

Desinformación y vacunas en redes: Comportamiento de los bulos en Twitter

Disinformation and vaccines on social networks: Behavior of
hoaxes on Twitter

José Manuel Noguera-Vivo.

Universidad Católica de Murcia. España.

jmnoguera@ucam.edu



María del Mar Grandío-Pérez.

Universidad de Murcia. España.

mgrandio@um.es



Guillermo Villar-Rodríguez.

Universidad Politécnica de Madrid. España.

guillermo.villar@upm.es



Alejandro Martín.

Universidad Politécnica de Madrid. España.

alejandro.martin@upm.es



David Camacho.

Universidad Politécnica de Madrid. España.

david.camacho@upm.es



Esta investigación presenta resultados parciales del proyecto I+D+I Ecosistemas de innovación en las industrias de la comunicación: Actores, tecnologías y configuraciones para la generación de innovación en contenido y comunicación (INNOVACOM), financiado por el Ministerio de Ciencia e Innovación (ref. PID2020-114007RB-I00). De igual modo, se ha realizado gracias al apoyo del Ministerio de Ciencia y Educación, en el marco de los proyectos FightDIS (PID2020-117263GB-I00, 2021-2024) y XAI-Disinfodemics (PLEC2021-007681, 2021-2023); así como gracias a la Comunidad Autónoma de Madrid, en el marco de CYNAMON (S2018/TCS-4566, 2018-2022); la Fundación BBVA, dentro de la convocatoria para equipos de investigación sobre SARS-CoV-2 y COVID-19, con el proyecto CIVIC: Caracterización Inteligente de la Veracidad de la Información relacionada con COVID-19 (2021-2022); la Comisión Europea, con el proyecto IBERIFIER (2021-2023): Iberian Digital Media Research and Fact-Checking Hub (CEF-TC-2020-2: 2020-EU-IA-0252), y la Barcelona Mobile World Capital Foundation, con el proyecto DisTrack: Tracking disinformation in Online Social Networks through Deep Natural Language Processing.

Cómo citar este artículo / Referencia normalizada.

Noguera-Vivo, J. M., Grandío-Pérez, M. del M., Villar-Rodríguez, G., Martín, A. y Camacho, D. (2023). Desinformación y vacunas en redes: Comportamiento de los bulos en Twitter. *Revista Latina de Comunicación Social*, 81, 44-62. <https://www.doi.org/10.4185/RLCS-2023-1820>

RESUMEN

Introducción: La desinformación antivacunas tiene un gran peligro por sus efectos tangibles en la sociedad. Existen investigaciones relevantes sobre tipologías de bulos, discursos negacionistas en redes o la popularidad de las vacunas, pero este estudio aporta una visión complementaria y pionera sobre el discurso antivacunas de COVID-19 en Twitter, centrada en el comportamiento de sus propagadores. **Metodología:** Dada una muestra inicial de un centenar de bulos (de diciembre de 2020 a septiembre de 2021) para la descarga de 200.246 tuits, se han filtrado mediante un algoritmo para la inferencia del lenguaje natural (NLI) alrededor de 36.000 tuits (N=36.292) que apoyan o desmienten la desinformación para analizar a sus difusores a través de sus métricas en la plataforma. **Resultados:** En números relativos, los resultados muestran, entre otros, más bulos con contenido original (no retuits) entre las cuentas con más seguidores y aquellas verificadas; más irrupción de desinformación frente a su objeción por cuentas creadas de 2013 a 2020, y la asociación del reconocimiento (mayor presencia en listas o muchos más seguidores que seguidos) a la preferencia por negar información falsa en lugar de aprobarla. **Discusión:** El artículo muestra cómo la tipología de las cuentas es un factor predictivo del comportamiento de usuarios que expanden desinformación. **Conclusiones:** Se revelan patrones similares de comportamiento del discurso antivacunas según indicadores de las cuentas de Twitter. El tamaño de la muestra y las técnicas empleadas dan una base sólida para otros estudios comparativos en desinformación sobre salud y en otros fenómenos en redes sociales.

PALABRAS CLAVE:

Desinformación; Bulos; Vacunas; Twitter; Inteligencia Artificial; Información de Salud; España.

ABSTRACT

Introduction: Anti-vaccine disinformation is highly dangerous due to its direct effects on society. Although there is relevant research on typologies of hoaxes, denialist discourses on networks or the popularity of vaccines, this study provides a complementary and pioneering vision about the anti-vaccine discourse of COVID-19 on Twitter, focused on its spreaders' behavior. **Methodology:** Given an initial sample of a hundred hoaxes (from December 2020 to September 2021) for the download of 200,246 tweets, around 36,000 tweets (N=36.292) that support or deny disinformation have been filtered through an algorithm for Natural Language Inference (NLI) to analyze their spreaders' through their metrics in the platform. **Results:** In relative numbers, the results show, among others, more hoaxes with original content (not retweets) among accounts with more followers and those verified; more irruption of disinformation opposed to its objection by accounts created between 2013 and 2020, and the association of the acknowledgement (more presence in lists or many more followers than followed users) to the preference for denying false information instead of approving it. **Discussion:** The article shows how the typology of the accounts can be a predictive factor about the behavior of users who spread disinformation. **Conclusions:** Similar behavioral patterns of anti-vaccine discourse are revealed according to the accounts' Twitter-related indicators. The size of the sample and the techniques used give a solid foundation for other comparative studies on disinformation about health and on other phenomena on social networks.

KEYWORDS:

Disinformation; Hoaxes; Vaccines; Twitter; Artificial Intelligence; Health Information; Spain.

1. Introducción

En septiembre de 2018, Nueva York sufrió un brote de sarampión que duró once meses y costó millones de dólares al sistema de salud. La lucha contra el brote tuvo una dificultad añadida: el discurso de los antivacunas emergió con fuerza en un contexto de desconfianza hacia las autoridades y un argumentario basado en la libertad de los padres y el rechazo a la industria farmacéutica (Zucker *et al.*, 2020). No era la primera vez que se registraban discursos antivacunas en redes sociales alrededor del sarampión, en especial en Twitter y Facebook (Deiner *et al.*, 2017), pero en ese año, Twitter registró un récord de narrativas antivacunas, superior incluso al vivido con la pandemia del coronavirus –un dato hallado en investigaciones preliminares para este estudio-. No estamos, por tanto, ante un fenómeno nuevo cuando hablamos del discurso de los antivacunas de COVID-19 en redes sociales, pero sí de una nueva oportunidad para monitorizar patrones de comportamiento de unos flujos de desinformación con graves consecuencias directas en la sociedad.

Desde el inicio de la pandemia en 2019, la desinformación ha sido difundida tanto por medios tradicionales como por redes sociales contribuyendo a generar determinados estados de opinión. Organismos Internacionales como la Organización Mundial de la Salud o la Unión Europea han establecido la lucha contra la desinformación como una de sus líneas prioritarias. Entendemos como desinformación la «información deliberadamente falsa, difundida por motivos económicos, ideológicos o por alguna otra razón» (Ireton y Posetti, 2018, p. 44). Dentro de este término encontramos un tipo específico de información, como son los bulos, que consisten en «mensajes falsos fabricados en las redes por usuarios y/o colectivos a fin de crear un determinado estado de opinión» (Aparici *et al.*, 2019, p. 3). En la actualidad, es tan visible este proceso de intoxicación informativa y tal es la percepción entre los usuarios de internet, que la desinformación es la principal preocupación en el mundo respecto al uso de redes y medios sociales, y más de la mitad de estos usuarios, un 53%, señala de forma explícita estar preocupado por este fenómeno (Knuutila *et al.*, 2020).

La desinformación circula con mucha facilidad en entornos online, especialmente, en redes sociales (Kouzy *et al.*, 2020). Investigaciones previas nos presentan algunas características de la tipología de bulos que circulan por la red sobre la COVID-19 (Salaverría *et al.*, 2020) o de cómo los creadores de desinformación –con ayuda de bots– son más prolíficos que quienes publican información veraz (Saby *et al.*, 2021). En concreto, personalidades públicas juegan un papel muy importante en la difusión de bulos sobre la COVID-19 al comprobarse que tiene un gran nivel de compromiso –*engagement*– en las redes sociales y convertirse en *super-spreaders*, especialmente en Twitter (Shahi *et al.*, 2021; Bodaghi y Oliveira, 2022). Un estudio basado en 38 millones de artículos sobre los temas más prominentes de desinformación de la COVID-19 demuestra que únicamente el 16,4% de la información difundida online había sido verificada (Evanega *et al.*, 2021).

La desinformación en torno a las vacunas está teniendo un elevado impacto de salud pública a nivel mundial. Recientes estudios han observado cómo la desinformación no solo ha fluido, sino que ha influido en la percepción que la gente tiene sobre la pandemia (Islam *et al.*, 2020; Kim *et al.*, 2020). En concreto y a través de estudios experimentales, existen evidencias de cómo la exposición a desinformación sobre las vacunas de la COVID-19 afecta directamente a la intención de vacunación (El-Mohandes *et al.*, 2021; Loomba *et al.*, 2021).

Twitter se presenta como una herramienta fundamental para la conversación pública en torno a las vacunas COVID-19 frente a otras plataformas, como ya ocurriera anteriormente en otras enfermedades (Zhou *et al.*, 2015; Surian *et al.*, 2016; Larrondo-Ureta *et al.*, 2021; López-Martín *et al.*, 2021). Sin embargo, el diálogo a veces no es completamente horizontal y, parafraseando una teoría clásica de redes (Granovetter, 1973), podría decirse que las lógicas de la comunicación en red actúan convirtiendo

algunos lazos débiles en ocasionales lazos fuertes, otorgándoles mayor influencia. Estudios previos han demostrado que los usuarios con cuentas verificadas mostraban un poder de difusión sobre las vacunas casi cincuenta veces superior a los usuarios no verificados (Carrasco-Polaino *et al.*, 2021). Las noticias falsas continúan siendo difundidas especialmente por Twitter y WhatsApp un año después del inicio del estado de alarma en España (Almansa-Martínez *et al.*, 2022) y existe un amplio bagaje de investigación sobre vacunas y contenido emocional (Blankenship *et al.*, 2018; Himelboim *et al.*, 2020; Kummervold *et al.*, 2021).

También encontramos análisis más cercanos al objeto de estudio de este trabajo, como algunas investigaciones sobre los discursos negacionistas alrededor de COVID-19 (Morel, 2021), sobre las dudas ante las vacunas (Nowak *et al.*, 2020; Thelwall *et al.*, 2021; Subbaraman, 2021) o sobre la popularidad de las vacunas y su nivel de controversia en Twitter (Carrasco-Polaino *et al.*, 2021). Sobre esta última temática, el presente trabajo ofrece una nueva perspectiva complementaria y se enfoca directamente en los bulos generados por usuarios antivacunas. El análisis de la fisonomía o morfología de las cuentas responsables de la producción y difusión de desinformación, en concreto alrededor del discurso de los antivacunas de COVID-19, ofrece una mayor perspectiva para comprender la naturaleza de estos flujos, así como una herramienta para predecir futuros comportamientos ante otros fenómenos similares de desinformación.

2. Objetivos

El objetivo general de esta investigación es averiguar si el tipo de cuenta en Twitter influye en el comportamiento de los flujos de desinformación del discurso antivacunas. A partir de este objetivo, surgen otros más específicos, derivados de las distintas variables objeto de estudio:

1. Identificar la forma mayoritaria de difusión de desinformación del discurso antivacunas en Twitter (creación de contenido original, citación de otro tuit, respuesta a otro tuit o retuit) en función del número de seguidores de la cuenta.
2. Detectar diferencias en las formas de difusión de desinformación entre cuentas verificadas y no verificadas.
3. Señalar pautas de comportamiento de difusión de desinformación en función del volumen de contenido publicado en la cuenta.
4. Comparar tendencias entre cuentas pro-bulo y las que rebaten esos mensajes, en función del año de creación de las cuentas de Twitter.
5. Descubrir si la presencia en listas públicas de Twitter puede ser un rasgo significativo de las cuentas que difunden desinformación o bien de las que niegan los bulos.
6. Localizar tendencias en el ratio de seguidos y seguidores en el comportamiento de las cuentas protagonistas de los flujos de desinformación en Twitter.

2.1. Hipótesis

La literatura reseñada nos ofrece evidencias científicas sobre las cuentas difusoras de bulos en Twitter no solo de su mayor actividad productiva respecto a otras cuentas, sino también de diferencias significativas en la propagación del bulo en relación a su nivel de *engagement* o compromiso en la Red. De ahí que asumamos las siguientes hipótesis de investigación:

H1. La morfología de las cuentas de Twitter permite distinguir diferentes patrones de comportamiento en la producción y difusión del discurso de los bulos antivacunas COVID-19.

H2. Variables como el número de seguidores de la cuenta, su carácter de cuenta verificada o no, el volumen de contenido que publica habitualmente, el año de creación de la cuenta, su presencia en listas públicas o su ratio de seguidos/seguidores, son factores que permiten identificar patrones de conducta particulares en los flujos de comunicación de dichos bulos en Twitter.

3. Metodología

Este estudio experimental requiere de una metodología mixta diseñada por un equipo multidisciplinar que aplica técnicas de investigación de la neurociencia (impacto emocional fisiológico), la psicología (impacto en la ansiedad y el estrés) y la comunicación (impacto cognitivo) a los sujetos de esta exploración.

3.1. Muestra, procedimiento e instrumentos

La recopilación, procesado y análisis de datos consta de cinco fases, inspirados en la metodología de FacTeR-Check (Martín *et al.*, 2021): 1) se recolectan los bulos (en este caso de antivacunas) a partir de los desmentidos de fact-checkers; 2) se crean de forma automática las consultas (*queries*) que permiten realizar las búsquedas de tuits sobre los bulos considerados; 3) se utiliza la API (del inglés *Application Programming Interface*) de Twitter que facilita la descarga de los tuits mediante las consultas previamente definidas; 4) se emplea una técnica de Inteligencia Artificial conocida como *Natural Language Inference* (NLI) (MacCartney, 2009) para filtrar entre aquellas publicaciones que afirman o contradicen el bulo; 5) y, por último, se descargan los datos de los usuarios de tales tuits relacionados con el bulo.

Las consultas automáticas en combinación con el NLI utilizan como tecnología de base los *transformers* (Vaswani *et al.*, 2017), un tipo de red neuronal que trabaja con representaciones vectoriales del texto que recogen las propiedades semánticas de las palabras y que tienen en cuenta el contexto en el que se producen. A partir de la similitud entre vectores (Huertas-García *et al.*, 2021a; Huertas-García *et al.*, 2021b) y la inferencia entre estos (Huertas-Tato *et al.*, 2021), los *transformers* han sido aplicados de forma exitosa en diferentes aplicaciones para la detección de desinformación.

En primer lugar, y para la selección de bulos, se han escogido los desmentidos de la fundación de *fact-checking* Maldita. La elección únicamente de Maldita como medio verificador con el que seleccionar los bulos sobre antivacunas viene dada por la significación de este medio digital. Desde que en 2014 los periodistas Clara Jiménez y Julio Montes abrieron un pequeño producto mínimo viable en forma de blog en Wordpress y cuenta de Twitter (Maldita Hemeroteca), este medio verificador no sólo fue el punto de inflexión para motivar la creación de otros medios verificadores en España, sino también un proyecto periodístico que no ha dejado de crecer. Desde 2017 forma parte de la International Fact Checking Network (IFCN) y en 2018 fue el único medio español seleccionado para estar en el Grupo de Alto Nivel sobre Fake News y Desinformación creado por la Unión Europea para asesorar en esta materia. En la actualidad siguen avalados por el sello de la IFCN (Mantzarlis, 2018).

De los bulos sobre vacunación, este trabajo ha explorado los ejemplos que van directamente en contra de la utilización de las vacunas para la COVID-19. Además, dentro del cribado, se han eliminado las afirmaciones falsas que daban problemas para identificar desinformación en Twitter a través de una consulta. De forma más concreta, los bulos eliminados han sido aquellos que contenían:

1. Contenidos expresados a partir de elementos multimedia (imagen, vídeo, audio) pero no de forma textual, ya que las consultas permiten localizar los tuits sólo en función de las palabras. Es decir, el texto del bulo debe contener el propio hecho falso.
2. Acusaciones falsas a un medio de haber publicado un tipo de información sobre las vacunas, por haber dos focos en el bulo, el medio cuestionado y la supuesta información emitida, que dificultan la creación de consultas y el etiquetado de tuits.
3. Afirmaciones ambiguas sacadas de contexto, en las que la verificación es una aclaración de todo lo sucedido, pero no una negación de ello.
4. Desinformación ya repetida en la selección, de tal forma que un mismo bulo no aparezca varias veces.

Una vez realizado el análisis y cribado de los bulos seleccionados, para cada pieza de desinformación verificada y seleccionada para este estudio se ha generado una consulta automática con sus palabras clave más relevantes, junto con los operadores lógicos que permiten definir la sintaxis de la búsqueda de tuits en la red social.

De esta manera, el bulo «Las vacunas contienen óxido de grafeno» daría lugar a la consulta «(grafeno vacunas óxido) OR (grafeno vacunas contienen) OR (grafeno óxido contienen) OR (vacunas óxido contienen)» y se emplearía para la búsqueda automática de información relacionada en la red social seleccionada. A partir de estas consultas y la API de Twitter se han extraído las publicaciones en Twitter entre enero de 2020 y noviembre de 2021, consistente en un total de 220.246 tuits a partir de 54 bulos, y se han almacenado en una base de datos (MongoDB).

A continuación, y dado que la inclusión de las palabras claves de la consulta no garantiza que los tuits contengan realmente desinformación, se infieren, mediante el mencionado método de NLI, las probabilidades de que su contenido corresponda a un *entailment* (el tuit dice lo mismo que el bulo), sea una *contradiction* (el tuit niega el bulo) o bien, sólo pueda establecerse que sea «neutral» (el algoritmo no puede determinar si el tuit lo afirma o lo contradice, por lo que, o bien no se refiere al bulo considerado, o bien no puede establecer una correlación entre el contenido del mismo y su significado semántico). Para el objeto de este estudio, se han seleccionado los ejemplos con más del 99% de *entailment* y *contradiction* para comparar entre quienes difunden la desinformación y quienes la contradicen, respectivamente.

Finalmente, se han recopilado los datos en la plataforma de los usuarios que han publicado o diseminado los tuits seleccionados. En concreto, las variables utilizadas de cada cuenta de Twitter para este artículo son: la fecha de su creación; su número de tuits, seguidores y seguidos, y si tiene el símbolo de verificado en la red social. Se han descartado aquellos ejemplos cuyas cuentas han sido suspendidas o eliminadas.

Así, la muestra final la conforman los datos de usuarios de 36.292 tuits (originales, retuits, respuestas y citas de otros tuits) procedentes de 49 bulos desmentidos por Maldita (de una selección original de 100 bulos antivacunas, detectados entre el 30 de diciembre de 2020 y el 9 de septiembre de 2021, un listado original disponible a través del correo de los autores). De esta cantidad, 15.513 publicaciones tienen más del 99% de *entailment* y 20.779 más del 99% de *contradiction*. Según la política de uso de datos de Twitter, los textos, métricas y usuarios de los tuits no pueden publicarse, pero sí sus ID (números de identificación en Twitter). Estos datos pueden obtenerse en *Figshare* [<https://figshare.com/s/399422354519a2a51fc5>] y a partir de las direcciones de correo de los autores de este artículo.

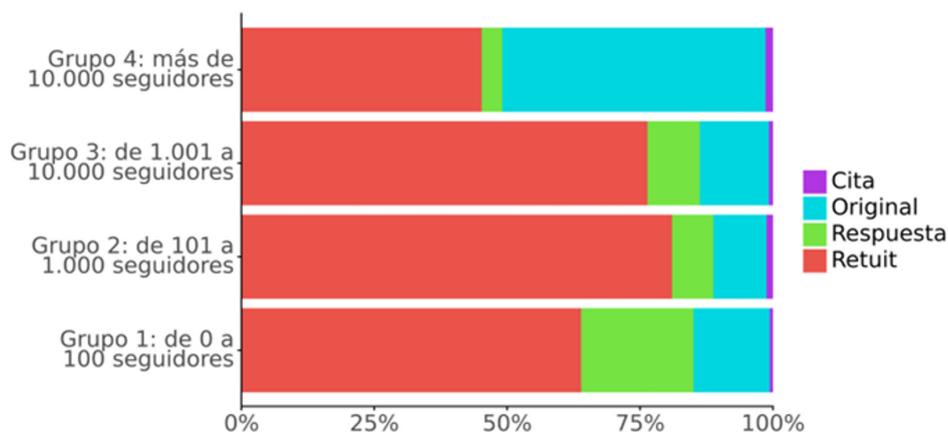
4. Resultados

Tal y como se refleja en las hipótesis, el objetivo general y los consecuentes objetivos específicos señalados al inicio, este apartado presenta los resultados más significativos en torno a las diferentes variables, agrupadas en tres bloques.

4.1. Según número de seguidores y cuentas verificadas

De las cuentas que apoyan la desinformación, es decir, que dicen lo mismo que el bulo (*entailment*), las que más desarrollan contenido propio (en azul) son las que acumulan mayor número de seguidores, ya que en este grupo encontramos hasta un 49,51% de contenido original. En el otro extremo, las que tienen un número menor de seguidores, hay un mayor uso del retuit para la difusión de desinformación (gráfico 1), siendo este recurso empleado hasta en un 63,9% de ocasiones dentro de este tipo de cuentas. Como se ha señalado, el grupo de usuarios que supera los diez mil seguidores es el que más contenido propio genera respecto a la desinformación. Es decir, la desinformación ha sido creada de manera original por ellos, sin provenir de ninguna otra información previa, tal y como ocurre con otras dinámicas en Twitter como el retuit directo o el tuit citado, que parten de una fuente previa. Por lo tanto, estos usuarios generan la desinformación en un porcentaje muy acusado respecto a los otros grupos de usuarios clasificados por sus seguidores. La tendencia del resto de grupos con menos seguidores es retuitear la información falsa sobre las vacunas, es decir, ser altavoces de una información que circula por la red y que ellos amplifican aún más, sin ningún tipo de reflexión ni comentarios personales.

Figura 1. Tipo de tuits (%) por grupos de mayor «entailment», según número de seguidores.



Fuente: Elaboración propia.

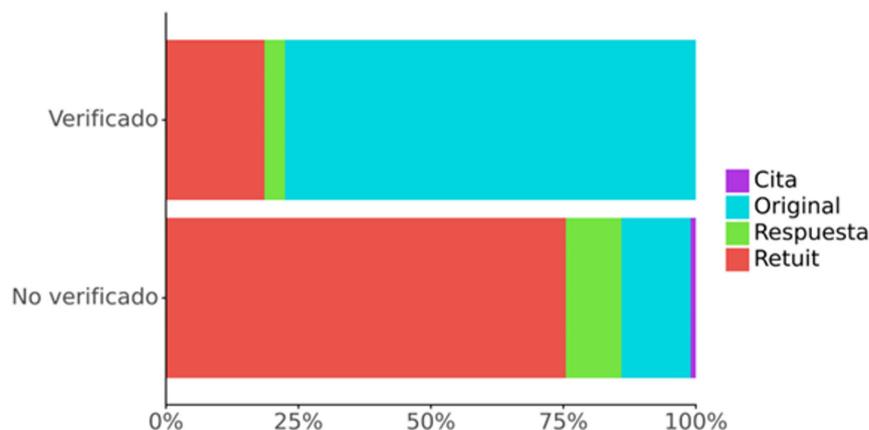
De hecho, otro dato interesante a destacar con el gráfico 1 es la poca difusión comentada que se hace de la desinformación (tuit citado, en morado en el gráfico), es decir, el escaso comentario o explicación sobre esa información que el usuario ofrece de manera personal cuando difunde ese contenido, algo que aparece de forma residual en todos los grupos, siempre con porcentajes menores al 2% (de mayor a menor número de seguidores, el tuit citado sólo aparece en porcentajes del 1,39%, 0,77%, 1,21% y 0,62%, respectivamente). Esto puede deberse a que la cita puede actuar en apoyo al tuit original de desinformación, sin necesidad de formular la desinformación como tal de nuevo porque ya viene expresada en el tuit que motiva su citación. Es decir, el usuario no necesitaría repetir la desinformación sino sólo mostrar su apoyo a un tuit que ya amplifica con el mero hecho de citarlo. Por otro lado, esto también nos recuerda la poca reflexión o lo impulsiva que es la difusión de contenido en esta red

social. En otras palabras, que hablamos de una red social con dificultades para dotar de contexto a la información (algo que Twitter trató de paliar en parte en diciembre de 2017, con la creación de los hilos).

Respecto a la capacidad de generar conversación o inicio de diálogo (respuestas a otros usuarios, en verde en el gráfico 1), llama la atención como la desinformación del grupo 1 (de 0 a 100 seguidores) ocurre más a través de la contestación que en otros grupos, ya que esta acción la encontramos hasta en un 21,15% de las ocasiones. Es, por tanto, el grupo de menos seguidores el que se relaciona en mayor medida con la desinformación a través de la respuesta a otros usuarios. Un intento de diálogo que cuanto menos habría que matizar, ya que este bajo número de seguidores puede indicarnos que estamos ante un grupo de cuentas formado mayoritariamente por *bots*.

Por su parte, el gráfico 2 confirma con datos una de las presunciones lógicas que podíamos tener respecto a la difusión de bulos antivacunas, como es la poca presencia de cuentas verificadas en los discursos, ya que cabe advertir (tratándose de un gráfico porcentual de bloques) que el bloque de cuentas verificadas opera sobre un total de apenas 129 unidades de análisis, mientras que el de cuentas no verificadas lo hace sobre 15.384 unidades. Llama la atención cómo las cuentas verificadas, tal vez por su propia condición de fuente creíble a priori, no utilizan en absoluto (0%) el recurso del tuit citado (en el gráfico 2, en morado), un recurso que sí es utilizado por las cuentas no verificadas, aunque es cierto que también de forma residual, sin apenas llegar al 1% (0,98%).

Figura 2. Tipo de tuits (%) de mayor «entailment», según tipo de cuenta (verificada o no).



Fuente: Elaboración propia.

Por otro lado, se puede ver bastante asimetría, ya que mientras que las cuentas verificadas tienen mucho contenido original (77,52%) y pocos retuits (18,6%) -y aún menos respuestas a usuarios-, en las no verificadas hay mayoría de retuits (75,47%) y poco contenido original (13,06%). En cualquier caso, habría que matizar, dado que hablamos de tuits con alto nivel de *entailment* (mensajes que no contradicen el bulo), que la diferencia respecto a unidades de análisis es lógica, ya que no es habitual que una cuenta verificada sea víctima de un bulo, si bien pueden existir estos casos ya que la condición de tuit verificado la otorga Twitter no sólo a cuentas de organismos e instituciones, sino también a personalidades públicas de cualquier ámbito.

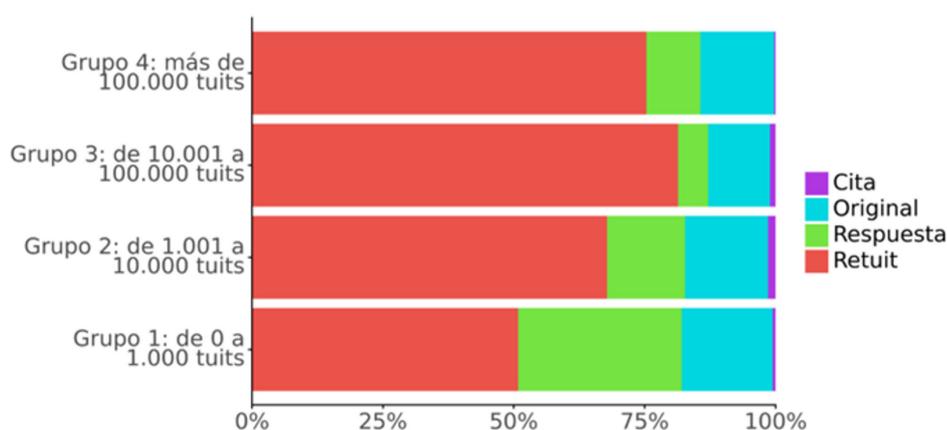
Paradójicamente y de forma peligrosa, dichas personalidades en ocasiones no es que no nieguen la desinformación, sino que incluso son actores activos en la difusión de bulos. Por ejemplo, en España

puede citarse el caso del famoso cantante Miguel Bosé, cuya cuenta fue cerrada por Twitter en agosto de 2020, al difundir información falsa precisamente sobre el coronavirus.

4.2. Según cantidad de tuits publicados y antigüedad de la cuenta

Otro aspecto para tener en cuenta, como muestra el gráfico 3, es la variable relacionada con el número de tuits que el usuario ha publicado, cuya división por grupos también presenta tendencias similares a los análisis de gráficos anteriores. En el caso del primer grupo, cuentas con menos de mil tuits publicados, podemos estar hablando tanto de cuentas de muy reciente creación como de otras más antiguas, aunque con poca actividad. Ambos tipos de perfil son los agrupados en ese primer grupo y el primer tipo (cuentas de reciente creación) puede incluir, lógicamente, a las cuentas falsas o protagonizadas por *bots*.

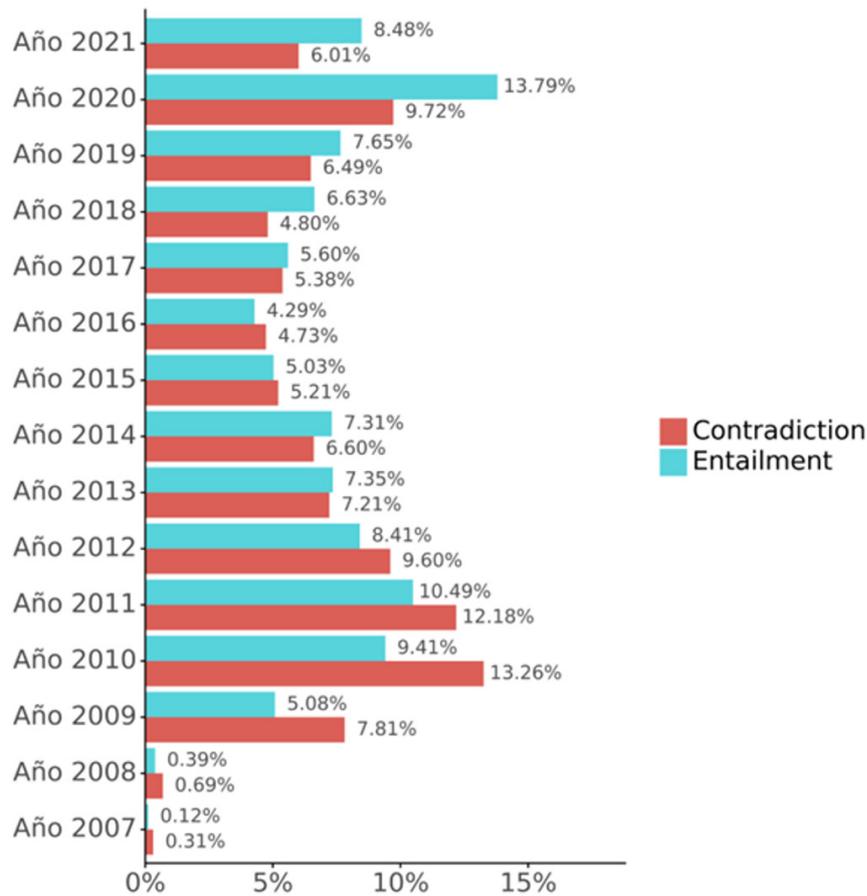
Figura 3. Tipo de tuits (%) por grupos de mayor «entailment», según número de tuits publicados.



Fuente: Elaboración propia.

Respecto a este primer grupo con poco contenido publicado, se trata del tipo de usuario que más recurre a la respuesta a otros usuarios como recurso para la difusión de desinformación, ya que encontramos este recurso empleado hasta en un 31,24% de ocasiones. En cualquier caso, si algo comparten los cuatro tipos de cuentas de este gráfico 3, es que el retuit es siempre la forma mayoritaria de difusión de bulos (de menor a mayor cantidad de contenido publicado, el retuit aparece en un 50,84%, 67,82%, 81,38% y 75,33%, respectivamente). Respecto a la distribución o tamaño de estas categorías, la menor es precisamente el comentado grupo 1, con un total de 893 ejemplos, mientras que el grupo más abundante es el número 3, con un total de 7.668 unidades (los grupos 2 y 4 son más parejos, con 3944 y 3008 unidades de análisis, respectivamente).

Por su parte, el gráfico 4 nos permite observar una panorámica de los usuarios a favor o a contracorriente de este tipo de bulos antivacunas durante los últimos quince años y presenta de forma visual unas tendencias significativas, tanto respecto a la presencia de *entailment* (discursos pro bulo) como después, en su comparativa con los tuits que niegan los bulos (*contradiction*).

Figura 4. *Entailment vs Contradiction de número de tuits, por año de creación del usuario.*

Fuente: Elaboración propia.

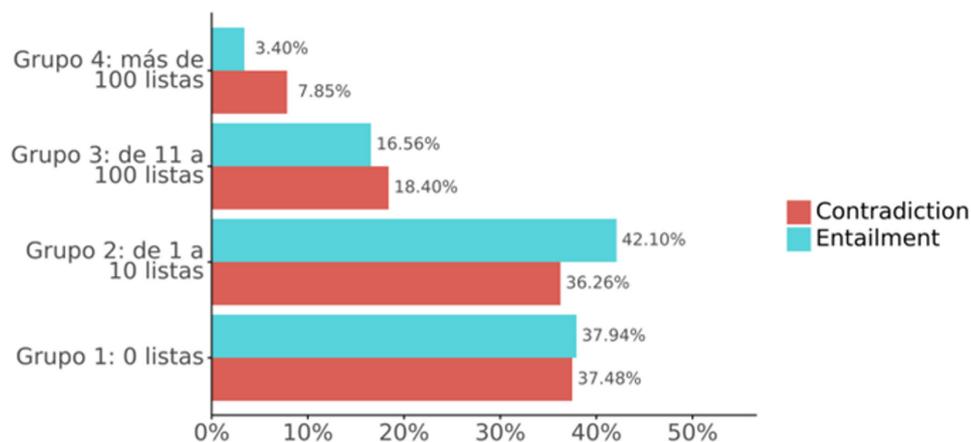
Si nos fijamos únicamente en la gráfica azul (*entailment*), el periodo de 2020 coincide con el mayor pico de creación de cuentas productoras de desinformación sobre el discurso antivacunas relacionado con COVID-19. Otra tendencia al respecto, iniciada incluso antes de la pandemia, es que desde 2013 hasta la fecha (con dos contadas excepciones, 2015 y 2016) estamos presenciando un periodo en el que cada año el número de cuentas creadas que apoyan los bulos (*entailment*) es mayor que el de cuentas creadas que los niegan (*contradiction*). Twitter, una red social que fue creada en 2006, tuvo tras su nacimiento seis años (2007-2012) donde la cantidad de cuentas creadas contra los bulos era siempre mayor que el número de cuentas pro-bulo. Sin embargo, en 2013 cambió la tendencia para tener a continuación (con la salvedad de 2015 y 2016) una predominancia de cuentas que apoyaban bulos antivacunas.

4.3. Según presencias en listas públicas y ratio de seguidores/seguídos

Como analiza el gráfico 5, la presencia en listas públicas de Twitter es otro de los aspectos que puede darnos pistas acerca de la naturaleza de las cuentas que participan en la producción y difusión de desinformación sobre las vacunas de COVID-19. Las listas son una herramienta nativa de Twitter con la que la comunidad reconoce una determinada identidad a un usuario concreto de esta red social. Una identidad que, por otro lado, no tiene por qué coincidir con la auto-denominación del usuario en cuestión (un usuario puede denominarse de una forma concreta en la biografía de Twitter, pero luego

ser calificado por la comunidad de un modo distinto al incluirlo en Listas). Y aquí radica el valor de las listas: como elemento exógeno configurador de nuestra identidad digital.

Figura 5. «Entailment» vs «Contradiction», según presencia en listas públicas de Twitter.



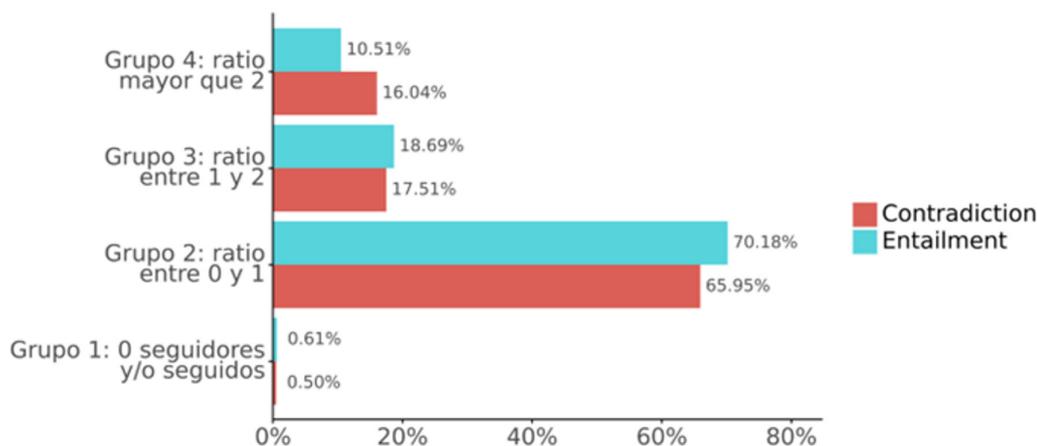
Fuente: Elaboración propia.

Cabría añadir que estar dentro de listas públicas en Twitter no es algo habitual de las cuentas, ya que muy pocos usuarios reciben este tipo de reconocimiento público por parte del resto de la comunidad. Esto explicaría los datos mayoritarios de la gráfica: la mayoría de cuentas, sean de *entailment* (37,94%) o de *contradiction* (37,48%) no están en listas o están en muy pocas (de una a diez listas). En este segundo supuesto hablamos de un 36,26% para *contradiction* y de hasta un 42,10% para *entailment*.

La comparativa avanza pareja y con ligera ventaja hacia *entailment* hasta que se llega a los grupos con más presencia en listas. Cuando hablamos de usuarios reconocidos por la comunidad de Twitter en más de un centenar de ocasiones a través de la herramienta de Listas (grupo 4), las cuentas que niegan la desinformación tienen una ventaja de casi cuatro puntos y medio (7,85% frente a 3,40%) sobre las que difunden desinformación. También ganan, aunque por menos (18,40% frente a 16,56%) en el grupo 3, el de la presencia en 11-100 listas. La lectura rápida de estos datos es que, a mayor presencia en listas públicas de Twitter, menos probabilidad de que estemos hablando de una cuenta que apoya la desinformación. Sin embargo, se trataría de una primera lectura de los datos porque en esta gráfica, al igual que ocurre en la gráfica 2 (comportamiento de cuentas verificadas), lo que a priori es un elemento que otorga credibilidad a una cuenta (aquí, las Listas) no siempre es así. Esto sería debido en este caso al particular comportamiento de ciertas personalidades públicas que, por su fama, lógicamente tendrán presencia en un alto número de listas públicas de Twitter.

Por último, en el gráfico 6 tenemos la oportunidad de saber si existe alguna tendencia especial de las cuentas que apoyan la desinformación en cuanto a su ratio de seguidores y seguidos. La oportunidad de saber cuántos seguidos necesitan para tener seguidores. En otras palabras, si se trata de cuentas que practican o buscan el *follow back* (es decir, que haya seguimiento mutuo) o que, por el contrario, tienen un reconocimiento genuino y orgánico, en tanto que natural, por la comunidad de Twitter. A mayor ratio, mayor reconocimiento orgánico (verdadero) tiene esa cuenta por parte del resto de usuarios, ya que no necesita practicar el citado *follow back* para acumular seguidores.

Figura 6. «Entailment» vs «Contradiction», según ratio seguidores y seguidos.



Fuente: Elaboración propia.

5. Discusión y Conclusiones

El discurso en redes sociales alrededor de la COVID-19, con sus correspondientes bulos, es uno de los ejemplos actuales más claros que tenemos sobre cómo la desinformación sanitaria puede desembocar en problemas tangibles de salud pública. A veces se trata de aglomeraciones en lugares señalados falsamente como centros de vacunación o, al contrario, de desinformaciones que debilitan los deseos de ser vacunado y repercuten en la asistencia de ciertos colectivos a los centros de vacunación reales. Entre estos dos extremos de falsedades, también encontramos multitud de variantes que otorgan falsos efectos secundarios a las vacunas y que, igualmente, influyen en la confianza sobre éstas y en la eficacia de las campañas de vacunación.

Este trabajo preliminar ha mostrado que conocer la morfología de las cuentas que protagonizan los flujos de desinformación puede ayudar a entender mejor los patrones de comportamiento de los discursos y bulos antivacunas. Si bien existe un nivel oculto de difusión de los bulos sobre el coronavirus, que se produce en redes de mensajería privada como WhatsApp (Salaverría *et al.*, 2020), el discurso online público y mediático de estos bulos tiene lugar sobre todo en redes sociales y, en particular, en Twitter. Por este motivo, el presente trabajo ha querido aportar una visión genérica que ofrece pistas para identificar algunas pautas de comportamiento del discurso antivacunas en esta red social. Ulteriores estudios, con similares discursos negacionistas, muestras y técnicas (la alta reproducibilidad de este artículo así lo permite), mostrarán hasta qué grado las tendencias que se han esbozado en este análisis pueden ser extrapoladas hacia otro tipo de bulos, desinformaciones sanitarias o discursos negacionistas.

Respecto a las hipótesis planteadas en esta investigación, se ha confirmado que hay comportamientos particulares relacionados con características de las cuentas de Twitter y que afectan tanto a la producción como a la difusión de desinformación sobre las vacunas para la COVID-19 (H1). Por otro lado, y respecto a la segunda hipótesis (H2), en el caso del atributo *entailment* (apoyo al bulo) hay patrones de comportamiento específicos vinculados al número de seguidores de la cuenta, la cantidad de tuits publicados y su condición de cuenta verificada o no. Por su parte, en el caso del atributo *contradiction*, la comparativa con su contrario nos permite ver otras pautas en grupos específicos de cuentas, sobre todo cuando los agrupamos según año de creación de las cuentas, su presencia en listas públicas de Twitter o su ratio de seguidores y seguidos.

En relación con el primer objetivo (forma mayoritaria de difusión de bulos antivacunas), la creación de contenido propio emerge como la forma preferida de desinformación por aquellas cuentas con mayor número de seguidores, mientras que, en el extremo opuesto, resulta curiosa la predominancia del retuit en aquellas cuentas de muy pocos seguidores y que nos recuerdan el papel propagandístico de los *bots*, muy posiblemente encuadrados en este grupo de cuentas.

El segundo objetivo remarca la importancia de interactuar con cuentas verificadas en periodos de desinformación, si bien los resultados muestran que este no es el modo preferido de diálogo por este tipo de cuentas. Cabe pensar que las cuentas verificadas, conocedoras de poseer una voz propia reconocida por la comunidad, optan sobre todo por el contenido de producción propia. La condición de cuenta verificada atañe también, como se ha mencionado en este artículo, a las personalidades públicas con un rol evidente de influencia sobre la sociedad. Y aunque la red social ha actuado en casos evidentes de desinformación a pesar de tratarse de cuentas verificadas (cancelación de la cuenta del cantante Miguel Bosé por parte de Twitter), lo cierto es que es anecdótica la presencia de este tipo de perfiles en los discursos sobre los bulos antivacunas, ya sea apoyándolos (*entailment*) o negándolos (*contradiction*).

La búsqueda del tercer objetivo (comportamiento de la cuenta en función de la cantidad de tuits publicados) nos remite también, al igual que el primer objetivo, a los debates sobre la propaganda de estado y la irrupción de los *bots* que, encuadrados en el grupo de cuentas con menos contenido publicado (junto a cuentas de reciente creación), optarían por la respuesta a otros usuarios como la forma preferida de difusión de la desinformación sobre vacunas. Los datos han mostrado que los últimos años en Twitter son de predominancia de las cuentas que apoyan la desinformación frente a las que la combaten. El contexto actual de polarización, discursos del odio y propaganda gubernamental, con su consecuente creación de *bots*, podría ser una de las explicaciones a esta tendencia detectada en los últimos años.

Respecto al cuarto objetivo, podemos trazar una clara evolución de la red social Twitter que, aunque en este caso va referida a los bulos antivacunas, indicaría que puede contener numerosas similitudes si lo intentamos extrapolar en futuros estudios a otros ámbitos. La tendencia general detectada consiste en una primera etapa, justo después del nacimiento de esta red (2007-2012), donde los usuarios que niegan los bulos antivacunas tenían más peso que aquellos que diseminan desinformación sobre este tema, mientras que a partir de 2013 y en especial de 2016, se abre una tendencia donde predomina la visibilidad del bulo (*entailment*) frente a las cuentas que lo niegan (*contradiction*). No parece un descubrimiento extraño si se pone en el contexto de los discursos del odio y la actual polarización en el escenario comunicativo (Pérez-Escolar y Noguera-Vivo, 2022), en especial en Twitter (Garimella y Weber, 2017; Yardi y Boyd, 2010). Una polarización donde las emociones juegan un papel clave (Döveling *et al.*, 2018) y determinan el alcance de la desinformación en el ámbito digital (Serrano-Puche, 2021).

Los dos últimos objetivos se plantearon en torno al análisis de elementos de identidad exógena, otorgada por el reconocimiento de la comunidad y no por el propio usuario, como son las listas públicas de Twitter y el número de seguidores como elemento de credibilidad, a partir de la ratio de seguidores y seguidos. Las redes sociales, y en particular Twitter, son un ejemplo de identidad mosaico (Caro-Castaño, 2015) donde la identidad digital se configura tanto por el usuario como por su comunidad de seguidores (Kietzmann *et al.*, 2011). Curiosamente, ambas variables han ofrecido resultados similares en el sentido de que, a mayor reconocimiento de la comunidad, hay menos probabilidad de que existan cuentas de apoyo a la desinformación. Tener un reconocimiento orgánico y natural por la comunidad (en el caso del último gráfico, una ratio elevada de seguidores/seguidos) aumenta las probabilidades de que estemos ante una cuenta que rebate la desinformación sobre las vacunas de COVID-19.

Este trabajo amplía los horizontes de la investigación sobre desinformación en redes sociales, a partir de un análisis que pone en relación la morfología de las cuentas de Twitter con determinados patrones de comportamiento online del discurso antivacunas y los bulos que genera.

6. Referencias

- 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>
- Aparici, R., García-Marín, D. y Rincón-Manzano, L. (2019). Noticias falsas, bulos y trending topics. Anatomía y estrategias de la desinformación en el conflicto catalán. *El Profesional de la Información*, 28. <https://doi.org/10.3145/epi.2019.may.13>
- Blankenship, E. B., Goff, M. E., Yin, J., Tse, Z. T. H., Fu, K. W., Liang, H. y Fung, I. C. H. (2018). Sentiment, contents, and retweets: a study of two vaccine-related Twitter datasets. *The Permanente Journal*, 22. <https://doi.org/10.7812/tpp/17-138>
- Bodaghi, A. y Oliveira, J. (2022). The theater of fake news spreading, who plays which role? A study on real graphs of spreading on Twitter. *Expert Systems with Applications*, 189, <https://doi.org/10.1016/j.eswa.2021.116110>
- Caro-Castaño, L. (2015). *La identidad mosaico como modo de subjetividad propio de las redes sociales digitales y sus formas de comunicación paramediáticas: La microcelebridad y la marca personal* [tesis doctoral, Universidad de Cádiz]. <https://bit.ly/3mJwShQ>
- Carrasco-Polaino, R., Martín-Cárdaba, M. y Villar-Cirujano, E. (2021). Citizen participation in Twitter: Anti-vaccine controversies in times of COVID-19. *Comunicar*, 69, 21-31. <https://doi.org/10.3916/C69-2021-02>
- Deiner, M. S., Fathy, C., Kim, J., Niemeyer, K., Ramirez, D., Ackley, S. F. y Porco, T. C. (2017). Facebook and Twitter vaccine sentiment in response to measles outbreaks. *Health Informatics Journal*, 25, 1116-1132. <https://doi.org/10.1177/1460458217740723>
- Döveling, K., Harju, A. y Sommer, D. (2018). From mediatized emotion to digital affect cultures: New technologies and global flows of emotion. *Social Media + Society*, 4, 1-11. <https://doi.org/10.1177/2056305117743141>
- El-Mohandes, A., White, T. M., Wyka, K., Rauh, L., Rabin, K., Kimball, S. H., Ratzan, S. C., & Lazarus, J. V. (2021). COVID-19 vaccine acceptance among adults in four major US metropolitan areas and nationwide. *Scientific Reports*, 11, <https://doi.org/10.1038/s41598-021-00794-6>
- Evanega, S. Lynas, M., Adams, J. y Smolenyak, K. (2021). Coronavirus misinformation: quantifying sources and themes in the COVID-19 infodemic. *JMIR*, 10. <https://bit.ly/3HoN1RM>
- Garimella, V. R. K. y Weber, I. (2017). A long-term analysis of polarization on Twitter. En: *Proceedings of the International AAAI Conference on Web and Social Media*, 1. <https://bit.ly/3mCmCb5>

- Granovetter, M. S. (1973). The strength of weak ties. *American journal of sociology*, 78(6), 1360-1380. <https://bit.ly/3sD5AgO>
- Himmelboim, I., Xiao, X., Lee, D. K. L., Wang, M. Y. y Borah, P. (2020). A social networks approach to understanding vaccine conversations on Twitter: Network clusters, sentiment, and certainty in HPV social networks. *Health Communication*, 35, 607-615. <https://doi.org/10.1080/10410236.2019.1573446>
- Huertas-García, Á., Huertas-Tato, J., Martín, A. y Camacho, D. (2021a). CIVIC-UPM at CheckThat! 2021: integration of transformers in misinformation detection and topic classification. *Faggioli*, 33. <https://bit.ly/3Jrhwi>
- Huertas-García, Á., Huertas-Tato, J., Martín, A. y Camacho, D. (2021b). *Countering Misinformation Through Semantic-Aware Multilingual Models*. International Conference on Intelligent Data Engineering and Automated Learning. Springer. <https://bit.ly/319DiiF>
- Huertas-Tato, J., Martín, A. y Camacho, D. (2021). SML: a new Semantic Embedding Alignment Transformer for efficient cross-lingual Natural Language Inference. *arXiv preprint*, arXiv:2103.09635. <https://bit.ly/3qtsbts>
- Ireton, Ch. y Posetti, J. (eds). (2018). *Journalism, 'fake news' and disinformation: Handbook for journalism education and training*. Unesco. <https://bit.ly/3mGXoZd>
- Islam, M.S., Sarkar, T., Khan, S., Mostofa Kamal, A.-H., Hasan, S. M. M., Kabir, A., Yeasmin, D., Islam, M. A., Amin Chowdhury, K. I., Anwar, K. S., Chughtai, A. A., & Seale, H. (2020). COVID-19-related infodemic and its impact on public health: a global social media analysis. *The American Journal of Tropical Medicine and Hygiene*, 103, 1621-1629. <https://dx.doi.org/10.4269%2Fajtmh.20-0812>
- Kietzmann, J. H., Hermkens, K., McCarthy, I. P. y Silvestre, B. S. (2011). Social media? Get serious! Understanding the functional building blocks of social media. *Business Horizons*, 54, 241-251. <https://bit.ly/3mGwujY>
- Kim, H. K., Ahn, J., Atkinson, L. y Kahlor, L.A. (2020). Effects of COVID-19 misinformation on information seeking, avoidance, and processing: a multi-country comparative study. *Science Commun*, 42. <https://doi.org/10.1177/1075547020959670>
- Knuutila, A., Neudert, L.-M. y Howard, P. (2020). Global fears of disinformation: Perceived Internet and Social Media Harms in 142 countries. *COMPROM Data Memo*, 8. <https://bit.ly/3FEoCXD>
- Kouzy, R., Abi Jaoude, J., Kraitem, A., El Alam, M. B., Karam, B., Adib, E., Zarka, J., Traboulsi, C., Akl, E. W. y Baddour, K. (2020). Coronavirus Goes Viral: Quantifying the COVID-19 Misinformation Epidemic on Twitter. *Cureus*, 12. <https://doi.org/10.7759/cureus.7255>
- Kummervold, P. E., Martin, S., Dada, S., Kilich, E., Denny, C., Paterson, P. y Larson, H. J. (2021). Categorizing Vaccine Confidence with a Transformer-Based Machine Learning Model: Analysis of Nuances of Vaccine Sentiment in Twitter Discourse. *JMIR Medical Informatics*, 9, <https://doi.org/10.2196/29584>
-

- Larrondo-Ureta, A., Fernández, S.-P., & Morales-i-Gras, J. (2021). Desinformación, vacunas y Covid-19. Análisis de la infodemia y la conversación digital en Twitter. *Revista Latina de Comunicación Social*, 79, 1-18. <https://doi.org/10.4185/RLCS-2021-1504>
- Loomba, S., de Figueiredo, A., Piatek, S. J., de Graaf, K. y Larson, H. J. (2021). Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA. *Nature Human Behaviour*, 5, 337-348. <https://doi.org/10.1038/s41562-021-01056-1>
- López-Martín, Á., Gómez-Calderón, B. y Córdoba-Cabús, A. (2021). Desinformación y verificación de datos. El caso de los bulos sobre la vacunación contra la Covid-19 en España. *Revista Ibérica de Sistemas e Tecnologías de Informação*, 431-443.
- MacCartney, B. (2009). *Natural language inference*. Stanford University. <https://bit.ly/3qsAtla>
- Mantzaris, A. (2018). Fact-checking 101. 85-100. En: Ireton, Ch. y Posetti, J. (eds). *Journalism, 'fake news' and disinformation: Handbook for journalism education and training*. UNESCO. <https://bit.ly/3pAWizZ>
- Martín, A., Huertas-Tato, J., Huertas-García, Á., Villar-Rodríguez, G. y Camacho, D. (2021). FacTeR-Check: Semi-automated fact-checking through Semantic Similarity and Natural Language Inference. *arXiv preprint, arXiv:2110.14532*. <https://bit.ly/32xKfux>
- Morel, A. (2021). Negationism of the COVID-19 and popular health education: to beyond the necropolitics. *Trabalho, Educação e Saúde*, 19. <https://doi.org/10.1590/1981-7746-sol00315>
- Noguera Vivo, J. M., Grandío Pérez, M. M., Villar Rodríguez, G., Martín, A. y Camacho, D. (2023). Desinformación y vacunas en redes: Comportamiento de los bulos en Twitter. *Revista Latina de Comunicación Social*, 81, 44-62. <https://doi.org/10.4185/RLCS-2023-1820>
- Nowak, S. A., Chen, C., Parker, A. M., Gidengil, C. A. y Matthews, L. J. (2020). Comparing covariation among vaccine hesitancy and broader beliefs within Twitter and survey data. *PloS One*, 15. <https://doi.org/10.1371/journal.pone.0239826>
- Pérez-Escolar, M. y Noguera-Vivo, J. M. (eds.) (2022). *Hate speech and polarization in participatory society*. Routledge. <https://bit.ly/3FE8EwB>
- Piqueiras Conlledo, P., De Marchis, G. y Cuesta Díaz, V. (2020). Análisis del contenido publicado en YouTube, Facebook e Internet sobre vacunas y anti-vacunas. *Revista de Comunicación y Salud*, 10(1), 67-90. [http://doi.org/10.35669/rcys.2020.10\(1\).67-90](http://doi.org/10.35669/rcys.2020.10(1).67-90)
- Saby, D., Philippe, O., Buslón, N., del Valle, J., Puig, O., Salaverría, R. y Rementeria, M. J. (2021) Twitter Analysis of COVID-19 Misinformation in Spain. En: Mohaisen, D. y Jin, R. (eds.) *Computational Data and Social Networks*. CSoNet 2021. *Lecture Notes in Computer Science*, 13116. https://doi.org/10.1007/978-3-030-91434-9_24
- 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. *El Profesional de la Información*, 29. <https://doi.org/10.3145/epi.2020.may.15>

- Sánchez Castillo, S., López Olano, C. y Peris Blanes, Á. (2023). Política, sanidad y desinformación: argumentos en Instagram de los partidos de extrema derecha europea sobre las vacunas contra la COVID-19. *Revista Latina de Comunicación Social*, 81, 210-229. <https://doi.org/10.4185/RLCS-2023-1870>
- Serrano-Puche, J. (2021). Digital desinformation and emotions: exploring the social risks of affective polarization. *International Review of Sociology*, 31, 231-245. <https://10.1080/03906701.2021.1947953>
- Shahi, G., Dirkson, A. y Majchrzak, T. (2021). An exploratory study of COVID-19 misinformation on Twitter. *Online Social Networks and Media*, 22, <https://doi.org/10.1016/j.osnem.2020.100104>
- Subbaraman, N. (2021). This COVID-vaccine designer is tackling vaccine hesitancy-in churches and on Twitter. *Nature*, 377-377. <https://doi.org/10.1038/d41586-021-00338-y>
- Surian, D., Nguyen, D. Q., Kennedy, G., Johnson, M., Coiera, E. y Dunn, A. G. (2016). Characterizing Twitter discussions about HPV vaccines using topic modeling and community detection. *Journal of Medical Internet Research*, 18. <https://doi.org/10.2196/jmir.6045>
- Thelwall, M., Kousha, K. y Thelwall, S. (2021). COVID-19 vaccine hesitancy on English-language Twitter. *El Profesional de la información*, 30. <https://doi.org/10.3145/epi.2021.mar.12>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N. y Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 5998-6008. <https://bit.ly/3pFbpYY>
- Yardi, S. y Boyd, D. (2010). Dynamic debates: An analysis of group polarization over time on twitter. *Bulletin of Science, Technology & Society*, 30, 316-327. <https://doi.org/10.1177%2F0270467610380011>
- Zhou, X., Coiera, E., Tsafnat, G., Arachi, D., Ong, M. S. y Dunn, A. G. (2015). Using social connection information to improve opinion mining: Identifying negative sentiment about HPV vaccines on Twitter. *MEDINFO*. <https://bit.ly/3mGQ2ow>
- Zucker, J., Rosen, J., Iwamoto, M., Arciuolo, R., Langdon-Embry, M., Vora, N., Rakeman, J., Isaac, B., Jean, A., Asfaw, M., Hawkins, S. Merrill, T., Kennelly, M., Maldin-Morgenthau, B., Daskalakis, D. y Barbot, O. (2020). Consequences of Undervaccination-Measles Outbreak, New York City, 2018–2019. *The New England Journal of Medicine*, 382, 1009-1017. <https://doi.org/10.1056/NEJMoa1912514>

7. Autores

José Manuel Noguera-Vivo

Universidad Católica de Murcia

Profesor Titular en la Universidad Católica de Murcia (UCAM), donde es investigador del área de Periodismo y dirige el departamento de Ciencias de la Comunicación. En esta misma institución, es el Investigador Principal del grupo de investigación “Comunicación, Política e Imagen”. Con más de medio de centenar de publicaciones sobre periodismo digital, participación y redes sociales,

su investigación se centra en la intersección entre medios, tecnología y sociedad. Ha sido becario postdoctoral en *The University of British Columbia* (Vancouver, Canadá), así como investigador invitado en varias universidades y congresos. En la actualidad es el responsable en España de la *Online News Association* (ONA). Entre sus últimas publicaciones, cabe destacar el libro *Hate Speech and Polarization in Participatory Society* (Routledge, 2022).

jmnoguera@ucam.edu

Índice H: 16

Orcid ID: <https://orcid.org/0000-0002-7189-7017>

Google Scholar: <https://scholar.google.com/citations?user=yQGsjxkAAAAJ&hl=en>

Scopus ID: <https://www.scopus.com/authid/detail.uri?authorId=36617884600>

María del Mar Grandío-Pérez

Universidad de Murcia

Profesora Titular de la Universidad Murcia, forma parte del Comité de Gestión español de la Acción COST europea INDCOR (*Interactive Narrative Design for Complexity Representations*, 2019-2023), como también lo fue con anterioridad de otra COST, *Transforming Audiences, Transforming Societies* (2010-2014). Premio de Transferencia del Conocimiento de la Universidad de Murcia en 2017, ha realizado estancias de investigación en *University of Missouri*, *University of California*, *University of Maryland* y *Georgetown*, entre otras instituciones. Sus líneas de investigación se centran en los contenidos de entretenimiento y los estudios de audiencias, teniendo publicaciones científicas en revistas como *The International Journal of Audience Research* o *Media Studies*, entre otras.

mgrandio@um.es

Índice H: 18

Orcid ID: <https://orcid.org/0000-0002-2577-4059>

Google Scholar: https://scholar.google.com/citations?hl=es&user=eV4aG_wAAAAJ

Scopus ID: <https://www.scopus.com/authid/detail.uri?authorId=57131060900>

Guillermo Villar-Rodríguez

Universidad Politécnica de Madrid

Investigador de posgrado asociado al proyecto sobre desinformación CIVIC (Caracterización Inteligente de la Veracidad de la Información sobre COVID) y doctorando dentro del área de computación en la Universidad Politécnica de Madrid. Tras obtener un máster propio en periodismo de datos, realizó estancias en el LAB de RTVE y en el diario El País, donde consolidó su especialización en el área. Premio Extraordinario de carrera en Periodismo, cursó un máster oficial de ciencia de datos y sociedad en Países Bajos para llevar las técnicas computacionales al ejercicio periodístico. Sus últimas publicaciones analizan las prácticas de verificación automatizadas mediante procesamiento del lenguaje natural y las formas de producción y difusión de los bulos alrededor de la COVID-19 en Twitter.

guillermo.villar@upm.es

Índice H: 1

Orcid ID: <https://orcid.org/0000-0001-7942-2879>

Google Scholar: <https://scholar.google.com/citations?hl=es&user=WGpTH9cAAAAJ>

Alejandro Martín

Universidad Politécnica de Madrid

Profesor Ayudante Doctor de la Universidad Politécnica de Madrid, sus principales áreas de interés son el aprendizaje profundo, el lenguaje modelado, la ciberseguridad y el procesamiento de lenguajes naturales (*natural language processing*). Ha sido investigador invitado en la Universidad de Kent (Reino Unido) y en la Universidad de Córdoba. Además de conferenciante, revisor y organizador de numerosos congresos internacionales, es el Investigador Principal del proyecto Caracterización Inteligente de la Veracidad de la Información relacionada con la COVID-19 (CIVIC), financiado por la Fundación BBVA dentro de la convocatoria para equipos de investigación sobre SARS-CoV-2 y COVID-19 (2021-2022). Ha publicado en revistas como *Communication & Society*, *Information Fusion* o *Applied Soft Computing*, entre otras.

alejandro.martin@upm.es

Índice H: 12

Orcid ID: <https://orcid.org/0000-0002-0800-7632>

Google Scholar: <https://scholar.google.es/citations?user=b3J9VRsAAAAJ&hl=es>

Scopus ID: <https://www.scopus.com/authid/detail.uri?authorId=57143182900>

David Camacho

Universidad Politécnica de Madrid

Catedrático del departamento de Sistemas Informáticos de la Universidad Politécnica de Madrid e Investigador Principal del grupo de investigación *Applied Intelligence and Data Analysis Group* (AIDA). Sus principales áreas de interés son la desinformación, el análisis de redes sociales, minería de datos, *Machine Learning* o inteligencia artificial, en particular el área específica de la inteligencia de enjambre (*swarm intelligence*). Entre artículos, libros y conferencias, acredita más de 300 publicaciones. Ha publicado en revistas científicas como *The Journal of Supercomputing*, *Information Fusion* o *The Journal of Ambient Intelligence and Humanized Computing*. Es parte del comité español de gestión de la Acción COST europea INDCOR (*Interactive Narratives for Complexity Representations*).

david.camacho@upm.es

Índice H: 36

Orcid ID: <https://orcid.org/0000-0002-5051-3475>

Google Scholar: <https://scholar.google.com/citations?user=fpf6EDAAAAAJ&hl=en>

Scopus ID: <https://www.scopus.com/authid/detail.uri?authorId=57191344221>