



Revista Interamericana de Psicología/Interamerican
Journal of Psychology

ISSN: 0034-9690

etorresrivera@gmail.com

Sociedad Interamericana de Psicología
Puerto Rico

Lobato, Roberto; Velandia-Morales, Andrea; Sánchez-Rodríguez,
Ángel; Montoya-Lozano, Mar; García-Sánchez [2], Efraín
El *fact-checking* en Twitter: Un análisis del hashtag #StopBulos
Revista Interamericana de Psicología/Interamerican Journal
of Psychology, vol. 55, núm. 2, 1371, 2021, Mayo-Agosto
Sociedad Interamericana de Psicología
San Juan, Puerto Rico

Disponible en: <https://www.redalyc.org/articulo.oa?id=28474830003>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica Redalyc

Red de Revistas Científicas de América Latina y el Caribe, España y Portugal
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso
abierto

El *fact-checking* en Twitter: Un análisis del hashtag #StopBulos

Roberto Lobato ^a, Andrea Velandia-Morales ^{b1}, Ángel Sánchez-Rodríguez ^c, Mar Montoya-Lozano ^b, & Efraín García-Sánchez ^d²

Fundación Euroárabe de Altos Estudios, Granada, España ^a; Universidad de Granada, Granada, España ^b; Universidad de Salamanca, Salamanca, España ^c; Universidad de Sao Paulo, Sao Paulo, Brasil ^d.

RESUMEN

La verificación de noticias falsas es una herramienta importante para mejorar la calidad de la información que circula en redes virtuales. Aunque existen distintas agencias de verificación de noticias, también encontramos algunas estrategias más informales como el uso del hashtag #StopBulos. Así, la presente investigación pretende realizar una caracterización del hashtag #StopBulos en Twitter como forma de verificación de la información y de control de la difusión de *fake news*. Los resultados mostraron que existía diversidad entre los usuarios y las temáticas de los tuits que incluían dicho hashtag, mientras que la principal función era desmentir *fake news*. No obstante, se encontró que aquellos que conseguían una mayor difusión eran los usuarios con mayor número de seguidores y carácter institucional. Se discuten las implicaciones del uso del hashtag #StopBulos como una herramienta que permite identificar información falsa en redes sociales.

Palabras Claves

fake news; posverdad; bulos; redes sociales; sociedades red.

ABSTRACT

The fact-checking is an important tool to improve the quality of the information that circulates in virtual networks. Although there are different fact-checking verification agencies, we also found some more informal strategies such as the use of the hashtag #StopBulos. Thus, this research aims to characterize the #StopBulos hashtag on Twitter as a way to verify information and control the spread of fake news. The results showed that there was diversity among users and the themes of the tweets that included this hashtag, while the main function was to deny fake news. However, it was found that those who achieved greater dissemination were the users with the largest number of followers and institutional character. The implications of using the #StopBulos hashtag as a tool to identify false information on social networks are discussed.

Keywords

fake news; post-truth; post-news; social media; network societies

¹ Correspondence about this article should be addressed **Andrea Velandia Morales:** avelandia@correo.ugr.es

² **Conflicts of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Fact-checking on Twitter: An analysis of the hashtag #StopBulos

La posverdad es un término acuñado en la década de los noventa y utilizado, principalmente, en contextos políticos (Grech, 2017). Esta se entiende como una distorsión deliberada de una realidad cuyo fin es manipular las creencias y emociones para influir en la opinión pública, así como en las actitudes sociales y políticas (Braun y Dodge, 2018). En el marco de la posverdad, los hechos cobran menos relevancia al momento de generar y formar opinión dando paso a las emociones y las creencias personales (Bakir y McStay, 2018; Rabin-Havt, 2016). Partiendo de esta definición, la posverdad puede moldear creencias, emociones y actitudes hacia un evento específico, lo que puede manifestarse a través de la polarización, los sesgos, el prejuicio y los comportamientos asociados (Harsin, 2018).

En la actualidad, la posverdad se materializa a través de *fake news*—o bulos por su nombre en español—siendo uno de los medios más efectivos a la hora de generar desinformación (Harsin, 2015). Las *fake news* son contenidos deliberadamente creados para tratar de persuadir y engañar a la población (Allcott y Gentzkow, 2017; Gelfert, 2018; Nelson, 2017; Schifferes et al., 2014; Walsh, 2010). Las *fake news* se publican en *websites* con estructuras similares a páginas de medios de comunicación tradicionales y se difunden utilizando las redes sociales virtuales (Nelson, 2017). Su contenido puede ser variado, desde incluir la sátira a suplantar fuentes o contenido falso (Grech, 2007; Wardle, 2017). Asimismo, entre sus efectos encontramos la desinformación, el partidismo, la polarización, la discriminación, la alienación o el cinismo (Balmas, 2014; Harsin, 2018).

Según el Estudio Anual de Redes Sociales realizado por IAB Span (2018) en el contexto español, el 85% de los internautas de 16-65 años utilizan redes sociales. Esto representa más de 25 millones de usuarios. Entre las redes sociales con más usuarios, por orden, se encuentran Facebook, Instagram y Twitter, además de otro tipo de redes de mensajería como WhatsApp. Entre los principales usos de las redes sociales se encuentran enviar mensajes privados o públicos a los contactos, seguir otras cuentas como medios de comunicación, partidos políticos y comentar asuntos de actualidad.

Ante este panorama, se hace evidente que las redes sociales pueden ser usadas como una herramienta de persuasión y control mediante la difusión de mensajes que atiendan a intereses de particulares (Castells, 2007). Concretamente, las redes virtuales

están siendo utilizadas, principalmente, para difundir *fake news* enfocadas en generar clics y retornos publicitarios (Barfar, 2019).

Ante el incremento en la producción y difusión de *fake news* a través de las redes virtuales, se han desarrollado estrategias, como el *fact-checking*, para verificar la información difundida en las noticias, lo que permite identificar contenidos que buscan desinformar. Así, el *fact-checking* ayuda a reducir la desinformación e incrementar la calidad de la información que consumen los usuarios en las redes virtuales (Molina-Cañabate y Magallón-Rosa, 2020). En este sentido, la presente investigación tiene por objetivo analizar uno de estas estrategias de *fact-checking*. Específicamente, se propone un análisis descriptivo del hashtag #StopBulos dentro de la red social de Twitter como un mecanismo de verificación y control de la difusión de noticias falsas.

Dinámica de las fake news

La dinámica de las *fake news* puede ser abordada desde la perspectiva de las sociedades en red (Castells 2006, 2007, 2009). Según Castells, la ciudadanía se agrupa en redes a través de las cuales se difunde la información de acuerdo a niveles de poder. Así, existen grupos con la capacidad de “influir de forma asimétrica en las decisiones de otros actores sociales, de modo que se favorezcan la voluntad, los intereses y los valores del actor que tiene el poder” (Castells, 2009: 33). Donde los que tienen el poder son aquellos con mayor capacidad para crear redes de información e incidir en las dinámicas de difusión.

Tradicionalmente las redes de comunicación han sido verticales, con flujos de información unidireccional (e.g., televisión y periódico en papel). En estas, quienes ostentaban el poder de la comunicación eran las entidades que poseían los medios de comunicación. Por el contrario, las redes sociales virtuales han facilitado un flujo de información horizontal, de manera que las personas pueden interactuar y crear noticias sin intermediarios (e.g., a través de redes sociales y blogs), generando nuevas formas de comunicación denominadas auto-comunicación de masas (Castells, 2007).

Esta nueva forma de comunicación incrementa la agencia de los individuos al ser tanto receptores como emisores de mensajes y permite mayor flexibilidad, instantaneidad y capacidad ilimitada para difundir cualquier tipo de información. Asimismo, este tipo de comunicación permite que el poder pueda distribuirse hacia nuevos agentes de comunicación, los cuales no obtienen su reconocimiento por la tradición o experiencia,

sino por otros indicadores, como, por ejemplo, la cantidad de seguidores (Himma-Kadakas, 2017).

Aunado a esto, las *fake news* pueden ser políticamente motivadas (e.g., elección presidencial en USA 2016) o creadas con fines publicitarios o de mercadeo (Chen y Cheng, 2020; Mehta et al., 2021). Por ejemplo, Twitter ha sido considerada como una metáfora de la democracia puesto que la información se transmite sin intermediarios ni editores, sumado a su inmediatez en la difusión del contenido (Aparicio Moreno, 2016). No obstante, Twitter y otras redes sociales también pueden minar la democracia, ya que allí, se encuentran también medios de información masivos (tradicionales), que puede usar el poder de difusión de las redes sociales para instigar la polarización y amplificar discursos autoritarios que ataquen o estigmaticen a diferentes grupos sociales (Gervais, 2015; Theocharis et al., 2016)

Estas nuevas dinámicas comunicativas facilitan la difusión de información y noticias, cuyos contenidos no siempre se verifican, ya sea porque se confía en la fuente debido a su reconocimiento público (e.g., número de seguidores) o por la imposibilidad de verificar todas las noticias dado el volumen y velocidad de exposición a la información (Himma-Kadakas, 2017). Teniendo en cuenta además la proliferación de máquinas de propaganda, algoritmos sesgados, robots (*bots*) usados en las redes sociales para viralizar cualquier tipo de información, y sumado a la falta de regulación a nivel tanto local como global, se genera un escenario propicio para la proliferación de las *fake news* (Mehta et al., 2021).

Las *fake news* no son un fenómeno nuevo, estas se pueden considerar la nueva retórica, utilizada tanto por líderes, como por gobiernos y medios de comunicación, con el propósito de atender a intereses particulares. Estas han ganado relevancia social al convertirse en una herramienta de desinformación usada para desestabilizar procesos democráticos. Las *fake news* pueden ser usadas para cambiar la agenda política, la opinión pública y para impulsar el cambio social (Iosifidis y Wheeler, 2018). Su posible influencia en el proceso electoral en EE.UU. (Allcott y Gentzkow, 2017) y en el Brexit (Bastos y Mercea, 2017) ha instaurado el miedo acerca del impacto en la determinación del voto de los ciudadanos, minando así los procesos democráticos. Asimismo, parecen estar relacionadas con el surgimiento de protestas (Tucker et al., 2017), el descarrilamiento de investigaciones policiales (Starbird et al., 2014) y el ciberacoso de opositores políticos (Munger, 2017). También se han utilizado para mantener discursos e intereses políticos, desestimando medios de comunicación y cuestionando prácticas periodísticas que buscan

divulgar información precisa, con el propósito de tergiversar la verdad y minar la confianza de los usuarios frente a fuentes de información confiables, lo cual dificulta determinar qué es y qué no es verdad (Brummette et al., 2018; Mehta et al., 2021; Van Duyn y Collier, 2019).

Investigaciones acerca del impacto e influencia de la propagación de contenido falso han mostrado que, cuando un contenido es falso, este se difunde mucho más rápido y tiene mayor alcance debido a su novedad y las emociones que despierta (e.g., sorpresa, disgusto) que cuando no lo es. De hecho, la información real tardaría seis veces más tiempo en llegar a 1,500 personas en comparación con la información falsa (Vosoughi, Roy y Aral, 2018). En consecuencia y a fin de frenar el impacto de las *fake news*, han surgido diversas estrategias que tratan de identificar las noticias con contenido falso y alertar a la audiencia sobre su veracidad, estas son denominadas *fact-checking* (Molina-Cañabate y Magallón-Rosa, 2020).

El fact-checking y el control de las fake news

Los receptores de las noticias son quienes les otorgan la validez. Este proceso, que se asemeja a la labor que realizan los periodistas, se denomina *citizen curation* o verificación ciudadana (Pedersen y Burnett, 2018). No obstante, por lo general, muchas personas no verifican la información a la hora de evaluar las noticias (Himma-Kadakas, 2017). De esta forma, la verificación de las noticias por parte de la ciudadanía está sujeta a diferentes sesgos cognitivos. Por ejemplo, el acceso y disponibilidad de información (Petty y Cacioppo, 1984), la existencia de un anclaje previo (confiar demasiado en la primera información ofrecida), un sesgo de confirmación (favorecer la información que confirma las propias creencias), una sobrestimación en la percepción del nivel de confianza o una motivación opuesta (Endsley, 2018). Asimismo, es de destacar la confianza depositada en uno mismo a la hora de verificar las noticias, puesto que se tiende a pensar que los demás son peores a la hora de reconocer las *fake news*, especialmente cuando no pertenecen a nuestro grupo social (Jang y Kim, 2018).

Ante el poder de desestabilizar y el fuerte impacto social de las *fake news*, el *fact-checking* surge como una medida para contrarrestar la desinformación que circula en las redes (Uscinski y Butler, 2013) y contribuye a la corrección de la desinformación por parte de la prensa en las sociedades democráticas (Kovach y Rosenstiel, 2010). El *fact-checking* se basa en la comprobación de los datos que aparecen en las noticias a fin de estimar su credibilidad (Molina-Cañabate y Magallón-Rosa, 2020). Según el Manual de

Verificación de Silverman (2016), algunos de los aspectos más relevantes a tener en cuenta a la hora de verificar una noticia residen en el emisor, en el contenido y en la consistencia con otras noticias de la misma temática. En cuanto al emisor de la noticia, es importante saber quién es el autor, si es un experto en la temática, cuáles son sus objetivos, qué medio de comunicación distribuye la noticia y que objetivos tiene dicho medio. Respecto al contenido, resulta imprescindible comprobar si la noticia se presenta de forma lógica, distingue entre hechos y valoraciones, está apoyada en hechos verificados, es de actualidad, presenta diferentes puntos de vista y si muestra un fuerte sesgo ideológico. Finalmente, en lo relativo a la consistencia con otras noticias, hay que comprobar si las noticias relacionadas son consistentes con la información que presenta la noticia a verificar y si la misma noticia aparece en otros medios.

Debido a los sesgos que intervienen en el consumo de información, a las diferencias en las habilidades cognitivas de las personas, tiempo y otros factores que dificultan la verificación de las noticias, algunos medios de comunicación han creado códigos éticos. Por ejemplo, el Código de Principios del IFCN (*Intenational Fact-Checking Network*), organismo comprometido con la excelencia en los procesos de verificación periodística a nivel mundial que se encarga de asegurar que los firmantes (*signatarios*) cumplan con una serie de principios deontológicos que garantizan la imparcialidad y profesionalidad de sus desmentidos. El mencionado código requiere el compromiso con el no partidismo y la equidad, con la transparencia de las fuentes, con la transparencia de la financiación y la organización, con la transparencia de la metodología, y con las correcciones abiertas y honestas. No obstante, la dinámica de las redes sociales en internet ha propiciado la creación de otros tipos de verificación auto-gestionados con la participación pública y abierta de diferentes actores sociales que desmienten las *fake news*. El hashtag #StopBulos, utilizado principalmente en Twitter, es una de estas tácticas.

El hashtag #StopBulos, de autoría indeterminada hasta donde nuestro conocimiento alcanza, puede considerarse como una iniciativa orgánica, es decir, que surge espontáneamente de la necesidad de visibilizar en las redes virtuales que el contenido de un mensaje es falso. Este hashtag es una herramienta usada por diversos usuarios de Twitter para resaltar la falsedad de ciertos contenidos, lo cual podría catalogarse como una respuesta de la ciudadanía y las instituciones para identificar rápidamente los bulos en las redes sociales virtuales. No obstante, aunque existen trabajos previos en los que se hace una descripción del proceso de *fact-checking* a nivel metodológico (Molina-Cañabate y Magallón-Rosa, 2020), es igual de relevante llevar a

cabo un análisis descriptivo y relacional de estas herramientas entendiéndolas como estrategias de respuesta ciudadana para la identificación de *fake news*. Por tanto, esta investigación busca analizar las dinámicas de verificación de información a través del uso del hashtag #StopBulos en la red social Twitter.

La presente investigación

El objetivo principal de esta investigación es caracterizar el uso del hashtag #StopBulos en Twitter como forma de verificación de la información y de control de la difusión de *fake news*. Para tal fin, se proponen dos objetivos específicos. El primero es identificar cuáles son los usuarios, temáticas, organizaciones y funciones de los tuits que incluyen el hashtag #StopBulos a fin de comprender quiénes usan este mecanismo y para qué lo usan. El segundo objetivo es identificar las redes de interacción a partir de los usuarios de Twitter que usaron el hashtag, con fin de discernir qué usuarios tienen una mayor repercusión y quiénes promueven más el uso de estas estrategias. Adicionalmente, se examina la evolución del uso del hashtag #stopbulos en tres momentos diferentes en 2019, 2020 y 2021, los cuales se caracterizan por coyunturas sociales específicas tales como la crisis del COVID-19 y la gestión de la vacuna.

Método

Corpus de datos

Se realizó la búsqueda de (re)tuits utilizando como palabra clave el hashtag #StopBulos y limitando los retuits al idioma español utilizando los paquetes twitterR (Gentry, 2015) y rtweet (Kearney, 2019) implementados en el software de análisis de datos R (R Core Team, 2019). Se hicieron tres búsquedas en tres momentos distintos utilizando la API de Twitter; en estas búsquedas se descargaron tanto los tuits originales como los retuits publicados en español. Un primer momento entre el 23 y el 31 de marzo de 2019 (N = 536), la segunda descarga entre el 12 y 20 de marzo de 2020 (N = 16869), y un tercer periodo entre el 12 y el 19 de mayo de 2021 (N = 421) (véase Figura 1). El número de tuits se corresponde con todos los publicados en esas fechas mientras que la extensión temporal abarca el máximo que permite la API de Twitter. Los scripts usados para descargar los tuits y exportarlos para su posterior análisis, así como las tablas con los datos del análisis de redes y demás material, se pueden encontrar en el Material Suplementario (<https://osf.io/dek57/>).

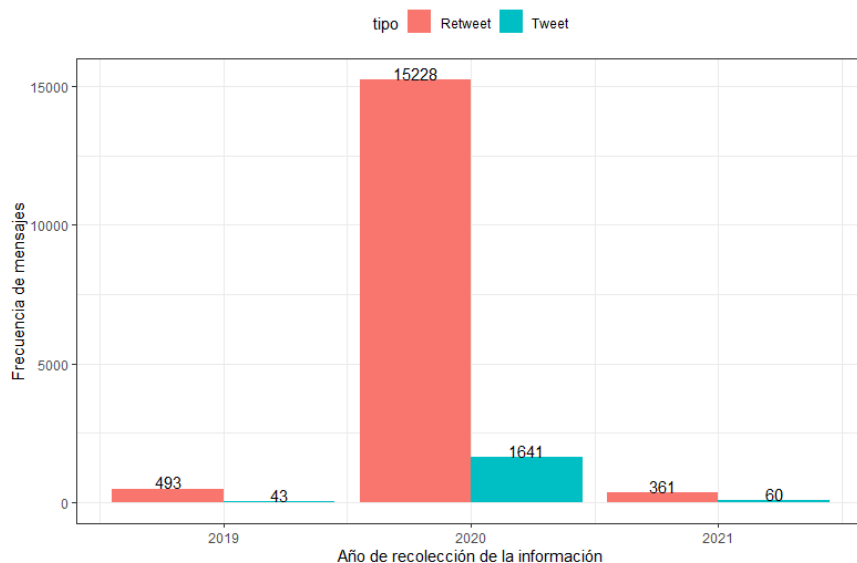


Figura 1. Frecuencia de los mensajes que usan el hashtag #StopBulos en los distintos momentos de descarga, divididos en tuits originales y retuits. *Fuente.* Elaboración Propia.

Estrategia Analítica

Se utilizaron dos estrategias analíticas diferentes pero complementarias para analizar el corpus de datos. Por un lado, se realizó análisis de contenido teniendo como base la Teoría Fundamentada (Glasser y Strauss, 1967). Este se llevó a cabo mediante estrategias de codificación abierta y axial (Strauss y Corbin, 2002). Con base en esta codificación, se hizo un análisis de frecuencias de las categorías y se caracterizaron los tuits que incluían el hashtag #StopBulos. Estos análisis fueron soportados en el software de análisis cualitativo Atlas.ti 8.

Posteriormente, se utilizaron técnicas de análisis de redes. Estos análisis permiten construir grafos basados en los principios de la Teoría Matricial (Diestel, 2010) y tienen dos componentes principales: los nodos y las aristas (o vértices). En este caso, los nodos representan a cada usuario de la muestra, mientras que las aristas muestran las relaciones entre esos usuarios, es decir, los retuits que realizaron los usuarios. Estos análisis fueron soportados en el software de análisis cualitativo Gephi 0.9.2.

Resultados

Análisis de contenido

Con el objetivo de identificar usuarios, temáticas, organizaciones y funciones, se realizó un análisis de contenido mediante el cual se crearon sistemas categoriales. En el primer sistema, se describen los usuarios de Twitter que usaron el hashtag #StopBulos. En el segundo sistema, se categorizan los contenidos de dichos tuits. En el tercer sistema,

se distinguen las estrategias de *fact-checking* asociadas a este hashtag. Finalmente, en el cuarto sistema, se engloban las principales funciones que se dan a este hashtag.

Descripción de usuarios de la red de difusión

De los 536 tuits descargados en el 2019, 43 tuits fueron mensajes originales, en torno a los cuales interactuaron 66 usuarios y se realizaron 493 retuits. Dado que algunos usuarios citaban a otros usuarios en los tuits que estos escribían o retuiteaban más de una vez los mensajes, el número de usuarios es mayor al número de mensajes.

Utilizando el modelo de Barreto-Galeano et al. (2019), se clasificaron los actores en diferentes niveles, identificando cuatro niveles diferentes (véase Tabla 1). En el primer nivel encontramos a la *sociedad* (56.06%), donde se ubican principalmente ciudadanos y organizaciones privadas (relacionadas con la educación y la salud). En el segundo nivel, encontramos los organismos dedicados al *fact-checking* (18.18%), es decir, organizaciones que se dedican a corroborar y desmentir los bulos que circulan en la red. Dentro de estos organismos, destacan organizaciones como VOST (Asociación Nacional de Voluntarios Digitales de Emergencias) y Maldita.es. El tercer nivel se denomina *grupo* (13.64%), este se compone de cuentas gubernamentales (e.g. La Policía Nacional, Junta de Andalucía, Gobierno de Gran Canaria), algunos hospitales, un partido político y una organización religiosa. Finalmente, en el cuarto nivel están los *medios de comunicación* tradicionales (e.g., El Mundo, RTVE) e independientes (12.12%).

Tabla 1
Caracterización de los actores por niveles

Nivel	Descripción	fr	%
Sociedad		37	56.06
	Ciudadanos	27	
	Empresas privadas	10	
Fact-checking		12	18.18
Grupos		9	13.64
	Gobierno	7	
	Grupos políticos	1	
	Organizaciones religiosas	1	
Medios de comunicación		8	12.12
Total		66	100

Contenidos asociados al hashtag #StopBulos

En cuanto al contenido que se asocia a este hashtag, es su mayoría son tuits que se refieren a bulos relacionados con salud y belleza (35%), en concreto cáncer (“El #cáncer es, sin duda, la enfermedad que más bulos genera. @aecc_es alerta y @maldita_ciencia lo desmiente una vez más: las dietas alcalinas no sirven para curar el cáncer. #Medicina #Salud #StopBulos”) y dietas (“#stopbulos El vino tinto, la cerveza y el chocolate negro alargan la vida”). Seguido de tuits que mencionan la importancia del buen uso y la gestión de las noticias, así como las consecuencias de los bulos (26%; e.g., “La seguridad se incrementa eliminando bulos. De ahí el hashtag #StopBulos. @MariaLuisaMoreo @vostSPAIN #IICongresoSeguridadISO @InterSecurObser”). En tercer lugar se encuentran los bulos relacionados con la inmigración y la xenofobia (16%), cuyos contenidos difunden información falsa sobre refugiados de diferentes grupos étnicos (e.g., “@MBSocialValue @jnbeuze Son refugiados albanos en el puerto italiano de Bari, el 8 de agosto de 1991, en el buque Vlora #stopbulos”).

Posteriormente está la seguridad ciudadana (14%), donde se usa el miedo como recurso, mencionando acciones relacionadas con el hurto, el secuestro y la estafa (e.g., “Comprueba las informaciones que compartes Proliferan noticias falsas de intentos de secuestro de niños por extraños -que algunos son capaces de describir, sin haberlos visto-. Alcanzan gran difusión en #RedesSociales. Ayuda a frenarlas. #StopBulos”). Con menciones del 5% y por debajo están los bulos que abordan temas económicos, políticos, sobre derechos de autor (e.g., “Si alguien intenta venderte una papeleta a nuestro nombre para un sorteo, NO COMPRES, es un engaño. Todos los actos que realizamos para recaudar fondos son anunciados por web y #RedesSociales y llevan SIEMPRE el sello de la hermandad. #StopBulos #SevillaHoy”) y gestación subrogada (e.g., “#ElPSOEMiente Hoy le ha vuelto a tocar a la señora Lastra, con la #GestaciónSubrogada Ella cree que tachando nuestra propuesta puede engañar a alguien... Aquí dejamos el documento original #STOPbulos”) (Tabla 2).

Tabla 2
Temática de los bulos

Nivel	Fr	%
Salud y belleza	15	35
Gestión, uso y consecuencias de los bulos	11	26
Inmigración y xenofobia	7	16
Seguridad ciudadana	6	14
Economía y política	2	5
Derechos de autor	1	2
Gestación subrogada	1	2
Total	43	100

Organizaciones de fact-checking asociadas a #StopBulos

Entre los usuarios que se mencionan en los mensajes que usan el hashtag #StopBulos, encontramos diferentes organizaciones dedicadas al *fact-checking*. En concreto, el 26% de los tuits incluían este tipo de menciones. Entre las organizaciones que se mencionaban, se identificaron tres tipos. En primer lugar, se mencionan organizaciones especializadas en desmentir bulos. Entre ellas, destaca Maldita.es (<https://maldita.es>), que cuenta con distintas secciones especializadas en diferentes temáticas, por ejemplo, sobre migración y ciencia. En segundo lugar, se nombran organizaciones dedicadas a promover un uso adecuado de las redes sociales, sobre todo en situaciones de emergencia, como VOST, la Asociación de Voluntarios Digitales de emergencias (<https://vost.es>). En tercer lugar, se alude a organizaciones especializadas en el manejo de la reputación en situaciones de crisis, que se centran en el manejo de bulos relacionados con empresas privadas o marcas, entre las que se destaca SOS Works (<https://sosworkstech.com>).

Función de los tuits que incluyen #StopBulos

De manera adicional, se pudieron identificar tres objetivos o funciones asociadas a los tuits analizados. Estas son: 1) desmentir *fake news*, 2) informar sobre el *fact-checking* y las consecuencias de no desmentir las *fake news*, y 3) deslegitimar a una persona o entidad.

Desmentir *fake news* fue el principal objetivo (48.84%) y representa a los tuits cuyo contenido advertía sobre distintas noticias, a priori, falsas (e.g., “#stopbulos El vino tinto, la cerveza y el chocolate negro alargan la vida”). Informar sobre el *fact-checking* fue la segunda más usada (37.21%). En estos tuits, el contenido se centraba en presentar

información sobre las consecuencias de creer noticias falsas y en facilitar recursos y organismos dedicados al *fact-checking* (e.g., “Comprueba las informaciones que compartes Proliferan noticias falsas de intentos de secuestro de niños por extraños -que algunos son capaces de describir, sin haberlos visto-. Alcanzan gran difusión en #RedesSociales. Ayuda a frenarlas. #StopBulos”). Por último, en menor medida, algunos tuits buscaban la deslegitimación de una persona o entidad (13.95%; e.g., “@elmundoes Terrorismo informativo. #StopBulos #fakenews”).

Redes de difusión e interacción asociadas a #StopBulos

En segundo lugar, a fin de identificar cómo son las relaciones entre los usuarios y estos tuits, se realizó un análisis de redes teniendo en cuenta el número de retuits o, en otras palabras, la capacidad de difusión. Este tipo de análisis permitió visualizar un grafo que contaba con 514 nodos y 618 aristas usando el algoritmo *Fruchterman Reingold* (véase Figura 2). Este algoritmo, que se basa en la gravitación de los nodos según la atracción o repulsión hacia otros nodos, muestra una red donde los nodos más relevantes son de mayor tamaño (mayor grado de entrada o retuits); los nodos con más conexiones (o retuits) se colocan en posiciones más centrales; y los nodos más relacionados entre sí se ubican más cerca el uno del otro (centralidad de intermediación).

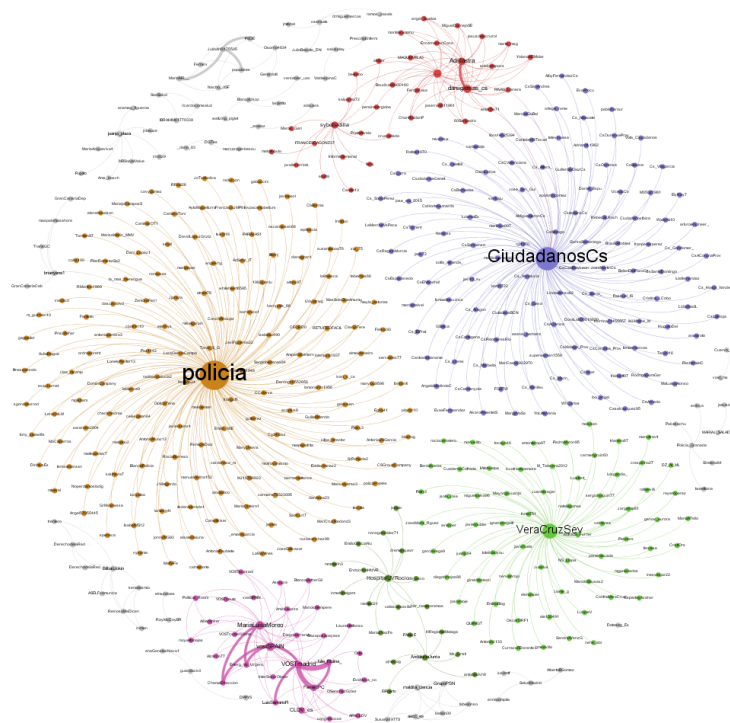


Figura 2. Grafo de los retuits de tuits que usaban el hashtag #StopBulos con el tamaño atendiendo al grado de entrada. *Nota:* en el grafo solo se muestran los nodos pertenecientes a aquellas comunidades que integraban más del 1% dentro de la red. *Fuente:* Elaboración Propia.

Posteriormente, se calcularon parámetros como la Densidad = .002, el Grado medio = 1.173, la Longitud media = 1.164, el Diámetro = 3, y la Modularidad = .083. En su conjunto, estos parámetros indicaron que el número de relaciones entre los usuarios de Twitter que conforman la red era bajo. Es decir, la tendencia de los usuarios era a realizar uno o dos retuits y casi siempre de un solo tuit. Asimismo, se usaron algunas medidas de centralidad para detectar los nodos con más influencia dentro de la red: el grado nodal de entrada, número de retuits que han recibido; el grado de intermediación (*betweeness*), que detecta los nodos que sirven de puente o, en otras palabras, el grado en que un nodo se encuentra a lo largo del camino más corto que conecta a otros en la red; y la centralidad de vector propio (*eigenvector*), influencia que tiene el nodo en la red. Todas las métricas para cada uno de los nodos que integran la red, junto a otras que no se han considerado relevantes para este estudio, se encuentran en el Material Suplementario.

Estas medidas de centralidad revelaron que los usuarios de los tuits con mayor grado nodal eran la Policía Nacional, la cuenta del partido político Ciudadanos y la cuenta de la cofradía Hermandad de la Vera Cruz. Atendiendo al grado de intermediación, VOST fue el usuario que mayor conexión dio a los diferentes tuits. Esto se debe a que ese usuario fue mencionado en tuits de otros usuarios y realizó retuits, por lo que conectaba diferentes comunidades. Finalmente, el usuario más influyente en toda la red (el que mayor centralidad de vector propio presentaba) fue la Policía Nacional.

Por último, con el objetivo de identificar distintas comunidades (o *clusters*) atendiendo a las interacciones entre los usuarios—los retuits—, se aplicó un algoritmo de modularidad que identifica diferentes grupos de nodos o comunidades atendiendo a la fuerza de sus relaciones. Como se puede observar en la Figura 2, aparecieron un total de seis comunidades, las cuales se encuentran diferenciadas por colores. Estas comunidades, con escasas conexiones entre ellos, reflejan el patrón de uso de este hashtag. Por lo general, un usuario publica un tuit y otros usuarios lo retuitean sin que los mismos usuarios retuiteen otros tuits con el mismo hashtag.

Evolución del uso del hashtag #StopBulos a lo largo del tiempo

Después de analizar las características de los mensajes y actores que han usado el hashtag #StopBulos en 2019, se hizo una evaluación de su uso en los dos años posteriores con el fin de identificar cambios y ver su evolución. Se analizó la frecuencia de uso del hashtag #StopBulos, los actores que lo utilizaron y los temas generales asociados. Para ello, se descargaron los tuits que utilizaron el hashtag #StopBulos durante una semana en

el año 2020 (N = 16869, desde el 12/03/2020 hasta el 20/03/2021) y una semana en el año 2021 (N = 421, desde el 12/05/2021 hasta el 19/05/2021).

En cuanto a la frecuencia de los mensajes, el uso del hashtag #StopBulos aumentó sustancialmente en el año 2020, con un crecimiento del 3007% con respecto al año 2019. Por otro lado, comparando 2021 con 2019, la frecuencia del 2021 es menor en un 22%. Además, se observa que la mayor movilización del hashtag se da a través de los retuits. En todos los años se observa que hay un número menor de mensajes originales (tuits), los cuales son amplificados a través de diferentes usuarios que comparten esta información en sus redes sociales. Así, los tuits originales representaron el 8% de todos los mensajes compartidos en 2019, el 10% de los mensajes compartidos en 2020 y el 16% de los mensajes compartidos en 2021.

Respecto a los temas asociados, se encuentra que el hashtag #StopBulos se usa de forma simultánea con otros hashtags que marcan las coyunturas sociales de cada momento. Así, en el año 2019, cuando el uso de este marcador era apenas incipiente (43 tuits originales), se observaba que era usado para marcar noticias falsas (#fakenews), eventos sociales y cuestiones relacionadas con la salud y la estética (#salud, #operacionbikini). En el año 2020, este marcador estuvo orientado principalmente a desmentir bulos sobre el coronavirus, el estado de alarma y la respuesta de las personas para cooperar (#estevirusloparamosunidos). Y en el año 2021 se mantiene la misma tendencia temática sobre el coronavirus y algunas medidas de protección, pero también aparece asociado a campañas de promoción de la igualdad (véase Figura 3).

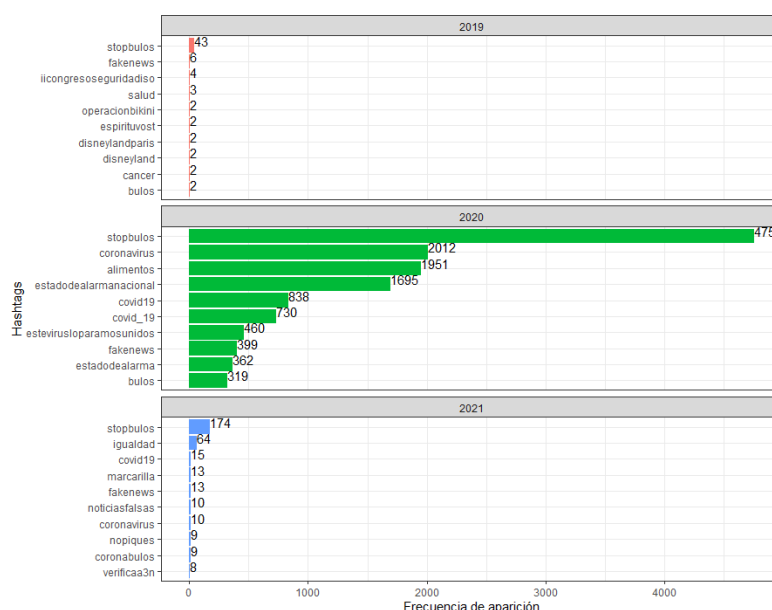


Figura 3. Hashtags asociados al hashtag #StopBulos en los tres años de observación. *Fuente.* Elaboración Propia.

Finalmente, en cuanto a los usuarios que usan el hashtag #StopBulos en sus mensajes originales, se encuentra que para el año 2020 y 2021, la mayoría fueron organizaciones civiles que procuran verificar información falsa (i.e., #vosteuskadi, #vostgalicia, #vostmadrid, #vostmurcia), medios de comunicación y miembros de la sociedad civil que usan el hashtag para señalar información falsa. En esta ocasión, la policía y las fuerzas de seguridad no tuvieron tanta presencia como en el año 2019 (véase Figura 4).

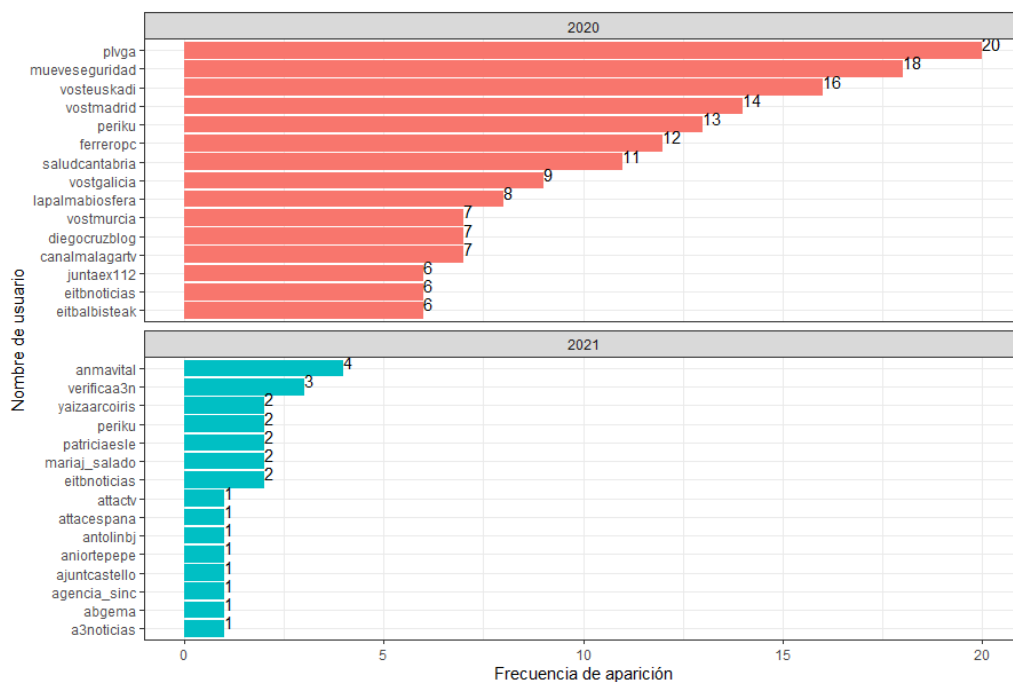


Figura 4. Frecuencia de aparición de los usuarios asociados al hashtag #StopBulos años 2020 y 2021.
Fuente. Elaboración Propia.

Discusión

En la presente investigación se propuso caracterizar el uso del hashtag #StopBulos en Twitter entendiéndolo como una estrategia de *fact-checking*. Atendiendo a este fin, se llevó a cabo un análisis de contenido, para identificar quiénes usan este mecanismo y para qué lo usan, y un análisis de redes a fin de encontrar qué usuarios tenían una mayor repercusión y quiénes promovían más el uso de estas estrategias.

En primer lugar, el análisis de contenido realizado para los tuits descargados en el 2019, se realizó agrupando los actores de los tuits en cuatro niveles: sociedad (ciudadanos y organizaciones privadas), organismos dedicados al *fact-checking*, grupos (entidades públicas, un partido político y una organización religiosa) y medios de comunicación. Asimismo, las temáticas de los tuits fueron diversas y se centraban en temas cotidianos

(e.g., salud y belleza). Respecto a las organizaciones que se incluían en los tuits, se encontraron diferentes herramientas y organismos dedicados al *fact-checking*. Finalmente, entre las funciones de dichos tuits destacaban desmentir *fake news* e informar sobre el *fact-checking* y, en menor medida, deslegitimar a personas o entidades.

En segundo lugar, el análisis de redes mostró la tendencia de los usuarios a realizar un retuit de los tuits con este hashtag; en otras palabras, las interacciones de segundo nivel eran mínimas, ya que cuando un usuario usa este hashtag, otros usuarios dan difusión retuiteándolo sin que existan otro tipo de interacciones. Por otro lado, los usuarios más influyentes fueron aquellos con más seguidores que se encuentran en la categoría grupos. Es de destacar que estos mismos usuarios se correspondían con instituciones con gran relevancia en la sociedad. Por su parte, las organizaciones de *fact-checking* actuaban como intermediadores conectando a los usuarios con los diferentes tuits.

En su conjunto, estos resultados parecen indicar que el uso del hashtag #StopBulos no está delimitado a ciertos perfiles de usuarios. Al contrario, la diversidad de usuarios y temáticas denota que cualquiera puede hacer uso de este mecanismo de *fact-checking*, lo cual está en línea con la perspectiva de la auto-comunicación de masas, según la cual cualquiera puede interactuar y crear noticias sin intermediarios y donde las redes de comunicación son horizontales (Castells 2006, 2007, 2009). No obstante, los resultados del análisis de redes ponen de manifiesto que, aunque cualquiera puede usar este mecanismo de *fact-checking*, solo aquellos usuarios con mayor número de seguidores alcanzan una gran difusión. Estos resultados coinciden con el estudio de Bakshy et al. (2011), en el que se encontró que la influencia en Twitter viene determinada por la influencia en el pasado y por el número de seguidores. Este resultado podría ser explicado porque cuantos más seguidores tiene un usuario, más confianza se le atribuye (Hofer y Aubert, 2013).

En el corpus de tuits analizados, solo una entidad pública, un partido político y una organización religiosa alcanzaron grandes cuotas al difundir su mensaje. Es de destacar que estas tres organizaciones representan el poder ejecutivo, el legislativo y la autoridad moral. Estos resultados coinciden con los que encontraron Vasoughi, Mohsenvand y Roy (2017), quienes comprobaron cómo ciertos organismos pueden controlar los rumores a través de Twitter. Por tanto, teniendo en cuenta los preceptos de la sociedad en red (Castells 2006, 2007, 2009), podemos afirmar que el *fact-checking* derivado del uso del hashtag #StopBulos, aunque tiene una ubicación abierta y pública,

está dominado por los poderosos; es decir, por quienes cuentan con redes más grandes y tienen la capacidad de hacer llegar su mensaje a más gente.

Por otra parte, nuestros resultados sugieren que el uso del hashtag #StopBulos está dedicado principalmente a desmentir noticias realmente falsas, lo que ayudaría a reducir la desinformación (Molina-Cañabate y Magallón-Rosa, 2020). En este sentido, los actores con mayor influencia en Twitter pueden jugar un papel relevante en la verificación de la información puesto que llegan a más usuarios. Incluso, esto podría contribuir a la mejora de los procesos democráticos al aumentar la transparencia y fomentar la deliberación de la ciudadanía, lo que se corresponde con el reciente concepto de ciber-democracia o, por el contrario, puede ser otra herramienta que utiliza la élite para aumentar sus ventajas democráticas (Aparicio Moreno, 2016).

Asimismo, el análisis de la evolución en el uso del hashtag #StopBulos evidencia su afianzamiento como mecanismo de *fact-checking*. En este sentido, la crisis sanitaria a nivel mundial, así como la proliferación de noticias falsas relacionadas con el virus del COVID-19 pudieron haber disparado significativamente el uso del hashtag #StopBulos para marcar los mensajes e información que en su momento difundían tratamientos milagrosos y teorías conspirativas sobre el origen del virus. Además, la descripción de los principales hashtags asociados a #StopBulos a través de estos tres años, muestra cómo este marcador se ha usado de forma progresiva a lo largo del tiempo y su uso es bastante sensible a las coyunturas sociales de cada momento. Por tanto, el uso del marcador #StopBulos para señalar noticias falsas ha sido apropiado por diferentes actores sociales, mostrando una forma de comunicación online que moviliza un discurso en torno a la verificación de la información, la confianza en las autoridades científicas y políticas, y la responsabilidad de producir y compartir información veraz que permita tomar mejores decisiones.

En este contexto, se hace evidente que este tipo de estrategias de *fact-checking* actúan como medidas al uso del poder y no como formas de contrapoder (Castells, 2007). Si bien estas estrategias facilitan a la población general mostrar su opinión respecto a contenidos que consideran *fake news*, si no cuentan con grandes redes de seguidores en estas redes virtuales no van a conseguir la difusión de dicha verificación. No obstante, los resultados sugieren que la principal función que se da a este hashtag es desmentir noticias realmente falsas en detrimento de deslegitimar a otros usuarios, lo cual presenta una doble cara. Por un lado, es positivo, ya que ayuda a reducir la desinformación por parte de entidades que cuentan con un alto grado de responsabilidad. Por otro lado, las

motivaciones de estos usuarios con gran cantidad de seguidores van a determinar, en gran medida, si se contribuye o se reduce la desinformación.

Brevemente, las redes sociales virtuales permiten una gestión de la información más horizontal como propone la auto-comunicación de masas (Castells, 2007). Sin embargo, en el caso del hashtag #StopBulos, las dinámicas parecen estar sujetas a las dinámicas de poder, de modo que son grupos que tradicionalmente cuentan con poder los que logran dar más difusión a sus mensajes. En consecuencia, la verificación de la información a través de este mecanismo y el control de la difusión de las *fake news* siguen sujetos, en cierta medida, a las dinámicas de poder que se daban en las redes verticales aunque dando paso a nuevos actores. En consecuencia, para que este mecanismo tenga una repercusión más equitativa por parte de los usuarios, es necesario crear previamente una cultura de *fact-checking* que capacite a dichos usuarios para contribuir y discutir las funciones de los tuits asociados a este mecanismo. En este sentido, se ha sugerido que el uso de herramientas de *fact-checking* por parte de los amigos, generan más atención y respuestas (Hannak et al., 2014).

Limitaciones y futuros estudios

A continuación, especificamos algunas de las limitaciones de las que el presente estudio adolece. En primer lugar, el limitado número de tuits descargados, los momentos específicos en los que se hacen las descargas y la imposibilidad de analizar toda la información que se genera a diario en las redes sociales no permiten generalizar, aunque este tampoco era el objetivo de este estudio. Tanto la muestra como la metodología usada nos permiten únicamente exponer esta investigación como algo contextual. Esto implica que no podemos asegurar que los tuits descargados en otra fecha mostrasen resultados similares, debido a la dependencia del contexto, el cual podría cambiar estos resultados de forma significativa.

En segundo lugar, aunque en este estudio se ha optado por indagar en el *fact-checking* mediante el uso del #StopBulos, este solo constituye una solución parcial a la propagación de las *fake news*. La mejora de esta estrategia no constituye una solución final a las *fake news* y otras formas, como el ya mencionado Código de Principios del IFCN, la propia verificación que realiza Twitter (Papanastasiou, 2020) o la detección de *fake news* basada en algoritmos (e.g., Atodiresei, Tănăsescu y Iftene, 2018), deben ser tenidos en cuenta en su conjunto a la hora de verificar la información y controlar las *fake news*.

Futuros estudios deberían profundizar en el contenido de los tuits que incluyen estas estrategias. En este sentido, analizar el lenguaje podría desentrañar los factores que contribuyen al alto impacto y difusión de contenido falso. A este respecto, Ott (2017) encontró que los tuits con un lenguaje simple, impulsivo e incívico eran preferido al hablar sobre ecología. Asimismo, estos resultados deberían acompañarse de otros estudios que profundizaran en los mecanismos psicológicos que llevan a las personas a aceptar y compartir este tipo de mensajes (e.g., factores individuales: Balmas, 2014; factores grupales: Jang y Kim, 2018). Además, para conocer el uso por parte de los actores con más influencia del hashtag #StopBulos, sería interesante un análisis de contenido que describiera el tipo de información que se desmiente y si estas responden al interés particular de los poderosos o si se usa de forma generalizada para reducir la desinformación. Así, en línea con la idea de Castells (2007), se podría discernir si la estrategia responde al ejercicio de poder o de contrapoder en redes.

Conclusiones

Esta investigación valida las ideas de la sociedad red en un contexto virtual, en concreto, Twitter. Además, plantea que algunas de las posibles estrategias de contrapoder usadas para detectar las *fake news* no sólo funcionan como tales, sino que siguen siendo estrategias al uso de los poderosos. Así, quienes controlan las redes controlan también la verificación de la información. A pesar de ello, parece que los usos que dieron al hashtag #StopBulos se centraban principalmente en desmentir *fake news* e informar sobre el *fact-checking*, lo cual estaría relacionado con la función con la que surge dicho hashtag.

Disponibilidad de datos depositados

Los datos de la presente investigación se encuentran depositados en la Open Science Framework - OSF.io en el link: <https://osf.io/dek57/>. Allí se puede acceder a:

1. Los scripts usados para descargar los tuits y los análisis de redes. Nombre del archivo: #StopBulos_analisis gephi_grado entrada.gephi
2. Las bases de datos para los análisis de redes. Nombre de los archivos: #StopBulos_bases datos tuits.xls y #StopBulos_Métricas análisis de redes.xls
3. Gráfico de redes. Nombre del archivo: #StopBulos_entrada_V2.png

Referencias

- Allcott, H., y Gentzkow, M. 2017. "Social media and fake news in the 2016 election". *Journal of Economic Perspectives* 31:211-236.
<https://doi.org/10.1257/jep.31.2.211>
- Aparicio Moreno, J. M. 2016. *Ciberdemocracia: El poder del discurso político-ideológico en Twitter en el contexto de la Comunidad Valenciana* (Doctoral dissertation, Universitat Politècnica de València).
- Atodiresei, C. S., Tănăselea, A., & Iftene, A. 2018. "Identifying fake news and fake users on Twitter". *Procedia Computer Science* 126:451-461.
<https://doi.org/10.1016/j.procs.2018.07.279>
- Bakir, V., & McStay, A. (2018). Fake news and the economy of emotions: Problems, causes, solutions. *Digital journalism*, 6(2), 154-175.
<https://doi.org/10.1080/21670811.2017.1345645>
- Bakshy, E., Hofman, J. M., Mason, W. A., y Watts, D. J. 2011, February. "Everyone's an influencer: Quantifying influence on twitter". Pp. 65-74 en *Proceedings of the fourth ACM international conference on Web search and data mining*. Hong Kong, China: Association for Computing Machinery.
<https://doi.org/10.1145/1935826.1935845>
- Balmas, M. 2014. "When fake news becomes real: Combined exposure to multiple news sources and political attitudes of inefficacy, alienation, and cynicism". *Communication Research* 41:430-454.
<https://doi.org/10.1177/0093650212453600>
- Barfar, A. 2019. "Cognitive and affective responses to political disinformation in Facebook". *Computers in Human Behavior* 101:173-179.
<https://doi.org/10.1016/j.chb.2019.07.026>
- Barreto-Galeano, M. I., Medina-Arboleda, I. F., Zambrano-Hernández, S., Sabucedo-Cameselle, J. M., Blanco-Abarca, A., y Maurice Lair, É. 2019. "Rhetoric, political ideology and the peace process in Colombia: A Twitter® Analysis". *Studies in Conflict & Terrorism*.
<https://doi.org/10.1080/1057610X.2019.1615256>
- Bastos, M. T., y Mercea, D. 2017. "The brexit botnet and user-generated hyperpartisan news". *Social Science Computer Review* 37:38-54.
<https://doi.org/10.1177/0894439317734157>
- Braun, K., y Dodge, J. 2018. "Critical policy studies and the politics of post-truth politics". *Critical Policy Studies*, 12(1):1-2.
<https://doi.org/10.1080/19460171.2018.1444498>
- Brummette, J., DiStaso, M., Vafeiadis, M., & Messner, M. 2018." Read all about it: The politicization of "fake news" on twitter". *Journalism & Mass Communication Quarterly*, 95(2):497-517.
<https://doi.org/10.1177/1077699018769906>
- Castells, M. 2006. "Informacionalismo, redes y sociedad red: Una propuesta teórica". Pp. 27-75 en *La sociedad red: Una visión global*, coordinado por M. Castells. Madrid: Alianza.
- Castells, M. 2007. "Communication, power and counter-power in the network society". *International Journal of Communication* 1:238-266.
- Castells, M. 2009. *Comunicación y poder*. Madrid: Alianza Editorial.
- Chen ZF., y Cheng, Y. 2020. "Consumer response to fake news about brands on social media: the effects of self-efficacy, media trust, and persuasion

- knowledge on brand trust". *J Prod Brand Manag* 29(2):188–198.
<https://doi.org/10.1108/JPBM-12-2018-2145>
- Diestel, R. 2010. *Graph Theory*. Berlin: Springer. <https://doi.org/10.1007/978-3-642-14279-6>
- Endsley, M. R. 2018. "Combating information attacks in the age of the internet: New challenges for cognitive engineering". *Human factors* 60:1081-1094.
<https://doi.org/10.1177/0018720818807357>
- Gelfert, A. 2018. "Fake news: A definition". *Informal Logic* 38:84-117.
<https://doi.org/10.22329/il.v38i1.5068>
- Gervais, B.T. (2015), "Incivility Online: Affective and Behavioral Reactions to Uncivil Political Posts in a Web-based Experiment", *Journal of Information Technology and Politics*, Routledge, Vol. 12 No. 2, pp. 167–185.
- Gentry, J. 2015. *Twitter: R based Twitter client*. R package.
- Glaser, B., y Strauss, A. 1967. *The discovery of Grounded Theory*. Chicago: Aldine.
- Grech, V. 2017. "Fake news and post-truth pronouncements in general and in early human development". *Early human development* 115:118-120.
<https://doi.org/10.1016/j.earlhumdev.2017.09.017>
- Hannak, A., Margolin, D., Keegan, B., y Weber, I. 2014, May. "Get back! you don't know me like that: The social mediation of fact checking interventions in twitter conversations". Pp. 187-196 en *Proceedings of the International AAAI Conference on Web and Social Media* (Vol. 8, No. 1). Michigan, USA: International AAAI Conference on Weblogs and Social Media (ICWSM). Consulta 26 de Mayo del 2021
(<https://ojs.aaai.org/index.php/ICWSM/article/view/14555>)
- Harsin, J. 2015. "Regimes of posttruth, postpolitics, and attention economies". *Communication, Culture & Critique* 8:327-333.
<https://doi.org/10.1111/cccr.12097>
- Harsin, J. 2018. "Post-truth populism: The French anti-gender theory movement and cross-cultural similarities". *Communication Culture & Critique* 11:35-52.
<https://doi.org/10.1093/ccc/tcx017>
- Himma-Kadakas, M. 2017. "Alternative facts and fake news entering journalistic content production cycle". *Cosmopolitan Civil Societies: An Interdisciplinary Journal* 9:25-40. <https://doi.org/10.5130/ccs.v9i2.5469>
- Hofer, M., y Aubert, V. 2013. "Perceived bridging and bonding social capital on Twitter: Differentiating between followers and followees". *Computers in Human Behavior* 29:2134-2142. https://doi.org/10.1007/978-3-642-12035-0_21
- IAB Spain (2018). *Estudio anual de redes sociales 2018*. Consulta 2 de Mayo del 2020 (https://iabspain.es/wp-content/uploads/estudio-redes-sociales-2018_vreducida.pdf) .
- Iosifidis, P., y Wheeler, M. 2018. "Modern political communication and web 2.0 in representative democracies". *Javnost-The public* 25:110-118.
<https://doi.org/10.1080/13183222.2018.1418962>
- Jang, S. M., y Kim, J. K. 2018. "Third person effects of fake news: Fake news regulation and media literacy interventions". *Computers in Human Behavior* 80:295-302. <https://doi.org/10.1016/j.chb.2017.11.034>
- Kearney, W. K. 2019. *rtweet: Collecting Twitter data*. R package.
- Mehta, D., Dwivedi, A., Patra, A., y Kumar, M. A. 2021. "A transformer-based architecture for fake news classification". *Social Network Analysis and Mining* 11(1):1-12. <https://doi.org/10.1007/s13278-021-00738-y>

- Molina-Cañabate, J. P., y Magallón-Rosa, R. 2020. “Desinformación y periodismo científico. El caso de Maldita Ciencia”. *Revista Mediterránea de Comunicación*.
- Munger, K. 2017. “Experimentally reducing partisan incivility on Twitter”. Working paper. Consultado 2 de Mayo del 2020 (kmunger.github.io/pdfs/jmp.pdf).
- Nelson, J. L. 2017. “Is ‘fake news’ a fake problem?” *Columbia Journalism Review* 31.
- Ott, B. L. 2017. “The age of Twitter: Donald J. Trump and the politics of debasement”. *Critical Studies in Media Communication* 34:59-68.
<https://doi.org/10.1080/15295036.2016.1266686>
- Papanastasiou, Y. 2020. “Fake news propagation and detection: A sequential model”. *Management Science*. <https://doi.org/10.1287/mnsc.2019.3295>
- Pedersen, S., y Burnett, S. 2018. “‘Citizen Curation’ in online Discussions of donald Trump’s presidency: Sharing the news on mumsnet”. *Digital Journalism* 6:545-562. <https://doi.org/10.1080/21670811.2017.1399806>
- Petty, R. E., y Cacioppo, J. T. 1984. “Source factors and The elaboration likelihood model of persuasion”. *Advances in Consumer Research* 11: 668-672.
[https://doi.org/10.1016/S0065-2601\(08\)60214-2](https://doi.org/10.1016/S0065-2601(08)60214-2)
- R Core Team. 2019. *R: A language and environment for statistical computing*. Viena, Austria: R Foundation for Statistical Computing.
- Rabin-Havt, A. (2016). *Lies, incorporated: The world of post-truth politics*. Anchor.
- Schifferes, S., Newman, N., Thurman, N., Corney, D., Göker, A., y Martin, C. 2014. “Identifying and verifying news through social media: Developing a user-centred tool for professional journalists”. *Digital Journalism* 2:406-418.
<https://doi.org/10.1080/21670811.2014.892747>
- Silverman, C. 2016. “Manual de verificación”. European Journalism Center. Consulta 2 de Mayo del 2020
(<http://verificationhandbook.com/downloads/manual.de.verificacion.pdf>).
- Starbird, K., Maddock, J., Orand, M., Achterman, P., y Mason, R. M. 2014. “Rumors, false flags, and digital vigilantes: Misinformation on Twitter after the 2013 Boston marathon bombing”. Pp. 654-662 en *iConference 2014 Proceedings*. <https://doi.org/10.9776/14308>
- Strauss, A., y Corbin, J. 2002. *Bases de la investigación cualitativa. Técnicas y procedimiento para desarrollar la teoría fundamentada*. Medellín: Editorial Universidad de Antioquia.
- Theocharis, Y., Barberá, P., Fazekas, Z., Popa, S.A. and Parnet, O. (2016), “A Bad Workman Blames His Tweets: The Consequences of Citizens’ Uncivil Twitter Use When Interacting With Party Candidates”, *Journal of Communication*, Vol. 66 No. 6, pp. 1007–1031.
- Tucker, J. A., Theocharis, Y., Roberts, M. E., y Barberá, P. 2017. “From liberation to turmoil: Social media and democracy”. *Journal of Democracy* 28:46-59. doi: 10.1353/jod.2017.0064
- Uscinski, J. E., y Butler, R. W. 2013. “The epistemology of fact checking”. *Critical Review* 25:162-180. <https://doi.org/10.1080/08913811.2013.843872>
- Vosoughi, S., Roy, D., y Aral, S. 2018. “The spread of true and false news online”. *Science* 359:1146-1151. doi: 10.1126/science.aap9559
- Van Duyn, E., y Collier, J. 2019. “Priming and fake news: The effects of elite discourse on evaluations of news media”. *Mass Communication & Society*, 22(1): 29–48. <https://doi.org/10.1080/15205436.2018.1511807>

- Walsh, M. J. B. 2010. "Librarians and controlling disinformation: Is multi-literacy instruction the answer?" *Library Review* 59:498-511.
<https://doi.org/10.1108/00242531011065091>
- Wardle, C. 2017. "Fake news. It's complicated". *First Draft News*, 16 de Febrero. Consulta 2 de Mayo del 2020 (<https://firstdraftnews.org/fake-news-complicated>).

Received: 2020-07-10

Accepted: 2021-09-28