

Galileo, una plataforma de datos para leer noticias en redes sociales

Galileo, a data platform for viewing news on social networks

Luis Cárcamo-Ulloa; Claudia Mellado; Carlos Blaña-Romero; Diego Sáez-Trumper

Note: This article can be read in its English original version on:
<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/86941>

Cómo citar este artículo.

Este artículo es una traducción. Por favor cite el original inglés:

Cárcamo-Ulloa, Luis; Mellado, Claudia; Blaña-Romero, Carlos; Sáez-Trumper, Diego (2022). "Galileo, a data platform for viewing news on social networks". *Profesional de la información*, v. 31, n. 5, e310512.

<https://doi.org/10.3145/epi.2022.sep.12>

Artículo recibido el 14-03-2022
Aceptación definitiva: 04-07-2022



Luis Cárcamo-Ulloa ✉
<https://orcid.org/0000-0003-0633-9606>

Universidad Austral de Chile
Instituto de Comunicación Social
Campus Isla Teja
Independencia 641, Valdivia, Chile
lcarcamo@uach.cl



Claudia Mellado
<https://orcid.org/0000-0002-9281-1526>

Pontificia Universidad Católica de Valparaíso
Escuela de Periodismo
Avenida Universidad 330
Campus Curauma, Valparaíso, Chile
claudia.mellado@pucv.cl



Carlos Blaña-Romero
<https://orcid.org/0000-0001-6987-5567>

Universidad Austral de Chile
Instituto de Informática
Campus Miraflores
Independencia 641, Valdivia, Chile
carlos.blana@uach.cl



Diego Sáez-Trumper
<https://orcid.org/0000-0002-7679-5423>

Fundación Wikimedia
Universitat Pompeu Fabra
Barcelona, España
dsaez-trumper@acm.org

Resumen

El presente artículo tiene por objetivo presentar *Galileo*, una plataforma de extracción y organización de datos de medios informativos en redes sociales. *Galileo* integra publicaciones realizadas en las principales redes sociales utilizadas en el ecosistema informativo, a saber, *Facebook*, *Twitter* e *Instagram*. Actualmente el sistema considera 97 medios informativos, de nueve países: Alemania, Brasil, Chile, Corea del Sur, España, Estados Unidos, Japón, México y Reino Unido. *Galileo* utiliza la API de *Twitter* y el servicio *CrowdTangle* para descargar las publicaciones de *Facebook* e *Instagram*. Estos datos son almacenados en una base de datos local, y son accesibles mediante una interfaz de usuario amigable que permite analizar diferentes características de los posteos tales como su contenido textual, popularidad del medio, y su dimensión temporal. *Galileo* es una herramienta para investigadores interesados en comprender los ciclos informativos y analizar los contenidos informativos en redes sociales.

Palabras clave

Noticias; Ciencia de datos; Visualización; Datos textuales; Periodismo; Medios sociales; Redes sociales; Plataformas; *Twitter*; *Facebook*; *Instagram*; *Galileo*.

Abstract

This article aims to introduce *Galileo*, a platform for extracting and organizing news media data on social networks. *Galileo* integrates publications made on the main social networks used in the information ecosystem, namely *Facebook*, *Twitter*, and *Instagram*. Currently, the system includes 97 media outlets from nine countries: Brazil, Chile, Germany, Japan, Mexico, South Korea, Spain, United Kingdom, and United States. *Galileo* uses a *Twitter* API and the service *CrowdTangle* to download *Facebook* and *Instagram* posts. This data is stored in a local database and can be accessed through

a user-friendly interface, which allows for the analysis of different characteristics of the posts, such as their text, source popularity, and temporal dimension. *Galileo* is a tool for researchers interested in understanding news cycles and analyzing news content on social networks.

Keywords

News; Visualisation; Data science; Textual data; Journalism; Social media; Social networks; Platforms; *Twitter*; *Facebook*; *Instagram*; *Galileo*.

Financiación

Los investigadores Luis Cárcamo-Ulloa y Claudia Mellado han sido beneficiarios de la *Agencia Nacional de Investigación y Desarrollo (ANID-Chile)*. En particular de las iniciativas *ANID-Covid 0172* y *ANID-Plu 210013*.

1. Introducción

Watts (2016; 2017) destaca la importancia de las Ciencias Sociales Computacionales (CSC) para investigar fenómenos complejos, en base a vastos volúmenes de datos. Los medios de comunicación generan a diario grandes cantidades de datos que reflejan el quehacer político, económico y cultural de las sociedades en que están inmersos. Por lo mismo, analizar grandes volúmenes de datos textuales que a diario aportan los medios es importante para comprender los sesgos y los encuadres parciales que pueden dañar las democracias (**Watts; Rothschild; Mobius**, 2021).

Dentro de los estudios de periodismo, la mayor parte de los análisis textuales que se realizan para investigación académica cuantitativa, se obtienen a través de análisis de contenido cuantitativos manuales. Pese a la revolución tecnológica y el nuevo ecosistema digital, la codificación automática de textos, sobre todo para medidas complejas y contenido latente, es un área relativamente reciente (**Hamborg; Donnay; Gipp**, 2019). Entusiastas y críticos han elaborado diferentes argumentos para subrayar las ventajas y desventajas del análisis automático de textos. Mientras algunos han argumentado que los métodos automáticos son muy prometedores, rápidos y objetivos (**Pereira et al.**, 2015); otros subrayan que por ejemplo los análisis de encuadres noticiosos requieren de un contexto semántico y pistas intertextuales exteriores a la noticia (**Baden**, 2018). Estos últimos aspectos parecen ser aún muy poco manejables para la inteligencia artificial.

Sin embargo, en términos globales, varias investigaciones concuerdan en que la enorme cantidad de información existente necesita de herramientas que hagan posible su obtención, procesamiento y análisis de forma oportuna (**Grimmer; Stewart**, 2013; **Lewis; Zamith; Hermida**, 2013; **Matthes; Kohring**, 2008; **Trilling; Jonkman**, 2018; **Van-Atteveldt; Peng**, 2018).

Los desarrollos para observar las interacciones en los ecosistemas informativos son variados y muy eficientemente usados, por ejemplo, para efectos del marketing digital. Sin embargo, para los profesionales y los analistas de la comunicación es difícil aún acceder de forma compilada, ordenada y eficiente a lo que los medios dicen sobre un tema en las redes sociales y/o en la web (**Cárcamo-Ulloa et al.**, 2017). Normalmente, para la obtención de datos se requiere de desarrollos específicos o *scripts* para extraer los datos dependiendo de los objetivos de análisis que se persigan (**Zhang; Boons; Batista-Navarro**, 2019)

Galileo es una plataforma elaborada en el marco del proyecto *Anid- Covid 0172* "Análisis y monitoreo automático del rol del periodismo y los medios en sus diferentes plataformas durante las fases de la crisis sanitaria provocada por la Covid-19 en nueve países de América, Europa y Asia". Esta iniciativa fue financiada por la *Agencia Nacional de Investigación y Desarrollo del Gobierno de Chile*. Cabe señalar que en el proyecto *Anid-Covid 0172* se integran investigadores de comunicación y periodismo de nueve países: Alemania, Brasil, Chile, Corea del Sur, España, Estados Unidos, Japón, México y Reino Unido.

El proyecto *Anid-Covid* tiene como objetivo principal diseñar e implementar estrategias de análisis y monitorización de los ciclos de atención mediática otorgada por plataformas tradicionales y redes sociales durante la cobertura de la Covid-19 y su relación con el rol que el periodismo cumple en la comunicación de crisis sanitarias. En este contexto, *Galileo* no solo se reduce a informaciones propias de la crisis sanitaria, sino a toda la producción informativa que un conjunto de medios publicó en redes sociales desde el inicio de la pandemia (enero de 2020) en adelante. Esta plataforma surge de la necesidad de crear una herramienta que permita a investigadores de comunicación acceder a los contenidos que los medios vierten en sus cuentas de redes sociales, especialmente considerando que la acción de plataformas como *Twitter* o *Facebook* concentran gran parte de las lectorías de noticias (**Salazar**, 2019; **Newman et al.**, 2021).

<http://www.galileo-jrp.org>

Las redes sociales de los medios informativos, como yacimiento de extracción de datos, presentan variaciones respecto de sus homólogos en sitios web. En este sentido, es importante señalar que:

- La cantidad de los datos textuales recogidos puede cambiar dependiendo de la cultura organizacional de los equipos de prensa o de las definiciones de negocio editorial¹ ya que los medios pueden no verter todos sus contenidos periodísticos a redes sociales, pero podemos suponer que seguramente publicarán lo más relevante.

- Los textos noticiosos en redes sociales son obviamente más sintéticos que las noticias de cuerpo extendido que ofrecen los medios en sus versiones web o tradicionales. Los posteos informativos en redes sociales se parecen más a lo que se conoce como “bajadas” o “leads” periodísticos. Sin embargo igualmente deben capturar al menos las preguntas básicas de “qué pasó”, “dónde” y “cuándo” ocurrieron los hechos narrados. Estas estructuras tradicionales resultan irrenunciables a la hora de componer una noticia y se responden al inicio de cualquier texto informativo.

1.1. Interés en los datos textuales

En las redes sociales circula una cantidad enorme de información y lo hace en fórmulas compactas que son consumidas por los usuarios. El post es la maximización de la síntesis informativa que desde siempre ha perseguido la eficacia del periodismo informativo y, a la vez, es también un nuevo género discursivo que los medios ensayan en sus redes sociales (Cárdenas-Neira, 2016; Raimondo-Anselmino; Sambrana; Cardoso, 2017).

Analizar grandes volúmenes de datos textuales que a diario aportan los medios es importante para comprender los sesgos y los encuadres parciales que pueden dañar las democracias

Ojo y Heravi, (2018) proponen analizar patrones y tipologías de narraciones informativas exitosas que pudieran orientar el futuro del periodismo de datos. Newman, Dutton y Blank (2019) llegaron a indicar que la ampliación de los escenarios informativos hacia las redes sociales permitiría -a medida que aumenta la información en línea y los individuos se conectan habitualmente a internet- la emergencia de un “quinto poder” basado en las actividades de los individuos en red que obtienen y distribuyen información. Por otro lado, en congruencia con la idea de filtro de burbuja propuesta por Parisier (2011), también se ha comprendido que por ejemplo el algoritmo de Facebook es menos proclive a suministrar a los individuos publicaciones de los medios de comunicación contrarios a las actitudes de los usuarios (Levy, 2021) y que por tanto los algoritmos de las redes sociales, al limitar la exposición a las noticias contrarias, aumentan la polarización. Todo esto nos invita a pensar que la investigación de las noticias que circulan en redes sociales resultan ser un fenómeno de particular interés.

Galileo centra su interés en los datos textuales, más que en las interacciones de los usuarios. Es una herramienta que permite gestionar volúmenes considerables de corpus textuales para realizar análisis de contenidos manuales, así como también aplicar técnicas de tratamiento automático de lenguaje para *topic modeling*, *word-embedding* y análisis de entidades (fuentes y actores).

La plataforma Galileo busca facilitar la exploración de datos poniendo a prueba modelos de análisis de datos textuales para la comprensión de los ecosistemas informativos en redes sociales. Desde la ciencia de datos se pueden aportar herramientas que permitan comprender al periodismo y los medios (Hamborg; Donnay; Gipp, 2019). El desafío central para la Ciencia de Datos aplicada al análisis de informaciones de prensa es entonces proyectar técnicas que contribuyan a la evaluación de la calidad de las noticias en un mundo que se caracteriza por una creciente entropía informacional (Cardon, 2018).

Existen experiencias relevantes en análisis informatizados de grandes volúmenes de datos en la Web y en los medios de comunicación. Así, por ejemplo, las técnicas informatizadas han permitido dar cuenta de una serie de sesgos en la Web, en general (Baeza-Yates, 2018), y en torno a personajes políticos aparecidos en la prensa, en particular (Sáez-Trumper, Castillo; Lalmas, 2013). Recientemente, incluso, se avanza hacia la identificación de patrones sutiles que determinan sesgos en los encuadres periodísticos (Morstatter *et al.*, 2018). Como exponen Schmitz-Weiss *et al.*, (2017), la tecnología ha eliminado las barreras de trabajo del periodismo pudiendo estudiar y analizar información que las propias noticias generan.

Los avances en lingüística computacional y tratamiento automático del lenguaje pueden aplicarse a datos textuales de noticias. Hoy en día se intenta avanzar en detección de *fake news* (Zhou *et al.*, 2019) y en calidad de la información (Rommannou *et al.*, 2020). Por lo mismo, con los datos textuales extraídos y organizados por la plataforma Galileo se puede proyectar la realización tanto de análisis de contenidos tradicionales como de procesamientos automatizados.

En la bibliografía encontramos trabajos a partir de identificación de entidades en corpus de prensa de 37 medios de prensa chilenos (Vernier; Cárcamo-Ulloa; Scheihing-García, 2017). Por su parte, Jiang *et al.* (2017) comparan las actitudes de los medios británicos frente al cambio climático durante un período de 10 años. El corpus de investigación de este último trabajo se compuso de 11.720 artículos periodísticos recopilados entre 2007 y 2016 desde cuatro periódicos del Reino Unido (*The Guardian*, *The Times*, *The Telegraph* y *The Independent*). Este trabajo se ocupa de análisis de sentimientos y *Latent Dirichlet Allocation* (LDA) para identificar tópicos (Blei; Ng; Jordan, 2003). Particularmente, el modelamiento de tópicos es una estrategia bastante extendida para grandes volúmenes de datos textuales de prensa (Li *et al.*, 2020).

1.2. Algunas herramientas para acceder a noticias del mundo

Existen buenas herramientas para acceder a corpus de noticias y estudiar los contenidos de los medios de comunicación. A modo de ejemplo, destacaremos dos plataformas.

- *MediaCloud* es una plataforma para analizar grandes medios de comunicación y tiene tres softwares fundamentales: *Explorer*, *Topic Mapper* y *Source Manager*. Incluye una base de datos de fuentes de más de 100 países. <https://mediacloud.org>



Figura 1. Interfaz de visualización de Newsmap

- Newsmap es una aplicación muy eficaz para leer las noticias del momento en cada país. Refleja visualmente el paisaje en constante cambio de Google News. La información se visualiza en forma de treemaps que componen proporcionalmente los tópicos más relevantes de cada país.
<https://newsmap-js.herokuapp.com>

Por otro lado, también existen aplicaciones de pago que desde el marketing digital se han creado para analizar y gestionar el trabajo de marcas en redes sociales y que también pueden usarse para seguir medios de comunicación social. Quizás las más conocidas sean Hoosuite, Buffer y Karma Fanpage.

Desde el punto de vista de la investigación en el análisis de contenidos periodísticos, podemos plantearnos algunos elementos de interés valorables para equipos académicos:

- Foco: la aplicación permite acceder a contenidos informativos organizados en el tiempo.
- Cobertura de medios y países: Rastrea medios relevantes y diversos de países para propiciar estudios comparativos.
- Accesibilidad: libre o registrada o de pago.
- Ductilidad de las consultas: permite al investigador consultar palabras clave y/o expresiones exactas, configurar -únicas o agregadas- con inclusores y exclusores (booleanos, por ejemplo).
- Ductilidad de las visualizaciones: permite al investigador configurar gráficos de composición, relación, comparación y distribución.
- Descarga de corpus maleable: tiene la posibilidad de descargar un archivo con los contenidos textuales y links referentes a las informaciones.

Tabla 1. Algunas características comparables

Factor	Galileo	MediaCloud	NewsMap
Foco	Redes sociales de medios/ descarga de datos de consulta	Medios en la Web/ descarga de datos de consulta	Google News (visualizaciones por país)
Cobertura	9 países	+100 países	+100 países
Accesibilidad	Previa solicitud	Registro	Abierta
Ductilidad de las consultas	Construcción agregada basada en booleanos	Construcción agregada basada en booleanos	Predefinidas por la actualidad de cada país
Ductilidad para las visualizaciones	Múltiples visualizaciones configurables por el usuario	Las definidas por la plataforma	Treemaps
Descarga de corpus	Descarga en csv o json	Descarga en csv	No tiene
Costos	Sin costo	Sin costo	Sin costo

2. Galileo: una opción centrada en el periodismo de redes sociales

Hasta la fecha, *Galileo* permite integrar datos minados a través de API's de *Facebook*, *Instagram* y *Twitter* que emiten 97 medios de comunicación provenientes de nueve países (ver listado Anexo).

<http://www.galileo-jrp.org>

Podríamos pensar que es una infraestructura informática similar a las usadas por servicios de marketing digital que buscan seguir cuentas de marcas en redes sociales, captar sus contenidos y sus interacciones. Sin embargo, su sentido principal es colaborar en la búsqueda de patrones en datos textuales de medios informativos que circulan en redes sociales. Así es como *Galileo* se concibe como una aplicación para extraer y filtrar noticias desde redes sociales.

En términos de gestión de información, *Galileo* puede definirse como una plataforma de gestión de datos unificada (UDMP, *Unified data management platform*). Es decir, un sistema informático centralizado que trabaja con grandes cantidades de datos estructurados (frecuencias y valores) y no estructurados (textos) y de fuentes diversas (*Facebook*, *Instagram* y *Twitter*) para recopilarlos, integrarlos, gestionarlos y visualizarlos.

2.1. Arquitectura informática

Uno de los problemas que enfrentan las plataformas que recolectan información variada (corpus textuales, valores e imágenes, entre otros) es la necesaria flexibilidad para sus posteriores usos. Si la información es almacenada en casillas estructuradas como bases de datos SQL, el usuario final tendrá las mismas restricciones que se planteó el programador al idear los encasillamientos. Particularmente, al trabajar con datos textuales, una excesiva estructuración reducirá las posibilidades de consultas de los investigadores. *Galileo* propone usar una estructuración flexible NoSQL que se sirve de *Elastic Stack*, una suite de productos de software de código abierto (licencia *Apache 2.0*) que permiten ingestar de manera segura datos de cualquier fuente y formato para luego trabajar productivamente ejecutando búsquedas, análisis y visualizaciones en tiempo real u offline según sea el caso.

<https://www.elastic.co>

El *crawler* (rastreador) de *Galileo* se conecta a la API de *CrowdTangle*² para extraer datos de *Facebook* e *Instagram* y en el caso de las publicaciones de *Twitter*, directamente a la API de esta red social.

Un *stack* es una combinación de herramientas, aplicaciones y servicios que se utilizan para crear una aplicación web o móvil. *Elastic Stack* comprende herramientas tales como *Elasticsearch*, *Logstash* y *Kibana* (*ELK Stack*) que ofrecen funciones especialmente flexibles para trabajar con datos poco o nada estructurados como pueden ser los textos de prensa y los posteos en redes sociales. La maleabilidad de *Elastic Stack* radica en contar con una base de datos *NoSQL* (no estructurados previamente), que construye índices a partir de un conjunto de datos. Un índice contiene mapeos a diversos tipos de datos estableciéndose como un espacio que organiza la información. Así, *Galileo* es una suerte de cajón de sastre que guarda los distintos tipos de datos sin una organización que los enclaustre, justamente para que las inquietudes de los usuarios/investigadores o consultas resulten eficientes, productivas y casi sin restricciones.

Después de recoger los datos, el mayor mérito de *Galileo* es ejecutar consultas (*queries*) y aportar los datos -que a priori no están estructurados- de una manera gestionable para realizar análisis de contenidos o discursivos. Durante estas operaciones, la suite *Elastic Stack* ofrece técnicamente tres herramientas muy eficientes:

- *Elasticsearch*, o componente principal del *Elastic Stack*, es un motor de analítica y análisis distribuido y abierto para todos los tipos de datos, incluidos textuales, numéricos, geoespaciales, estructurados y no estructurados, conocido por sus API REST simples, naturaleza distribuida, velocidad y escalabilidad.
- *Logstash* es un *pipeline* de procesamiento de datos gratuito y abierto del lado del servidor que permite la ingesta de datos de una multitud de fuentes, los transforma y luego los almacena e indexa en *Elasticsearch*.
- *Kibana* es una aplicación de *front-end* o interfaz gratuita y abierta que se encuentra sobre el *Elastic Stack* y proporciona capacidades de búsqueda sobre los datos indexados en *Elasticsearch*, opciones de visualización de la información y creación de paneles con estas visualizaciones.

Finalmente, y no menos importante, es la posibilidad de exportar los corpus de datos construido en base a las consultas. *Elastic ELK* permite exportar los datos en archivos .csv que luego pueden ser fácilmente leídos como hojas de cálculo u

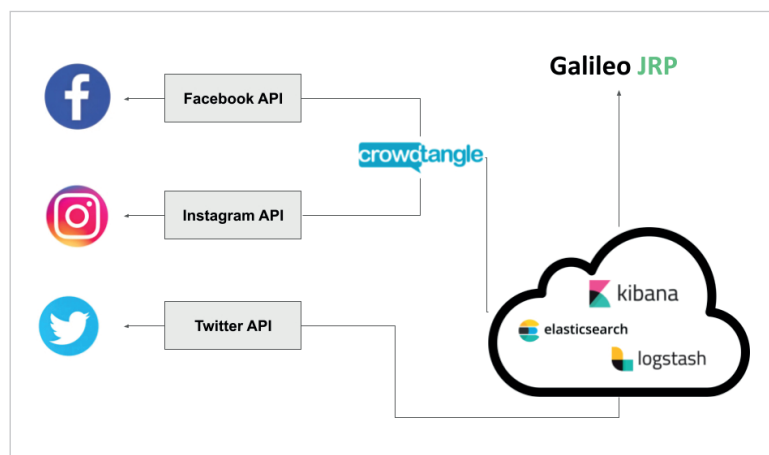


Figura 2. Arquitectura Informática de *Galileo*

otras aplicaciones que permiten elaborar análisis automatizados, tales como *Jupyter*, *Rproject* o *SpaCy* (Guo et al., 2022). El uso de esta combinación de componentes (ELK) como suite de herramienta ha encontrado múltiples adaptaciones. Así, encontramos aplicaciones para el control de aforos en recintos masivos (Cecchet et al., 2020); optimización de procesos en computación de alto rendimiento (Underwod, 2017); y organización de datos textuales de medios (Cárcamo-Ulloa et al., 2017; Vernier; Cárcamo-Ulloa; Scheihing-García, 2017).

2.2. Posibilidades para la investigación en periodismo

Galileo ofrece y pone a disposición de comunidades de investigadores más de 10 millones de posteos informativos. En este contexto, ofrece oportunidades para hacer investigación con grandes volúmenes de datos textuales de prensa, identificar patrones y comparar culturas periodísticas (Mellado et al., 2021b).

Los datos minados y disponibles hasta el momento corresponden a los posteos vertidos a *Facebook* (3.794.460), *Twitter* (6.474.089) e *Instagram* (369.960) entre el 1 de enero de 2020 y el 31 diciembre de 2021. En paralelo, el *crawler* detrás de *Galileo* diariamente sigue excavando datos desde las cuentas de redes sociales de los medios que forman parte del proyecto (ver listado Anexo).

Tabla 2. Medios seguidos por cada país

País	N° de medios	Lengua	Origen medial				Promedio diario de posts		
			Radio	TV	Prensa	Online	Twitter	Facebook	Instagram
Chile	10	Español	2	4	2	2	1.310,67	985,77	133,38
España	16	Español	4	4	4	4	1.838,53	1.014,89	55,66
México	12	Español	3	2	5	2	1.369,27	842,13	38,45
Estados Unidos	11	Inglés	2	4	3	2	863,78	600,73	93,79
Reino Unido	15	Inglés	4	4	4	3	1.417,09	698,25	42,36
Brasil	8	Portugués	2	2	2	2	765,47	397,99	96,85
Corea del Sur	8	Coreano	2	2	2	2	333,71	223,61	5,79
Japón	9	Japonés	2	3	2	2	591,39	172,93	5,80
Alemania	8	Alemán	2	2	3	1	378,70	261,59	34,69

Como podemos apreciar en la tabla 2, *Twitter* es la red social preferente para hacer circular noticias a diario por parte de los principales medios de los nueve países que participan en el proyecto. Sin embargo, la cantidad de posts ejecutados en *Facebook* es también muy abundante principalmente en países de habla hispana. *Instagram* ocupa un tercer lugar muy a la distancia de las otras dos redes sociales. Así mismo pueden observarse variaciones en el uso de redes sociales entre países, lo que refuerza la idea de que las culturas periodísticas ejecutan en forma diferenciada tal o cual red social (Mellado; Hermida, 2021). Por ejemplo, en España el uso más intensivo es en *Facebook*, mientras que en los otros ocho países el uso más frecuente sigue siendo en *Twitter*, que tradicionalmente se ha caracterizado por referenciar a la élite política opinante en cada país. En estudios de comunicación se han privilegiado largamente los estudios con *Twitter* (Arcila-Calderón; Barredo-Ibáñez; Castro, 2017; Hermida, 2010). Esto seguramente se debe, además del uso tradicional de esta red para asuntos de interés informativo, a aplicaciones o plugins que transparentaron el acceso a la API de *Twitter* y la descarga de conjuntos de textos asociados a los hashtags populares.

Tabla 3. Registros disponibles por país y red social

País	Twitter		Facebook		Instagram		Total	%
	Posts	%	Posts	%	Posts	%		
Chile	956.786	15%	719.609	19%	97.371	27%	1.773.766	17%
España	1.342.124	21%	740.869	20%	40.633	11%	2.123.626	20%
México	999.570	15%	614.753	16%	28.072	8%	1.642.395	15%
Estados Unidos	630.556	10%	438.535	12%	68.470	19%	1.137.561	11%
Reino Unido	1.034.479	16%	509.722	13%	30.925	9%	1.575.126	15%
Brasil	558.794	9%	290.533	8%	70.701	19%	920.028	9%
Corea del Sur	243.609	4%	163.235	4%	4.228	1%	411.072	4%
Japón	431.717	7%	126.242	3%	4.235	1%	562.194	5%
Alemania	276.454	4%	190.962	5%	25.325	7%	492.741	5%
Total	6.474.089	100%	3.794.460	100%	369.960	100%	10.638.509	100%

En términos generales Galileo contiene 10.638.509 (tabla 3) posts de medios de comunicación, siendo el ecosistema español el que aporta el 20% del conjunto de datos, seguido de Chile con 17, México y Reino Unido con un 15% (cada uno), Estados Unidos con 11% y Brasil con un 9%.

2.3. Funcionamiento

Galileo permite elaborar consultas a partir de palabras clave (figura 3). Se puede definir la consulta [1] por a) un conjunto de palabras clave; b) por “expresiones exactas”; o c) combinando con operadores booleanos [AND, OR, NOT]

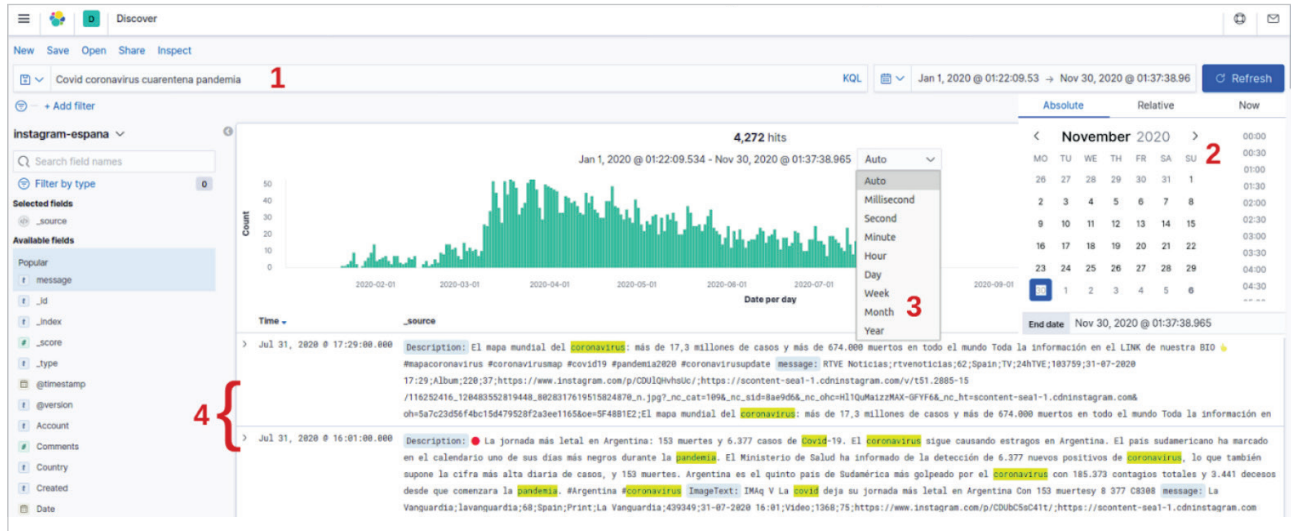


Figura 3. Búsqueda y filtrado de datos en base a palabras clave

La plataforma también permite determinar la fecha de entrada y salida de datos para la consulta [2] y permite elegir la periodicidad de las barras del histograma para visualizar la distribución en el tiempo de los datos consultados en forma diaria, semanal o mensual [3].

Los datos que resultan de la consulta gestionada por el motor *Elastic Search* pueden ser organizados y descargados para continuar nuevos análisis de contenidos. Para ello es necesario organizar los datos, que aún mantienen la organización en Json [4], antes de ejecutar la descarga.

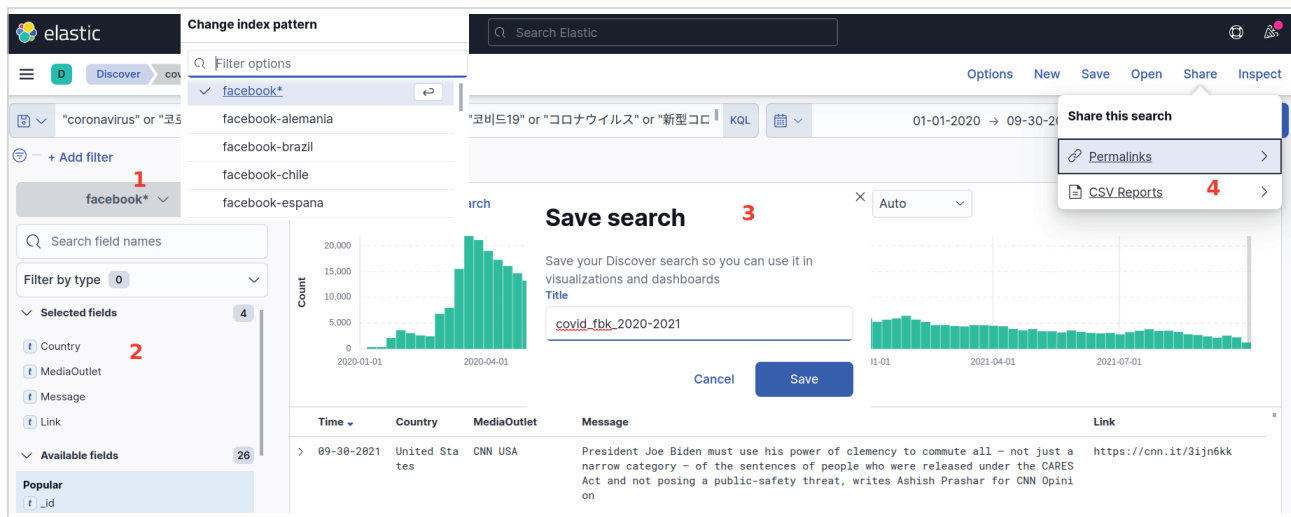


Figura 4. Ordenar, guardar y descargar los datos en Galileo

A la hora de ordenar, guardar y descargar los datos (figura 4), el usuario tiene varias opciones y debe elegir seleccionar un índice (*Facebook*, *Twitter* o *Instagram*) o un sub-índice (cada red social y cada país) [1], seleccionar los campos que se desea descargar (fecha, país, medio, mensaje y url) [2].

Los campos resultan variables dependiendo de cada red social, pero en cada caso se pueden definir los principales campos que componen el posteo: nombre del medio, textos y url para ir al mensaje en su origen.

Tabla 4. Campos del posteo

Tipo de post	Campos descargables
Facebook	Mensaje; texto del titular (link) que conduce a la noticia; nombre del medio; fecha; país; url al post original; seguidores en el momento de la publicación.
Instagram	Descripción (mensaje); texto de la imagen; nombre del medio; fecha; país; url al post original; seguidores en el momento de la publicación.
Twitter	Mensaje; nombre del medio; fecha, país, url al post original, link que conduce a la noticia; seguidores en el momento de la publicación.

Una vez ordenada la salida de datos, tal y como la requerimos, se debe guardar los resultados de la consulta [3] y descargar el corpus en formato csv [4], en el que cada columna se separa por punto y coma (;).

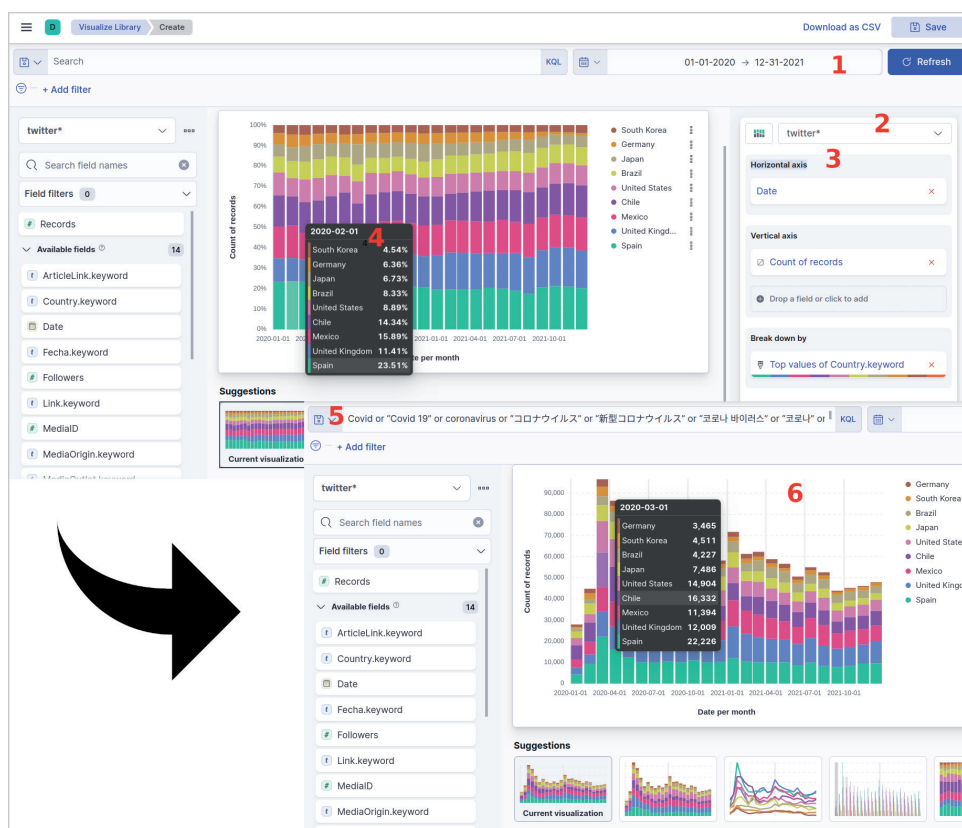


Figura 5. Visualizar datos con Kibana

Aunque en las figuras anteriores ya se aprecia la visualización por defecto de histogramas, la figura 5, muestra cómo el constructor de visualizaciones Kibana permite configurar fácilmente distintos tipos de gráficos. Tras definir el período de consulta [1], el panel de la derecha [2] permite elegir la red social sobre la que se quiere ejecutar la visualización y [3] definir los ejes X e Y del gráfico y sus variables. Al pasar el cursor sobre el gráfico [4] se despliega información flotante que permite comprender con detalle los valores en este caso para cada país. Si se aplica una consulta multilingüe [5] el gráfico se redibuja [6] con solo clicar “refresh”.

3. Potencialidades para la exploración de tópicos con Galileo

Como una forma de mostrar las potencialidades de la plataforma Galileo, en esta sección describiremos una exploración hecha sobre la cantidad de apariciones en prensa de un conjunto de palabras clave sobre “cambio climático” en los últimos 2 años³.

3.1. El caso cambio climático

Se consultó en Galileo la presencia de expresiones exactas vinculadas al cambio climático en cinco lenguas:

- Español: “cambio climático” or “calentamiento global”.
- Inglés: “climate change” or “global warming”.
- Portugués: “aquecimento global” or “alterações climáticas”.
- Japonés: “地球温暖化” or “気候変動”.
- Koreano: “지구 온난화” or “기후 변화”.
- Alemán: “globale Erwärmung” or “Klimawandel”.

Tabla 5. Extracción de datos sobre cambio climático

Países	Twitter	Facebook	Instagram	Total
Reino Unido	5.145	2.035	596	7.776
Estados Unidos	3.781	2.834	663	7.278
Chile	1.767	2.275	645	4.687
España	2.252	1.815	311	4.378
México	1.523	1.536	127	3.186
Alemania	961	870	616	2.447
Japón	770	308	51	1.129
Brasil	210	243	96	549
Corea del Sur	12	13	5	30
Total	16.421	11.929	3.110	31.460

El resultado permite extraer más de 31.460 publicaciones para analizar. Los países anglosajones son los que mayormente se refieren al cambio climático, seguidos de Chile, España y México. Para Alemania, Japón y Corea del Sur las menciones descienden bastante y es seguro que se pueden encontrar mejores resultados afinando consultas país por país incorporando palabras clave alternativas o variaciones lingüísticas que puedan sugerir especialistas locales.

4. Conclusiones

La enorme cantidad de datos textuales que circulan a diario en redes sociales es una gran oportunidad para hacer investigación en comunicación y periodismo (Arcila-Calderón; Barredo-Ibáñez; Castro, 2017). Los medios de comunicación han trasladado sus contenidos a estas plataformas para sostener o mejorar la circulación de sus noticias abriendo nuevas oportunidades de investigación para comprender los ecosistemas informativos. Este tipo de investigación requiere de metodologías y herramientas adecuadas para obtener y tratar dicha información (Hamborg; Donnay; Gipp, 2019). Galileo facilita la obtención de *datasets* textuales de medios informativos que circulan en redes sociales. Dichos corpus pueden ser analizados en base a procesos de tratamiento automático del lenguaje tales como: tokenización y lematización, entre otros. Los dataset resultan compatibles para entornos tales como *Jupyter*, *Rproject* y *SpaCy*, entre otros.

La prueba experimental de la plataforma resultó estable y desde enero de 2020 a diciembre de 2021 no reporta dificultades de acceso a las fuentes de información (redes sociales de los medios) ni en el apoyo a investigaciones que hacen ciencia de datos de noticias tales como los trabajos de Mellado *et al.* (2021a; 2021b).

Los investigadores interesados en usar Galileo pueden seguir los ciclos de noticias en nueve países y examinar tópicos relevantes y la presencia de entidades específicas. Evidentemente se sigue a un número restringido de medios, pero hemos cuidado atender a un conjunto relevante para cada país. Galileo puede ser útil también para ONGs o servicios públicos que hacen seguimientos de cómo se presentan problemáticas o causas en los medios de comunicación de masas.

Por su parte, a los creadores de contenidos, Galileo les puede servir como herramienta para comprender qué narrativas y/o qué formatos narrativos dominan frente a un tema, permite hacer algo así como ingeniería inversa en el proceso de desarrollo de informaciones o como proponen Ojo y Heravi, (2018) analizar patrones y tipos de narraciones informativas exitosas.

La arquitectura de Galileo es simple y robusta. Se puede conectar directamente a la API de Twitter e indirectamente a Facebook e Instagram a través de la API de CrowTangle. Galileo se propone como un *know how* que permite instalar y configurar una plataforma de seguimiento de cuentas de redes sociales abiertas públicas (no sigue perfiles personales) y un recurso para la obtención de corpus de redes sociales de medios de comunicación para estudiar fenómenos relevantes como la polarización (Levy, 2021), las *fake news* (Cárcamo-Ulloa *et al.*, 2021) y los sesgos de confirmación (Ling, 2020).

5. Proyecciones y limitaciones

A modo de proyecciones el equipo de investigadores ofrece un usuario “*invitado_epi*” y una contraseña “*pass_epi*” que permite abiertamente probar Galileo⁴, ejecutando consultas y obteniendo corpus que atiendan a sus intereses. Así mismo se abre una invitación a solicitar un acceso académico y conformar la comunidad de usuarios de Galileo.

La configuración de una comunidad de usuarios, permitirá—en adelante—extender las capacidades de Galileo para incorporar nuevos ecosistemas informativos de redes sociales y facilitar la extracción de datos para investigación. Es decir, puede incorporar nuevos países (más allá de los nueve países que actualmente recopila) o aumentar la cantidad de medios que se siguen en cada país, generando potenciales bases de datos para múltiples temas de investigación en Comunicación y Periodismo.

Galileo ofrece y pone a disposición de comunidades de investigadores más de 10 millones de posteos informativos. Ofrece oportunidades para hacer investigación con grandes volúmenes de datos textuales de prensa, identificar patrones y comparar culturas periodísticas

Entre sus limitaciones podemos sincerar que la herramienta requiere desarrollo aún para ofrecer:

- una solución automática para solicitar la incorporación de nuevos medios de comunicación y otros países;
- *Galileo* no hace *scraping* o raspado de contenidos sobre los websites originales de los medios de comunicación y con ello limita los corpus de trabajo a los posteos como unidad de análisis;
- aunque *Galileo* es capaz de contener los textos transcritos de las publicaciones en redes sociales, no almacena directamente las imágenes asociadas. Todas estas limitaciones se transforman en desafíos para próximos desarrollos.

Galileo permite a investigadores de Comunicación acceder a los contenidos que los medios vierten en sus cuentas de redes sociales, especialmente considerando que la acción de plataformas como *Twitter* o *Facebook* concentran gran parte de las lectorías de noticias

6. Notas

1. Por ejemplo, hay medios como *EMOL.com* de Chile que desde 2019 en adelante adoptó la política de publicar pocas noticias en redes sociales privilegiando la suscripción de lectores en su web.
 2. *CrowdTangle Team* (2020). *CrowdTangle*. Facebook, Menlo Park, California, Estados Unidos.
 3. Compartimos en forma de datos abiertos los datos de esta extracción para el uso de la comunidad académica en ciencias de la comunicación en:
https://github.com/luisarcamo/cc_9c_2Y
 4. Acceso a:
<http://kibana.galileo-jrp.org/app/discover#/>
- También se puede llenar el formulario disponible en
<https://forms.gle/VgJ1BAyuga1nTn4v5>
para formar parte de la comunidad de investigadores.

7. Referencias

- Arcila-Calderón, Carlos; Barredo-Ibáñez, Daniel; Castro, Cosette** (2017). *Análítica y visualización de datos en Twitter*. Barcelona: Colección Comunicación. Editorial UOC. ISBN: 978 84 9116 960 4
- Baden, Christian** (2018). "Reconstructing frames from intertextual news discourse: A semantic network approach to news framing analysis". In: D'Angelo, Paul (ed.). *Doing news framing analysis II: Empirical and theoretical perspectives*. New York: Routledge, pp. 3-26. ISBN: 978 1 315642239
- Baeza-Yates, Ricardo** (2018). "Bias on the web". *Communications of the ACM*, v. 61, n. 6, pp. 54-61.
<https://doi.org/10.1145/3209581>
- Blei, David M.; Ng, Andrew Y.; Jordan, Michael I.** (2003). "Latent Dirichlet allocation". *Journal of machine learning research*, v. 3, pp. 993-1022.
<https://dl.acm.org/doi/10.5555/944919.944937>
- Cárcamo-Ulloa, Luis; Cárdenas-Neira, Camila; Sáez-Trumper, Diego; Toural-Bran, Carlos** (2021). "Fake news en Chile y España: ¿Cómo los medios nos hablan de noticias falsas?". *Journal of Iberian and Latin American research*, v. 26, n. 3, pp. 320-337.
<https://doi.org/10.1080/13260219.2020.1909849>
- Cárcamo-Ulloa, Luis; Vernier, Matthieu; Scheihing-García, Eliana; Aravena, Matías; Pérez, Javier** (2017). "Sophia una herramienta para la construcción y análisis de casos noticiosos en la enseñanza del periodismo". *Nuevas ideas en informática educativa*, v. 13, pp. 667-672. ISBN: 978 956 19 1043 0
<http://www.tise.cl/volumen13/TISE2017/96.pdf>
- Cárdenas-Neira, Camila** (2016). "Representación online del movimiento estudiantil chileno: Reapropiación de noticias en Facebook". *Estudios filológicos* n. 58, pp. 25-49.
<https://doi.org/10.4067/S0071-17132016000200002>
- Cardon, Dominique** (2018). *Con qué sueñan los algoritmos: Nuestras vidas en los tiempos de los big data*. Madrid: Dado Ediciones. ISBN: 978 849 450 728 1
- Cecchet, Emmanuel; Acharya, Amrita; Molom-Ochir, Tergel; Trivedi, Ameer; Shenoy, Prashant** (2020). "WiFiMon: a mobility analytics platform for building occupancy monitoring and contact tracing using wifi sensing: poster abstract". In: *Proceedings of the 18th conference on embedded networked sensor systems, SenSys'20*, pp. 792-793.
<https://doi.org/10.1145/3384419.3430598>

- Grimmer, Justin; Stewart, Brandon M.** (2013). "Text as data: The promise and pitfalls of automatic content analysis methods for political texts". *Political analysis*, v. 21, n. 3, pp. 267-297.
<https://doi.org/10.1093/pan%2Fmps028>
- Guo, Lei; Su, Chao; Paik, Sejin; Bhatia, Vibhu; Prasad-Akavoor, Vidya; Gao, Ge; Betke, Margrit; Wijaya, Derry** (2022). "Proposing an open-sourced tool for computational framing analysis of multilingual data". *Digital journalism*, first online.
<https://doi.org/10.1080/21670811.2022.2031241>
- Hamborg, Felix; Donnay, Karsten; Gipp, Bela** (2019). "Automated identification of media bias in news articles: an interdisciplinary literature review". *International journal on digital libraries*, n. 20, pp. 391-415.
<https://doi.org/10.1007/s00799-018-0261-y>
- Hermida, Alfred** (2010). "Twittering the news. The emergence of ambient journalism". *Journalism practice*, v. 4, n. 3, pp. 297-308.
<https://doi.org/10.1080/17512781003640703>
- Jiang, Ye; Song, Xingyi, Harrison, Jackie; Quegan, Shaun; Maynard, Diana** (2017). "Comparing attitudes to climate change in the media using sentiment analysis based on latent Dirichlet allocation". In: *Proceedings of the 2017 EMNLP workshop: Natural language processing meets journalism*, pp. 25-30.
<https://doi.org/10.18653/v1/W17-4205>
- Levy, Ro'ee** (2021). "Social media, news consumption, and polarization: evidence from a field experiment". *American economic review*, v. 111, n. 3, pp. 831-870.
<https://doi.org/10.1257/aer.20191777>
- Lewis, Seth C.; Zamith, Rodrigo; Hermida, Alfred** (2013). "Content analysis in an era of big data: a hybrid approach to computational and manual methods". *Journal of broadcasting & electronic media*, v. 57, n. 1, pp. 34-52.
<https://doi.org/10.1080/08838151.2012.761702>
- Li, Yue; Nair, Pratheeksha; Wen, Zhi; Chafi, Imane; Okhmatovskaia, Anya; Powell, Guido; Shen, Yannan; Buckeridge, David** (2020). "Global surveillance of Covid-19 by mining news media using a multi-source dynamic embedded topic model". In: *Proceedings of the 11th ACM international conference on bioinformatics, computational biology and health informatics, BCB'20*, n. 34.
<https://doi.org/10.1145/3388440.3412418>
- Ling, Rich** (2020) "Confirmation bias in the era of mobile news consumption: the social and psychological dimensions". *Digital journalism*, v. 8, n. 5, pp. 596-604.
<https://doi.org/10.1080/21670811.2020.1766987>
- Matthes, Jörg; Kohring, Mathias** (2008). "The content analysis of media frames: Toward improving reliability and validity". *Journal of communication*, v. 58, n. 2, pp. 258-279.
<https://doi.org/10.1111/j.1460-2466.2008.00384.x>
- Mellado, Claudia; Cárcamo-Ulloa, Luis; Alfaro, Amaranta; Inai, Darla; Isbej, José** (2021a). "Fuentes informativas en tiempos de Covid-19: Cómo los medios en Chile narraron la pandemia a través de sus redes sociales". *Profesional de la información*, v. 30, n. 4, e300421.
<https://doi.org/10.3145/epi.2021.jul.21>
- Mellado, Claudia; Hallin, Daniel; Cárcamo-Ulloa, Luis; Alfaro, Rodrigo; Jackson, Daniel; Humanes, María-Luisa; Márquez-Ramírez, Mireya; Mick, Jacques; Mothes, Cornelia; Lin, Christi-I-Hsuan; Lee, Misook; Alfaro, Amaranta; Isbej, José; Ramos, Andrés** (2021b). "Sourcing pandemic news: A cross-national computational analysis of mainstream media coverage of Covid-19 on Facebook, Twitter, and Instagram". *Digital journalism*, v. 9, n. 9, pp. 1261-1285.
<https://doi.org/10.1080/21670811.2021.1942114>
- Mellado, Claudia; Hermida, Alfred** (2021). "The promoter, celebrity, and joker roles in journalists' social media performance". *Social media + society*, v. 7, n. 1.
<https://doi.org/10.1177/2056305121990643>
- Morstatter, Fred; Wu, Liang; Yavanoglu, Uraz; Corman, Stephen R.; Liu, Huan** (2018). "Identifying framing bias in online news". *ACM Transactions on social computing*, v. 1, n. 2.
<https://doi.org/10.1145/3204948>
- Newman, Nic; Dutton, William H.; Blank, Grant** (2012). "Social media in the changing ecology of news: the fourth and fifth estates in Britain". *International journal of internet science*, v. 7, n. 1, pp. 6-22.
https://www.ijis.net/ijis7_1/ijis7_1_newman_et_al.pdf
- Newman, Nic; Fletcher, Richard; Schulz, Anne; Andi, Singe; Robertson, Craig T.; Nielsen, Rasmus-Kleis** (2021). *Digital news report 2021*. Reuters Institute for the Study of Journalism.
https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2021-06/Digital_News_Report_2021_FINAL.pdf

- Ojo, Adegboyega; Heravi, Bahareh** (2018). "Patterns in award winning data storytelling. Story types, enabling tools and competences". *Digital journalism*, v. 6, n. 6, pp. 693-718.
<https://doi.org/10.1080/21670811.2017.1403291>
- Pariser, Ely** (2011). *The filter bubble: How the new personalized web is changing what we read and how we think*. Penguin. ISBN: 978 0 143121237
- Pereira, Moisés; Cardeal-Pádua, Flavio-Luis; Machado-Pereira, Adriano-César; Silva, Giani-David; Benevenuto-de-Souza, Fabricio** (2015). "Multimodal sentiment analysis for automatic estimation of polarity tension of TV news in TV newscasts videos". In: *Proceedings of the 21st Brazilian symposium on multimedia and the web, WebMedia'15*, pp. 157-160.
<https://doi.org/10.1145/2820426.2820461>
- Raimondo-Anselmino, Natalia; Sambrana, Alejandro; Cardoso, Ana-Laura** (2017). "Medios tradicionales y redes sociales en internet: un análisis de los posts compartidos por los diarios argentinos *Clarín* y *La Nación* en Facebook (2010-2015)". *Astrolabio*, n. 19, pp. 32-68.
<https://revistas.unc.edu.ar/index.php/astrolabio/article/view/17787>
- Romanou, Angelika; Smeros, Panayiotis; Castillo, Carlos; Aberer, Karl** (2020). "Scilens news platform: A system for real-time evaluation of news articles". In: *Proceedings of the VLDB endowment*, v. 13, n. 12, pp. 2969-2972.
<https://doi.org/10.14778/3415478.3415521>
- Sáez-Trumper, Diego; Castillo, Carlos; Lalmas, Mounia** (2013). "Social media news communities: gatekeeping, coverage, and statement bias". In: *Proceedings of the 22nd ACM international conference on information & knowledge management*, pp. 1679-1684.
<https://doi.org/10.1145/2505515.2505623>
- Salazar, Diego** (2019). *No hemos entendido nada: Qué ocurre cuando dejamos el futuro de la prensa a merced de un algoritmo*. Editorial Debate. ISBN: 978 84 17636258
- Schmitz-Weiss, Amy; De-Macedo-Higgins-Joyce, Vanessa; Saldaña, Magdalena; Calmon-Alves, Rosental** (2017). "Latin American investigative journalism education: Learning practices, learning gaps". *Journalism & mass communication educator*, v. 72, n. 3, pp. 334-348.
<https://doi.org/10.1177/1077695817711611>
- Trilling, Damian; Jonkman, Jeroen G. F.** (2018). "Scaling up content analysis". *Communication methods and measures*, v. 12, n. 2-3, pp. 158-174.
<https://doi.org/10.1080/19312458.2018.1447655>
- Underwood, Richard** (2017). "Building bridges: The system administration tools and techniques used to deploy bridges". In: *Proceedings of the practice and experience in advanced research computing 2017 on sustainability, success and impact, PEARC17*, article n. 5.
<https://doi.org/10.1145/3093338.3093339>
- Van-Atteveldt, Wouter; Peng, Tai-Quan** (2018). "When communication meets computation: Opportunities, challenges, and pitfalls in computational communication science". *Communication methods and measures*, v. 12, n. 2-3, pp. 81-92.
<https://doi.org/10.1080/19312458.2018.1458084>
- Vernier, Mathieu; Cárcamo-Ulloa, Luis; Scheihing-García, Eliana** (2017). "Diagnóstico de la estrategia editorial de medios informativos chilenos en *Twitter* mediante un clasificador de noticias automatizado". *Revista austral de ciencias sociales*, n. 30, pp. 183-201.
<https://doi.org/10.4206/rev.austral.cienc.soc.2016.n30-09>
- Watts, Duncan J.** (2016). "Computational social science: Exciting progress and future challenges". In: *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, KDD'16*, p. 419.
<https://doi.org/10.1145/2939672.2945366>
- Watts, Duncan J.** (2017). "Should social science be more solution-oriented?". *Nature human behaviour*, v. 1, artículo n. 15.
<https://doi.org/10.1038/s41562-016-0015>
- Watts, Duncan J.; Rothschild, David M.; Mobius, Markus** (2021). "Measuring the news and its impact on democracy". In: Scheufele, Dietram (ed.). *Proceedings of the National Academy of Sciences*, v. 118, n. 15.
<https://doi.org/10.1073/pnas.1912443118>
- Zhang, Hao; Boons, Frank; Batista-Navarro, Riza** (2019). "Whose story is it anyway? Automatic extraction of accounts from news articles". *Information processing & management*, v. 56, n. 5, pp. 1837-1848.
<https://doi.org/10.1016/j.ipm.2019.02.012>
- Zhou, Xinyi; Zafarani, Reza; Shu, Kai; Liu, Huan** (2019). "Fake news: fundamental theories, detection strategies and challenges". In: *Proceedings of the 12th ACM international conference on web search and data mining, WSDM '19*, pp. 836-837.
<https://doi.org/10.1145/3289600.3291382>

8. Anexo

País (Medios)	Medio	Origen medial	Lengua	Extracción de datos		
				Twitter	Facebook	Instagram
Corea del Sur 8 medios	KBS	TV	Coreano	Sí	Sí	Sí
	JTBC	TV	Coreano	Sí	Sí	Sí
	CBS South Korea	Radio	Coreano	Sí	Sí	No
	TBS Radio	Radio	Coreano	Sí	Sí	Sí
	Chosun	Print	Coreano	Sí	Sí	Sí
	Hankyoreh	Print	Coreano	Sí	Sí	No
	Dailian	Online	Coreano	Sí	Sí	Sí
	OhMyNews	Online	Coreano	Sí	Sí	Sí
Japón 9	NHK	TV	Japonés	Sí	Sí	Sí
	TV Asahi	TV	Japonés	Sí	Sí	Sí
	Nihon TV	TV	Japonés	Sí	Sí	Sí
	NHK radio	Radio	Japonés	Sí	Sí	Sí
	Bunka Housou	Radio	Japonés	Sí	Sí	No
	The Yomiuri Shinbum	Print	Japonés	Sí	Sí	Sí
	The Asahi Shinbum	Print	Japonés	Sí	Sí	Sí
	Huffpost Japan	Online	Japonés	Sí	Sí	Sí
	Buzzfeed Japan	Online	Japonés	Sí	Sí	Sí
Chile 10	Mega	TV	Español	Sí	Sí	Sí
	TVN	TV	Español	Sí	Sí	Sí
	Canal 13	TV	Español	Sí	Sí	Sí
	CNN Chile	TV	Español	Sí	Sí	Sí
	Biobio.cl	Radio	Español	Sí	Sí	Sí
	Emol.com	Online	Español	Sí	Sí	Sí
	Elmostrador.cl	Online	Español	Sí	Sí	Sí
	Cooperativa	Radio	Español	Sí	Sí	Sí
	La Tercera	Print	Español	Sí	Sí	Sí
	Las Últimas Noticias	Print	Español	Sí	Sí	Sí
México 12	Televisa	TV	Español	Sí	Sí	Sí
	Canal Once	TV	Español	Sí	Sí	No
	Radio Fórmula	Radio	Español	Sí	Sí	Sí
	Aristegui Noticias	Radio	Español	Sí	Sí	Sí
	Instituto Mexicano de la Radio	Radio	Español	Sí	Sí	No
	La Jornada	Print	Español	Sí	Sí	Sí
	Reforma	Print	Español	Sí	Sí	Sí
	El Financiero	Print	Español	Sí	Sí	Sí
	La Prensa	Print	Español	Sí	Sí	Sí
	Animal Político	Online	Español	Sí	Sí	Sí
	El Universal Online México	Print	Español	Sí	Sí	Sí
	UnoTV	Online	Español	Sí	Sí	Sí
Brasil 8	Record	TV	Portugués	Sí	Sí	Sí
	CBN	Radio	Portugués	Sí	Sí	Sí
	Band	Radio	Portugués	Sí	Sí	Sí
	Folha de S. Paulo	Print	Portugués	Sí	Sí	Sí
	O Globo	TV	Portugués	Sí	Sí	Sí
	O Estado de S. Paulo	Print	Portugués	Sí	Sí	Sí
	R7	Online	Portugués	Sí	Sí	Sí
	G1	Online	Portugués	Sí	Sí	No

Estados Unidos 11	NBC	TV	Inglés	Sí	Sí	Sí
	CNN USA	TV	Inglés	Sí	Sí	Sí
	KABC	TV	Inglés	Sí	Sí	Sí
	NPR	Radio	Inglés	Sí	Sí	Sí
	CBS US	Radio	Inglés	Sí	Sí	Sí
	New York Times	Print	Inglés	Sí	Sí	Sí
	USA Today	Print	Inglés	Sí	Sí	Sí
	Los Angeles Times	Print	Inglés	Sí	Sí	Sí
	Buzzfeed US	Online	Inglés	Sí	Sí	Sí
	Fox US	TV	Inglés	Sí	Sí	Sí
	Huffington Post US	Online	Inglés	Sí	Sí	Sí
España 16	Telecinco	TV	Español	Sí	Sí	Sí
	Antena 3	TV	Español	Sí	Sí	Sí
	La Sexta	TV	Español	Sí	Sí	Sí
	TVE	TV	Español	Sí	Sí	Sí
	SER	Radio	Español	Sí	Sí	Sí
	COPE	Radio	Español	Sí	Sí	Sí
	Onda Cero	Radio	Español	Sí	Sí	Sí
	Radio Nacional de España	Radio	Español	Sí	Sí	Sí
	El País	Print	Español	Sí	Sí	Sí
	La Vanguardia	Print	Español	Sí	Sí	Sí
	El Mundo	Print	Español	Sí	Sí	Sí
	ABC Spain	Print	Español	Sí	Sí	Sí
	El Confidencial	Online	Español	Sí	Sí	Sí
	OK Diario	Online	Español	Sí	Sí	Sí
	Eldiario.es	Online	Español	Sí	Sí	Sí
	Huffpost ES	Online	Español	Sí	Sí	Sí
Alemania 8	ARD	TV	Alemán	Sí	Sí	Sí
	RTL	TV	Alemán	Sí	Sí	Sí
	Deutschland-funk	Radio	Alemán	Sí	Sí	Sí
	Klassik Radio	Radio	Alemán	Sí	Sí	Sí
	BILD	Print	Alemán	Sí	Sí	Sí
	Frankfurter Allgemeine Zeitung	Print	Alemán	Sí	Sí	Sí
	Süddeutsche Zeitung	Print	Alemán	Sí	Sí	Sí
Spiegel Online	Print	Alemán	Sí	Sí	Sí	
Reino Unido 15	BBC News	TV	Inglés	Sí	Sí	Sí
	Channel 4 News	TV	Inglés	Sí	Sí	Sí
	Sky News	TV	Inglés	Sí	Sí	Sí
	ITV News	TV	Inglés	Sí	Sí	Sí
	BBC Radio 4	Radio	Inglés	Sí	Sí	No
	BBC Radio 2	Radio	Inglés	Sí	Sí	Sí
	TalkSport	Radio	Inglés	Sí	Sí	Sí
	ClasSic FM	Radio	Inglés	Sí	Sí	Sí
	The Daily Telegraph	Print	Inglés	Sí	Sí	Sí
	The Guardian	Print	Inglés	Sí	Sí	Sí
	The Daily Mirror	Print	Inglés	Sí	Sí	Sí
	The Sun	Print	Inglés	Sí	Sí	Sí
	Mail online	Online	Inglés	Sí	Sí	Sí
	Huffpost UK	Online	Inglés	Sí	Sí	Sí
	Buzzfeed UK	Online	Inglés	Sí	Sí	Sí