

Pandemia, politización y odio: características de la desinformación en España

David Blanco Herrero

University of Amsterdam  

Carlos Arcila Calderón

Universidad de Salamanca  

Margarita Tovar Torrealba

Universidad de Salamanca  

<https://dx.doi.org/10.5209/emp.96593>

Recibido: 20 de junio de 2024 / Aceptado: 4 de septiembre de 2024

ES Resumen. A pesar de los intentos previos de analizar y combatir la desinformación, siguen resultando escasas las iniciativas capaces de medir el fenómeno de manera amplia y sistemática. Una de las aproximaciones más productivas se ha basado en el estudio de los desmentidos realizados por agencias de verificación. Centrados en el caso español, se ha utilizado un instrumento previamente validado para realizar un análisis de contenido de 4245 desmentidos realizados por Maldito Bulo (perteneciente a la iniciativa Maldita.es, uno de los *fact-checkers* de referencia en España) entre 2017 y 2022. Se ha identificado la presencia predominante de temáticas sociales, aunque con la política como materia subyacente y transversal, reforzando el rol polarizador de la desinformación. Además, se observó que la mitad de los contenidos desmentidos podían utilizarse para fomentar el odio o el rechazo hacia determinados colectivos. Entre los formatos y plataformas de difusión, destacan los contenidos textuales y Twitter. Respecto a la evolución temporal, junto a una mayor diversidad de formatos y plataformas, se confirma la gran disrupción de la pandemia, tanto en temáticas como en cantidad de desmentidos, sobre todo en 2020. También se observa una mayor presencia de contenidos que pueden fomentar el odio en los dos últimos años de la muestra.

Palabras clave: Desinformación, discurso de odio, *fact-checking*, polarización política, redes sociales.

ENG Pandemic, politization and hatred: features of disinformation in Spain

Abstract. Despite previous attempts to analyze and combat disinformation, initiatives capable of measuring the phenomenon in a broad and systematic way remain scarce. One of the most productive approaches has been based on the study of de debunks published by the main fact-checking agencies. Focusing on the Spanish case, a quantitative content analysis has been carried out on 4,245 fact-checks made by Maldito Bulo (a branch of Maldita.es, one of the main fact-checking initiatives in Spain) between 2017 and 2022 using a previously validated instrument. The predominant presence of social themes has been identified, although with politics as an underlying and transversal theme in a large majority of content, reinforcing the polarizing role of misinformation. Furthermore, it was observed that half of the debunked content could be used to promote hatred or rejection towards certain groups. Among formats and platforms for the spread of hoaxes, text and Twitter are predominant. Regarding the temporal evolution, together with a greater diversity of formats and platforms, the great disruption that marked the pandemic is confirmed, both in topics as in the amount of debunks, especially in 2020. It was also observed a greater presence of content that could spread hatred in the last two years of the sample.

Keywords: Disinformation, hate speech, fact-checking, political polarization, social media.

Cómo citar: Blanco Herrero, D., Arcila Calderón, C. y Tovar Torrealba, M. (2024). Pandemia, politización y odio: características de la desinformación en España. *Estudios sobre el Mensaje Periodístico*, 30(3), 503-515. <https://dx.doi.org/10.5209/emp.96593>

1. Introducción

Han sido numerosos los estudios que han centrado su atención en el análisis de los contenidos falsos y de odio, abordando tanto su detección como el examen de sus atributos, dado que estas aproximaciones han representado los principales avances en el conocimiento de la materia. El análisis del discurso de odio ha sido objeto de una tradición más extensa en este sentido, y aunque el progreso ha sido más notable en el ámbito anglosajón, se han registrado avances significativos también en el contexto español. La mayoría de las investigaciones se han dedicado a analizar la presencia y características de determinados discursos de odio o rechazo en diversos medios. Paz *et al.* (2020) destacan en su revisión sistemática de la literatura que los temas más recurrentemente abordados son los discursos racistas o xenófobos, seguidos de la islamofobia y la homofobia. Además, observaron que los análisis recientes sobre discurso de odio tienden a enfocarse en las redes sociales, las plataformas más relevantes para la propagación del odio (Cammaerts, 2009). Añaden que los estudios sobre discurso de odio en el ámbito de la comunicación están predominantemente centrados en Internet y las redes sociales. En particular, las plataformas que reciben mayor atención a nivel internacional son Facebook (Farkas *et al.*, 2018) y, especialmente, Twitter (Arcila-Calderón *et al.*, 2022).

Si bien en el caso del odio existen mediciones sistemáticas más completas, la cuantificación del fenómeno de la desinformación aún resulta desafiante. Aunque algunos trabajos han intentado modelar la desinformación de manera extensa (González-Bailón *et al.*, 2023; Vosoughi *et al.*, 2018), este tipo de mediciones son escasas y a menudo se limitan al ámbito anglosajón. De hecho, Chen *et al.* (2023) señalan en su metainvestigación que la literatura previa sobre desinformación no ha abordado suficientemente visiones amplias que permitan comprender los factores que explican el problema en su totalidad. En el contexto español la tendencia más común ha sido el estudio de la información falsa generada en torno a eventos específicos: por ejemplo, Gutiérrez-Coba *et al.* (2020), Montesi (2020) y Salaverría *et al.* (2020) se enfocaron en la pandemia de la COVID-19, mientras que Cantón Correa y Galindo Calvo (2019) dirigieron su atención hacia un hecho mucho más concreto, como fue el incendio en la catedral de Notre Dame en abril de 2019.

Aunque algunos estudios ya han aportado conocimientos significativos en esta área, en España aún no existen investigaciones que hayan abordado el fenómeno de la desinformación en su diversidad temática y a lo largo de un periodo de tiempo extenso. Y aún menos que lo hayan hecho prestando atención a su interacción con los discursos de odio, cuya propagación a través de contenidos falsos agrega nuevas dimensiones al problema (Schwarz y Holnburger, 2019). Con la intención de llenar este vacío de conocimiento, este trabajo tiene como objetivo principal (OP) identificar las características de la desinformación diseminada en España entre 2017 y 2022 a partir de los desmentidos de una de las principales agencias de verificación, prestando especial atención a los contenidos falsos que puedan ser utilizados para difundir odio. Para ello se realizará un análisis de

contenido cuantitativo de 4245 desmentidos llevados a cabo por la iniciativa de *fact-checking* Maldito Buló entre 2017 y 2022.

2. Planteamiento de la investigación

Como se ha señalado, hasta donde alcanza nuestro conocimiento, no existen investigaciones que hayan abordado de manera integral y sistemática la comprensión de la desinformación propagada en España. Los esfuerzos principales se han centrado en áreas específicas en lugar de adoptar enfoques globales, lo que ha dificultado las comparaciones tanto en términos temáticos como temporales. Además, aunque algunos trabajos particulares han examinado cómo la desinformación puede ser utilizada para difundir odio (Narváez-Llinares y Pérez-Rufi, 2022), no se ha verificado si el odio es un fenómeno transversal o se limita a temas específicos.

Un aspecto crucial al analizar los fenómenos que nos ocupan es entender las temáticas que abordan. En relación con la desinformación en España, se han realizado trabajos centrados en la salud (Salaverría *et al.*, 2020), el cambio climático (Fernández-Castrillo y Magallón-Rosa, 2023) o la migración (Narváez-Llinares y Pérez-Rufi, 2022). Sin embargo, estos estudios no permiten evaluar si estas temáticas son predominantes en términos generales ni facilitan comparaciones. Trabajos como los de Almansa-Martínez *et al.* (2022) o Gamir-Ríos y Tarullo (2022), que emplean una metodología similar a la que se utilizará en esta investigación, sí lograron establecer comparaciones. No obstante, estos estudios se centran en un periodo específico, como fue la pandemia de la COVID-19, por lo que es necesario obtener una visión más amplia sobre la desinformación en España. Así, se plantea la primera pregunta de investigación:

PI1: ¿Cuáles son los temas predominantes en la información falsa difundida en España entre 2017 y 2022?

Por otra parte, comprender con qué formatos y a través de qué plataformas se difunde la desinformación es esencial para diseñar estrategias de lucha contra sus efectos. De hecho, la desinformación se ha venido estudiando eminentemente como problemática textual, pero aspectos como los *deep-fakes* y las imágenes generadas por Inteligencia Artificial (Yang *et al.*, 2019; Zhang *et al.*, 2019) están añadiendo nuevas capas al problema. A su vez, la investigación sobre el tema se ha centrado en redes como Facebook (Allcott y Gentzkow, 2017; Farkas *et al.*, 2018) o Twitter (Grinberg *et al.*, 2019), si bien otras como YouTube (Calvo *et al.*, 2022), WhatsApp (Moreno-Castro *et al.*, 2021), TikTok (Alonso-López *et al.*, 2021) o Telegram (Rogers, 2020) han ganado relevancia en los últimos años. Así, para poder evaluar y comparar los formatos y las plataformas más utilizadas para la difusión de desinformación, se plantea la siguiente pregunta de investigación:

PI2: ¿A través de qué formatos y plataformas se ha difundido la información falsa en España entre 2017 y 2022?

Las dos preguntas previas aluden al fenómeno de la desinformación en su conjunto. Sin embargo,

conviene prestar especial atención al uso que se hace de los contenidos falsos para promover narrativas de odio. A pesar de haberse mostrado que este uso es frecuente y preocupante (Molina Cañabate y Magallón Rosa, 2018; Narváez-Llinares y Pérez-Rufi, 2022), no se ha contrastado su dimensión real ni se conoce la proporción de mensajes falsos que difunden odio. También es importante identificar si existen características particulares que diferencien los contenidos falsos que difunden odio de aquellos que no lo hacen. Para esto se propone la siguiente pregunta de investigación, con dos subpreguntas:

PI3.1: ¿Qué proporción de la información falsa difundida en España entre 2017 y 2022 se puede usar para promover discursos de odio?

PI3.2: ¿Existen diferencias en la temática predominante y los formatos y plataformas de propagación entre la información falsa difundida en España entre 2017 y 2022 que puede promover discursos de odio y la que no?

Las preguntas anteriores ofrecen una visión general de la desinformación a lo largo de un periodo de seis años. Sin embargo, en este periodo se han producido hechos de gran relevancia informativa y, con ello, desinformativa. El ejemplo más claro puede ser la pandemia, que causó una disrupción y un flujo de desinformación sin precedentes (Almansa-Martínez *et al.*, 2022; Gamir-Ríos y Tarullo, 2022; Salaverría *et al.*, 2020), pero cabe también mencionar la importancia que ha tenido la popularización de nuevas plataformas, como TikTok (Alonso-López *et al.*, 2021). Además, eventos noticiosos de alto impacto, como las convocatorias electorales (Grinberg *et al.*, 2019), implican picos de desinformación que deben ser analizados. Para conocer la evolución temporal de la desinformación a lo largo de estos seis años se plantea la siguiente pregunta de investigación:

PI4: ¿Han experimentado la temática predominante los formatos y plataformas de propagación y la presencia de odio cambios temporales entre 2017 y 2022?

Estas preguntas conectan con el marco teórico de perspectivas clásicas. Así, comprender qué temas se tratan en los mensajes desinformativos es clave en la teoría de *agenda-setting* (McCombs y Shaw, 1972), mientras que abordar los enfoques que utilizan de forma más amplia —aunque no es una pregunta de investigación concreta, el análisis permite extraer lecturas en este sentido—, conecta con la teoría del *framing* (Goffman, 1974). La presencia de odio en la información falsa también entronca con las teorías del prejuicio (Allport, 1954) y con la del cultivo (Gerbner y Gross, 1976), pues el uso de la desinformación permite fomentar el miedo para propagar el odio. Serán estas bases teóricas las que se tengan en cuenta al dar respuesta a las anteriores preguntas de investigación en la sección de resultados.

3. Metodología

3.1. Justificación y muestra

Para el estudio de los desmentidos se siguió la técnica del análisis de contenido cuantitativo. Este análisis

no se realiza de manera directa, pues la cantidad de contenidos falsos existentes resulta cambiante e inabarcable. De hecho, Bruns *et al.* (2022) argumentan que medir la prevalencia real o percibida de la desinformación es difícil y los resultados dependen de los métodos utilizados y de otra serie de factores que no siempre pueden controlarse en la investigación. Así, la aproximación seguida utiliza la actividad de una de las principales plataformas de verificación en España, Maldita.es, como base para un estudio articulado y sistematizado que permita medir el fenómeno a través de los casos de desinformación más relevantes por peligrosidad y viralidad, criterios que utiliza para llevar a cabo sus verificaciones. Esta estrategia es la que han seguido numerosos trabajos. Los más semejantes, por su aproximación y metodología, son los de Almansa-Martínez *et al.* (2022) y Gamir-Ríos y Tarullo (2022). Otros ejemplos que también han partido de los desmentidos de agencias de verificación para abordar sus objetos de estudio y que se toman como referencia son los de Fernández-Castrillo y Magallón-Rosa (2023), Gutiérrez-Coba *et al.* (2020), Montesi (2020), Narváez-Llinares y Pérez-Rufi (2022) o Salaverría *et al.* (2020). También a nivel internacional existen ejemplos de esta aproximación metodológica (Shahi *et al.*, 2021).

Siguiendo esta línea, se optó por analizar todos los desmentidos realizados por Maldito Buló, la rama centrada en el desmentido de contenidos falsos de la Fundación Maldita.es, referente dentro de las agencias de verificación en España. El proceso de verificación, evaluado por la International Fact-Checking Network (IFCN), de cuyo Código de Principios es firmante, se basa en dos criterios: viralidad —y, con ello, propagación— y peligrosidad —riesgo para personas o convivencia—. Aunque se consideró la inclusión de Newtral en el estudio, tras la prueba piloto se descartó esta opción para poner el foco exclusivamente en Maldito Buló dada la frecuente repetición de contenidos en ambas plataformas, además de por la mayor cantidad de contenidos desmentidos por Maldito Buló desde su fundación y por su mayor nivel de detalle en los desmentidos.

Para el análisis se identificaron todos los desmentidos de Maldito Buló, desde el más antiguo, datado el 13 de enero de 2017, hasta llegar al 31 de diciembre de 2021. El corpus incluido en el estudio alcanzó los 5610 contenidos, si bien de estos deben excluirse, por no constituir unidades de análisis, publicaciones como reportajes o análisis, recopilaciones, publicaciones en idiomas distintos al español o contenidos repetidos. Se prescindió también de aquellos desmentidos centrados en timos, campañas de *phishing* o similares, puesto que no constituyen estrategias de desinformación, sino métodos que buscan estafar a personas u obtener un beneficio y que no guardan relación con las dimensiones comunicativas, periodísticas o sociales a las que se está haciendo referencia en esta investigación. También se obviaron los desmentidos del tipo «Qué sabemos», contenidos que no pueden calificarse como falsos y sobre los que se aporta contexto, entre los que se encuentran en ocasiones asuntos verdaderos o sobre los que no hay información suficiente. Con esto, se obtuvo una muestra compuesta por 3516 desmentidos de contenidos falsos.

Posteriormente, y obviando de antemano los contenidos con las características que se acaban de mencionar, se recogieron los desmentidos realizados por Maldito Bulo a lo largo de 2022, por lo que aumentó la muestra en 729 casos. Esta recogida adicional de datos permitió incorporar en el análisis más ejemplos, incluidos los relacionados con otro evento de gran peso desinformativo, como la invasión rusa de Ucrania. Con este añadido se obtuvo una muestra definitiva de 4245 casos.

3.2. Prueba piloto

Para el diseño metodológico se llevó a cabo, en primer lugar, una prueba piloto entre mayo y junio de 2021 con 1072 respuestas, incluyendo 115 doblemente codificadas para poder realizar las pruebas interjueces. Los desmentidos utilizados habían sido publicados por Maldito Bulo y por Newtral durante el año 2020.

El trabajo se realizó con codificadores previamente entrenados, a los que se facilitó un listado de desmentidos, el libro de códigos y la hoja de codificación. Para evaluar la fiabilidad de las variables se calcularon los valores α de Krippendorff con la macro KALPHA (Hayes y Krippendorff, 2007). El valor medio fue de 0,70. El cuestionario, así como la tabla con los valores α de todas las variables se puede encontrar como material complementario en el siguiente enlace: https://osf.io/sm287/?view_only=077caf4b09bc43f0b4a239c33680cac7

Durante este proceso se llevó a cabo también una validación externa por expertos, con dos personas además de los autores del trabajo. En conjunto, la evaluación de la calidad del instrumento fue considerada satisfactoria. No obstante, las observaciones realizadas por los expertos y la prueba interjueces permitieron corregir deficiencias y adaptar el proceso de entrenamiento y el instrumento (la herramienta definitiva está disponible en https://osf.io/sm287/?view_only=077caf4b09bc43f0b4a239c33680cac7; se puede observar que el libro de códigos incluye más variables de las que se detallan en este estudio). También se decidió utilizar exclusivamente Maldito Bulo, pues sus desmentidos aportaban más detalles sobre la información falsa, al tiempo que se describía también con más profundidad el proceso de verificación.

3.3. Procedimiento

Tras la prueba piloto, se procedió en primer lugar a la identificación y recogida de todos los contenidos recopilados en la web de Maldito Bulo. Se descargó cada una de las URLs para su posterior análisis. Antes de llevar a cabo la codificación se realizó un proceso de entrenamiento con el equipo de cuatro codificadores. La codificación definitiva se llevó a cabo entre enero y agosto de 2022. Tras esto, entre enero y marzo de 2023, se completó el análisis de los contenidos correspondientes al año 2022.

Para garantizar la fiabilidad de la codificación se realizó un nuevo análisis interjueces con una muestra de 274 casos (6,5 % de la muestra), incluyendo datos doblemente analizados de todos los años y de los cuatro codificadores. Los valores en este caso alcanzaron un α de Krippendorff medio

de 0,77, que puede considerarse adecuado al superar la barrera del 0,7 (Igartua, 2006; Neuendorf, 2002). En el material suplementario [https://osf.io/sm287/?view_only=077caf4b09bc43f0b4a239c33680cac7] se detallan los valores α de Krippendorff de todas las variables.

3.4. Medidas

El libro de códigos incluía mayor cantidad de variables de las que se detallan en este estudio. A continuación se enumeran únicamente aquellas presentes en esta investigación. Salvo en las que se indica lo contrario, cada una de las opciones se mide de modo independiente como variables dicotómicas, evaluado su presencia (1) o ausencia (0):

1. Fecha de publicación del desmentido.
2. Materia: Se podía indicar más de una, pero siempre que fuera un elemento central del contenido, no era suficiente con la inclusión de actores o elementos contextuales. Las opciones se extrajeron de una búsqueda exploratoria previa y fueron: Política; Social; Economía; Ciencia y tecnología; Salud; Otros temas (que deberían indicarse específicamente).
3. Formato: Se podían marcar todas las opciones que fueran pertinentes de las siguientes: Textual; Noticias o titulares de medios de comunicación; Cadenas de WhatsApp o de otras redes; Imagen o meme; Vídeo o audio; Entrevistas, ruedas de prensa u otras formas de declaraciones públicas; Otros.
4. Redes sociales en las que tuvo presencia un contenido, independientemente de si se trataba del origen o de una difusión posterior. Se podían marcar más de una de las siguientes opciones: El contenido no tiene presencia en redes sociales (marcar esta opción era incompatible con las demás); El contenido se distribuye en redes sociales, pero no se identifica ninguna concreta (marcar esta opción era incompatible con las demás); Facebook; Twitter; Instagram; YouTube; LinkedIn; Telegram; TikTok; WhatsApp; Otras (que deberían indicarse específicamente).
5. Presencia de odio: Variable dicotómica que busca identificar si el contenido pudiera fomentar el odio o el rechazo a algún colectivo. Siguiendo trabajos previos sobre rechazo (Arcila-Calderón *et al.*, 2020; Miró Llinares, 2016), no se entiende el fomento del odio únicamente como ataques frontales o directos, sino que se incluyen también alusiones más sutiles o la criminalización o ridiculización de grupos.

4. Resultados

4.1. Temáticas

Los contenidos desmentidos tienen como temática dominante la social, que incluye todo tipo de cuestiones de convivencia, como delincuencia o educación, y que se observa en 1707 contenidos (40,2 %). Le sigue la temática política, con 1410 casos (33,2 %), que fue solo identificada cuando los protagonistas de la información falsa eran personajes o partidos políticos o cuando trataban sobre aspectos relacionados con el parlamentarismo o las elecciones. Esta categoría, no obstante, resulta

transversal, pues aparece de forma subyacente en prácticamente la totalidad de contenidos. La salud estuvo presente como temática central en 877 casos (20,7 %), sobre todo a raíz de la pandemia de COVID-19. La economía fue la materia dominante en 220 casos (5,2 %), mientras que la ciencia y la tecnología lo fueron en 152 (el 3,6 %). Junto a estas, existen materias relacionadas con clima y catástrofes naturales, con cultura, con personajes públicos, con deporte con religión o con curiosidades en general, que se detectaron entre la categoría de Otros, que alcanzó 530 casos (el 12,5 %). Sobre estas categorías conviene apuntar que, aunque se buscaba la temática principal, lo habitual es que más de una estuviera presente y, aunque se intentó siempre primar solo una de ellas, en ocasiones se seleccionó más de una, lo que se traduce en un porcentaje total que supera el 100 %.

4.2. Formatos y plataformas de difusión

Centrados en el formato, 1619 (el 38,1 %) desmentidos aluden a contenidos textuales, 619 (14,6 %) incluyen noticias o titulares de apariencia periodística, 302 (7,1 %) son cadenas de WhatsApp o de otras plataformas de mensajería, 1212 (28,6 %) se difunden a través de imágenes o memes, 797 (18,8 %) tienen formato audiovisual, ya sea en forma de vídeo o de audio, y 22 (0,5 %) parten de contenidos originados en contextos ajenos a las redes y

los medios, como entrevistas y ruedas de prensa o creencias populares. Cabe señalar aquí que hay una parte minoritaria de los contenidos de los que no se conoce el formato original, pues no está especificado en el desmentido. Por otro lado, es importante añadir que es frecuente que los contenidos tengan más de un formato. El ejemplo más frecuente es el de un vídeo o imagen, manipulados o no, que se interpretan de manera desinformadora en un texto que los acompaña.

Por otro lado, si se evalúa la plataforma de difusión, se encuentra que la red social con mayor presencia de contenidos falsos desmentidos es Twitter, presente en 1867 casos (44,0 %). Le siguen a una distancia considerable Facebook, con 731 casos (17,2 %) y WhatsApp, con 599 (14,1 %). Por su parte, TikTok aparece en 80 casos (1,9 %), Telegram en 78 (1,8 %), YouTube en 74 (1,7 %), e Instagram en 51 (1,2 %). Por otro lado, en 505 desmentidos (11,9 %) no se mencionan redes sociales, lo que implica que no se tiene constancia de su difusión en redes, aunque no necesariamente que no se haya difundido en ellas, y en 825 casos más (19,4 %) no se mencionan redes concretas, pero sí se indica que se ha difundido a través de estas plataformas. A modo de ejemplo, la Figura 1 ilustra un desmentido codificado con materia sanitaria, formato audiovisual y textual, y difundido en Twitter y Facebook.

Figura 1. Ejemplo de desmentido analizado.



Fuente: Maldito Buló (<https://maldita.es/malditobulo/20221216/ceo-pfizer-farmaceutica-reducir-poblacion-mundial-montaje/>).

4.3. Presencia de odio

Ya en lo relativo a discursos de odio difundidos a través de noticias falsas, podemos responder a la PI3.1 indicando que 2033 contenidos (el 47,9 % de todos los contenidos falsos) se podrían utilizar para fomentar odio o rechazo. También resulta de interés comprender si existen diferencias entre aquellos que se podrían utilizar con este fin y aquellos en los que la desinformación no lleva asociado el rechazo a un grupo concreto, como se cuestionaba en la PI3.2.

Antes de profundizar en estas diferencias, conviene puntualizar que la presencia de odio en un contenido desmentido no se ha medido de forma

directa, es decir, no se mide si un mensaje tiene odio como variable lingüística, como hacen la mayoría de los detectores de odio. En su lugar se ha evaluado si un mensaje pudiera utilizarse para incitar al odio hacia un colectivo, incluso aunque no lo haga de forma directa o explícita. Así, se mide la posible utilización de un contenido como parte de narrativas de odio, sin evaluar la intención del mensaje, sino su posible utilización. La Figura 2 muestra un ejemplo de cómo, si bien el texto no incluye un lenguaje claramente peyorativo, sí que distribuye una narrativa falsa que podría usarse para promover el rechazo a un colectivo.

Figura 2. Ejemplo de un contenido falso que podría utilizarse para difundir narrativas de odio.



Fuente: Maldito Bulo (<https://maldita.es/malditobulo/20200820/no-un-magrebi-menor-de-edad-no-recibe-740-euros-los-menores-de-edad-no-pueden-optar-a-ayudas-publicas/>).

Con esto, las comparaciones entre los contenidos falsos que podrían difundir odio y los que no se resumen en la Tabla 1, donde se observa que los contenidos textuales y las imágenes parecen ser utilizados en mayor medida para trasladar rechazo u odio, mientras que las cadenas de WhatsApp u otros sistemas de mensajería tienen menor presencia de odio. También se aprecia que los contenidos que podrían fomentar el odio hacia otros colectivos se dan más frecuentemente cuando se abordan materias políticas o sociales, y menos cuando se trata de ciencia y tecnología o de salud. Por último, en los contenidos que han tenido presencia en Twitter es más frecuente el fomento del rechazo, mientras que en todas las demás plataformas, sobre todo WhatsApp, es más probable lo contrario.

Tabla 1. Diferencias entre los contenidos desmentidos que pueden fomentar el odio y los que no.

Red social	¿Puede fomentar odio?		t de Student
	Sí (N=2033)	No (N=2212)	
Formato Textual***	M=0,43; DT=0,495 (N=874)	M=0,34; DT=0,473 (N=745)	t(4171,751)=6,254, p<0,001, d=0,19
Formato Noticia o titular	M=0,14; DT=0,348 (N=286)	M=0,15; DT=0,358 (N=333)	t(4243)=-0,919, p=0,182, d=0,03
Formato Cadena de WhatsApp***	M=0,03; DT=0,173 (N=63)	M=0,11; DT=0,311 (N=239)	t(3523,646)=-10,087, p<0,001, d=0,30
Formato Imagen o meme**	M=0,31; DT=0,461 (N=623)	M=0,27; DT=0,442 (N=589)	t(4176,261)=2,892, p<0,01, d=0,10
Formato Vídeo o audio	M=0,19; DT=0,390 (N=381)	M=0,19; DT=0,391 (N=416)	t(4243)=-0,055, p=0,478, d=0,00
Formato Otra forma de declaraciones públicas	M=0,00; DT=0,066 (N=9)	M=0,01; DT=0,076 (N=13)	t(4243)=-0,657, p=0,256, d=0,02
Formato Otros**	M=0,00; DT=0,000 (N=0)	M=0,00; DT=0,060 (N=8)	t(2211,000)=-2,833, p<0,01, d=0,08
Materia Política***	M=0,48; DT=0,500 (N=980)	M=0,19; DT=0,396 (N=430)	t(3868,503)=20,668, p<0,001, d=0,64
Materia Social***	M=0,48; DT=0,500 (N=984)	M=0,33; DT=0,469 (N=723)	t(4152,655)=10,538, p<0,001, d=0,32
Materia Economía*	M=0,04; DT=0,207 (N=91)	M=0,06; DT=0,234 (N=129)	t(4236,023)=-2,001, p<0,05, d=0,06
Materia Ciencia y tecnología***	M=0,01; DT=0,101 (N=21)	M=0,06; DT=0,236 (N=131)	t(3049,704)=-8,893, p<0,001, d=0,27
Materia Salud***	M=0,09; DT=0,287 (N=184)	M=0,31; DT=0,464 (N=693)	t(3731,457)=-18,977, p<0,001, d=0,57
Materia Otra***	M=0,06; DT=0,244 (N=129)	M=0,18; DT=0,385 (N=401)	t(3777,109)=-12,003, p<0,001, d=0,36
Red No se mencionan redes sociales**	M=0,10; DT=0,303 (N=208)	M=0,13; DT=0,341 (N=297)	t(4238,309)=-3,232, p<0,01, d=0,10
Red No se mencionan redes sociales concretas**	M=0,18; DT=0,381 (N=358)	M=0,21; DT=0,408 (N=467)	t(4241,989)=-2,892, p<0,01, d=0,09
Red Facebook**	M=0,16; DT=0,362 (N=316)	M=0,19; DT=0,390 (N=415)	t(4242,595)=-2,785, p<0,01, d=0,09
Red Twitter***	M=0,54; DT=0,498 (N=1103)	M=0,35; DT=0,476 (N=764)	t(4171,521)=13,162, p<0,001, d=0,41
Red Instagram	M=0,01; DT=0,099 (N=20)	M=0,01; DT=0,118 (N=31)	t(4209,110)=-1,257, p=0,104, d=0,04

Red social	¿Puede fomentar odio?		t de Student
	Sí (N=2033)	No (N=2212)	
Red YouTube**	M=0,01; DT=0,108 (N=24)	M=0,02; DT=0,149 (N=50)	t(4032,943)=-2,722, p<0,01, d=0,08
Red LinkedIn	M=0,00; DT=0,00 (N=0)	M=0,00; DT=0,00 (N=0)	-
Red Telegram**	M=0,01; DT=0,115 (N=27)	M=0,02; DT=0,150 (N=51)	t(4105,966)=-2,397, p<0,01, d=0,07
Red TikTok	M=0,02; DT=0,128 (N=34)	M=0,02; DT=0,143 (N=46)	t(4240,871)=-0,979, p=0,164, d=0,03
Red WhatsApp***	M=0,09; DT=0,289 (N=187)	M=0,19; DT=0,389 (N=412)	t(4066,543)=-9,003, p<0,001, d=0,27
Red Otra	M=0,00; DT=0,066 (N=9)	M=0,01; DT=0,079 (N=14)	t(4243)=-0,843, p=0,200, d=0,03

+, p<0,1; *, p<0,05; **, p<0,01; ***, p<0,001.

Fuente: elaboración propia.

Por otro lado, también se realizaron tablas cruzadas para contrastar si existen diferencias en la presencia de odio entre los distintos años, observaron diferencias significativas de acuerdo al estadístico Chi-cuadrado de Pearson [$\chi^2(5, n=4245)=96,596, p<0,001$]. La Tabla 2 muestra las diferencias en la presencia de odio entre los distintos años, de manera

que en 2020 la presencia de contenidos que podrían fomentar el odio hacia ciertos grupos fue significativamente menor, en parte por la gran presencia de contenidos sobre la pandemia en ese año, que no suelen llevar asociado ese rechazo. Por su parte, en los años 2021 y 2022 sucede lo contrario, y la presencia de contenidos que fomentan el rechazo se acrecienta.

Tabla 2. Diferencias temporales en la presencia de contenidos falsos que podrían fomentar el odio.

	¿Podría utilizarse este contenido para fomentar el odio o el rechazo hacia algún colectivo?		
	Sí	No	
2017	Recuento	143	t(4171,751)=6,254, p<0,001, d=0,19
	Recuento esperado	151,8	t(4243)=-0,919, p=0,182, d=0,03
	Residuo corregido	-1	t(3523,646)=-10,087, p<0,001, d=0,30
2018	Recuento	161	t(4176,261)=2,892, p<0,01, d=0,10
	Recuento esperado	183,4	t(4243)=-0,055, p=0,478, d=0,00
	Residuo corregido	-2,4*	t(4243)=-0,657, p=0,256, d=0,02
2019	Recuento	376	t(2211,000)=-2,833, p<0,01, d=0,08
	Recuento esperado	355,4	t(3868,503)=20,668, p<0,001, d=0,64
	Residuo corregido	1,7	t(4152,655)=10,538, p<0,001, d=0,32
2020	Recuento	419	t(4236,023)=-2,001, p<0,05, d=0,06
	Recuento esperado	534,9	t(3049,704)=-8,893, p<0,001, d=0,27
	Residuo corregido	-8,1*	t(3731,457)=-18,977, p<0,001, d=0,57
2021	Recuento	510	t(3777,109)=-12,003, p<0,001, d=0,36
	Recuento esperado	452,6	t(4238,309)=-3,232, p<0,01, d=0,10
	Residuo corregido	4,2*	t(4241,989)=-2,892, p<0,01, d=0,09
2022	Recuento	424	t(4242,595)=-2,785, p<0,01, d=0,09
	Recuento esperado	354,9	t(4171,521)=13,162, p<0,001, d=0,41
	Residuo corregido	5,6*	t(4209,110)=-1,257, p=0,104, d=0,04

* Valores con diferencias estadísticamente significativas.

Fuente: elaboración propia.

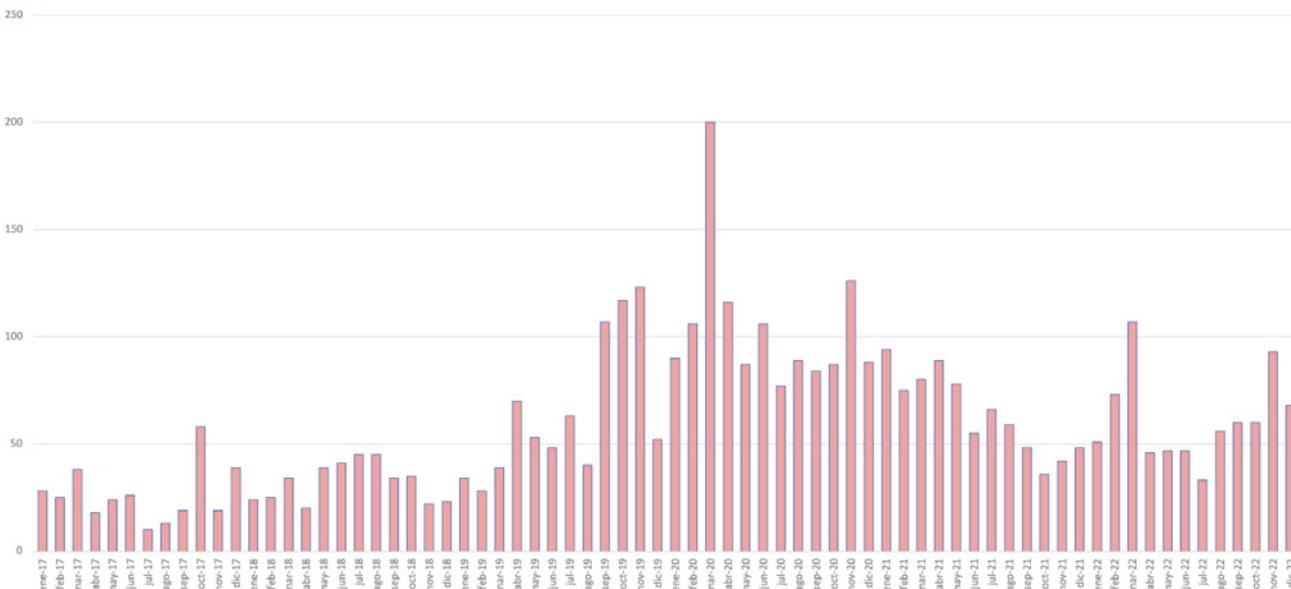
4.4. Evolución temporal

La desinformación es un proceso cambiante, frecuentemente relacionado con sucesos de actualidad. Es por eso que en la PI4 se planteaba cómo había evolucionado la desinformación en España entre 2017 —momento en el que comienza la actividad de Maldito Bulo— y 2022. La Figura 3 refleja el número de desmentidos cada uno de los meses presentes en el análisis, entre enero de 2017 y diciembre de 2022. De los 4245 desmentidos analizados, 317 se publicaron en 2017 (el 7,5 %), 383 en 2018 (9,0 %), 742 en 2019 (17,5 %), 1117 en 2020 (26,3 %), 945 en 2021 (22,3 %) y 741 corresponden a 2022 (17,5 %). La mayor actividad se sitúa en 2020, año de mayor impacto social de la pandemia y el confinamiento. En

concreto, se observa un claro pico de actividad en marzo de 2020, coincidiendo con la declaración de la COVID-19 como pandemia y el comienzo de las restricciones en España.

En general, se puede aludir a una cierta tendencia creciente durante los cuatro primeros años, que puede derivarse tanto del crecimiento propio de la actividad de una agencia que comienza a operar en enero de 2017, así como la creciente relevancia e interés por el fenómeno. El decrecimiento de los años 2021 y 2022 puede explicarse en parte por una diferente aproximación de Maldita.es, en parte diversificando su actividad, aunque también por la eliminación de contenidos repetidos durante el proceso de análisis:

Figura 3. Distribución temporal de los desmentidos analizados entre 2017 y 2022.



Fuente: elaboración propia.

Además, la realización de pruebas ANOVA de un factor muestra diferencias significativas en numerosas variables entre los seis años analizados. Los resultados de la prueba se pueden ver detalladas en el material suplementario [https://osf.io/sm287/?view_only=077caf4b09bc43f0b4a239c33680cac7].

En la mayoría de los casos existen diferencias estadísticamente significativas entre los distintos años. Esto puede deberse en parte al elevado volumen de casos, puesto que los tamaños del efecto η^2 son casi marginales en muchas ocasiones. La variable con unas diferencias más marcadas es la presencia de temas relacionados con la salud en los contenidos verificados, que alcanzó un máximo en 2020, cuando estuvo presente en el 39 % de los casos, decreciendo en 2021 al 28 % y en 2022 al 12 %, tras haber tenido valores mucho más reducidos los tres primeros años analizados. Esto también ayuda a explicar el hecho de que durante 2020 la presencia de contenidos que podrían utilizarse para promover el odio decreciera a 38 %, muy alejado del 57 % de 2022, y que se deriva del hecho de que los bulos sobre la pandemia tienen una menor carga de rechazo, pues abordan en mayor medida aspectos sanitarios:

datos sobre infectados y fallecidos, remedios, vacunas, origen del virus, etc.

Otra variable que llama la atención es el volumen de contenidos falsos con formato de noticia o titular, que va reduciéndose desde un 38 % en 2017 hasta un 8 % en 2022. En la misma línea, los medios o periodistas han pasado de ser responsables del 36 % de los contenidos desmentidos en 2017 a serlo de un 13 % en 2020 y 2022. Esto puede deberse a la propia actividad de verificación de Maldito Bulo, que ha dejado de centrarse en formatos y orígenes más tradicionales o periodísticos para prestar atención a otros más diversos y presentes en las redes sociales. En este sentido, se puede ver que el formato audiovisual ha crecido progresivamente, pues se localiza en el 9 % de los contenidos analizados en 2017 y en un 27 % de los desmentidos de 2022. De la misma forma, en 2017 (27 %) y 2018 (28 %) era más habitual encontrar contenidos que no habían sido difundidos en redes, mientras que eso solo sucedió en un 1 % y un 4 % de los casos en 2021 y 2022, respectivamente.

Esta diversificación de contenidos también se aprecia en dos redes sociales que han comenzado a ganar peso en fechas recientes. De hecho, hay

una ausencia total de casos en los primeros años en TikTok y Telegram, pero la primera estuvo presente ya en un 4 % de los desmentidos de 2021 y en un 5 % de 2022, mientras que Telegram alcanzó el 2 % en 2021 y el 8 % en 2022.

5. Discusión de resultados y conclusiones

La primera observación que debe realizarse es que la temática social es predominante en la mayoría de las informaciones desmentidas. Más allá de esto, es relevante también el gran peso que tiene la política en la desinformación. Así, aunque es el segundo bloque temático que más veces aparece de manera predominante, su presencia es transversal y subyace en la mayoría del resto de los contenidos. A ello hay que sumar la relevancia de algunos fenómenos concretos de gran calado y con gran circulación de desinformación en estos años: el conflicto catalán, las distintas convocatorias electorales en España y en otros países (sobre todo EE.UU.) e incluso la invasión rusa de Ucrania, que en numerosos casos incorporaba una fuerte dimensión ideológica. Esto coincide con los numerosos trabajos previos que han vinculado la desinformación con cuestiones de política, ya sea de ámbito nacional (Del-Fresno-García y Manfredi, 2018) o internacional (Bastos y Mercea, 2019). En total, es importante seguir reflexionando sobre esta dimensión política de la desinformación porque Vosoughi *et al.* (2018) encontraron que los contenidos falsos sobre este tema se difunden más. Además, esta materia es relevante porque puede tener un impacto directo sobre «el carácter y la calidad de nuestras democracias» (Van Aelst *et al.*, 2017, p. 4). Como ejemplo de esto, se observó que esa presencia subyacente y transversal de la política se hace presente en cómo las noticias falsas relacionadas con la COVID-19 son susceptibles de promover debates antisistema (Gutiérrez-Coba *et al.*, 2020).

Precisamente la pandemia ha tenido también un rol determinante en la desinformación que ha circulado en España en los últimos años. Es el tercer tema que más veces ha aparecido como predominante en los desmentidos analizados. Aunque no es una materia ajena a las narrativas desinformadoras (Montero-Liberona y Halpern, 2019), a raíz de la pandemia pasó a ser central y a marcar la agenda y el discurso durante muchos meses. Esto no resulta sorprendente, pues la conexión entre la pandemia y la desinformación ha sido ampliamente estudiada en estos años (Bechmann, 2020; Salaverría *et al.*, 2020), pero sí es importante porque la desinformación sobre la COVID-19 ha sido relacionada con otro tipo de narrativas falsas y de odio. Por ejemplo, se han encontrado vínculos entre discursos desinformadores sobre las vacunas y narrativas de extrema derecha (Calvo *et al.*, 2022). Como defienden estos autores,

la COVID-19 resulta clave en la investigación en comunicación sobre desinformación, no solo por el papel central de esta pandemia en los últimos tiempos, sino por la trascendencia que las informaciones falsas pueden tener para el aumento del escepticismo, la desconfianza en los sistemas políticos y el fomento de la división social. (2022, p. 226)

Con respecto a los formatos y plataformas más habituales, predominan el formato textual y Twitter, respectivamente. Es frecuente que se combinen más de un formato, sobre todo cuando el texto complementa una imagen o vídeo. Esto tiene especial relevancia, porque cuando la fotografía o el vídeo acompañan a un noticia falsa, la complementan, ofreciendo la ilusión de poder comprobar los hechos, ya que se entiende que estos contenidos audiovisuales son más factuales —algo que las nuevas técnicas de IA generativa están desmintiendo—, haciendo que las historias falsas sean más difíciles de contrarrestar (Weidner *et al.*, 2020). Es decir, cuando una información está acompañada por elementos visuales, como fotos y vídeos, la credibilidad del mensaje aumenta (Lin *et al.*, 2012).

Sobre las plataformas, es cierto que Twitter ha recibido una gran cantidad de atención mediática y académica como foco de desinformación (Grinberg *et al.*, 2019), algo que podría estar acentuándose en estos momentos, tras la eliminación de algunos controles y de la moderación que se realizaba en la plataforma desde su adquisición por Elon Musk (Hickey *et al.*, 2023). Sin embargo, este predominio de Twitter también puede deberse a la propia actividad de la agencia de verificación, que en las imágenes que ilustran el desmentido es más propicia a poner capturas de textos de Twitter que de otras redes, lo que garantiza la presencia del contenido falso en esta red, lo que no significa que al mismo tiempo el bulo no esté difundiéndose en otras plataformas. En cualquier caso, el predominio de Twitter puede deberse también a la importancia de esta red social para el sector periodístico (Bruns y Burgess, 2012), sus atributos y capacidad de diseminación de contenidos (Ausserhofer y Maireder, 2013; Moragas-Fernández *et al.*, 2019) o su capacidad para difundir mensajes con odio (Miró Llinares, 2016).

Por otra parte, una de las principales observaciones de esta investigación confirma la conexión entre desinformación y discurso de odio. Prácticamente la mitad de las informaciones falsas desmentidas pueden usarse para promover el rechazo, algo que refuerza la frecuente y preocupante combinación de ambos fenómenos (Narváez-Llinares y Pérez-Rufi, 2022; Schwarz y Holnburger, 2019).

El análisis más pormenorizado de los contenidos que podrían utilizarse para fomentar el rechazo ha permitido observar que estos mensajes utilizan en mayor medida formatos textuales o visuales, que se dan en mayor medida en Twitter, y que la temática social y política son las más habituales. Sobre la política ya se ha indicado su transversalidad y la presencia de rechazo político incluso cuando la materia central de la información falsa no está directamente relacionada con el tema. Sobre la materia social hay que señalar que aquí se incluyen cuestiones de convivencia o delincuencia, de ahí que haya una fuerte asociación entre esto y la abundancia de discurso de odio. Ello se debe a que gran parte de las narrativas de rechazo, sobre todo a personas migrantes y refugiadas, se dan en conexión con marcos de criminalidad, terrorismo o carga social, algo que ya se había observado en trabajos pasados que han modelado las características del discurso de odio hacia migrantes y refugiados (Amores, 2022). Esto, que conecta con las teorías del *framing* y del cultivo, podrá

ser desarrollado en futuros análisis, que podrán profundizar precisamente en estos marcos para comprobar de manera específica las características de las narrativas de rechazo que se apoyan en información falsa, así como sus efectos en el largo plazo.

Por otro lado, aunque ya se ha aludido a la pandemia y a la cuestión sanitaria como la tercera materia más presente, hay que destacar la disrupción que supuso la crisis del coronavirus, que dio lugar al pico más abultado de desmentidos en marzo de 2020 y a un incremento de la actividad en ese año. Además, muchas de las diferencias temporales observadas se derivan en parte de esta actividad, que marcó los años 2020 y, en menor medida, 2021. Más allá de los efectos derivados de la pandemia, es posible encontrar otros patrones temporales de interés, como la creciente atención por redes como Telegram o TikTok o por formatos audiovisuales, en detrimento de los formatos puramente textuales y periodísticos.

De manera más general, podemos indicar que el análisis de contenido comparte varias lecturas con algunos trabajos previos (Almansa-Martínez *et al.*, 2022; Gamir-Ríos y Tarullo, 2022). El principal atractivo de la investigación no ha sido lo sorprendente de sus descubrimientos, sino su capacidad de ir más allá tanto en temática como en cantidad de casos analizados, así como en amplitud temporal, lo que ha permitido observaciones más consistentes. Sin embargo, como se ha argumentado en la metodología, nuestro análisis puede tener limitaciones, ya que no es un estudio de toda la desinformación que circula en España, sino de la desmentida por esta agencia de verificación. Dadas las garantías que ofrecen los códigos de principios de la IFCN y de la EFCSN, de

los que Maldita.es es signataria, se considera que esta aproximación es adecuada, pues su actividad se centra en los casos más difundidos y peligrosos, y así lo demuestra la gran cantidad de trabajos previos que han usado esta estrategia para analizar la desinformación (Almansa-Martínez *et al.*, 2022; Fernández-Castrillo y Magallón-Rosa, 2023; Gamir-Ríos y Tarullo, 2022; Gutiérrez-Coba *et al.*, 2020; Narváez-Llinares y Pérez-Rufi, 2022; Salaverría *et al.*, 2020). En cualquier caso, es posible que los propios sesgos de la agencia o el tipo de información que ofrece —pensada para desmentir los bulos de manera efectiva, no para ofrecer datos para el análisis a la academia— estén limitando algunas de las conclusiones.

Finalmente, cabe señalar como línea de trabajo futura la identificación de los actores que están detrás de la creación y propagación de los contenidos falsos. Aquí reside uno de los principales desafíos de la desinformación en la red, pues el propio anonimato impide conocer de dónde procede una información. Sin embargo, la aproximación que aquí se utiliza añade complejidad a esta dimensión, dado que la agencia de verificación tiende a no difundir el origen de un contenido si no es relevante para contextualizarlo o para realizar el desmentido. Así, futuros estudios con otras estrategias metodológicas deberán explorar esta cuestión.

6. Financiación y apoyos

Estudio derivado de la tesis doctoral de David Blanco-Herrero, financiada por el Ministerio de Universidades a través de la ayuda FPU18/01455 del Subprograma de Formación de Profesorado Universitario (FPU).

7. Contribución de autores

Conceptualización	Ideas; formulación o evolución de los objetivos y metas generales de la investigación.	Autores 1 y 2
Curación de datos	Actividades de gestión para anotar (producir metadatos), depurar datos y mantener los datos de la investigación (incluido el código de software, cuando sea necesario para interpretar los propios datos) para su uso inicial y su posterior reutilización.	Autores 1 y 3
Análisis formal	Aplicación de técnicas estadísticas, matemáticas, computacionales u otras técnicas formales para analizar o sintetizar datos de estudio.	Autor 1
Adquisición de fondos	Adquisición del apoyo financiero para el proyecto que conduce a esta publicación.	Autores 1 y 2
Investigación	Realización de una investigación y proceso de investigación, realizando específicamente los experimentos, o la recolección de datos/evidencia.	Autores 1 y 3
Metodología	Desarrollo o diseño de la metodología; creación de modelos.	Autor 1
Administración del proyecto	Responsabilidad de gestión y coordinación de la planificación y ejecución de la actividad de investigación.	Autores 1 y 2
Recursos	Suministro de materiales de estudio, reactivos, materiales, pacientes, muestras de laboratorio, animales, instrumentación, recursos informáticos u otras herramientas de análisis.	Autores 1, 2 y 3
Software	Programación, desarrollo de software; diseño de programas informáticos; implementación del código informático y de los algoritmos de apoyo; prueba de los componentes de código existentes.	Autores 1, 2 y 3
Supervisión	Responsabilidad de supervisión y liderazgo en la planificación y ejecución de actividades de investigación, incluyendo la tutoría externa al equipo central.	Autores 1 y 2

Validación	Verificación, ya sea como parte de la actividad o por separado, de la replicabilidad/reproducción general de los resultados/experimentos y otros productos de la investigación.	Autor 1
Visualización	Preparación, creación y/o presentación del trabajo publicado, específicamente la visualización/presentación de datos.	Autor 1
Redacción / Borrador original	Preparación, creación y/o presentación del trabajo publicado, específicamente la redacción del borrador inicial (incluyendo la traducción sustantiva).	Autor 1
Redacción / Revisión y edición	Preparación, creación y/o presentación del trabajo publicado por los miembros del grupo de investigación original, específicamente revisión crítica, comentario o revisión, incluidas las etapas previas o posteriores a la publicación.	Autores 1, 2 y 3

8. Referencias bibliográficas

- Amores, J. J. (2022). *De la representación al odio. Desarrollo de nuevas estrategias para entender los discursos de odio hacia los migrantes y refugiados: detectando, midiendo y analizando los encuadres visuales, las actitudes y los mensajes de rechazo a la migración* [Tesis doctoral]. Universidad de Salamanca. <http://bit.ly/4crwzB>
- Allcott, H. y Gentzkow, M. (2017). Social Media and Fake News in the 2016 Election. *Journal of Economic Perspectives*, 31(2), 211-236. <https://doi.org/10.1257/jep.31.2.211>
- Allport, G. (1954). *The Nature of Prejudice*. Addison-Wesley.
- Almansa-Martínez, A., Fernández-Torres, M. J. y Rodríguez-Fernández, L. (2022). Desinformación en España un año después de la COVID-19. Análisis de las verificaciones de Newtral y Maldita. *Revista Latina De Comunicación Social*, 80, 183-200. <https://doi.org/10.4185/RLCS-2022-1538>
- Alonso-López, N., Sidorenko-Bautista, P. y Giacometti, F. (2021). Beyond challenges and viral dance moves: TikTok as a vehicle for disinformation and fact-checking in Spain, Portugal, Brazil, and the USA. *Anàlisi: Quaderns de Comunicació i Cultura*, 64, 65-84. <https://doi.org/10.5565/rev/analisi.3411>
- Arcila-Calderón, C., Blanco-Herrero, D. y Valdez-Apolo, M. B. (2020). Rechazo y discurso de odio en Twitter: análisis de contenido de los tuits sobre migrantes y refugiados en español. *REIS: Revista Española de Investigaciones Sociológicas*, 172, 21-40. <https://doi.org/10.5477/cis/reis.172.21>
- Arcila-Calderón, C., Sánchez-Holgado, P., Quintana-Moreno, C., Amores, J. J. y Blanco-Herrero, D. (2022). Discurso de odio y aceptación social hacia migrantes en Europa: Análisis de tuits con geolocalización. *Comunicar*, 30(71). <https://doi.org/10.3916/C71-2022-02>
- Ausserhofer, J. y Maireder, A. (2013). National politics on Twitter. Structures and topics of a networked public sphere. *Information, Communication & Society*, 16(3), 291-314. <https://doi.org/10.1080/1369118X.2012.756050>
- Bastos, M. T. y Mercea, D. (2019). The Brexit Botnet and User-Generated Hyperpartisan News. *Social Science Computer Review*, 37(1), 38-54. <https://www.doi.org/10.1177/0894439317734157>
- Bechmann, A. (2020). Tackling disinformation and infodemics demands media policy changes. *Digital Journalism*, 8(6), 855-863. <https://doi.org/10.1080/21670811.2020.1773887>
- Bruns, A. y Burgess, J. (2012). Researching news discussion on Twitter. *Journalism Studies*, 13(5-6), 801-814. <https://doi.org/10.1080/1461670X.2012.664428>
- Bruns, H., Dessart, F. J. y Pantazi, M. (2022). *COVID-19 Misinformation: Preparing for future crises. An overview of the early behavioural sciences literature*. Publications office of the European Union.
- Calvo, D., Cano-Orón, L. y Llorca-Abad, G. (2022). Desinformación sobre la vacuna de la COVID-19 en YouTube. Análisis de una red de visionado. *Communication & Society*, 35(2), 223-238. <https://doi.org/10.15581/003.35.2.223-238>
- Cammaerts, B. (2009). Radical pluralism and free speech in online public spaces. *International Journal of Cultural Studies*, 12(6), 555-575. <https://doi.org/10.1177/1367877909342479>
- Cantón Correa, F. J. y Galindo Calvo, P. (2019). Posverdad, redes sociales e islamofobia en Europa. Un estudio de caso: el incendio de Notre Dame. *RAEIC, Revista de la Asociación Española de Investigación de la Comunicación*, 6(12), 35-57. <https://doi.org/10.24137/raeic.6.12.8>
- Chen, S., Xiao, L. y Kumar, A. (2022). Spread of misinformation on social media: What contributes to it and how to combat it. *Computers in Human Behavior*, 141, 107643. <https://doi.org/10.1016/j.chb.2022.107643>
- Del-Fresno-García, M. y Manfredi, J.-L. (2018). Politics, hackers and partisan networking. Misinformation, national utility and free election in the Catalan independence movement. *El Profesional de la Información*, 27(6), 1225-1238. <https://doi.org/10.3145/epi.2018.nov.06>
- Farkas, J., Schou, J. y Neumayer, C. (2018). Cloaked Facebook pages: Exploring fake Islamist propaganda in social media. *New Media & Society*, 20(5), 1850-1867. <https://doi.org/10.1177/1461444817707759>
- Fernández-Castrillo, C. y Magallón-Rosa, R. (2023). El periodismo especializado ante el obstruccionismo climático. El caso de Maldito Clima. *Revista Mediterránea De Comunicación*, 14(2). <https://www.doi.org/10.14198/MEDCOM.24101>
- Gamir-Ríos, J. y Tarullo, R. (2022). Características de la desinformación en redes sociales. Estudio comparado de los engaños desmentidos en Argentina y España durante el 2020. *Contratexto*,

- 37, 203-228. <https://doi.org/10.26439/contratexto2022.n0375343>
- Gerbner, G. y Gross, L. (1976). Living with television: The violence profile. *Journal of Communication*, 26(2), 172-199. <https://doi.org/10.1111/j.1460-2466.1976.tb01397.x>
- Goffmann, E. (1974). *Frame Analysis: An Essay on the Organization of Experience*. Harvard University Press.
- González-Bailón, S., Lazer, D., Barberá, P., Zhang, M., Allcott, H., Brown, T. ... y Tucker, J. A. (2023). Asymmetric ideological segregation in exposure to political news on Facebook. *Science*, 381(6656), 392-398. <https://doi.org/10.1126/science.ade7138>
- Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B. y Lazer, D. (2019). Fake news on Twitter during the 2016 US presidential election. *Science*, 363(6425), 374-378. <https://doi.org/10.1126/science.aau2706>
- Gutiérrez-Coba, L. M., Coba-Gutiérrez, P. y Gómez-Díaz, J. A. (2020). Noticias falsas y desinformación sobre el Covid-19: análisis comparativo de seis países iberoamericanos. *Revista Latina de Comunicación Social*, 78, 237-264. <https://www.doi.org/10.4185/RLCS-2020-1476>
- Hayes, A. F. y Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, 1(1), 77-89. <https://doi.org/10.1080/19312450709336664>
- Hickey, D., Schmitz, M., Fessler, D., Smaldino, P. E., Muric, G. y Burghardt, K. (2023). Auditing elon musk's impact on hate speech and bots. En Y.-R. Lin, M. Cha y D. Quercia (Eds.), *Proceedings of the 17th International AAAI Conference on Web and Social Media* (pp. 1133-1137). AAAI. <https://doi.org/10.1609/icwsm.v17i1.22222>
- Igartua, J. J. (2006). *Métodos Cuantitativos de Investigación en Comunicación*. Bosch.
- Lin, T. M., Lu, K. Y. y Wu, J. J. (2012). The effects of visual information in eWOM communication. *Journal of Research in Interactive Marketing*, 6(1), 7-26. <https://doi.org/10.1108/17505931211241341>
- McCombs, M. E. y Shaw, D. L. (1972). The Agenda-Setting Function of Mass Media. *Public Opinion Quarterly*, 36(2), 176-187. <https://doi.org/10.1086/267990>
- Miró Llinares, F. (2016). Taxonomía de la comunicación violenta y el discurso del odio en Internet. *IDP. Revista de Internet, Derecho y Política*, 22, 82-107.
- Molina Cañabate, J. P. y Magallón Rosa, R. (2019). Procedimientos para verificar y desmontar informaciones falsas basadas en el discurso del odio. El caso de Maldita Migración. *RAEIC, Revista de la Asociación Española de Investigación de la Comunicación*, 6(12), 95-122. <https://doi.org/10.24137/raeic.6.12.5>
- Montero-Liberona, C. y Halpern, D. (2019). Factores que influyen en compartir noticias falsas de salud online. *El profesional de la información*, 28(3), e280317. <https://doi.org/10.3145/epi.2019.may.17>
- Montesi, M. (2020). Understanding fake news during the Covid-19 health crisis from the perspective of information behaviour: the case of Spain. *Journal of Librarianship and Information Science*, 53(3). <https://doi.org/10.1177/0961000620949653>
- Moragas-Fernández, C. M., Grau-Masot, J.-M. y Capdevila-Gómez, A. (2019). Articulación de la influencia en Twitter ante el anuncio de la Ley del referéndum en Cataluña. *El Profesional de la Información*, 28(3), e280320. <https://doi.org/10.3145/epi.2019.may.20>
- Moreno-Castro, C., Vengut-Climent, E., Cano-Orón, L. y Mendoza-Poudereux, I. (2021). Estudio exploratorio de los bulos difundidos por WhatsApp en España para prevenir o curar la COVID-19. *Gaceta Sanitaria*, 35(6), 534-540. <https://doi.org/10.1016/j.gaceta.2020.07.008>
- Narváez-Llinares, Á. y Pérez-Rufí, J.P. (2022). Fake news y desinformación sobre migración en España: prácticas del discurso xenófobo en redes sociales y medios online según la plataforma Maldita Migración. *Estudios sobre el Mensaje Periodístico*, 28(4), 841-854. <https://doi.org/10.5209/esmp.82845>
- Neuendorf, K.A. (2002). *The Content Analysis Guidebook*. Sage.
- Paz, M. A., Montero-Díaz, J. y Moreno-Delgado, A. (2020). Hate speech: A systematized review. *Sage Open*, 10(4), 1-12. <https://doi.org/10.1177/2158244020973022>
- Rogers, R. (2020). Deplatforming: following extreme internet celebrities to Telegram and alternative social media. *European Journal of Communication*, 35(3), 213-229. <https://doi.org/10.1177/0267323120922066>
- Salaverría, R., Buslón, N., López-Pan, F., León, B., López-Goñi, I. y Erviti, M.-C. (2020). Desinformación en tiempos de pandemia: tipología de los bulos sobre la Covid-19. *Profesional de la Información*, 29(3). <https://doi.org/10.3145/epi.2020.may.15>
- Schwarz, K. y Holnburger, J. (2019). Disinformation: what role does disinformation play for hate speech and extremism on the internet and what measures have social media companies taken to combat it? En J. Baldauf, J. Ebner y J. Guhl (Eds.), *Hate Speech and Radicalisation Online. The OCCI Research Report* (pp. 35-43). ISD.
- Shahi G. K., Dirkson, A. y Majchrzak, T. A. (2021) An exploratory study of COVID-19 misinformation on Twitter. *Online Social Networks and Media*, 22, 100104. <https://doi.org/10.1016/j.osnem.2020.100104>
- Van Aelst, P., Strömbäck, J., Aalberg, T., Esser, F., de Vreese, C., Matthes, J., Hopmann, D., Salgado, S., Hubé, N., Stepińska, A., Papathanassopoulos, S., Berganza, R., Legnante, G., Reinemann, C., Sheaffer, T. y Stanyer, J. (2017). Political communication in a high-choice media environment: A challenge for democracy? *Annals of the International Communication Association*, 41(1), 3-27. <https://doi.org/10.1080/23808985.2017.1288551>
- Vosoughi, S., Roy, D. y Aral, S. (2018). The spread of true and false news online. *Science*, 359, 1146-1151. <https://doi.org/10.1126/science.aap9559>
- Weidner, K., Beuk, F. y Bal, A. (2020). Fake news and the willingness to share: a schemer schema and confirmatory bias perspective. *Journal of Product & Brand Management*, 29(2), 180-187. <https://doi.org/10.1108/JPBM-12-2018-2155>

- Yang, X., Li, Y. y Lyu, S. (2019). Exposing deep fakes using inconsistent head poses. En *ICASSP 2019-2019 IEEE international conference on acoustics, speech and signal processing (ICASSP)* (pp. 8261-8265). IEEE. <https://doi.org/10.1109/ICASSP.2019.8683164>
- Zhang, X., Karaman, S. y Chang, S. F. (2019a). Detecting and simulating artifacts in GAN fake images. En *2019 IEEE international workshop on information forensics and security (WIFS)*. IEEE. <https://doi.org/10.1109/WIFS47025.2019.9035107>

David Blanco Herrero. Investigador posdoctoral en la Amsterdam School of Communication Research (AS-CoR) de la Universidad de Amsterdam (Países Bajos) desde 2023. Obtuvo su doctorado en la Universidad de Salamanca (España) con una beca FPU. Es también miembro del grupo de investigación Observatorio de Contenidos Audiovisuales de la Universidad de Salamanca y editor del *Anuario Electrónico de Estudios en Comunicación Social «Disertaciones»*. Desde 2018, ha realizado una extensa investigación en el campo de la desinformación y el discurso de odio, así como en otras vertientes del periodismo y la comunicación, generalmente ligado a las redes sociales, pero también abordando campos como la comunicación política, la inteligencia artificial o el análisis audiovisual. ORCID: <https://orcid.org/0000-0002-7414-2998>

Carlos Arcila Calderón. Es Profesor Titular del Departamento de Sociología y Comunicación de la Universidad de Salamanca (España). Miembro del Observatorio de Contenidos Audiovisuales (OCA) y profesor del Doctorado en Formación en la Sociedad del Conocimiento. Editor-in-chief de la revista *Disertaciones*. Doctor Europeo en «Comunicación, Cambio Social y Desarrollo» por la Universidad Complutense de Madrid. Máster en Data Science y Máster en Periodismo, ambos por la Universidad Rey Juan Carlos (URJC). Su línea de investigación está centrada en el discurso de odio, la adopción de TIC, el estudio de medios y redes sociales, los métodos computacionales y el *big data*. ORCID: <https://orcid.org/0000-0002-2636-2849>

Margarita Tovar Torrealba. Es responsable de Calidad y Programación en la Asociación Universitaria Iberoamericana de Postgrado. Ha colaborado como investigadora en la Universidad de Salamanca en proyectos nacionales e internacionales, como HummingBird o FakeDetector: Desarrollo y evaluación de un prototipo de detección automática de noticias falsas online, con la Fundación General de la Universidad de Salamanca, Radio Usal, el Magazine Audiovisual Su+ de la Facultad de Ciencias Sociales y con el Observatorio del Cambio del Cambio Climático y la Transición Ecológica. ORCID: <https://orcid.org/0009-0004-1692-8572>