

El *fact-checking* como estrategia global para contener la desinformación¹

Ignacio Blanco-Alfonso², María-Ángeles Chaparro-Domínguez³ y Rafael Repiso⁴

Recibido: 1 de marzo de 2021 / Aceptado: 2 de julio de 2021

Resumen. Las organizaciones de fact-checking están imponiendo unos sistemas de verificación que deben ser evaluados para diseñar futuras estrategias de contención de las fake news. El objetivo de este artículo es describir dichos sistemas para ahondar en la comprensión de los mensajes dudosos para la ciudadanía. Para ello, se han analizado 2.894 registros de la base de datos de Newtral correspondientes a 2018 y 2019 al completo. En concreto, se ha examinado el porcentaje de contenido falso y verdadero, las fuentes de difusión y de verificación, el tiempo empleado en la comprobación y los temas de los mensajes verificados. Se concluye que hay más mensajes fraudulentos que auténticos y los medios de comunicación son la principal fuente de difusión de mensajes dudosos que fueron verificados. Además, las fuentes gubernamentales son la primera fuente de verificación, prevalece la credibilidad sobre la inmediatez en el proceso de fact-checking y la política es el área temática que genera más bulos.

Palabras clave: Fact-checking; noticias falsas; verificación; Newtral; desinformación.

[en] Fact-checking as a global strategy to fight against disinformation

Abstract. Fact-checking services are establishing a system of verification that must be evaluated in order to design future strategies to stop the spread of fake news. The aim of this article is to describe that process to deepen the understanding of dubious messages for citizens. To this end, 2,894 records from 2018 and 2019 from the Newtral database have been analysed in full. More specifically, the percentage of true and false content, the sources of dissemination and verification, the time spent on verification, and the topics of the verified messages were examined. The main findings point to the fact that there are more messages that are fraudulent than authentic, and that the media are the main source for the dissemination of dubious messages that were verified. Moreover, the findings also show that governments are the main source of verification, credibility prevails over immediacy in the fact-checking process, and politics is the topic that generates the most hoaxes.

Keywords: Fact-checking; fake news; verification; Newtral; disinformation.

Sumario. 1. Introducción. Contexto y marco teórico. 1.1. La batalla mundial contra los bulos: el fact-checking. 1.2. Newtral, la startup española de fact-checking. 2. Objetivos, preguntas de investigación e hipótesis. 3. Metodología. 4. Resultados. 5. Discusión y conclusiones. 6. Referencias bibliográficas.

Cómo citar: Blanco-Alfonso, I., Chaparro-Domínguez, M.Á., & Repiso, R. (2021). El *fact-checking* como estrategia global para contener la desinformación. *Estudios sobre el Mensaje Periodístico* 27 (3), 779-791.

1. Introducción. Contexto y marco teórico

Las múltiples iniciativas institucionales de alcance internacional para combatir la propagación de las *fake news* son la muestra más evidente de la enorme amenaza social que este fenómeno ha adquirido en nuestro tiempo. Desde que en 2013 el Foro Económico Mundial (2013) alertara de que la propagación online de la desinformación sería uno de los diez grandes problemas que afrontaría el mundo a partir de 2014, se han producido actuaciones supranacionales de enorme envergadura que así lo atestiguan. Dos

ejemplos los encontramos, por un lado, en la creación en marzo de 2015 de la *EU East StratCom Task Force* en el seno del Servicio de Acción Exterior del Consejo Europeo (2015) para combatir las campañas de desinformación procedentes de Rusia y, por otro, en la convocatoria del *High Level Expert Group on Fake News and Online Disinformation* en enero de 2018 por parte de la Comisión Europea (2018) para establecer la estrategia comunitaria de lucha contra la desinformación.

Si bien la investigación disponible vincula claramente la proliferación de las *fake news* con los acon-

¹ Este artículo forma parte del proyecto “Nuevos escenarios de vulnerabilidad digital: alfabetización mediática para una sociedad inclusiva” (PRO-VULDIG-2), financiado por la Comunidad de Madrid y el Fondo Social Europeo (ref. H2019/HUM-5775), del que el primer autor es Investigador Principal.

² Universidad CEU San Pablo (España)
E-mail: iblanco@ceu.es

³ Universidad Complutense de Madrid
E-mail: ma.chaparro@ucm.es

⁴ Universidad Internacional de la Rioja (España)
E-mail: rafael.repiso@gmail.com

tecimientos electorales del Brexit y la victoria Donald Trump en 2016 (Blanco-Alfonso et al., 2019), no se puede ignorar que la irrupción del coronavirus ha producido una exarcebación de la investigación en este campo (Salaverría et al., 2020; Aguado-Guadalupe y Bernaola-Serrano, 2020; Pérez-Dasilva et al., 2020).

Ya en los primeros compases de la crisis sanitaria, en marzo de 2020, la entonces directora de la *International Fact-Checking Network* (IFCN), Cristina Tardáguila, alertaba de la “gigantesca demanda de información confiable” que estaban recibiendo las organizaciones de *fact-checking* (Tardáguila, 2020). El grupo Digilab de la Universidad Ramón Llull detectó en abril de 2020 que el 78% de los ciudadanos se informaba más que antes de la pandemia, y que el 52% consumía información tres o cuatro veces al día cuando antes del estado de alarma solo el 28% de la ciudadanía tenía este hábito de consumo (Masip et al., 2020, p. 5-6).

Este incremento de la demanada de información se desarrolló de forma paralela a la “crisis de infodemia” denunciada en septiembre de 2020 por la OMS, que alertaba de una “sobrereabundancia de información [...] y de intentos deliberados por difundir información errónea para socavar la respuesta de salud pública y promover otros intereses de determinados grupos o personas” (OMS, 2020).

Sin embargo, hay que recordar que, antes de la crisis de infodemia provocada por la alerta sanitaria, ya se había comenzado a describir el fenómeno de las *fake news* como una amenaza transversal y multidisciplinar, enmarcada en un contexto social de guerra híbrida o guerra informativa (Boyd, 2017; Bradshaw; Howard, 2017; Colom, 2018) y de posverdad, es decir, un contexto donde los hechos demostrados han dejado de ser relevantes en la conformación de la opinión pública (D’Ancona, 2017; Ball, 2017; McIntyre, 2018). De este modo, las *fake news* no son percibidas como un fenómeno informacional aislado o coyuntural provocado por la crisis del coronavirus, sino como “una experiencia colectiva de desconfianza” que ha favorecido la emergencia de una “cultura de las *fake news*” (Magallón, 2019, p. 23).

En un informe muy anterior al coronavirus auspiciado por la Comisión Europea, Wardle y Derakhshan (2017) alertaban de “una contaminación de la información a escala mundial” y de la existencia de una “compleja red de motivaciones para crear, difundir y consumir estos mensajes contaminados”, además de “una gran variedad de técnicas para amplificar el contenido de las noticias falsas”.

Estas primeras investigaciones ya apuntaban a la necesidad de deslindar el concepto de *fake news*, pues pronto se detectó la naturaleza polimórfica del fenómeno. “Uno de los errores más comunes desde el mundo del periodismo –apunta Magallón– es utilizar ese concepto como imán semántico para intentar describir procesos mediáticos, informativos, tecnológicos y sociales muy diferentes: la lucha contra la difusión de contenidos falsos en redes sociales y plataformas, la publicidad segmentada, la utilización de plataformas para operaciones de influencia e injerencia extranjera, la amplificación de discursos del odio y propagandísticos a través de tro-

les y bots, así como contenidos de clickbait que buscan optimizar el consumo de las redes sociales” (Magallón, 2020, p. 16).

Desde el punto de vista operativo, Wardle (2017) ha señalado la necesidad de descomponer el análisis de la desinformación en tres grandes apartados: los diferentes tipos de contenido creados y difundidos; las motivaciones de quienes los crean, y las formas en que son difundidos. Además, dado que la expresión *fake news* se ha revelado problemática, cualquier aproximación al fenómeno debe partir de la distinción de estos tres tipos de desórdenes informativos: *mis-information*: información falsa que se difunde sin la intención de causar daño; *dis-information*: información falsa que se difunde conscientemente para causar daño, y *mal-information*: información auténtica que, sin embargo, es utilizada para causar daño, o que, siendo privada, se hace pública para provocar algún perjuicio.

A partir de aquí, Wardle (2017) ha propuesto siete tipos de mensajes vinculados con alguno de estos desórdenes informativos: sátira o parodia, contenido engañoso, contenido impostor, contenido fabricado, conexión falsa, contexto falso y contenido manipulado. Esta tipología y otras similares (Ruchansky; Seo; Liu, 2017; Tandoc-Jr.; Lim; Ling, 2018) han desencadenado una línea de investigación centrada en el análisis del mensaje que está permitiendo conocer ciertas cualidades lingüísticas de las *fake news*, como su preferencia por palabras relacionadas con el sexo, la muerte, la aniedad y, en general, con un lenguaje emocional, frente a las noticias auténticas que presentan una mayor proporción de palabras relacionadas con los negocios y la economía (Shariatmadari, 2019; Torabi-Asr; Taboada, 2019).

En cuanto a la investigación sobre la forma en que son propagados los mensajes engañosos, observamos, en primera instancia, que el fenómeno de las *fake news* no es un problema de los periodistas porque sean ellos los principales promotores sociales de información; antes al contrario, cuando examinamos los tipos de mensajes fraudulentos que son creados y difundidos por la red y cuando analizamos las intenciones de sus creadores llegamos a una comprensión práctica del problema: el auge de la desinformación está directamente relacionado con el sistema de comunicación entre pares establecido por las redes sociales (Wardle, 2017), favorecido eventualmente por la pérdida de mediación social del periodismo profesional (Bell, 2016) y que, en último caso, ha impactado negativamente en la confianza general de los ciudadanos hacia toda la información que circula por la esfera pública.

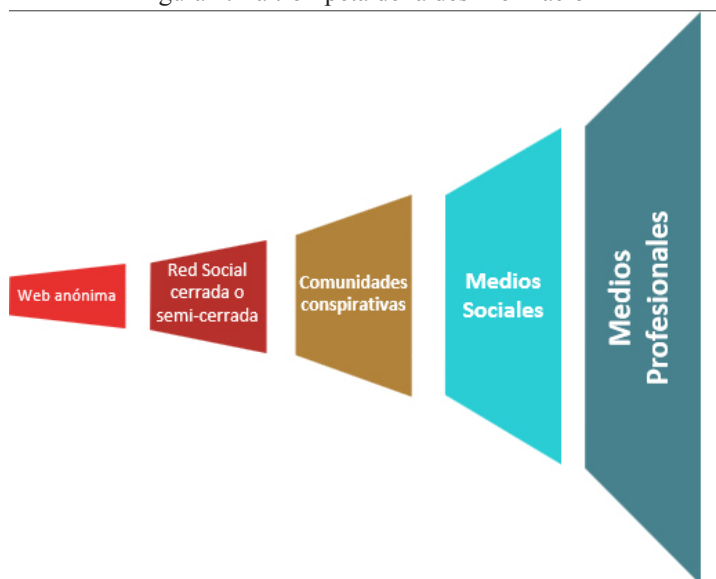
En relación con esta idea, cabe recordar que el último *Digital News Report* muestra que tan solo un 38% de los encuestados asegura confiar en las noticias, cuatro puntos menos de confianza respecto a 2019 (*Reuters Institute*, 2020). En el caso español, la confianza del público ha caído 15 puntos en 5 años, al pasar del 51% en 2017 al 36% en 2020 (Negredo et al., 2020). No obstante, cabe precisar que la confianza hacia las noticias difiere considerablemente cuando se observa el tipo de canal, pues “aquellos encuestados que declaran informarse a través de los medios más tradicionales

(televisión, radio o diarios impresos) se fían más de las noticias (45%) que los que optan por las ediciones digitales de los periódicos (39%), de las radios (30%), por los medios nativos digitales (29%) o las redes sociales (25%). Es decir, aquellos que se fían de las noticias, tienden a informarse más por medios tradicionales *–legacy media–* que por los nuevos canales y medios online” (Negredo et al. 2020).

Dicho esto, si bien los medios de comunicación tradicionales atesoran más confianza del público frente a los nativos digitales, es un hecho que pueden verse eventualmente implicados en la cadena de propagación de una información fraudulenta. En la esfera digital, los contenidos falsos se propagan siguiendo un esquema que se ha venido a denominar “trompeta de la desinformación” (Wardle, 2018) y que describe un movimiento

creciente de la difusión social del bulo (figura 1), también descrito como una “cascada de rumores” (Vosoughi; Roy; Aral, 2018). Este suele nacer en sitio web anónimo o minoritario, se mueve a grupos cerrados o semi-cerrados de Twitter, Facebook o WhatsApp; de ahí salta a comunidades de conspiración en foros, y luego a redes sociales abiertas como Twitter, Facebook e Instagram. Según Wardle (2018), llegado a este punto, “a menudo se traslada a los medios profesionales [...] cuando una información o contenido falso está incrustado en un artículo o citado en una historia sin la verificación adecuada [...] o cuando se publica una desacreditación o se expone la fuente principal de una conspiración”. Cualquiera que sea el caso, los promotores habrán conseguido su principal objetivo que es la máxima difusión del bulo una vez que llega a los medios de comunicación social.

Figura 1. La trompeta de la desinformación



Traducido de Wardle (2018).

El hecho de que los medios profesionales se conviertan eventualmente en amplificadores –acaso involuntarios– de información fraudulenta está en la raíz de la proliferación de los servicios de *fact-checking* (Cherubini y Graves, 2016). Esta nueva actividad profesional, que entronca con una larga tradición periodística de verificación (Vizoso y Vázquez–Herro, 2019), culmina con la creación de la IFCN, ubicada en el *Poynter Institute for Media Studies* (Florida, Estados Unidos), que nace en 2015 con la intención de dar soporte a las organizaciones asociadas y cuya actividad permite dibujar el contexto actual en el que se desarrolla el *fact-checking*.

1.1. La batalla mundial contra los bulos: el *fact-checking*

Los servicios de *fact-checking* han proiliferado por todo el mundo en los últimos años. La serie histórica de *Duke Reporters’ Lab*, que se remonta hasta 2014, muestra una curva en permanente crecimiento de proyectos censados, activos y repartidos por 102 países. La mayoría se sitúa en Europa (98) y Asia (89),

seguidos de Norteamérica (74), Suramérica (42), África (34) y Australia (5). En junio de 2021, *Duke Reporters’ Lab* tiene censados 342 servicios activos de *fact-checking*.

Dos factores principales explican la proliferación de este tipo de servicios de verificación: el aumento de la difusión de desinformación en plataformas como Facebook o Google, que están pidiendo ayuda a los *fact-checkers* para combatirlo, y un mayor número de alianzas entre medios y entidades de verificación en materia de *fact-checking*, especialmente durante las elecciones (Stencel y Luther, 2020). De esas 342 entidades, 87 son miembros activos de la IFCN, ya que han firmado un código de cinco principios que les compromete a ser imparciales, no partidistas, transparentes en cuanto a sus fuentes, financiación, organización y metodología de trabajo y a tener una política de correcciones abierta y honesta (IFCN, 2021a).

La IFCN llevó a cabo en 2019 una encuesta en la que participaron las entidades firmantes de su código de principios con el fin de elaborar una radiografía mundial del sector (IFCN, 2019). Según los resulta-

dos de la encuesta, aunque casi la mitad de las entidades no tiene ánimo de lucro, está creciendo el número de *fact-checkers* que sí que buscan un beneficio económico. Además, la mayor parte de las empresas tiene una plantilla de más de cinco trabajadores a jornada completa. En cuanto a su nacimiento, la proliferación de estas entidades se produjo entre 2015 y 2017.

España cuenta actualmente con cuatro entidades como integrantes activas de la IFCN: Newtral, que fue la primera española en formar parte de la red, en junio de 2017 (por aquel entonces con su programa de televisión El Objetivo, ya que no estaba creada la empresa actual); Maldita.es, que se integró en el 7 de junio de 2018; EFE Verifica, que fue aceptada el 16 de febrero de 2020, y Verificat, desde el 23 de noviembre de 2020 (IFCN, 2021b).

1.2. Newtral, la *startup* española de *fact-checking*

La presente investigación se ha realizado sobre la base de datos de la *startup* española Newtral, fundada en enero de 2018 por la periodista Ana Pastor. En realidad se trata de una *media startup* audiovisual que está a medio camino entre el periodismo y la tecnología. Sus tres áreas de negocio son: la verificación de datos y el *fact-checking*, la producción de programas de televisión y las nuevas narrativas en las redes sociales y la investigación periodística basada en la Inteligencia Artificial (Newtral, n.d.a).

Según su fundadora, que es la única accionista de la empresa, “nuestra aspiración es ser la referencia en verificación” y para lograrlo se han inspirado en el trabajo de otras organizaciones como Lupa, Chequeado o PolitiFact (Muela y Esplá, 2019). En Newtral trabajan 71 personas de perfiles diversos como periodistas, programadores, documentalistas o diseñadores. De todos ellos, en torno a una decena se dedica a la verificación de datos y el *fact-checking* (Newtral, n.d. b).

En cuanto a su modelo de negocio, según Pastor, una parte fundamental de sus ingresos se obtiene de la televisión, gracias a la emisión de sus producciones audiovisuales, como son los programas El Objetivo y Dónde estabas entonces, emitidos ambos en La Sexta. Los beneficios que obtienen los reinvierten en la empresa en I+D+i. Además, están en vías de autofinanciación de la parte digital y técnica mediante distintas fórmulas, como son los acuerdos con empresas (Muela y Esplá, 2019).

En lo que se refiere a su metodología como *fact-checkers*, como el resto de miembros de la IFCN, en Newtral siguen un protocolo de verificación estandarizado y su trabajo es auditado periódicamente. Verifican declaraciones de políticos de diferentes partidos y bulos que extraen de medios de comunicación y otras plataformas como las redes sociales. El criterio de selección que siguen es el interés periodístico. Una vez seleccionado el contenido que se va a verificar, lo contrastan y contextualizan con datos públicos y oficiales, expertos y departamentos de comunicación de los partidos o de otras instituciones. Tras concluir el

proceso de verificación, este tiene que ser validado por tres miembros del equipo de Newtral antes de ser publicado en la web (Ufarte-Ruiz et al., 2020).

Según Pastor (Muela y Esplá, 2019), en su trabajo sobre *fact-checking* han priorizado tres grandes fases desde que comenzaron a trabajar:

- Escucha: captan los bulos que les llegan a través de las redes sociales y WhatsApp.
- Captación: archivan toda la información del proceso de verificación de cada bulo en una gran base de datos sobre la que trabajar después.
- Detección: gracias a la Data Science Linguistic, es decir, la inteligencia artificial aplicada al lenguaje, buscan detectar ciertas mentiras de manera automática.

Sobre la selección de las verificaciones, también hay que señalar que Newtral se plantea una relación colaborativa con su audiencia, en el sentido de que invita a los usuarios a que alerten a los *fact-checkers* sobre nuevos bulos por medio de consultas a través del canal de WhatsApp. La propia plataforma indica en su sitio web que Newtral es “el primer servicio de verificación bajo demanda vía WhatsApp de España” y que “a través del número +34627280815 cualquier persona puede enviarnos mensajes sospechosos, bulos, fakes, enlaces, imágenes o vídeos que haya recibido y que quiera comprobar si son verdaderos o contienen información falsa. También recibimos consultas sobre declaraciones políticas. Nuestro equipo de verificación lo investigará y te contestará a través del mismo chat en WhatsApp” (n.d. c).

En marzo de 2019, Newtral se convirtió en uno de los *fact-checkers* elegidos por Facebook para combatir las noticias falsas en su plataforma. De este modo, la red social incorporó a España en su programa externo de verificación de datos, en el que entonces participaban otros 26 países. Además de Newtral, en el caso español Facebook también contrató los servicios de Maldita y la agencia France Presse (Pérez, 2019).

Un mes más tarde se lanzó en España el proyecto Comprobado, en el que Newtral participó junto con otros 15 medios, como Europa Press, RTVE, *Público* o *El Confidencial*. El objetivo del proyecto, coordinado por First Draft y Maldita, era acabar con la desinformación en la campaña de las elecciones generales del 28 de abril (First Draft, 2020).

2. Objetivos, preguntas de investigación e hipótesis

La experiencia acumulada por los servicios de *fact-checking* representa un ámbito de estudio ineludible para la actividad científica. En los últimos años, las instituciones públicas y privadas no han dejado de alentar y financiar el desarrollo de investigaciones básicas y aplicadas que permitan avanzar en el conocimiento disponible sobre la gran amenaza social que representa la desinformación para la estabilidad de

la democracia. Sirva como ejemplo el consorcio internacional Fandango (<https://fandango-project.eu>), proyecto financiado por la Unión Europea dentro del programa H2020 e integrado por entidades públicas y privadas de varias naciones.

En este sentido, el presente artículo pretende contribuir a esta línea de investigación observando ciertos parámetros relacionados con la difusión de bulos, con el objetivo de describir las características constitutivas del proceso de *fact-checking* y valorar su eficacia para desenmascararlos.

Partiendo de este objetivo, nos hemos formulado cuatro preguntas de investigación que promueven sendas hipótesis referidas tanto a los mensajes sometidos a verificación como a los procedimientos empleados por los servicios de *fact-checking*:

Sobre los mensajes sometidos a verificación:

P_1. ¿Qué porcentaje de los mensajes verificados por Newtral es finalmente identificado como contenido falso (bulo), verdad a medias o verídico?

H_1. A la vista de la creciente desconfianza de la ciudadanía hacia toda la información que circula por la esfera pública, es razonable pensar que existe un porcentaje elevado de mensajes auténticos de los que, sin embargo, desconfían los ciudadanos.

P_2. ¿Dónde fueron vistos y/o leídos los mensajes dudosos que los ciudadanos enviaron a Newtral?

H_2. Existe una correlación entre el creciente consumo de noticias a través de las redes sociales y el volumen de mensajes dudosos recibido por los usuarios de las redes.

P_3. ¿Qué proporción se da entre los temas prominentes de la agenda informativa y el número de mensajes dudosos recibidos en Newtral?

H_3. En la medida en que los contenidos políticos son prominentes en la agenda informativa española (López-Rabadán y Casero-Ripollés, 2012), cabe pensar en una prominencia similar de los bulos relacionados con la política.

Sobre el procedimiento de los servicios de *fact-checking*:

P_4. ¿Qué fuentes de verificación son utilizadas por los *fact-checkers* de Newtral para comprobar la veracidad de los mensajes recibidos de sus usuarios y cuánto tiempo se tardó en el resolver el proceso?

H_4. Se parte de la hipótesis de que el principio de credibilidad prima sobre el de inmediatez. En la medida en que la credibilidad de una información se basa en la fiabilidad de las fuentes consultadas, los servicios de *fact-checking* recurren preferentemente a fuentes primarias y/u oficiales para fundamentar sus desmentidos, a pesar de que ello pueda ralentizar el proceso y la publicación posterior.

3. Metodología

Para poder desarrollar esta investigación, se ha realizado un análisis de contenido cuantitativo a partir de la base de datos de Newtral, cuyo acceso se solicitó en 2020. El *dataset* inicial facilitado por esta organización contaba con 6.437 registros de consultas bajo demanda que abarcaban 24 meses, incluyendo 2018 y 2019 al completo.

Desde el punto de vista metodológico, se decidió prescindir de los datos de 2020, pues a partir de marzo, tras la irrupción del coronavirus, el número de mensajes dudosos relacionados con la pandemia se multiplicó por 16, lo que, en palabras del director de contenidos de Newtral, Joaquín Ortega, obligó a la empresa “a reorganizar todo el equipo de verificación y centrarlo en afrontar el volumen de solicitudes de gente que necesitaba certezas en uno de los momentos de mayor incertidumbre vividos en las últimas décadas” (Aguado-Guadalupe y Bernaola-Serrano, 2020, p. 290).

Para los fines de esta investigación, era fundamental disponer de un corpus estable y no hipertrofiado por un fenómeno tan absolutamente extraordinario como la pandemia sanitaria. Dado que el objetivo, como se ha explicado, consiste en describir las características constitutivas del proceso de *fact-checking* y su eficacia para desenmascarar los bulos, necesitábamos analizar diacrónicamente un corpus de mensajes suficientemente amplio y no sometido al estrés de la infodemia.

En la fase de limpieza de los datos hubo que eliminar los siguientes registros:

- Consultas sin pregunta y con una respuesta de Newtral que impide deducir la pregunta.
- Consultas sin respuesta de Newtral.
- Consultas que no son sobre posibles bulos.
- Consultas repetidas.

Tras esa depuración, el *dataset* definitivo lo conformó un corpus de 2.894 consultas (tabla 1).

Tabla 1. Información básica del dataset

Consultas registradas	2.894
Consultas con imágenes agregadas	936 (32%)
Usuarios únicos (autores de consultas)	1.788
Verificadores de <i>Newtral</i>	23
Errores en el <i>dataset</i>	175 consultas con errores de fechas

Elaboración propia

Posteriormente, se añadieron tres variables a las ya existentes en la matriz de Newtral, que eran: consulta de los usuarios, respuesta de Newtral, tema de la consulta, archivos adjuntos, palabras clave, fecha y hora de envío de la consulta, fecha y hora de envío de la respuesta de Newtral e identificación numérica del usuario que envía la consulta. Las tres nuevas variables incorporadas fueron:

- Resultado de la verificación: se establecieron cuatro posibles respuestas: bulo (cuando el contenido verificado es falso), no es bulo (cuando el contenido verificado es verídico), verdad a medias (cuando el contenido verificado tiene parte de verdad y parte de mentira) y no se sabe (cuando Newtral indica que no ha podido verificar el contenido porque las fuentes consultadas no le han respondido o porque la información no se puede verificar

al tratarse de un contenido de opinión o de una exclusiva con fuentes propias publicada por algún medio de comunicación).

- Fuente del contenido a verificar: es la fuente de origen, en la que el usuario ha visto, escuchado o leído un contenido dudoso que envía a Newtral para que lo verifique.
- Fuente principal de verificación: son las fuentes a las que acuden los verificadores de Newtral para saber si las consultas que les han enviado los usuarios son o no bulos.

Para llevar a cabo la categorización de los tipos de fuentes del contenido a verificar y de las fuentes consultadas por los *fact-checkers* para realizar la verificación, se siguió la clasificación de Segado-Boj, Chaparro-Domínguez y Díaz-Campo (2018), aunque adaptada al objeto de estudio de la investigación (tabla 2).

Tabla 2. Tipología de fuentes consultadas por los fact-checkers para la verificación

1.	Fuentes gubernamentales (en todos los niveles)
2.	Fuentes políticas (no gubernamentales)
3.	Fuerzas y cuerpos de seguridad del Estado y sanitarios
4.	Fuentes empresariales
5.	Fuentes académicas (profesores universitarios, científicos, revistas académicas, universidades, centros de investigación, libros, expertos)
6.	Asociaciones, ONG, colectivos, fundaciones, sindicatos, organizaciones
7.	Famosos (no políticos)
8.	Medios de comunicación y periodistas
9.	Medios satíricos
10.	Plataformas de bulos
11.	Fuentes judiciales y legislativas
12.	Cuenta de <i>Twitter</i> que no es ninguna de las fuentes anteriores
13.	Cuenta de <i>Facebook</i> que no es ninguna de las fuentes anteriores
14.	Redes sociales en general
15.	Mensaje de <i>Whastapp</i> sin fuente identificada
16.	Otras fuentes
17.	Ninguna fuente

Elaboración propia

Además de la incorporación de las variables anteriores, se llevó a cabo la recodificación de la variable “tema” facilitada por Newtral, para tratar de reducir la atomización de sus datos. Tras esta acción, los temas resultantes fueron los siguientes (por orden alfabético): alimentación, Cataluña, ciencia, consumo, cultura, deporte, economía, educación, empleo, empresas, famosos, feminismo, infancia, inmigración, internacional, leyes, medioambiente, monarquía, otros, política, religión, sanidad, seguridad, sucesos, tecnología, transportes, violencia de género y vivienda.

La limpieza de los datos del dataset, la incorporación de las tres nuevas variables y la recodificación de la variable “tema” fueron realizadas por uno de los autores del estudio, por lo que no fue necesario calcular ningún índice de fiabilidad interjueces.

Para la visualización de los datos, se han utilizado los softwares Excel, Tableau e Inkscape. Para los gráficos de temporalización del tiempo de respuesta en minutos no se han tenido en cuenta las consultas

que contenían algún tipo de error. El tiempo de verificación de cada consulta comienza cuando Newtral registra la entrada de la consulta y concluye cuando uno de los verificadores de Newtral envía el resultado de la verificación al usuario que envió la consulta.

4. Resultados

De las 2.894 consultas bajo demanda recibidas en Newtral, casi la mitad fueron clasificadas como bulos (41,2%), mientras que algo más de un tercio (32,5%) fueron identificados como contenido verdadero. El resto se corresponde con verdades a medias (12,89%) y contenidos que no se sabe si son ciertos o falsos (13,37%) (tabla 3) (P_1). Se confirma, por tanto, la hipótesis 1 de la investigación, que consideraba que existe un porcentaje elevado de mensajes auténticos de los que, sin embargo, desconfían los ciudadanos, pues un tercio de las consultas lo son.

La fuente de origen más frecuente de los mensajes dudosos enviados por los usuarios a Newtral fueron los medios de comunicación (840 registros, un 29,1% del corpus), seguida de Whatsapp (592 registros, 20,4% del corpus) y Twitter (484 registros, un 16,7% del corpus). En 267 registros (9,2%), los usuarios no identificaron la fuente de origen del mensaje (P_2). Se observa, por tanto, el alto número de mensajes dudosos que tienen como origen las redes sociales en sentido amplio (incluyendo WhatsApp, Twitter, Facebook y el resto de redes), lo que confirma la segunda hipótesis del estudio.

En cuanto al peso de los bulos en relación con las distintas fuentes de los mensajes sometidos a verificación, los contenidos falsos son mayoritarios en los medios satíricos (94,12%) y se sitúan en torno a la mitad de

los casos en los contenidos que provienen de las redes sociales en general (53,33%), Facebook (50%), WhatsApp (49,92%) o sin ninguna fuente clara (48,69%).

Con respecto al tipo de medio de comunicación que supuso el origen de la consulta, los medios tradicionales, con 262 casos, registraron 56 bulos y 130 contenidos verídicos, mientras que los medios online no tradicionales, con 565 casos, sumaron 195 bulos y 195 contenidos verídicos. Los periodistas fueron una fuente minoritaria de origen de los contenidos dudosos, con 14 casos.

Tabla 3. Distribución de mensajes verificados según fuente de origen y resultado de verificación

Fuente de origen	Resultado de la verificación				
	Total	Bulo	Verdad a medias	No es bulo	No se sabe
Medio satírico	17 (0,58%)	94,12%	0,00%	0,00%	5,88%
RR.SS. en general	15 (0,51%)	53,33%	6,67%	40,00%	0,00%
Facebook	166 (5,7%)	50,00%	12,05%	26,51%	11,45%
Whatsapp	592 (20,4%)	49,92%	13,37%	25,55%	11,17%
Ninguna fuente	267 (9,2%)	48,69%	9,74%	31,46%	10,11%
Twitter	484 (16,7%)	38,02%	12,40%	34,50%	15,08%
Políticos	215 (7,4%)	35,35%	17,21%	34,88%	12,56%
Gobierno	21 (0,72%)	33,33%	14,29%	42,86%	9,52%
Medio	840 (29,1%)	30,20%	14,27%	39,24%	16,29%
Asociación	21 (0,72%)	23,81%	14,29%	38,10%	23,81%
Otros	256 (8,8%)	52,73%	9,38%	26,17%	11,72%
TOTAL	2894	41,22%	12,89%	32,52%	13,37%

Elaboración propia

En cuanto a las fuentes principales consultadas por los *fact-checkers* de Newtral en el proceso de verificación (P_4), se han detectado 15 categorías. La más frecuente son las fuentes gubernamentales (29,1%), seguida de los medios de comunicación (19,2%) y los políticos que no pertenecen al gobierno (10,4%) (tabla 4).

Tras el proceso de verificación, se observa cómo los contenidos falsos son mayoritarios tras haber consultado medios satíricos, plataformas de bulos (100% en los dos casos), Facebook (83,3%), fuentes empresariales (57,8%) y Twitter (57,1%).

Tabla 4. Distribución de consultas según fuente principal de verificación y resultado

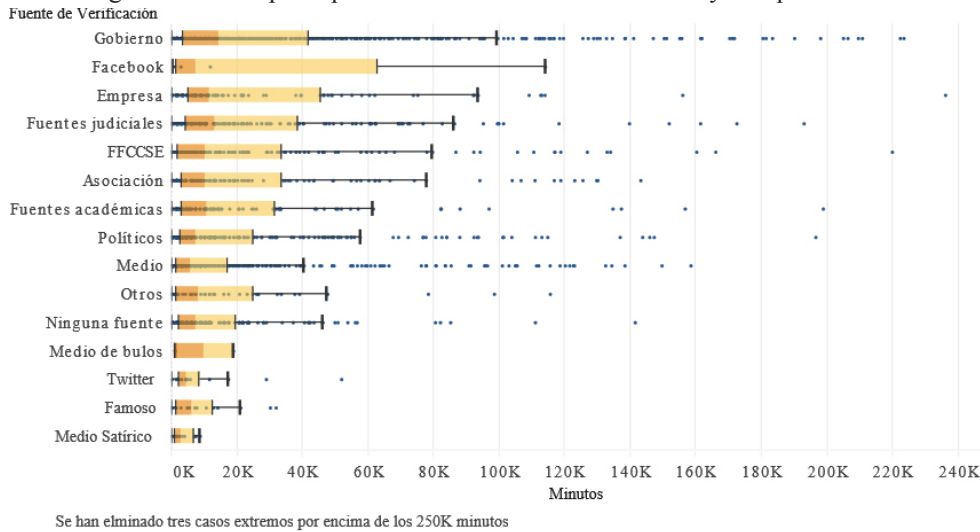
Fuente de verificación	Resultado de la verificación				
	Total	Bulo	Verdad a medias	No es bulo	No se sabe
Gobierno	842	47,39%	13,66%	27,43%	11,52%
Medio	558	34,83%	13,46%	42,55%	9,16%
Políticos	301	37,87%	11,96%	41,86%	8,31%
Fuentes judiciales	242	43,39%	16,12%	30,17%	10,33%
Ninguna fuente	203	24,14%	4,93%	25,12%	45,81%
FFCCSE	187	35,29%	13,37%	31,02%	20,32%
Asociación	170	34,12%	16,47%	37,65%	11,76%
Fuentes académicas	132	46,21%	13,64%	28,03%	12,12%
Empresa	116	57,76%	12,07%	21,55%	8,62%
Otros	69	40,58%	14,49%	30,43%	14,49%
Famoso	23	43,48%	13,04%	39,13%	4,35%
Medio satírico	22	100,00%	0,00%	0,00%	0,00%
Twitter	21	57,14%	0,00%	38,10%	4,76%
Facebook	6	83,33%	0,00%	16,67%	0,00%
Medio de bulos	2	100,00%	0,00%	0,00%	0,00%

Elaboración propia

Las fuentes de verificación que han permitido a los *fact-checkers* de Newtral obtener una verificación más rápida son los medios satíricos y Twitter, con una mediana de 2.858 y 4.458 minutos (2 y 3 días aproximadamente) empleados por cada consulta, respectivamente (figura 2). En el extremo opuesto se sitúan las fuentes académicas, judiciales y gu-

bernamentales, que poseen una mediana de 10.563, 13.154 y 14.252 minutos por consulta, respectivamente (P_4). Esto confirma la cuarta hipótesis de la investigación, al observar cómo los periodistas recurren sobre todo a fuentes primarias (gubernamentales, principalmente) aunque eso pueda ralentizar su trabajo.

Figura 2. Fuente principal de verificación de las consultas y tiempo de resolución



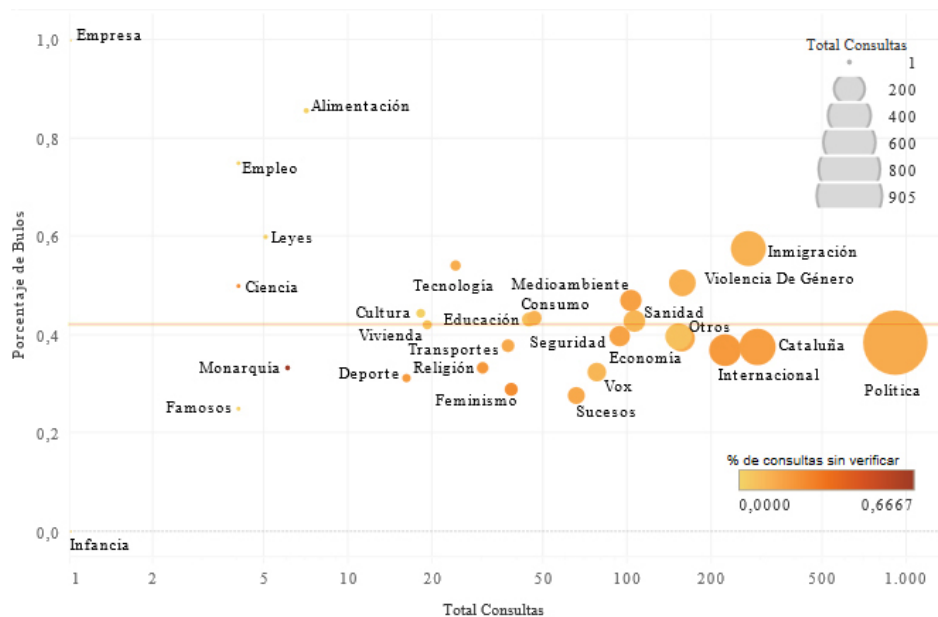
Elaboración propia

En la figura 3, se aprecian los temas con relación a tres variables: el promedio de bulos (eje Y), el número de consultas (eje X y tamaño de las burbujas) y el porcentaje de noticias sin verificar (color de las burbujas). En cuanto a los temas de las consultas enviadas a Newtral (P_3), casi un tercio de los registros (31%) trata sobre política, seguido de Cataluña (10%), inmigración (9,3%), asuntos internacionales (7,7%) y violencia de género (5,4%) (figura 3). Se

confirma, por tanto, la tercera hipótesis del estudio, que sostenía la prominencia de la temática política en las consultas enviadas a Newtral.

Los temas que mayor presencia de bulos registran son las consultas sobre empresas, alimentación y empleo, con un 100% (solo se registró una consulta), 85% y 75% de contenidos falsos, respectivamente. El promedio de bulos en todas las temáticas se sitúa en un 41% de las consultas remitidas a Newtral.

Figura 3. Distribución de consultas según áreas y porcentaje de bulos

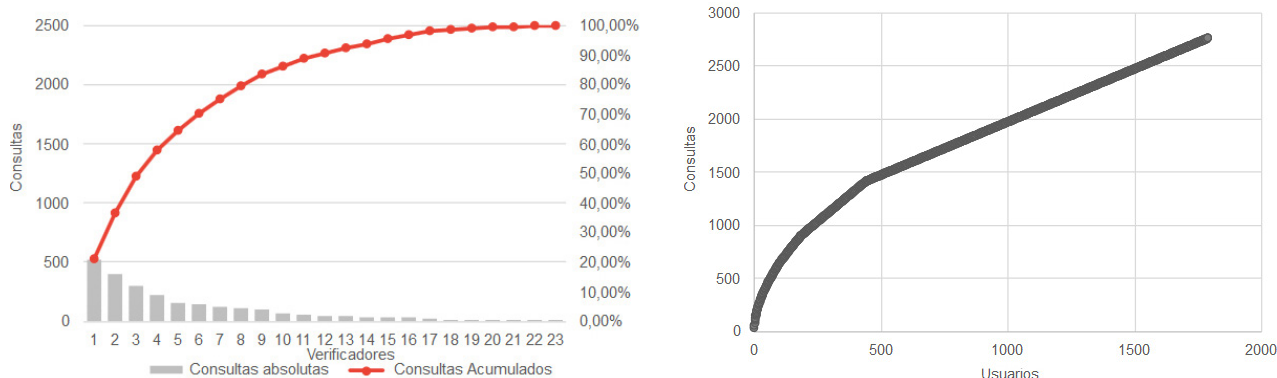


Elaboración propia

En lo que se refiere al trabajo de verificación de los periodistas de Newtral, de los 23 profesionales que durante el periodo de estudio desempeñaron su labor profesional en este ámbito, cinco de ellos resolvieron más de la mitad de las consultas realizadas (un 64,3%) (figura 4a).

Por otro lado, casi la mitad de las consultas (1.347 mensajes) fueron remitidas a Newtral por autores únicos, es decir, que solamente hicieron un único envío (figura 4b). Asimismo, 105 personas (un 5,8% de los autores) realizaron cuatro o más consultas (el 23,97% del total).

Figura 4. Distribución de consultas enviadas a Newtral según verificadores (figura A) y autores del envío (figura B)



Elaboración propia

5. Discusión y conclusiones

Los resultados de esta investigación permiten aportar conocimiento nuevo al fenómeno de las *fake news* y las estrategias de contención de la desinformación por parte de las instituciones de *fact-checking*. La principal contribución es, sin duda, la proyección de los resultados obtenidos sobre un universo tan extenso de contenidos verificados, al abarcar dos años completos (2018 y 2019) sobre un corpus de 2.894 mensajes, frente a anteriores estudios similares basados en muestras mucho más reducidas y en periodos de tiempo más cortos. Por lo tanto, se puede afirmar que los resultados obtenidos con esta investigación ofrecen una imagen cuantitativa y cualitativa mucho más precisa de la disponible hasta ahora sobre este fenómeno.

En cuanto a la clasificación de los mensajes verificados por Newtral en el periodo de estudio, se comprueba una preponderancia de contenidos falsos (41,22%) frente a contenidos verdaderos (32,52%). Este dato pone de relieve la desconfianza generalizada de los ciudadanos hacia la información que recibe, en el sentido de que los usuarios dudaron de la veracidad de 941 mensajes que, finalmente, resultaron ser verdaderos. Este resultado confirma la primera hipótesis de nuestra investigación, y permite evidenciar que la actual crisis de desinformación está provocando un importante aumento de la desconfianza de los ciudadanos hacia la información pública en general, en sintonía con los datos globales de desconfianza que periódicamente proporciona el Reuters Institute.

Por otra parte, de los resultados de nuestra investigación se puede inferir una correlación inquietante entre el número de contenidos dudosos y la confianza de los usuarios en los medios de comunicación. Dicha inferencia se fundamenta en el hecho de que del total de mensajes dudosos enviados a Newtral,

840 (el 29,1%) tenían como fuente original los medios de comunicación. Este dato confirma las tesis de la “trompeta de la desinformación” y del “efecto cascada” mencionados en el marco teórico, pues la mayoría de los mensajes dudosos llegaron a los usuarios a través de los medios de comunicación, que se convierten en el último eslabón de la cadena de transmisión de un número importante de bulos.

No obstante, en lo que atañe a nuestro estudio y en aras de una mejor comprensión de los datos, hemos segregado los medios de comunicación profesionales de los medios satíricos; y, dentro de los primeros, hemos distinguido entre los medios tradicionales (*legacy media*) y los medios nativos digitales, en línea con la metodología del Digital News Report, que ha demostrado diferencias a favor de la confianza del público hacia los medios tradicionales en detrimento de los medios online.

Efectivamente, al introducir este factor de discriminación, observamos diferencias significativas entre los medios tradicionales, que fueron la fuente de propagación de mensajes dudosos para la audiencia en un 31% de los casos, y los medios online, que lo fueron en el 67% de los casos, más del doble. Además, tras el proceso de verificación, el 21% de los mensajes dudosos transmitidos por los medios tradicionales se demostraron falsos, porcentaje que, sin embargo, alcanzó al 50% de los mensajes dudosos transmitidos por los medios online.

Estos datos evidencian que, desde el punto de vista científico, no se puede obviar la distinción entre los mensajes dudosos recibidos a través de los llamados “medios de referencia” (Imbert y Vidal-Beneyto, 1986; Labio-Bernal, 2008; Palau-Sampio, 2016), o a través de los medios sensacionalistas y los identificados como “medios de desinformación” especializados en distribuir *junk news* o noticias basura (Rodríguez-Fernández, 2020). Esta cuestión queda

planteada para futuras investigaciones que, sin duda, redundarán en beneficio de los medios periodísticos que realizan una labor rigurosa y profesional, y cuya credibilidad se puede ver eventualmente comprometida al aparecer clasificados junto con medios de dudosa calidad periodística. En este sentido, nuestra investigación viene a confirmar los resultados de otros estudios sobre, por un lado, la proliferación de *fake news* en contextos periodísticos profesionales (García-Galera et al., 2020; Del Hoyo-Hurtado et al., 2020) y, por otro, sobre la necesidad de aumentar las competencias de los periodistas en la lucha contra la desinformación (Rodríguez-Fernández, 2019).

En cuanto al papel de las redes sociales en la proliferación de contenidos falsos, la investigación confirma hallazgos previos que sitúan a WhatsApp como la red social privada a través de la cual los ciudadanos reciben más contenidos dudosos (Resende et al., 2019). En nuestra investigación, el 20,4% del corpus (592 mensajes) fueron vistos por los usuarios en WhatsApp, y de ellos, casi la mitad (49,92%) resultaron falsos.

Un porcentaje muy parecido de bulos se detectó en Facebook (50%) y algo menos en Twitter (38%), red que, sin embargo, se convirtió en la tercera fuente original de los mensajes sometidos a verificación. Si bien ambas tasas confirman nuestra segunda hipótesis, sería interesante replicar esta investigación con muestras de contenidos verificados a partir de 2020, año en el que, debido a la infodemia, ambas plataformas han implementado medidas específicas para contener la proliferación de la desinformación. Una futura comparación de estos resultados con los disponibles a partir de 2021 permitiría valorar la eficacia de dichas medidas.

También resulta interesante el dato de que cerca del 13% de los contenidos haya sido clasificado como verdad a medias, pues como se afirmó al comienzo del artículo, uno de los problemas de las *fake news* es la dificultad para deslindar la mentira de los mensajes que, basándose en información verdadera, producen contenidos oblicuos, descontextualizados y confusos para el ciudadano.

En cuanto a los temas dominantes, se confirma nuestra tercera hipótesis de que la política es el terreno que más contenidos dudosos produce entre los ciudadanos. En consonancia con estudios previos que demuestran la preponderancia de la política en la agenda periodística española (López-Rabadán y Casero-Ripollés, 2012, p. 488), nuestra investigación muestra que el 31,6% de las consultas fue sobre temas políticos, a los que cabría sumar el 10,1% de mensajes relacionados con Cataluña si se asume —a priori— que estos puedan pertenecer a la misma esfera temática.

Además, este resultado es consistente con el hallado por Bernal-Triviño y Clares-Gavilán en una investigación efectuada sobre los bulos desmentidos por Maldita.es en 2018. Según estas autoras, el tema que más bulos generó fue la política, con un 35% del total de casos analizados, seguido de la inmigración

y el racismo con un 12% (Bernal-Triviño y Clares-Gavilán, 2019, p. 4). Cabe añadir que también en nuestro estudio la inmigración se ha revelado, tras la política, como el tema que más consultas provoca entre los usuarios de Newtral.

Obviamente, la segregación de mensajes en categorías temáticas diferenciadas ofrece una perspectiva más variada y rica de la naturaleza de las *fake news* y puede servir para orientar futuras investigaciones específicas sobre la tematización de los bulos. Para una comprensión integral del fenómeno es necesario conocer qué temas destacan sobre los demás por el volumen de bulos que reciben, como la inmigración y la violencia de género. Sobre este último caso en particular, hay que subrayar que más de la mitad de las consultas analizadas (50,6%) desvelaron mensajes falsos.

Esto es relevante si se tiene en cuenta el cruce de resultados de nuestra investigación con los resultados obtenidos por Bernal-Triviño y Clares-Gavilán, cruce que muestra una decantación significativa de los bulos hacia contenidos con una evidente carga ideológica, lo que nos permite afirmar, con dichas autoras, que “los bulos existen porque generan un beneficio económico (clicks-publicidad), pero sobre todo por un interés ideológico cada vez más organizado” (Bernal-Triviño y Clares-Gavilán, 2019, p. 4).

Otro resultado de la investigación que merece discusión es el referido a las fuentes consultadas por los *fact-checkers* durante el proceso de verificación. Al cruzar las variables de “fuente de verificación” y “tiempo de resolución”, los datos confirman la cuarta hipótesis de que prima la credibilidad de la fuente sobre la inmediatez del proceso de verificación.

Las fuentes gubernamentales, que son las más consultadas por los *fact-checkers* para verificar la autenticidad de un mensaje, son las que a su vez necesitan más tiempo de resolución (14.252 minutos, es decir, 237,5 horas de promedio). También resulta llamativamente extenso el tiempo promedio empleado en la consulta de fuentes judiciales (219 horas) y las fuentes académicas (176 horas). Ahora bien, así como estos tres tipos de fuentes aportan credibilidad a la verificación, es indudable que el acceso a ellas y el tiempo de respuesta dilatarán el tiempo necesario para resolver los casos. En el polo opuesto encontramos la verificación de mensajes dudosos cuya fuente es un medio satírico; en estos casos, la verificación por parte del *fact-checker* se resolvió en menos de dos días de promedio (47,6 horas)

Por lo tanto, si establecemos una relación entre, por un lado, el contenido de los mensajes dudosos, cuyo tema prominente es la política, y, por otro, la fuente de difusión que, como hemos visto, mayoritariamente es un medio de comunicación, ambos factores explicarían que el Gobierno y los medios hayan sido las fuentes de verificación más recurrentes para los *fact-checkers* durante el periodo investigado.

Esta característica detectada en nuestro estudio permite concluir que el *fact-checking* bascula sobre una interdependencia entre fuentes oficiales (gu-

bernamentales, judiciales, etc.), medios de comunicación y verificadores. Los tres elementos de este ecosistema se necesitan mutuamente para contener la amenaza de la desinformación: los *fact-checkers*, como agentes independientes que tratan de desmentar la mentira; los medios de comunicación, como principal canal de difusión por el que los usuarios reciben mensajes dudosos, aunque muchas

veces sean víctimas de la suplantación fraudulenta de sus cabeceras o marcas; y las fuentes oficiales, en la medida en que de ellos depende la exactitud de los datos y hechos que circulan por la esfera pública. De la veracidad de la información dependen, nada menos, que la calidad de la democracia y el pleno ejercicio de los derechos civiles por parte de la ciudadanía.

Referencias

- Aguado-Guadalupe, G., & Bernaola-Serrano, I. (2020). Verificación en la infodemia de la Covid-19. El caso de Newtral. *Revista Latina de Comunicación Social*, 78, 289-308. <https://www.doi.org/10.4185/RLCS-2020-1478>
- Ball, J. (2017). *Post-Truth. How Bullshit Conquered the World*. Biteback Publishing Limited.
- Bell, E. (2016). Facebook is eating the world. *Columbia Journalism Review*, 7 marzo. <https://bit.ly/3xcRS3x>
- Bernal-Triviño, A., & Clares-Gavilán, J. (2019). Uso del móvil y las redes sociales como canales de verificación de fake news. El caso de Maldita.es. *El profesional de la información*, 28(3), e280312. <https://doi.org/10.3145/epi.2019.may.12>
- Blanco-Alfonso, I., García-Galera, M. C., & Tejedor-Calco, S. (2019). El impacto de las fake news en la investigación en Ciencias Sociales. Revisión bibliográfica sistematizada. *Historia y Comunicación Social*, 24(2), 449-469. <https://doi.org/10.5209/hics.66290>
- Boyd, D. (2017). The Information War has begun. *Apophenia*, 27 enero. <https://bit.ly/3w9Uu0V>
- Bradshaw, S., & Howard, P. R. (2017). *Troops, Trolls and Troublemakers: A Global Inventory of Organized Social Media Manipulation*. Computational Propaganda Research Project, Universidad de Oxford. <https://bit.ly/3w9UuxX>
- Cherubini, F., & Graves, L. (2016). *The rise of Fact-Checking sites in Europe*. Digital News Project 2016. Reuters Institute for the Study of Journalism. <https://bit.ly/3dvmVj9>
- Colom, G. (2018). Contextualizando la guerra híbrida. En: *Ciber Elcano. Informe mensual de ciberseguridad*, 32, enero, 5-9. <https://bit.ly/3dyKN5G>
- ColombiaCheck (2018). *Quiénes somos*. <https://bit.ly/3An8LuC>
- Comisión Europea (2018). *Final report of the High Level Expert Group on Fake News and Online Disinformation*. <https://bit.ly/3dwhaC8>
- Consejo Europeo (2015). *Questions and Answers about the East StratCom Task Force*. <https://bit.ly/363TznV>
- D'Ancona, M. (2017). *Post-Truth. The new war on truth and how to fight back*. Penguin Random House.
- Del Hoyo-Hurtado, M., García-Galera, M. C., & Blanco-Alfonso, I. (2020). Desinformación y erosión de la credibilidad periodística en el contexto de las noticias falsas. Estudio de caso. *Estudios sobre el Mensaje Periodístico*, 26(4), 1399-1409. <http://dx.doi.org/10.5209/esmp.70238>
- Duke Reporters' Lab (2020). *Fact-checking*. <https://bit.ly/3yinSDJ>
- First Draft (2020). *Comprobado: Preguntas frecuentes*. <https://bit.ly/3dwT6yI>
- Foro Económico Mundial (2013). *The rapid spread of misinformation online*. <https://bit.ly/366YrZO>
- García-Galera, M. C., Del Hoyo-Hurtado, M., & Blanco-Alfonso, I. (2020). Desinformación e intención comunicativa: una propuesta de clasificación de fake news producidas en entornos periodísticos profesionales. *Revista Mediterránea de Comunicación*, 11(2), 105-118. <https://www.doi.org/10.14198/MEDCOM2020.11.2.16>
- IFCN (2019). *State of the Fact-Checkers 2019*. <https://bit.ly/3ybZT8V>
- IFCN (2021a). *The commitments of the code of principles*. <https://bit.ly/3dyo5uw>
- IFCN (2021b). *Verified signatories of the IFCN code of principles*. <https://bit.ly/3yiMk0t>
- Imbert, G.; Vidal-Beneyto, J. (eds.) (1986). *El País o la referencia dominante*. Mitre.
- Labio-Bernal, A. (2008). Periodismo de entretenimiento: la trivialización de la prensa de referencia. *Estudios sobre el Mensaje Periodístico*, 14, 435-447. <https://bit.ly/3ybZXpb>
- López-Rabadán, P., & Casero-Ripollés, A. (2012). La evolución de la agenda mediática española (1980-2010). Un análisis longitudinal de la portada de la prensa de referencia. *Revista Latina de Comunicación Social*, 67, 470-493. <https://www.doi.org/10.4185/RLCS-2012-964>
- Lupa (2015). *O que é a Agência Lupa?* <https://bit.ly/3An90Wy>
- Magallón-Rosa, R. (2019). *Unfaking news. Cómo combatir la desinformación*. Pirámide.
- Magallón-Rosa, R. (2020). La nueva infonormalidad: no pienses en fake news, piensa en desinformación. *Cuadernos de periodistas*, 40, 12-21. <https://bit.ly/2ToPqbu>
- Maldita (2020). Maldita.es se une a LatamChequea, un esfuerzo de 22 medios de comunicación hispanoparlantes en 15 países para luchar contra la desinformación sobre coronavirus en Latinoamérica y España. *Maldita*, 2 abril. <https://bit.ly/3dRQ9cz>
- Masip, P. et al (2020). *El consumo de información durante el confinamiento por el coronavirus: medios, desinformación y memes*. Digilab. Grupo de investigación de la Universidad Ramón Llull. <https://bit.ly/2ToPyYw>
- McIntyre, L. (2018). *Post-truth*. Massachusetts Institute of Technology Press.

- Muela, C., & Esplá, D. (2019). Cazadores de fake news: Conocemos Newtral con Ana Pastor. *Xataka*, 10 febrero. <https://bit.ly/3qHaKWc>
- Negredo, S. et al. (2020). *Digital News Report España 2020*. DIGITAL UNAV, Center for Internet Studies and Digital Life, Universidad de Navarra. <https://bit.ly/3qGSrQP>
- Newtral (n.d.a). *Quiénes somos*. <https://bit.ly/3qDjTit>
- Newtral (n.d.b). *Equipo*. <https://bit.ly/3hbpUjm>
- Newtral (n.d.c). Servicio de verificación de WhatsApp. <https://bit.ly/3ho0IKx>
- OMS (2020). *Gestión de la infodemia sobre la COVID-19: Promover comportamientos saludables y mitigar los daños derivados de la información incorrecta y falsa*. Declaración. <https://bit.ly/3x74MQC>
- Palau-Sampio, D. (2016). Metamorfosis de la prensa de referencia en el contexto digital: clickbaity estrategias de tabloide en Elpais.com. *Communication & Society*, 29(2), 63-80. <https://doi.org/10.15581/003.29.2.63-79>
- Pérez, J. (2019). Facebook ficha a Newtral y Maldita para combatir las noticias falsas. *El País*, 12 marzo. <https://bit.ly/3h8KeSg>
- Pérez-Dasilva, J. A., Meso-Ayerdi, K., & Mendiguren-Galdospín, T. (2020). Fake news y coronavirus: detección de los principales actores y tendencias a través del análisis de las conversaciones en Twitter. *El profesional de la información*, 29(3), e290308. <https://doi.org/10.3145/epi.2020.may.08>
- Resende, G., Melo, P., Reis, J.C.S., Vasconcelos, M., Almeida, J. M., & Benevenuto, F. (2019). Analyzing Textual (Mis) Information Shared in WhatsApp Groups. En: *WebSci '19: Proceedings of the 10th ACM Conference on Web Science*, 225-234. <https://doi.org/10.1145/3292522.3326029>
- Reuters Institute (2020). *Digital News Report*. <https://bit.ly/3dziAvn>
- Rodríguez-Fernández, L. (2019). Desinformación: retos profesionales para el sector de la comunicación. *El Profesional de la Información*, 28(3), e280306. <https://doi.org/10.3145/epi.2019.may.06>
- Rodríguez-Fernández, L. (2020). Junk News y medios de desinformación en la campaña electoral del 10-N. *Revista Mediterránea de Comunicación*, 11(2), 71-83. <https://doi.org/10.14198/MEDCOM2020.11.2.19>
- Ruchansky, N., Seo, S., & Liu, Y. (2017). CSI: A Hybrid Deep Model for Fake News Detection. En: *CIKM '17 Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, pp. 797-806. <https://doi.org/10.1145/3132847.3132877>
- Salaverría, R., Buslón, N., López-Pan, F., León, B., López-Goñi, I., & Erviti, MC. (2020). Desinformación en tiempos de pandemia: tipología de los bulos sobre la Covid-19. *El Profesional de la Información*, 29(3), e290315. <https://doi.org/10.3145/epi.2020.may.15>
- Segado-Boj, F., Chaparro-Domínguez, M.A., & Díaz-Del Campo, J. (2018). Información científica en Argentina, España y México: fuentes, recursos multimedia y participación de los lectores en los diarios online. *Estudios sobre el Mensaje Periodístico*, 24(1), 397-412. <https://doi.org/10.5209/ESMP.59957>
- Shariatmadari, D. (2019). Could language be the key to detecting fake news? *The Guardian*, 2 septiembre. <https://bit.ly/3xdW1UV>
- Stencel, M., & Luther, J. (2020). Update: 237 fact-checkers in nearly 80 countries... and counting. *Duke Reporters' Lab*, 3 abril. <https://bit.ly/2UY3OYR>
- Tandoc-Jr., E.D., Lim, Z.W., & Ling, R. (2018). Defining Fake News. A typology of scholarly definitions. *Journal of Digital Journalism*, 6(2), 137-153. <https://doi.org/10.1080/21670811.2017.1360143>
- Tardáguila, C. (2020). The demand for COVID-19 facts on WhatsApp is skyrocketing. *Poynter*, 26 marzo. <https://bit.ly/3xaIFst>
- Torabi-Asr, F., & Taboada, M. (2019). Big Data and quality data for fake news and misinformation detection. *Big Data & Society*, 23 de mayo. <https://doi.org/10.1177/2053951719843310>
- Ufarte-Ruiz, M. J., Anzera, G., & Murcia-Verdú, F.J. (2020). Plataformas independientes de fact-checking en España e Italia. Características, organización y método. *Revista Mediterránea de Comunicación*, 11(2), 23-39. <https://www.doi.org/10.14198/MEDCOM2020.11.2.3>
- Vizoso, A., & Vázquez-Herrero, J. (2019). Fact-checking platforms in Spanish. Features, organisation and method. *Communication & Society*, 32(1), 127-142. <https://doi.org/10.15581/003.32.1.127-144>
- Vosougg, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146-1151. <https://www.doi.org/10.1126/science.aap9559>
- Wardle, C. (2017). Noticias falsas. Es complicado. *First Draft*, 14 marzo. <https://bit.ly/3y7OlnC>
- Wardle, C. (2018). 5 Lessons for Reporting in an Age of Disinformation. *Fisrt Draft Footnotes*, 28 diciembre. <https://bit.ly/3jMozBh>
- Wardle, C., & Derakhshan, H. (2017). *Information disorder: Toward an interdisciplinary framework for research and policy making*. Consejo de Europa. <https://bit.ly/3dwTMUM>

Ignacio Blanco-Alfonso

Catedrático de Periodismo de la Universidad CEU San Pablo, acreditado por ANECA y por ACAP, y con tres sexenios reconocidos por CNEAI. Director del Máster Universitario en Periodismo Cultural, del Máster en

Verificación digital, Fact-Checking y Periodismo de Datos (CEU & Newtral), y de *Doxa Comunicación*. Coordinador de Investigación y Posgrado de la Facultad de Humanidades y Ciencias de la Comunicación, e IP del proyecto “Nuevos escenarios de vulnerabilidad digital: alfabetización mediática para una sociedad inclusiva” (PROVULDIG-2. Ref. H2019/HUM5775, 2020-2022), financiado por la Comunidad de Madrid y el Fondo Social Europeo. Ha sido visiting scholar en las universidades de Génova (Italia) en 2007, del Pacífico (Lima, Perú) en 2015 y de Cambridge (Reino Unido) en 2017. ORCID: <https://orcid.org/0000-0002-2595-464X>

María-Ángeles Chaparro-Domínguez

Profesora Ayudante Doctora en la Universidad Complutense de Madrid. Cuenta con la acreditación de Profesora Titular de Universidad de ANECA y con un sexenio de investigación (2009-2014) reconocido por la CNEAI. Ha sido investigadora invitada en el centro SINTEF de Oslo (Noruega) y ha publicado una treintena de artículos en revistas como *Journalism Practice*, *Global Media and Communication*, *Young* o *Nordicom Review*. En los últimos años ha participado como investigadora en cuatro proyectos I+D+i financiados por el Ministerio de Ciencia e Innovación. ORCID: <https://orcid.org/0000-0001-7571-388X>

Rafael Repiso

Profesor Titular (Nivel 1) en la Universidad Internacional de La Rioja. Co-IP del Grupo PROCOMM y socio fundador de la SpinOff EC3metrics. ORCID: <https://rcid.org/0000-0002-2803-7505>