Jamil-Almeida; De-Aquino, Jakson-Alves; Miola, Edna** (2014a). "Congressmen in the age of social network sites: Brazilian representatives and Twitter use". *First Monday*, v. 19, n. 5.
<https://doi.org/10.5210/fm.v19i5.5022>
- Marques, Francisco-Paulo-Jamil-Almeida; De-Aquino, Jakson-Alves; Miola, Edna** (2014b). "Parlamentares, representação política e redes sociais digitais". *Opinião pública*, v. 20, n. 2, pp. 178-203.
<https://doi.org/10.1590/1807-01912014202178>
- Massuchin, Michele-Goulart; Campos-Domínguez, Eva** (2016). "Elecciones de Brasil en 2014: Las webs de los candidatos presidenciales como herramienta de campaña electoral". *Observatorio (OBS*)*, v. 10, n. 1, pp. 83-106.
<http://www.scielo.pt/pdf/obs/v10n1/v10n1a05.pdf>
- Matassi, Mora; Boczkowski, Pablo J.** (2020). Redes sociales en Iberoamérica. Artículo de revisión. *El profesional de la información*, v. 29, n. 1, e290104.
<https://doi.org/10.3145/epi.2020.ene.04>
- Metag, Julia; Marcinkowski, Frank** (2012). "Strategic, structural, and individual determinants of online campaigning in German elections". *Policy & internet*, v. 4, n. 3-4, pp. 136-158.
<https://doi.org/10.1002/poi.3.14>
- Mitoto, Isabel-Batista; Massuchin, Michele-Goulart; De-Carvalho, Fernanda-Cavassana** (2017). "Debate político-eleitoral no Facebook: Os comentários do público em posts jornalísticos na eleição presidencial de 2014". *Opinião pública*, v. 23, n. 2, pp. 459-484.
<https://doi.org/10.1590/1807-01912017232459>
- Montes-de-Oca-López, Juan-Carlos; Sandoval-Almazán, Rodrigo** (2019). "Assessing social media platforms in the independent presidential candidates race in Mexico in 2018". *25th Americas conference on information systems (Amcis)*.
<https://aisel.aisnet.org/amcis2019>
- Mussi-Reyero, Tomás; Beiró, Mariano G.; Álvarez-Hamelin, J. Ignacio; Hernández, Laura; Kotzinos, Dimitris** (2021). "Evolution of the political opinion landscape during electoral periods". *EPJ Data science*, v. 10, n. 1, article 31.
<https://doi.org/10.1140/epjds/s13688-021-00285-8>
- Neihouser, Marie** (2021). "Representation style and digital communication: The case of the online presence of French MPs elected in 2017". *The journal of legislative studies*, v. 29, n. 1, pp. 78-95.
<https://doi.org/10.1080/13572334.2021.1941544>

- Neihouser, Marie; Tremblay-Antoine, Camille** (2021). "The literature on the parliamentarians' digital communication: A scoping review", preprint.
<https://doi.org/10.13140/RG.2.2.31633.15202>
- Novais, Rui-Alexandre; De-Araújo, Viviane-Afonso** (2012). "1.0 mais do que 2.0? O uso tradicional de suportes inovadores nas campanhas eleitorais". *Sociedade e cultura*, v. 15, n. 1.
<https://doi.org/10.5216/sec.v15i1.20669>
- Obholzer, Lukas; Daniel, William T.** (2016). "An online electoral connection? How electoral systems condition representatives' social media use". *European Union politics*, v. 17, n. 3, pp. 387-407.
<https://doi.org/10.1177/1465116516630149>
- Olivares-Fernández, Gastón; Cárdenas, Juan-Pablo; Losada, Juan-Carlos; Borondo, Javier** (2019). "Opinion polarization during a dichotomous electoral process". *Complexity*, v. 2019, article ID 5854037.
<https://doi.org/10.1155/2019/5854037>
- Oliveira, Daniel-José-Silva; Bermejo, Paulo-Henrique-de-Souza; Dos-Santos, Pâmela-Aparecida** (2017). "Can social media reveal the preferences of voters? A comparison between sentiment analysis and traditional opinion polls". *Journal of information technology & politics*, v. 14, n. 1, pp. 34-45.
<https://doi.org/10.1080/19331681.2016.1214094>
- Oliveira, Lucasa S.; Vaz-De-Melo, Pedro O. S.; Amaral, Marcelo S.; Pinho, José-Antônio G.** (2018). "When politicians talk about politics: Identifying political tweets of Brazilian congressmen". In: *The 12th international AAAI conference on web and social media (Icwsml)*, pp. 664-667.
- Ortiz-Espinoza, Ángeles; Espejel-Trujillo, Angelina** (2021). "Análisis de sentimiento en las campañas presidenciales México 2018 / Sentiment analysis in presidential campaigns Mexico 2018". *Revista española de investigaciones sociológicas*, n. 173, pp. 79-100.
<https://doi.org/10.5477/cis/reis.173.79>
- Passos, Néelson R. S.; Rodrigues, Ariel F.; Macedo, Hendrik T.; Prado, Bruno O. P.; Da-Silva, Gilton J. F.; Matos, Leonardo N.** (2019). "Open data extraction, transformation, and loading as a tool for supporting 2018 elections' voters". *Proceedings of the XV Brazilian symposium on information systems*, article 23, 8 pp.
<https://doi.org/10.1145/3330204.3330232>
- Pedro-Carañana, Joan; Alvarado-Vivas, Sergio; López-López, Juan S.** (2020). "Agenda-setting and power relations during the 2018 Colombian election campaign on Twitter". *The journal of international communication*, v. 26, n. 2, pp. 260-280.
<https://doi.org/10.1080/13216597.2020.1806900>
- Pérez-Salazar, Gabriel** (2019). "El meme en internet como texto digital: Caracterización y usos sociales en procesos electorales". *Texto livre: Linguagem e tecnologia*, v. 12, n. 1.
<https://doi.org/10.17851/1983-3652.12.1.1-18>
- Peterson, Rolfe-Daus** (2012). "To tweet or not to tweet: Exploring the determinants of early adoption of Twitter by House members in the 111th Congress". *Social science journal*, v. 49, n. 4, pp. 430-438.
<https://doi.org/10.1016/j.soscij.2012.07.002>
- Puertas-Hidalgo, Rosario; Carpio-Jiménez, Lilia; Suing, Abel** (2019). "Political communication and social networks. Twitter: Comparative analysis of the presidencies of the republic of Latin America". *14th Iberian conference on information systems and technologies (Cisti)*, 6 pp.
<https://doi.org/10.23919/CISTI.2019.8760878>
- Rauchfleisch, Adrian; Metag, Julia** (2016). "The special case of Switzerland: Swiss politicians on Twitter". *New media & society*, v. 18, n. 10, pp. 2413-2431.
<https://doi.org/10.1177/1461444815586982>
- Rauchfleisch, Adrian; Metag, Julia** (2020). "Beyond normalization and equalization on Twitter: Politicians' Twitter use during non-election times and influences of media attention". *Journal of applied journalism & media studies*, v. 9, n. 2, pp. 169-189.
https://doi.org/10.1386/ajms_00021_1
- Recuero, Raquel; Bonow-Soares, Felipe; Gruzd, Anatoliy** (2020). "Hyperpartisanship, disinformation and political conversations on Twitter: The Brazilian presidential election of 2018". In: *Fourteenth international AAAI conference on web and social media*, pp. 569-578.
<https://ojs.aaai.org/index.php/ICWSM/article/view/7324>
- Recuero, Raquel; Zago, Gabriela; Bastos, Marco T.** (2015). "Twitter in political campaigns: The Brazilian 2014 presidential election". In: *The Routledge companion to social media and politics*, chapter 37, pp. 518-530.
<https://doi.org/10.4324/9781315716299>

- Reis-Longhi, Carla; Santos-Oliveira, Ivanilce** (2020). "Elecciones 2018 en Brasil: Análisis de las imágenes en Facebook de los candidatos en la segunda vuelta". *Estudios sobre el mensaje periodístico*, v. 26, n. 4, pp. 1455-1469.
<https://doi.org/10.5209/esmp.70143>
- Ribeiro-Ferreira, Ricardo** (2018). "Rede de mentiras: a propagação de fake news na pré-campanha presidencial brasileira". *Observatorio (OBS*)*, v. 12, n. 5, pp. 139-162.
<https://doi.org/10.15847/obsOBS12520181272>
- Rodríguez-Fidalgo, María-Isabel; Ruiz-Paz, Yanira; Paíno-Ambrosio, Adriana** (2019). "#EleccionesMéxico2018: Post election analysis of news media's tweets and citizens' feedback". *Communication & society*, v. 32, n. 4, pp. 193-205.
<https://doi.org/10.15581/003.32.4.193-205>
- Rofrío, Daniel; Ruiz, Anacaren; Sosebee, Erin; Raza, Qasim; Bashir, Adnan; Crandall, Jed; Sandoval, Ramiro** (2019). "Presidential elections in Ecuador: Bot presence in Twitter". *Sixth International conference on edemocracy & e-government (Icedeg)*, pp. 218-223.
<https://doi.org/10.1109/ICEDEG.2019.8734426>
- Rogers, Everett M.** (1962). *Diffusion of innovations*. The Free Press. ISBN: 978 0 598411044
- Ruano, Luis-Eduardo; López, Juan-Camilo; Mosquera, Jonathan-Felipe** (2018). "La política y lo político en Twitter: Análisis del discurso de los candidatos presidenciales de Colombia". *Risti. Revista ibérica de sistemas e tecnologías de informação*, n. 28, pp. 57-71.
<https://doi.org/10.17013/risti.28.57-71>
- Ruiz-Rojas, Melissa; Boguslavskaya, Vera V.** (2018). "The use of new technologies by women in politics in the 2018 elections in Russia and Colombia". *IEEE Communication strategies in digital society workshop (ComSDS)*, pp. 46-48.
<https://doi.org/10.1109/COMSDS.2018.8354984>
- Sandberg, Linn A. C.; Öhberg, Patrik** (2017). "The role of gender in online campaigning: Swedish candidates' motives and use of social media during the European election 2014". *Journal of information technology & politics*, v. 14, n. 4, pp. 314-333.
<https://doi.org/10.1080/19331681.2017.1369918>
- Sandoval, Rodrigo; Torres-Matus, Rodolfo; Nava-Rogel, Rosa** (2012). "Twitter in Mexican politics: Messages to people or candidates?". In: *Amcis 2012 proceedings*, pp. 363-373.
<https://aisel.aisnet.org/amcis2012/proceedings/EGovernment/18>
- Santana, Brenda-Salene; Vanin, Aline-Aver** (2020). "Detecting group beliefs related to 2018's Brazilian elections in tweets: A combined study on modeling topics and sentiment analysis". In: *Proceedings of the Workshop on digital humanities and natural language processing (DHANDLP)*, co-located with *International conference on the computational processing of portuguese (Propor)*, pp. 11-21.
<https://doi.org/10.48550/arXiv.2006.00490>
- Santana, Luis E.; Huerta-Cánepa, Gonzalo** (2019). "¿Son bots? Automatización en redes sociales durante las elecciones presidenciales de Chile 2017". *Cuadernos.info*, n. 44, pp. 61-77.
<https://doi.org/10.7764/cdi.44.1629>
- Santander, Pedro; Elórtogui, Claudio; Buzzo, Camila** (2020). "Twitter, presidential debates and attention economy: A symbiosis between television audience and social media users during campaign season". *Communication & society*, v. 33, n. 3, pp. 51-65.
<https://doi.org/10.15581/003.33.3.51-65>
- Santos, Gustavo-Ferreira** (2020). "Social media, disinformation, and regulation of the electoral process: A study based on 2018 Brazilian election experience". *Revista de investigações constitucionais*, v. 7, n. 2, pp. 429-449.
<https://doi.org/10.5380/rinc.v7i2.71057>
- Scherpereel, John A.; Wohlgemuth, Jerry; Lievens, Audrey** (2018). "Does institutional setting affect legislators' use of Twitter?". *Policy & internet*, v. 10, n. 1, pp. 43-60.
<https://doi.org/10.1002/poi3.156>
- Scherpereel, John A.; Wohlgemuth, Jerry; Schmelzinger, Margaret** (2017). "The adoption and use of Twitter as a representational tool among members of the European Parliament". *European politics and society*, v. 18, n. 2, pp. 111-127.
<https://doi.org/10.1080/23745118.2016.1151125>
- Segado-Boj, Francisco; Díaz-Campo, Jesús; Lloves-Sobrado, Beatriz** (2015). "Líderes latinoamericanos en Twitter. Viejas costumbres para nuevos medios en tiempos de crisis políticas". *Revista latina de comunicación social*, n. 70, pp. 156-173.
<https://doi.org/10.4185/RLCS-2015-1040>

- Snijders, Tom A. B.; Bosker, Roel J.** (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. Sage. ISBN: 978 1 849202015
- Soares, Felipe-Bonow; Recuero, Raquel; Zago, Gabriela** (2019). "Asymmetric polarization on Twitter and the 2018 Brazilian presidential elections". In: *Proceedings of the 10th International conference on social media and society*, pp. 67-76. <https://doi.org/10.1145/3328529.3328546>
- Southern, Rosalynd** (2015). "Is web 2.0 providing a voice for outsiders? A comparison of personal web site and social media use by candidates at the 2010 UK general election". *Journal of information technology & politics*, v. 12, n. 1. <https://doi.org/10.1080/19331681.2014.972603>
- Southern, Rosalynd; Lee, Benjamin J.** (2019). "Politics as usual? Assessing the extent and content of candidate-level online campaigning at the 2015 UK general election". *Journal of elections, public opinion & parties*, v. 29, n. 2, pp. 179-198. <https://doi.org/10.1080/17457289.2018.1515209>
- Strandberg, Kim** (2009). "Online campaigning: An opening for the outsiders? An analysis of Finnish parliamentary candidates' websites in the 2003 election campaign". *New media & society*, v. 11, n. 5, pp. 835-854. <https://doi.org/10.1177/1461444809105355>
- Strandberg, Kim** (2013). "A social media revolution or just a case of history repeating itself? The use of social media in the 2011 Finnish parliamentary elections". *New media & society*, v. 15, n. 8, pp. 1329-1347. <https://doi.org/10.1177/1461444812470612>
- Straus, Jacob R.; Glassman, Matthew-Eric; Shogan, Colleen J.; Smelcer, Susan-Navarro** (2013). "Communicating in 140 characters or less: Congressional adoption of Twitter in the 111th Congress". *PS Political science & politics*, v. 46, n. 1, pp. 60-66. <https://doi.org/10.1017/s1049096512001242>
- Sudulich, Maria-Laura; Wall, Matthew** (2009). "Keeping up with the Murphys? Candidate cyber-campaigning in the 2007 Irish general election". *Parliamentary affairs*, v. 62, n. 3, pp. 456-475. <https://doi.org/10.1093/pa/gsp008>
- Sullivan, Katherine V. R.** (2021). "The gendered digital turn: Canadian mayors on social media". *Information polity*, v. 26, n. 2, pp. 157-171. <https://doi.org/10.3233/ip-200301>
- Teixeira, Carlos; Kurtz, Gabriela; Leuck, Lorenzo; Sanvido, Pedro; Scherer, Joana; Tietzmann, Roberto; Manssour, Isabel; Silveira, Milene** (2019). "Polls, plans and tweets: An analysis of the candidates' discourses during the 2018 Brazilian presidential election". In: *Proceedings of the 20th Annual international conference on digital government research*, pp. 439-444. <https://doi.org/10.1145/3325112.3329882>
- Tomaz, Raíssa-Mendes; Tomaz, Jerzui-Mendes-Torres** (2020). "The Brazilian presidential election of 2018 and the relationship between technology and democracy in Latin America". *Journal of information, communication and ethics in society*, v. 18, n. 4, pp. 497-509. <https://doi.org/10.1108/JICES-12-2019-0134>
- United Nations Development Programme* (2022). *Human development reports - Development index (HDI)*. <https://hdr.undp.org/data-center/human-development-index#/indicies/HDI>
- Van-Vliet, Livia; Törnberg, Petter; Uitermark, Justus** (2020). "The Twitter parliamentarian database: Analyzing Twitter politics across 26 countries". *Plos one*, v. 15, n. 9, e0237073. <https://doi.org/10.1371/journal.pone.0237073>
- Vélez-Loor, Cecibel-Monserrate; Gonzales-Córdova, Mayra-Paola** (2021). "Twitter and political communication: 20/21 presidential pre-campaign and campaign's analysis in Ecuador". *16th Iberian conference on information systems and technologies (Cisti)*, Chaves, Portugal. <https://doi.org/10.23919/CISTI52073.2021.9476426>
- Vergeer, Maurice; Hermans, Liesbeth** (2013). "Campaigning on Twitter: Microblogging and online social networking as campaign tools in the 2010 general elections in the Netherlands". *Journal of computer-mediated communication*, v. 18, n. 4, pp. 399-419. <https://doi.org/10.1111/jcc4.12023>
- Vergeer, Maurice; Hermans, Liesbeth; Cunha, Carlos** (2013). "Web campaigning in the 2009 European Parliament elections: A cross-national comparative analysis". *New media & society*, v. 15, n. 1. <https://doi.org/10.1177/1461444812457337>

Vergeer, Maurice; Hermans, Liesbeth; Sams, Steven (2011). "Is the voter only a tweet away? Micro blogging during the 2009 European Parliament election campaign in the Netherlands". *First Monday*, v. 16, n. 8.

<https://doi.org/10.5210/fm.v16i8.3540>

Waisbord, Silvio; Amado, Adriana (2017). "Populist communication by digital means: Presidential Twitter in Latin America". *Information, communication & society*, v. 20, n. 9, pp. 1330-1346.

<https://doi.org/10.1080/1369118X.2017.1328521>

Walton, S. Courtney; Rice, Ronald E. (2013). "Mediated disclosure on Twitter: The roles of gender and identity in boundary impermeability, valence, disclosure, and stage". *Computers in human behavior*, v. 29, n. 4, pp. 1465-1474.

<https://doi.org/10.1016/j.chb.2013.01.033>

Welp, Yanina; Marzuca, Alejandra (2014). "South American politics in the information age. A study of political parties and MPs on the net in Argentina, Paraguay and Uruguay". In: A. Breuer; Y. Welp (eds.). *Digital technologies for democratic governance in Latin America: Opportunities and risks*, 1st ed., Routledge, pp. 88-108. ISBN: 978 0 203361986

<https://doi.org/10.4324/9780203361986>

Welp, Yanina; Marzuca, Alejandra (2016). "Presencia de partidos políticos y diputados en Internet en Argentina, Paraguay y Uruguay". *Perfiles latinoamericanos*, v. 24, n. 47, pp. 199-224.

<https://doi.org/10.18504/pl2447-011-2016>

Zumárraga-Espinosa, Marcos; Reyes-Valenzuela, Carlos; Carofilis-Cedeño, Cynthia (2017). "¿Verdad o ficción? El uso político de las redes sociales en la participación política offline en las elecciones presidenciales en Ecuador". *Análisis político*, v. 30, n. 91, pp. 130-145.

<https://doi.org/10.15446/anpol.v30n91.70268>

iralis®

International Registry for Authors:
Links to Identify Scientists

es:

- una **guía** para los autores hispanos para que firmen sus trabajos en el formato internacional usual
- una **base de datos** que registra las variantes de firma usadas por cada autor en diferentes épocas
- un **buscador** que usa automáticamente todas las variantes registradas

¡Ni un día más sin decidir tu firma!

Inscríbete en <http://iralis.org>

**El formato de
firma internacional**

Con el apoyo de:

Profesional de la
información



The metaverse: updating the Internet (Web 3.0) or just a new development for immersive videogames?

Jean-Paul Simon

Nota: Este artículo se puede leer en español en:
<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/87390>

Recommended citation:

Simon, Jean-Paul (2023). "The metaverse: updating the Internet (Web 3.0) or just a new development for immersive videogames?". *Profesional de la información*, v. 32, n. 3, e320317.

<https://doi.org/10.3145/epi.2023.may.17>

Manuscript received on 15th April 2023

Accepted on 12th May 2023



Jean-Paul Simon

<https://orcid.org/0000-0001-7331-4797>

JPS Public Policy Consulting

San Vicente 14

41002 Sevilla, Spain

jpsmultimedia@hotmail.com

Abstract

Are we witnessing the return of virtual worlds such as *Deuxième Monde* (1997) or *Second Life* (2003), boosted and enhanced by technologies? Or is it the coming of the next generation of the Internet (Web 3.0)? Or is it just a marketing re-packaging of virtual reality markets that up to now did not deliver as expected? This paper attempts to provide answers to these questions. It introduces the notion of the metaverse, looks at its definition(s), and describes its key elements, thereby outlining the metaverse ecosystem. The article also attempts to draw lessons from the pioneering experiences of former virtual worlds, and thus examines some case studies from the video game industry. In conclusion, we investigate the metaverse's potential constraints (energy/environment, cost of hardware and bandwidth, lack of business models, regulation) and opportunities, and reveal the challenges ahead for its widespread adoption.

Keywords

Metaverse; Virtual worlds; Video games; Web 3.0; Challenges; Opportunities; Evolution; Technologies; Trends; Future; Immersive; Markets.

1. Introduction

Immersive technologies have been characterised by a strong growth over the last years, mostly in the field of videogames; this trend has even been strengthened by the deployment of new technologies such as 5G (Simon, 2019). However, in 2021, a rise in sales of non-fungible tokens (NFTs) as well as announcements from Big Tech players indicating their interest and investment in the space, especially when Facebook re-branded itself as Meta in October 2021, triggered a lot of attention and buzz around a notion that has been around for over 30 years. Indeed, the term Metaverse was pushed forward in the 1992 Neal Stephenson's science-fiction book *Snow Crash*¹.

Metaverse became the stuff of rejuvenated corporate agendas, not only of Meta. Some consultancies have been forecasting that the metaverse industry (whatever the vagueness of its definition, and the remits of the industry) could reach \$USD 800 billion as soon as 2024 vs. about \$USD 500 billion in 2020 (Bloomberg Intelligence, 2021). McKinsey (2022) mentioned a potential of up to \$USD 5 trillion in value by 2030. Citi came with an even more optimistic forecast, stating that, depending on the definition, the positive contribution of the metaverse (Citi GPS, 2022, p. 4) could vary between \$USD 8 and 13 trillion, but \$USD 1 up to 2 trillion under a narrower definition. This was probably highly optimistic especially in the backdrop of what happened to Meta since: a whopping lost of value of \$USD 500 billion over 2022 (Majithia, 2022).



However, optimistic as they may have been, these predictions were predicated on the huge success of video-games supposedly part of this category, such as *Fortnite* or *Roblox*. Video game revenue, indeed, account for half of *Bloomberg's* estimate. Over the last twenty years, video game companies have been investing in virtual and immersive environments with technologies such as augmented or virtual reality. The Chinese video game behemoth *Tencent*, already holding shares (40%) of *Epic Games* (the editor of *Fortnite*) and *Roblox*, has also launched a musical metaverse *Tmeland*. The firm claimed having registered around a hundred trade marks as of September 2021 such as “QQ Metaverse”, “QQ Music Metaverse” and “Kings Metaverse” (Chan, 2022). In November 2021, *China Mobile*, *China Unicom* and *China Telecom* partnered with several tech companies to form *China's Metaverse Industry Committee* (Ye, 2021). One year later, in November 2022, the South Korean telco, *SK Telecom*², launched its *Ifland* platform in 49 international markets aiming at becoming a leading global social metaverse space (Donkin, 2022).

Deloitte (2022) is rather cautious noting that:

“Some observers of recent trends are circumspect about the potential opportunities related to the metaverse, casting recent startup activity as tantamount to the dot-com surge of the late 1990s”.

The apparently inflated figures of the future market are merely the output of an addition of already existing sub-markets (see Figure 2) that may or may not develop in a coordinated fashion and expand within this new ecosystem. Nevertheless, whatever the risks involved, in terms of opportunities as stressed by *McKinsey* (2022, p. 57):

“the metaverse is too big for companies to ignore”.

It also already became a goal of public policies^{3,4}. For instance, the metaverse is among the priorities set up in a post-covid recovery plan in South Korea, part of the *Digital New Deal* agenda, coupled with a 400 million euros strategic plan for the new metaverse industries for the year 2022 (Basdevant; François; Ronfard, 2022, p. 63). On November 1, 2022, the Chinese *MIIIT* jointly released a 4-year action regarded as China's first national-level policy that supports the metaverse development in the country (Interesse, 2022). In September 2022, Thierry Breton, the European Commissioner for the Internal Market, sketched

“a metaverse centred on Europe's values and rules” (Breton, 2022).

The French government commissioned a study on the topic (Basdevant; François; Ronfard, 2022).

The first part of this paper aims at introducing the notion, looking at it(s) definition(s) and describing its key elements, thereby presenting the metaverse ecosystem. The second part tries to draw the lessons from the pioneering experiences of former virtual worlds. It looks at some case studies from the video games industry. The third conclusive part investigates the constraints (energy/environment, cost of the hardware and bandwidth, lack of business models, regulation) and the opportunities. It reveals the challenges ahead for a widespread adoption.

2. Looking for the metaverse

“Metaverse is a portmanteau of meta, meaning transcendent, and verse, from universe” (Zyda, 2021).

If most experts do not agree on a definition of metaverse, they certainly agree that there is no single definition:

“The metaverse is still being defined, both literally and figuratively” (*Bloomberg*, 2022).

One of the most comprehensive studies⁵ of the notion still states:

“An agreed upon definition of the term metaverse within the literature has yet to be agreed on” (Dwivedi et al., 2022).

By the same token, most but not all will also categorize the metaverse as “the next iteration of the Internet” or Web.3⁶, mostly the successor of the mobile Internet, designed to bring the digital and physical worlds together.

3. Defining the metaverse

The metaverse may be tentatively defined as a network of multiuser virtual spaces in 3D, interconnected, interoperable, immersive and persistent, merging physical reality with digital virtuality. At the same time, it should be stressed that most experts also agree that we do not have such a metaverse yet: the existing proto-metaverses available today are neither interconnected, nor interoperable. They are not immersive or persistent⁷ either and are, at best, a spatialised network of virtual experiences (Lamarche-Toloza, 2022).

Besides, according to *Citi* (2022, p. 5):

“In the current state, the internet infrastructure is unsuitable for building a fully-immersive content streaming Metaverse environment, that enables users to go seamlessly from one experience to another”.

“Over the last twenty years, video game companies have been investing in virtual and immersive environments with technologies such as augmented or virtual reality”

“If most experts do not agree on a definition of metaverse, they certainly agree that there is no single definition: “The metaverse is still being defined, both literally and figuratively” (*Bloomberg*, 2022)

In other words, significant investments have still to be made in each of the major components of the metaverse, for instance in next-generation chips, servers, and networking hardware⁸.

The metaverse relies on five major components⁹ and four pillars. The components are:

1) Hardware: devices for access and interface. It encompasses connected devices like mobile phones, PCs, and gaming consoles but also new metaverse-focused hardware (headsets for VR, smart glasses for AR), or haptics to bring the sense of touch.

2) Infrastructure (network and computing): 5G and low latency network, cloud, and edge infrastructure. The metaverse requires compute and processing infrastructure that can support both big data flows and low latency.

3) Content and applications: all the various types of software and content, including gaming, third party content, UGC, developer and creator content.

4) Communities: various use cases with many individuals/users who interact and socialise within the platform and also across applications/platforms (Crédit Suisse, 2022, p. 8).

5) Enabling systems to “settle” transactions for participation, content creation or direct commerce, among which payment and security / identity.

The four pillars are the following:

I. 3D is key for the spatialisation of the web¹⁰ (real time 3D design engines –such as *Unity* or *Unreal*– and 3D models), as well as other visualisation tools (such as avatar development).

II. Formats of extended reality¹¹ (XR: VR, AR, mixed reality and other alternative forms of immersive applications) enable linking the physical and the virtual world.

III. Mass production of the requested content implied using artificial intelligence.

IV. The economy of the metaverse leans on the combination of the blockchain technology¹² and of NFTs, the latter create scarcity and hence value in the virtual worlds, they provide proof of ownership for metaverse-based property.

These pillars are important but not necessary conditions for a metaverse. As Figure 1 illustrates, some of the metaverses, for instance, relies on blockchain or VR helmet but others such as *Roblox* do not.

“The metaverse may be tentatively defined as a network of multiuser virtual spaces in 3D, interconnected, interoperable, immersive, and persistent, merging physical reality with digital virtuality. At the same time, it should be stressed that most experts also agree that we do not have such a metaverse yet: the existing proto-metaverses available today are neither interconnected, nor interoperable”



Figure 1. Mapping some metaverses according to technologies. Source: adapted from Basdevant, François and Ronfard (2022), p. 38.

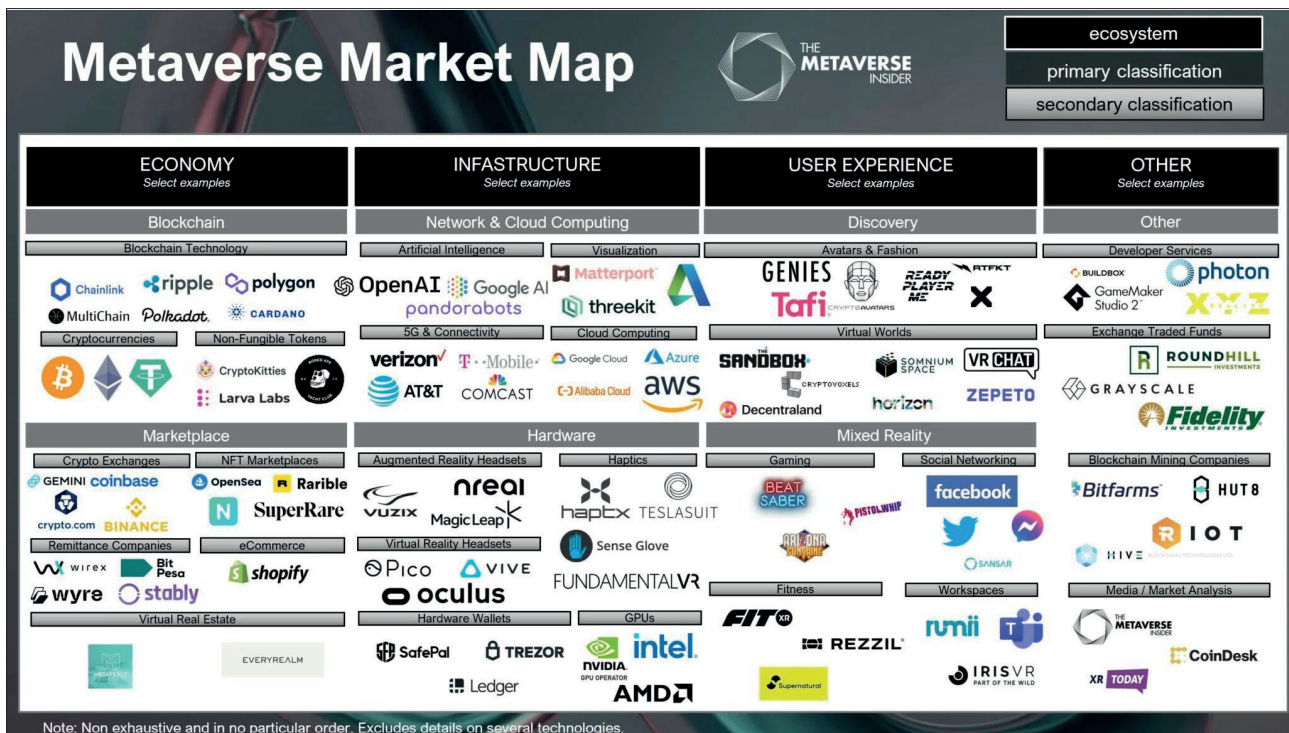


Figure 2. The metaverse market map. Source: Hussain (2022).

4. The metaverse ecosystem: verticals and expected use cases

Still bringing together already existing segments, some experts offer a map of the metaverse (Hussain, 2022; Newzoo, 2022; Radoff, 2021) under their own classification of components. Figure 2 reveals that the sector seems to be rather crowded¹³ with various firms acting in the different sub-segments. It looks more like a mosaic of non-interoperable sub-segments. This points to issues of competition, coordination and standardisation.

Today, the most popular way to experience the metaverse is via a video game played on a virtual reality (VR) headset. Recently, virtual events, performances, and shows have all grown in popularity. Figure 2 shows some examples of use cases but most of them remain by and large anecdotal. For instance, Balenciaga released virtual fashion brands in *Fortnite*, and a digital Gucci “bag” was sold for over \$4,000 on *Roblox* in May 2021 (Cbinsights, 2022). Disney registered a patent known as “virtual world simulator in a real place,” and was planning building a theme park in the metaverse (Citi, 2022, p. 31). However, the new metaverse unit was closed down in 2023. A handful of virtual real estate companies are buying, reselling, developing, and renting virtual properties in decentralized virtual worlds. The major players in virtual real estate are: *Decentraland*, *Sandbox*, *Somnium Space*, and *Cryptovoxels* (Mileva, 2022). Similar companies were already active in *Second Life*¹⁴. *Microsoft Mesh* and *Facebook Horizons Workrooms* offer office applications, *Infosys* introduced the *Infosys* metaverse foundry a platform to navigate technologies such as XR, DLT (distributed ledger technology), 5G, AI, IoT... In education, some game-based learning in virtual worlds have already been introduced (Mystakidis, 2022), and other current developments are happening¹⁵.

Some consultancies may claim that:

“Yet the metaverse appears to be much more than “just entertainment”” (Mind the bridge, 2022).

Nevertheless, as stressed earlier, most of these verticals are still to be developed and the bulk of market consists of games with live entertainment and social media making up the remainder, but these two use cases may just represent additional opportunities for game makers. The same consultancy acknowledges that, even of the case of South Korea, presented as a leading country:



Figure 3. Example of use cases. Source: Citi (2022, p. 4).

“Digital media and gaming applications are the most prominent (and obvious) uses of “metaverse” tech, ranging from 3D content creation, to AR, audio/video, and streaming and broadcasting experiences” (*Mind the bridge*, 2022).

Figure 4 shows a more sobering map of the ecosystem. Therefore, in the next section, we will take a closer look at the game industry and other predecessors of the metaverse.

5. Pioneers: Virtual environments and immersive games

Virtual environments (*Deuxième Monde*, *Habbo Hotel*, *Second Life*) and immersive games (such as *Fortnite*, *Roblox* and *VRChat*) as antecedents of the metaverse may offer some insights into the potential socio-economic impact of a fully functional persistent cross platform metaverse:

“The gaming industry has played a fundamental role in shaping the Metaverse and it is likely to continue doing so with immersive elements like 3D avatars, building new virtual worlds, and observations as a gameplay” (*Citi*, 2022, p. 36).

As early as 2006, during the “Metaverse Roadmap Summit”, experts described the metaverse as

“the meeting of video games and web 2.0” (**Basdevant; François; Ronfard**, 2022, p. 33).

5.1. Communities and virtual worlds¹⁶

In 1989, *Habitat* was the first virtual world platform, a 2D graphical interface enabled cartoon-like avatars to walk around and communicate with chat bubbles. The second wave of social virtual reality systems followed in the 1990s and 2000s. Platforms such as *Traveler*, *Croquet*, *ActiveWorlds*, *There*, *Blue Mars*, *Second Life* and *Open Simulator* used client-server architecture and integrated a graphical user interface and multimedia communication. These communities began as iterations on traditional online chat by providing users with agency and some form of embodiment of their avatar.

CitySpace was one of the first proto-metaverses, active from 1993 on to 1996. The following year, Alain Le Diberder from *Canal+ Multimedia* and a French studio, *Cryo Interactive*, launched the first French proto-metaverse: *Deuxième Monde*, a virtual copy of Paris where users could use chats to communicate with avatars (**Lamy**, 2022). It remained active for five years until being closed down in 2002 for lack of profitability as it was supposed to be funded through advertising and the sales of virtual shops (**Galibert**, 2003). As of 2000, a Finnish company, *Sulake*, opened up an online community, focusing on teens and young adults, under the guise on a virtual hotel *Habbo* where users could visit public areas (restaurants, cinemas and clubs) and create guest rooms. It accumulated 316 million avatars and is still active today (**Partleton**, 2020) with 800,000 monthly active users across 115 countries.

These early services paved the way for the introduction, by *Linden Lab*¹⁷, of *Second Life*¹⁸, a virtual online world, in 2003. Users could create avatars to interact with places, objects and other avatars through chat, IM or voice. By 2013, with 36 million accounts created and 1 million monthly active users it reached revenue of \$USD 3.2 billion from on-site, in-world transactions. The service also used a virtual currency to buy, sell, rent or trade goods and services. Users could participate to the production of news. In addition to magazines, blogs, news bureaus, podcasts and television stations, three newspapers, the “Alphaville Herald”, the “Metaverse Messenger” and the “Second Life Newspaper” were thriving. With 100,000 regular readers, the “Metaverse Messenger” became the most widely read newspaper (**Brennen et al.**, 2010). Although *Second Life* declined, as it did not manage to take the mobile turn among other elements, it is still active claiming 750,000 monthly active users on the platform and \$USD 650 million in annual transactions. As noted by Philip Rosedale, *Second Life*:

“is probably the longest running experiment in the possibilities of a metaverse-like experience” (quoted by **Gent**, 2021).

One of the issues faced by the community was the lack of a good user interface so that its players could easily move through and interact with its various user-built worlds. On the positive side, *Second Life* ushered in the use of a virtual currency for in-world transactions (Linden Dollars), the lessons were grasped by social media providers such as *Facebook* or *Tencent*. Asian game companies pioneered in-games virtual items. More recently, decentralized virtual worlds have been providing similar experiences to virtual worlds (see Box 1).

Today, the most popular way to experience the metaverse is via a video game played on a virtual reality (VR) headset. Recently, virtual events, performances, and shows have all grown in popularity

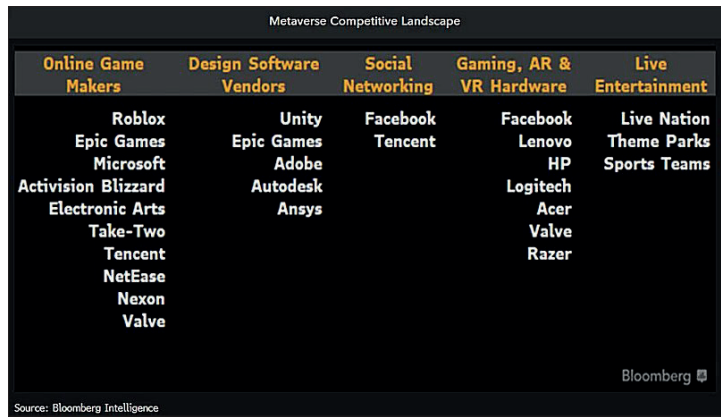


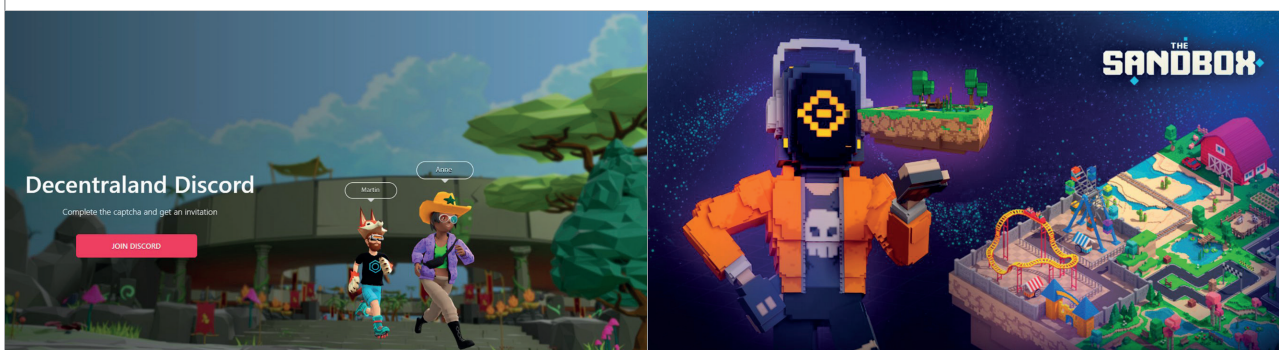
Figure 4. The metaverse competitive landscape. Source: Bloomberg Intelligence (2022).

Box 1. Decentralized virtual worlds

Decentralized virtual worlds such as *Decentraland*, *The Sandbox*, *Cryptovoxels* and *Somnium Space* are usually considered as metaverses or as the more recent proto-metaverses. They offer similar experiences to virtual worlds, but, unlike virtual worlds, are built using blockchain technologies. Transactions are all based on each world's unique cryptocurrency. Furthermore, in-world items or land are traded as NFTs, which act as decentralized proof-of-ownership certificates for digital assets. In contrast, the immersive metaverse is more concerned with how people will interact and experience the virtual world. Their business models differ, as well, from virtual worlds and games: they generate revenue from the sale of virtual land, crypto, and other digital assets as opposed to taking a percentage of the profits generated by in-world creators. The two decentralized worlds seem to be the closer to a community-owned, community-governed, open metaverse.

Decentraland and *The Sandbox* are two of the most popular decentralized worlds. *Decentraland* (<https://decentraland.org>), started off in 2017 and is a completely immersive 3D metaverse. It focuses on enabling players to purchase plots of land alongside exploring the massive virtual universe. MANA is the in-world cryptocurrency of the platform (based on the *Ethereum* blockchain technology). The company claims its world is owned by its users.

Sandbox, founded in 2011 by Arthur Madrid and Sébastien Borget, (<https://www.sandbox.game/en>), has been operating since 2012 as a blockchain-based virtual world and pivoted to a 3D gaming platform in 2017. SAND is the utility token for conducting all transactions on the platform. The platform claims that 70,000,000 worlds have been created; and that over 100,000 worlds are created every day. The is presented as an ecosystem for players and creators, consisting of three main components: a voxel* editor (to empower creators to design 3D voxel objects), a marketplace, and the game itself.



* A portmanteau of volume and pixel used in 3D computer graphics.

Source: compiled by author from *Cbinsights* (2022), *Howell* (2022), and *The Sandbox Whitepaper* (2020).
https://installers.sandbox.game/The_Sandbox_Whitepaper_2020.pdf

There.com, also launched in 2003, was very similar virtual world to *Second Life*, with its own digital currency called The-rebucks. *Messinger et al.* (2008) gave

“one estimate, 20 to 30 million people regularly participated in virtual worlds in 2006, spending an average of almost twenty-two hours per week within these spaces.”

Virtual worlds differ from massively multiplayer online role-playing games as they are not finalized (no goals to reach) and focus on interaction among users and the exploration of the virtual world. Lastly, these virtual worlds can be looked upon as some kind of field trials, of specific R&D as it was noted for *Deuxième Monde* (*Galibert*, 2003).

5.2. Games spearheading the development

Back in the late 1990s, online games were called “persistent worlds (PW),” which refers to continuously and steadily running online worlds, states Steven Ma (Senior Vice President at *Tencent*) (*Ma*, 2022). Gaming companies have been since the early 2000 racing to offer more unique experiences to their users, opening up multiuser virtual worlds and introducing innovative business models based on in-game transactions (*Simon*, 2021). Indeed, in the 2000s, the rise of multiplayer gaming and the launch of MMORPGs (massively multiplayer online games) like *EverQuest* (1999), *Eve Online* (2003), and the *World of Warcraft* (*Blizzard*, 2004) offered an opportunity for developers to test the concept of what will later be called the metaverse. In this section we take a quick look at some of the games that frequently labelled “metaverse” or given as examples of the current metaverse.

Released in 2017 by *Epic Games*, *Fortnite* is among the most successful Free-to-Play (F2P) multiplayer games. *Fortnite* is distributed as three game modes: *Fortnite*, “Save the World” is a player-versus-environment cooperative game, *Fortnite* “Battle Royale” is a player-versus-player game for up to 100 players, and *Fortnite* “Creative” is a sandbox game. *Fortnite* introduced live in-game events: in 2020, Travis Scott’s (an American rapper) “virtual concert” on drew 12.3 million concurrent players (*Lange*, 2020).

Virtual environments (*Deuxième Monde*, *Habbo Hotel*, *Second Life*) and immersive games (such as *Fortnite*, *Roblox* and *VRChat*) as antecedents of the metaverse may offer some insights into the potential socio-economic impact of a fully functional persistent cross platform metaverse

Minecraft and *Roblox* are the leading sandbox games where players can create their own path through the games. Both became some of the world most popular games, *Roblox* with 49.5 million DAU (daily active users) and *Minecraft* with 141 million MAU (monthly active users) (Newzoo, 2022, p. 37). Unlike other video games, but like virtual worlds, sandbox games do not have set objectives thereby granting users some space for creativity. Sandbox games, rather than focusing on linear gameplay and the requirement to complete levels, allow players to freely explore and build environments. Players therefore enjoy a large amount of freedom in choosing how to play the game. The two key industry players dominate the genre of creator games with voxel graphics. *Roblox* is free to play, while *Minecraft* charges around \$30 for the starter pack. Both games offer add-ons and extras.

Gaming companies have been since the early 2000 racing to offer more unique experiences to their users, opening up multiuser virtual worlds and introducing innovative business models based on in-game transactions (Simon, 2021).

*Roblox*¹⁹, launched in 2006, is an online gaming platform and storefront that allows users to develop, play user-created games and capture value. Like with *Second Life*, users are represented via their avatars, they are free to buy, sell, and create digital in-game items as well as participate in both real-life and virtual events. Most games in *Roblox* are multiplayer. *Meep City*, published by *Roblox* in February 2016, was the first *Roblox* game to surpass 1 billion player visits.

Minecraft was released in 2011 by the Swedish studio *Mojang*. Players can explore a pixelated world made from blocks, and voxel graphics-based 3D worlds, in which they can extract raw materials, craft in-game items and build virtually everything. *Minecraft* focuses more on solo play. *Minecraft* is reported to have been sold in over 200 million copies with over 125 million monthly active players (Henningson, 2022). *Minecraft Earth* implements augmented reality. The *Xbox Minecraft Marketplace* allows users to sell the items, models and mini-games they have created to other users, giving the world its own functioning economy.

Chinese companies such as *Tencent* and *NetEase* are increasingly investing into metaverse games. *NetEase*, China's second-largest publisher of online games after *Tencent*, has also invested in the metaverse social network *Imvu*, and some *Sandbox* games. *Tencent* has invested in *Roblox* and holds 40% of the shares of *Epic Games*. The Chinese behemoth is probably already leading in the field through its major portfolio²⁰ as shown in Figure 5.

Leading video games companies have been shopping around to complete their portfolios. Against that backdrop, it is easier to understand the acquisition strategy deployed by *Microsoft* to keep on growing a gaming empire: acquisition of *Mojang* in 2014 for \$USD 2.5 billion, and the planned²¹ acquisition of *Activision Blizzard* \$USD 68.7 billion as of January 2022. *Microsoft* was a leading player in the field of video games in any case. In sharp contrast, although *Facebook* was instrumental for the success of *Zynga*, the company never capitalized on videogames²² which may put the company in a weaker position compared to other major players. If *Fortnite* attracts 250 million active players a month, *Meta's Horizon Worlds* target of 500,000 users by the end of 2022 was not achieved: the current figure is nearer 200,000 (Rose, 2022).

Nevertheless, these strategies only reveal that these companies are trying to diversify their user's experience, offering additional products and services and expecting new streams of revenue. Up-to-now, the new streams of revenue remain marginal. For instance, *Tencent* posted decreased revenues from music and games-related live streaming services (Tencent, 2022, p. 10). The business models of the metaverse related services are still uncertain. The longevity of *Habbo* or *Second Life* testify, notwithstanding, of the ability to serve a niche market.

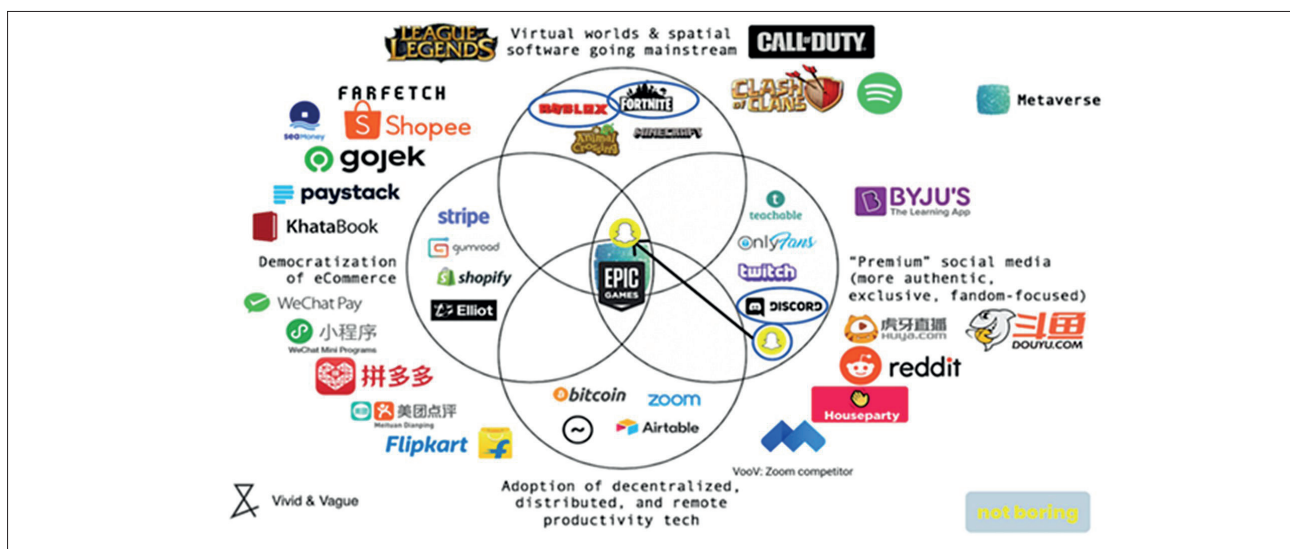


Figure 5. *Tencent's metaverse*.
Source: McCormick (2020), based on Geffen (2020).

An early attempt, Sony's 2008 metaverse *PlayStation Home* that allowed registered *PlayStation Network* users to create an in-game avatar and interact with others in an online social space, was a failure. It remained in beta all the way until its closure in 2015²³. Still on the negative side, pure metaverse players have not been so far fortunate. *Roblox* has seen its stock price plummet by 27% across the last year. Besides, *Roblox* is not a profitable company with a net loss of over \$162 million in Q1 2022 alone (*Newzoo*, 2022, p. 18).

There is still a long way to go before mass adoption, simply seen from a demand side

Play-to-earn gaming has been heralded as

“the job board for the metaverse” (**Brambilla-Hall; Baier-Lentz**, 2021).

The blockchain game *Axie Infinity*, a card collectible title, that allows players to earn tokens and trade in-game assets in real world digital exchanges has been

“the poster child of the blockchain gaming movement” (*Newzoo*, 2022, p. 29)

and grew up very quickly. It turned out that its in-game economy was not sustainable. By the same token, video game companies currently trying to get into NFTs are facing

“constant fan and media backlash” (**Van-der-Velde**, 2021),

as illustrated by the difficult introduction of *Ubisoft Quartz*²⁴ marketplaces. On the opposite, *Minecraft* took a strong position against the introduction of

“NFTs and other blockchain technologies creates digital ownership based on scarcity and exclusion, which does not align with values of creative inclusion and playing together”

<https://www.minecraft.net/en-us/article/minecraft-and-nfts>

6. Conclusion: challenges ahead

We concentrated so far on the supply side but these examples of the negative reactions from already active players indicate that there is still a long way to go before mass adoption, simply seen from a demand side and notwithstanding other issues. The new streams of activities may take some time to deliver: for instance, media fusion activity remains low with only 6% of consumers watching music events within a game, and merchandise sales within live streamed events is also low at 4% consumer penetration (**Mulligan**, 2022). Still on the demand side, corporations, right now, may not be up to embrace the metaverse.

Furthermore, it is likely that distinct media vertical preferences will remain strong. Rosedale stresses that:

“most adults are not yet comfortable engaging with new people, or engaging socially, in a multi-player context online” (quoted by **Gent**, 2021).

From another perspective, it means that one cannot expect the mobile video games boom to be easily reproduced; especially as casual games and now hyper-casual²⁵ are dominating. It is the major trend on the consumer side and it should be stressed that both kind of games are characterized by simple rules and lack of commitment required, do not need a major time investment to play, in contrast to more complex hardcore games targeted at hobbyist gamers. For instance, single player games are the most popular among Indian users and India is the world's biggest mobile game market in downloads (but not in value). For India, LatAm or Africa accessibility and affordability will be key. India is a leading country for software services and videogames but the metaverse may be quite far off. Indian big tech (*Infosys*, *Reliance*, and *Tata*) will focus on business applications (*IND*, 2022) with the possible exception of the video game company *Nazara*.

This also means that there is risk of a new digital divide without and within countries, as well as between generations. There is a potential disenfranchisement of sections of the population unable to access the necessary infrastructure or to acquire the devices needed to access the metaverse. The demand for bandwidth internet traffic will probably grow exponentially. The supply of electricity is still a problem in some countries and has been a barrier for the development of the Internet. And the new technologies are rather greedy in terms of consumption which raises issues of sustainability: the yearly amount of electricity consumed by the *Ethereum* network is estimated to be equivalent to the annual energy consumption of countries like Peru or Qatar (*Newzoo*, 2022, p. 44); avatars consume as much electricity as Brazilians (**Carr**, 2006).

The increasing number of layers and complexity of involved technologies create opportunities but generate a considerable number of uncertainties, according to *Deloitte* (2022): new cyber vulnerabilities and risks related to digital identity and fraud, new challenges in areas such as trust, reputational risk, disinformation, harassment²⁶ and even mental health concerns. Some abuses are already being pointed at, for instance, unregulated use of child labour force by *Roblox* (**Parkin**, 2022). The investments in edge computing, next-generation connectivity, software,

The increasing number of layers and complexity of involved technologies create opportunities but generate a considerable amount of uncertainties

hardware, and talent to support shifts to virtual reality may be significant. Standardization at different levels is also an issue, even if the industry²⁷ is starting addressing the issue, for instance with the *Metaverse Standards Forum* to develop industry guidelines ensuring immersive VR worlds are compatible. However, other technologies (like 3D modelling, volumetric video, and geospatial data) will need to be more tightly integrated.

Video games companies and social media are well positioned to take a leading role in developing of a metaverse, but they may favour closed, proprietary metaverses thereby facing competition challenges especially in the new regulatory environment

Assuming that the metaverse becomes the new iteration of the internet, it will mostly likely attract greater scrutiny from global regulators, policymakers, and governments: issues such as anti-money laundering rules for exchanges and wallets, the use of decentralized finance (DeFi), crypto assets, taxes and property rights will come to the fore. As we have seen, video games companies and social media are well positioned to take a leading role in developing of a metaverse, but they may favour closed, proprietary metaverses thereby facing competition challenges especially in the new regulatory environment. Nevertheless, gamers may be reluctant to jump in the band wagon, as the videogames editor of *The Guardian*, **MacDonald** (2023), puts it:

“The more I hear about the metaverse, the less I want to do with it.”

For the moment, we seem to be quite far from a utopian ideal “open metaverse” that would be community-owned, community-governed, and a freely interoperable version. We are left with competing, non-interoperable proto-metaverses. Berners-Lee, deems blockchain protocols to be a barrier to an effectively decentralized web (**Landymore**, 2022).

Meta allegedly invested more than \$USD 100 billion on research and development and product development in the sector (\$USD 15 billion in 2021 alone) (**Hern**, 2022). The output has been rather disappointing with losses stacking up, especially with *Reality Labs*, its metaverse department, slowing revenue growth.

“Facebook’s metaverse doesn’t have legs” concludes **Hern** (2022).

Maybe the name was stupid! *Meta* has been trying to dominate the field with the “Zuckerverse” (**Mohammed**, 2021) but there are major players competing, so the jury is still out.

7. Notes

1. According to *McKinsey* (2022, p. 12), it even dates back to 1978 with *MUD1*, the first multiplayer real-time virtual world. MUDs (Multi-User Dungeons) were inspired by the role-playing board game *Dungeons & Dragons*.
2. In May 2022, *SK Telecom* partnered with *Deutsche Telekom* to bring its *ifland* metaverse platform to Europe.
3. On November 1, 2022, the *MIIT* jointly released a 4-year action regarded as China’s first national-level policy that supports the metaverse development in the country (**Interesse**, *China Briefing*, November 2022).
4. For a presentation of the strategies of China, Saudi, Arabia, South Korea, and the United Arab Emirates, see **Kshetri** (2023), “National Metaverse Strategies”.
5. An international collective work of 42 authors. For another comprehensive academic presentation see **Mystakidis** (2022). See, as well, the special issue of *Méta-Media* (**Bremme**, 2021): “Métavers et métamedias. Un 3e chapitre d’Internet”.
6. A notion introduced in 2014 by one of the founders of the block chain *Ethereum*, Gavin Wood.
7. With the exception of some massively multiplayer online role-playing games.
8. *Intel* claims that the metaverse will necessitate a 1000x increase in computational efficiency, including advancements in 5G and hybrid edge-cloud infrastructures (*Cbinsights*, 2022).
9. We rely mostly on *Citi* (2022, p. 3) and *Credit Suisse* (2022, p. 4). *McKinsey* considers ten layers that fall into four categories: content and experiences [(content, applications, and virtual worlds), platforms (access and discoveries, creators/3D development platforms), infrastructure and hardware (devices, OS and accessories, infrastructure), and enablers (security/privacy, identity, payment and monetization)] (*McKinsey*, 2022, p. 16).
10. Integrated spatialized voice services first appeared as application for networked virtual environments, e.g., *Second Life*. With a spatialized voice service, people can identify who is talking if there are several participants in the vicinity
11. *XR: Extended reality* a term referring to all real-and-virtual combined environments and human machine interaction generated by computer technology and wearables. It encompasses AR, VR, mixed reality, and other forms of alternate, expanded, or immersive reality applications, including those not yet invented.

Augmented reality is a view of the real-world environment whose elements are supplemented and enhanced by computer-generated sensory input such as sound, video, or graphics.

Virtual reality is an immersive multimedia or computer simulated environment which allows to interact with it.

Mixed reality (also hybrid reality) is the merging of real and virtual worlds to produce new environments where physical and digital objects co-exist and interact in real time.

12. Most blockchain technologies can be divided into four primary categories: private, public, hybrid, and consortium (Hussain, 2022).
13. Radoff's (2021) map is even more overcrowded and Newzoo (2022, pp. 5-6) messier.
14. In 2006, a creator in the virtual world *Second Life*, Anshe Chung made million by buying virtual real estate, redeveloping it with *Second Life*'s creator tools, and renting it out to other *Second Life* inhabitants (Cbinsights, 2022).
15. See also:
<https://olc.worldbank.org/about-olc/education-meets-the-metaverse-reimagining-the-future-of-learning>
16. We follow *Credit Suisse*, p. 8, and Narin (2021).
17. Founded in 1999 by Philip Rosedale:
<https://www.lindenlab.com>
18. See Box 1 in Seekins (2022, p. 18).
19. See the presentation in Newzoo (2022, pp. 16-18).
20. See also the impressive list of Tencent ownership provided Chan (2022).
21. Still to be approved by competition authorities. As of May 2023, the UK's competition regulator has blocked the acquisition.
22. Revenue derived from its commercial relationships with Zynga (publisher of world hits such as "Farmville") were around 15% in 2011, but have been decreasing since to become almost unnoticeable.
22. The presentation at *Meta Connect* was ironically described as "PlayStation Home with worse graphics":
<https://www.playstationlifestyle.net/2022/10/12/mark-zuckerberg-metaverse-playstation-home-ps-home-twitter>
24. Digits (in-game NFTs) that enable to put items on sale to other eligible players at a price set by the player:
<https://quartz.ubisoft.com>
25. "Hypercasual game" or "instant game" are extremely easy-to-learn games that require no download, being played in an existing app like a web browser or messaging app.
26. Sexual harassment did plague *Habbo* and *Second World*.
27. Launched by technology and telecoms giants including *Meta Platforms*, *Huawei*, and *Qualcomm*.

8. References

- Basdevant, Adrien; François, Camille; Ronfard, Rémi** (2022). *Métavers: "10 leviers d'actions" pour une politique française et européenne*. Rapport, Gouvernement, Mission exploratoire sur les métavers.
<https://www.vie-publique.fr/rapport/286878-mission-exploratoire-sur-les-metavers>
- Bloomberg Intelligence* (2021). *Metaverse may be \$800 billion market, next tech platform*.
<https://www.bloomberg.com/professional/blog/metaverse-may-be-800-billion-market-next-tech-platform>
- Brambilla-Hall, Stefan; Baier-Lentz, Moritz** (2021). *What play-to-earn gaming can tell us about the future of the digital economy - and the metaverse*. World Economic Forum.
<https://www.weforum.org/agenda/2021/11/what-play-to-earn-games-mean-for-the-economy-and-metaverse>
- Bremme, Kati** (2021). "Métavers et métamedias. Un 3^e chapitre d'Internet". *Méta-media*, n. 20.
<https://www.meta-media.fr/wp-content/uploads/sites/33/2021/12/METAMEDIA-20.pdf>
- Brennen, Bonnie; Dela-Cerna, Erika** (2010). "Journalism in Second Life". *Journalism studies*, v. 11, n. 4, pp. 546-554.
https://epublications.marquette.edu/cgi/viewcontent.cgi?article=1065&context=comm_fac
- Breton, Thierry** (2022). "People, technologies & infrastructure - Europe's plan to thrive in the metaverse". *Blog of Commissioner Thierry Breton*.
https://ec.europa.eu/commission/presscorner/detail/en/STATEMENT_22_5525
- Carr, Nicholas** (2006). "Avatars consume as much electricity as Brazilians". *Rough type*.
<https://www.roughype.com/?s=Avatars+consume+as+much+electricity+as+Brazilians+>
- Chan, Khee-Hoon** (2022). *An overview of Tencent's burgeoning games dominance*.
<https://www.polygon.com/22949530/tencent-the-worlds-biggest-video-game-company>
- Cbinsights* (2022). "The metaverse could be tech's next trillion-dollar opportunity: These are the companies making it a reality". *Research briefs*.
<https://www.cbinsights.com/research/metaverse-market-map>
- Citi GPS* (2022). *Metaverse and money. Decrypting the future*. Global Perspectives & Solutions.
<https://goo.su/Feh5nlq>
- Credit Suisse* (2022). *Metaverse: A guide to the next-gen Internet*.
<https://www.credit-suisse.com/media/assets/corporate/docs/about-us/media/media-release/2022/03/metaverse-14032022.pdf>

- Deloitte (2022). *On the board's agenda | US What's all the buzz about the metaverse?*
<https://www2.deloitte.com/content/dam/Deloitte/us/Documents/center-for-board-effectiveness/us-cbe-otba-whats-all-the-buzz-about-the-metaverse.pdf>
- Donkin, Chris (2022). "SKT stakes claim for leading role in global metaverse". *Mobile World Live*, November 23.
<https://www.mobileworldlive.com/apps/news-apps/skt-stakes-claim-for-leading-role-in-global-metaverse>
- Dwivedi, Yogesh-Kumar; Hughes, Laurie; Baabdullah, Abdullah M.; Ribeiro-Navarrete, Samuel; Giannakis, Mihalis; Al-Debei, Mutaz M.; Dennehy, Denis; Metri, Bhimaraya; Buhalis, Dimitrios; Cheung, Christy M. K.; Conboy, Kieran et al., (2022). "Metaverse beyond the hype: Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy". *International Journal of Information Management*, v. 66.
<https://doi.org/10.1016/j.ijinfomgt.2022.102542>
- Galibert, Olivier (2003). "Vie et mort du *Deuxième Monde*. Etude critique d'une communauté virtuelle ludique". *Les cahiers du numérique*, v. 4, pp. 195-207.
<https://www.cairn.info/revue-les-cahiers-du-numerique-2003-2-page-195.htm>
- Geffen, Marc (2020). *The minimum viable metaverse*.
<https://vividandvague.substack.com/p/the-minimum-viable-metaverse>
- Gent, Edd (2022). "What can the metaverse learn from Second Life?". *IEEE Spectrum*, v. 59, n. 1, pp. 19-19.
<https://spectrum.ieee.org/metaverse-second-life>
- Howell, James (2022). *Decentraland (MANA) vs Sandbox (SAND). Which metaverse is the best?*
<https://101blockchains.com/decentraland-vs-sandbox>
- Henningson, Joakim (2022). "The history of Minecraft".
<https://www.redbull.com/se-en/history-of-minecraft>
- Hern, Alex (2022). "Meta shares dip is proof metaverse plan never really had legs". *The Guardian*, 27 October.
<https://www.theguardian.com/technology/2022/oct/27/metaverse-shares-dip-is-proof-metaverse-plan-never-really-had-legs-facebook>
- Hussain, Wazarat-Ali (2022). "The metaverse market map – the ecosystem of our emerging virtual worlds", May 16.
<https://metaverseinsider.tech/2022/05/16/the-metaverse-ecosystem-a-market-map-for-our-emerging-virtual-worlds>
- IND Money (2022). "Best metaverse stocks in India: What are their future plans?". November 8.
<https://www.indmoney.com/articles/stocks/best-metaverse-stocks-in-India>
- Interesse, Giulia (2022). *China's debut in the metaverse: trends to watch (updated)*. China briefing.
<https://www.china-briefing.com/news/metaverse-in-china-trends>
- Johar, Amiya (2022). "Tech players form metaverse standards group".
<https://www.mobileworldlive.com/featured-content/top-three/tech-players-form-metaverse-standards-group>
- Kshetri, Nir (2023). "National metaverse strategies". *Computer*, v. 56, n. 2, pp. 137-142.
<https://doi.org/10.1109/MC.2022.3227681>
<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=10042106>
- Lamarche-Toloza, Álvaro (2022). *Les réalités du Métavers. Derrière la ruée vers l'or. Les réalités du Métavers et le rôle de la blockchain dans la virtualisation du monde*.
<https://www.clubopenprospective.org/copie-de-2022-s%C3%A9ance-2>
- Lamy, Corentin (2022). "Le deuxième monde' de Canal+, dès 1997: le pionnier français du métavers". *Le Monde*, 10 juin.
https://www.lemonde.fr/pixels/visuel/2022/06/10/le-deuxieme-monde-de-canal-des-1997-le-pionnier-francais-du-metavers_6129767_4408996.html
- Landymore, Frank (2022). *The guy who invented the World Wide Web thinks we should ignore web3*.
<https://futurism.com/the-byte/guy-invented-world-wide-web-ignore-web3>
- Lange, Jeva (2020). *Travis Scott's 'virtual concert' was the biggest Fortnite event ever*.
<https://theweek.com/speedreads/910804/travis-scotts-virtual-concert-biggest-fortnite-event-ever>
- Ma, Steven (2022). *Metaverse adds new dimensions to games*.
<https://www.tencent.com/en-us/articles/2201300.html>
- MacDonald, Keza (2023). "I've seen the metaverse – and I don't want it". *The Guardian*, January 25.
<https://www.theguardian.com/games/2022/jan/25/ive-seen-the-metaverse-and-i-dont-want-it>
- Majithia, Kavitha (2022). *Meta value takes \$65B hit as Q3 disappoints*.
<https://www.mobileworldlive.com/featured-content/home-banner/meta-value-takes-65b-hit-as-q3-disappoints>

- McKinsey* (2022). *Value creation in the metaverse. The real business of the virtual world.*
<https://www.mckinsey.com/business-functions/growth-marketing-and-sales/our-insights/value-creation-in-the-metaverse>
- McCormick, Packy** (2020). "Tencent's dreams. Part II: Investing in the metaverse". *Not boring newsletter.*
<https://seekingalpha.com/article/4371548-tencents-dreams-part-ii-investing-in-metaverse>
- Messinger, Paul R.; Stroulia, Eleni; Lyons, Kelly** (2008). "A typology of virtual worlds: Historical overview and future directions". *Virtual worlds research: Past, present and future*, v. 1, n. 1.
<https://jvwr-ojs-utexas.tdl.org/jvwr/article/view/291>
- Mileva, Geri** (2022). "The ultimate guide to metaverse virtual real estate".
<https://influencermarketinghub.com/metaverse-virtual-real-estate>
- Mind the bridge* (2022). *Tech scaleup South Korea. Report 2022.*
https://dl.airtable.com/.attachments/1638d910ea6d650fa661e9e03f4d901b/8e60ac0e/2022_MTB_TechScaleupSouthKorea.pdf
- Mohammed, Sid** (2021). *Zuckerverse: why we should vote with our feet and stay away from Facebook.*
<https://blogs.lse.ac.uk/businessreview/2021/12/10/zuckerverse-why-we-should-vote-with-our-feet-and-stay-away-from-facebook>
- Mulligan, Tim** (2022). *Metaverse. Future scenarios for smart TVs.*
<https://www.midiaresearch.com/reports/metaverse-future-scenarios-for-smart-tvs>
- Mystakidis, Stylianos** (2022). "Metaverse". *Encyclopedia 2022*, v. 2, pp. 486-497.
<https://doi.org/10.3390/encyclopedia2010031>
- Narin, Gökçe** (2021). "A content analysis of the metaverse articles". *Journal of metaverse*, v. 1, n. 1, pp. 17-24 .
<https://dergipark.org.tr/en/pub/jmv/issue/67581/1051382>
- Newzoo* (2022). "The metaverse, blockchain gaming, and NFTs: Navigating the Internet's uncharted waters". *Newzoo trend report.*
https://resources.newzoo.com/hubfs/Reports/Metaverse/2022_Newzoo_Metaverse_Blockchain_NFT_Report.pdf
- Parkin, Simon** (2022). "The trouble with Roblox, the video game empire built on child labour".
<https://www.theguardian.com/games/2022/jan/09/the-trouble-with-roblox-the-video-game-empire-built-on-child-labour>
- Partleton, Kayleigh** (2020). "Sulake's networking title Habbo celebrates its 20th anniversary".
<https://www.pocketgamer.biz/news/74772/sulakes-networking-title-habbo-celebrates-its-20th-anniversary>
- Radoff, Jon** (2021). "Market map of the metaverse".
<https://medium.com/building-the-metaverse/market-map-of-the-metaverse-8ae0cde89696>
- Rose, Steve** (2022). "'The metaverse will be our slow death!' Is Facebook losing its \$100bn gamble on virtual reality?". *The Guardian*, December 7.
<https://www.theguardian.com/technology/2022/dec/07/metaverse-slow-death-facebook-losing-100bn-gamble-virtual-reality-mark-zuckerberg>
- Seekins, Russell** (2022). "Meeting of worlds. Impacts of the metaverse". *Intermedia*, v. 50, n. 1, pp. 18-21.
<https://www.iicom.org/intermedia/vol-50-issue-1/meeting-of-worlds-impacts-of-the-metaverse>
- Simon, Jean-Paul** (2019). "5G: Fuel for the media?". *Intermedia*, v. 47, n. 2, pp. 16-19.
<https://www.iicom.org/intermedia/intermedia-jul-2019/5g-fuel-for-the-media>
- Simon, Jean-Paul** (2021). "Free-to-play model". In: Wolf, Mark J. P., *Encyclopaedia of video games: The culture, technology, and art of gaming*, 2nd ed., Westport: Greenwood, pp. 357-359. ISBN: 978 1 440870200
- Tencent* (2022). *2022 Interim report.*
<https://static.www.tencent.com/uploads/2022/09/01/9865e43220e76fe82356b91d3b2a70e6.PDF>
- Van-der-Velde, Issy** (2021), *Report: Ubisoft Quartz NFTs aren't selling despite lots of trades.*
<https://www.thegamer.com/report-ubisoft-quartz-nfts-arent-selling-despite-lots-of-trades>
- Ye, Josh** (2021). *Why Chinese tech giants are embracing the metaverse despite state media warnings.*
<https://www.scmp.com/tech/tech-trends/article/3156562/why-chinese-tech-giants-are-embracing-metaverse-despite-state>
- Zyda, Michael** (2022). "Let's rename everything 'the metaverse!'". *Computer*, v. 55, n. 3, pp. 124-129.
<https://doi.org/10.1109/MC.2021.3130480>

PUBLICACIONES PROFESIONAL DE LA INFORMACIÓN

REVISTA PROFESIONAL DE LA INFORMACIÓN

Open access journal / Revista en acceso abierto

<https://revista.profesionaldelainformacion.com>

ANUARIO THINKEPI

Open access journal / Revista en acceso abierto

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

MANUAL SCIMAGO DE REVISTAS CIENTÍFICAS. CREACIÓN, GESTIÓN Y PUBLICACIÓN

Tomàs Baiget

a) Impreso, rústica: 29€ + 4% IVA + gastos de envío

b) Pdf: gratis / free:

<https://www.profesionaldelainformacion.com/producto/manual-scimago-de-revistas-cientificas-creacion-gestion-y-publicacion>

COLECCIÓN LIBROS EPI SCHOLAR (Editorial UOC)

- 1. Cibermetría. Midiendo el espacio red 20,00€
Enrique Orduña-Malea e Isidro F. Aguillo
- 2. La web social como nuevo medio de comunicación y evaluación científica 21,00€
Amalia Mas-Bleda e Isidro F. Aguillo
- 3. SEO. Introducción a la disciplina del posicionamiento en buscadores 20,00€
Jorge Serrano Cobos
- 4. Métricas de la web social para bibliotecas 26,00€
Nieves González-Fernández-Villavicencio
- 5. Calidad en sitios web. Métodos de análisis general, e-commerce, imágenes, hemerotecas y turismo 19,50€
Rafael Pedraza-Jiménez, Lluís Codina, Javier Guallar
- 6. Gestión de datos de investigación 17,00€
Fabiano Couto Corrêa
- 7. Desafío a la investigación estándar en comunicación. Crítica y alternativas 20,00€
Manuel Goyanes
- 8. Bibliotecas integradas 15,00€
Pablo Parra-Valero
- 9. Gestión de información personal. Integración mediante el correo electrónico 19,00€
Jorge Franganillo
- 10. Datos abiertos enlazados de archivos, bibliotecas y museos en la web 15,00€
Carlos Henrique Marcondes
- 11. Turismo rural online. Páginas web y redes sociales 18,00€
Gersón Beltrán
- 12. Documentación musical 16,00€
Cristina Martí-Martínez
- 13. Formatos digitales. Propiedades técnicas y contextos de uso 20,00€
Jorge Franganillo

<https://www.profesionaldelainformacion.com/producto/libros-epi-scholar>

Profesional de la
información

ANUARIO
Think
EPI

Para publicaciones en papel se añadirán los gastos de envío

**Information & orders /
Información y pedidos**
Isabel Olea

epi.iolea@gmail.com
Tel.: +34-608 491 521

LIBROS PROFESIONAL DE LA INFORMACIÓN

(Editorial UOC)

- 55. Patentes 13,00 €
Gema Velayos-Ortega, Rosana López Carreño
- 54. Gestión de la comunicación científica. 25 ideas para mejorar tu investigación 13,00 €
José-Manuel Noguera-Vivo
- 53. Espacios de biblioteca: presente y futuro 15,00 €
José-Pablo Gallo-León
- 52. Competencias mediáticas e informacionales: juego y aprendo 15,00 €
Felicidad Campal y Ana Ordás-García
- 51. La biblioteca pública tras la pandemia del coronavirus 12,00 €
Evelio Martínez-Cañadas
- 50. Video para redes sociales. Guía de iniciación en la publicación y difusión 14,00 €
Juan-José Boté
- 49. Biblioeconomía de guerrilla 13,00 €
Evelio Martínez-Cañadas
- 48. Partituras electrónicas. Lectoescritura musical en el ámbito digital ... 14,00 €
Javier Merchán-Sánchez-Jara
- 47. Makerspaces y bibliotecas 15,00 €
Julio Alonso-Arévalo
- 46. Bibliotecas ciudadanas. Espacios de desarrollo y participación 12,00 €
João De-Sousa-Guerreiro y María-Ramona Domínguez-Sanjurjo
- 45. Gamificación en bibliotecas. El juego como inspiración 15,00 €
Ana Ordás
- 44. Biblioteca: el servicio público con corazón 14,00 €
Magdalena Ortiz-Macías
- 43. La Hora del Cuento en las bibliotecas. Claves para su organización... 11,50 €
Joan de Boer
- 42. Curación de contenidos para bibliotecas 11,50 €
Evelio Martínez-Cañadas
- 41. Plan de contenidos para medios sociales 11,50 €
Eva Sanagustín
- 40. Legislación de archivos 12,50 €
Vicente Giménez-Chornet
- 39. Gestión de proyectos paso a paso 14,50 €
Candela Ollé y Berta Cerezuela
- 38. Fuentes de información. Guía básica y nueva clasificación 13,00 €
Rosana López-Carreño
- 37. La medición en plataformas sociales 11,00 €
Adrián Segovia
- 36. Big data. Gestión y explotación de grandes volúmenes de datos 15,00 €
Montserrat García-Alsina
- 35. Estudios de usuarios de información 12,00 €
Isabel Villaseñor-Rodríguez
- 34. Autopublicar. Los nuevos circuitos para autores e investigadores.. 12,00 €
Almudena Mangas-Vega
- 33. Lectura digital infantil. Dispositivos, aplicaciones y contenidos.... 13,00 €
Araceli García-Rodríguez y Raquel Gómez Díaz
- 32. Geolocalización online. La importancia del dónde 12,00 €
Gersón Beltrán
- 31. Biblioteca pública. Mientras llega el futuro 14,50 €
Fernando Juárez-Urquijo
- 30. Geobibliotecas 11,50 €
Estefanía Aguilar-Moreno y Carlos Granell-Canut
- 29. Los sexenios de investigación 12,00 €
Álvaro Cabezas-Clavijo y Daniel Torres-Salinas
- 28. Los impresos antiguos 12,50 €
Jon Zabala
- 27. La intranet social 13,50 €
Ana Carrillo Pozas
- 26. Innovación en bibliotecas 13,50 €
Isabel Riaza

<http://www.profesionaldelainformacion.com>

<http://www.thinkepi.net>

Esta revista *Profesional de la información (EPI)* ya no acepta suscripciones. Desde el 1 de enero de 2023 *EPI* es una revista 100% Open Access, es decir, todos los contenidos publicados estarán en acceso abierto.

Si desea recibir el sumario de la revista cuando se publique un nuevo número, por favor regístrese aquí:

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

NORMAS PARA LOS AUTORES

Profesional de la información tiene dos secciones principales:

ARTÍCULOS: Trabajos de investigación y temas analizados en profundidad.

ANÁLISIS: Experiencias, estudios de casos, análisis de productos, etc.

Las contribuciones han de ser originales e inéditas, no pueden haberse publicado previamente en soporte papel o electrónico.

El texto ha de enviarse en Word, rtf u odt. Las tablas deberán ir pegadas en el mismo documento. Todos los materiales gráficos (diagramas, fotografías, capturas de pantalla, etc.) deben pegarse en el Word y además enviarse en ficheros independientes (en formatos xls, jpg, pdf, etc.). Las imágenes jpg deben tener una resolución de al menos 300 pp (unos 200 KB cada una).

El texto debe presentarse completamente plano, sin autoformatos ni automatismos de Word (subsecciones, viñetas, citas enlazadas, pies de página, sangrías, tabulaciones, colores, etc.), pero debe seguir el estilo de EPI en cuanto a **negritas** (nombres de los autores citados), *cursivas* (instituciones, títulos de revista, marcas) y mayúsculas. Los urls deben estar sin hipervínculo.

Las citas bibliográficas en el texto se realizarán de la forma: (**Apellido**, año) o (**ApellidoAutor1**; **ApellidoAutor2**, año).

La redacción debe ser concisa y precisa, evitando la retórica.

Los trabajos deben incluir: a) título, b) resumen de 250-300 palabras, c) 14-15 palabras clave, d) title, e) abstract de 250-300 palabras, f) 14-15 keywords.

Aparte se incluirá el nombre de los autores, su lugar de trabajo y dirección (postal y electrónica), su foto tipo carnet superior a 100 KB en jpg, un currículum de unas 70 palabras, y su orcid.

EVALUACIÓN

Los trabajos son revisados según el sistema tradicional "peer review" en doble ciego por al menos dos expertos en el tema, del Consejo Asesor de la revista y/o externos. La revista se compromete a informar del resultado a los autores.

REFERENCIAS BIBLIOGRÁFICAS

Ordenadas alfabéticamente por autor, se limitarán a las obras citadas en el artículo. No se acepta bibliografía de relleno.

Artículos de una publicación periódica:

Apellido, Nombre; Apellido2, Nombre2 (año). "Título del artículo". *Título de la publicación periódica*, v., n., pp. xx-yy.

Dirección url iniciada en nueva línea sin barra y sin punto finales
<https://doi.org/10.xxxx/doi>

Ponencia presentada en un congreso:

Apellido, Nombre; Apellido2, Nombre2 (año). "Título de ponencia". En: *Nombre del congreso*, pp. xx-yy.

Dirección url iniciada en nueva línea sin barra y sin punto finales
<https://doi.org/10.xxxx/doi>

Monografías e informes:

Apellido, Nombre; Apellido2, Nombre2 (año). *Título del trabajo*. Lugar de publicación: editor, ISBN: espacios, sin guiones y sin punto final

Dirección url iniciada en nueva línea sin barra y sin punto finales
<https://doi.org/10.xxxx/doi>

Capítulo de una monografía:

Apellido, Nombre; Apellido2, Nombre2 (año). "Título del capítulo". En: **Apellido, Nombre; Apellido2, Nombre2**. *Título de la monografía*. Lugar de publicación: editor, pp. xx-yy. ISBN: espacios, sin guiones ni punto final

Dirección url iniciada en nueva línea sin barra y sin punto finales

Recurso online:

Apellido, Nombre; Apellido2, Nombre2 (año). *Título del recurso (sin fecha de la consulta)*.

Dirección url iniciada en nueva línea sin barra y sin punto finales

Todas las contribuciones se tienen que enviar a la plataforma OJS de EPI

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

Previamente los autores deben registrarse en:

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