


Tipos de campaña Astroturfing de contenidos desinformativos y polarizados en tiempos de pandemia en España

Types of Astroturfing campaigns of disinformative and polarised content in times of pandemic in Spain

Tipos de campanhas Astroturfing de conteúdo desinformativo e polarizado em tempos de pandemia em Espanha

Sergio Arce-García^{1*} 

Elías Said-Hung^{2**} 

Daria Mottareale-Calvanese^{2***} 

¹ Escuela Superior de Ingeniería y Tecnología (ESIT) de la Universidad Internacional de la Rioja (UNIR), Logroño, La Rioja, España

² Facultad de Educación de la Universidad Internacional de la Rioja (UNIR), Logroño, La Rioja, España

* Profesor Contratado Doctor de la Escuela Superior de Ingeniería y Tecnología (ESIT) de la Universidad Internacional de la Rioja (UNIR), Logroño, La Rioja, España

** Profesor titular de la Facultad de Educación de la Universidad Internacional de la Rioja (UNIR), Logroño, La Rioja, España

*** Investigadora y profesora en la Facultad de Educación de la Universidad Internacional de la Rioja (UNIR), Logroño, La Rioja, España

Recibido: 04/04/2022; Revisado: 19/05/2022; Aceptado: 14/07/2022; Preprint: 16/08/2022; Publicado: 01/01/2023



Para citar este artículo: Arce-García, Sergio; Said-Hung, Elías; & Mottareale-Calvanese, Daria. (2023). Tipos de campaña Astroturfing de contenidos desinformativos y polarizados en tiempos de pandemia en España. *ICONO 14. Revista Científica De Comunicación Y Tecnologías Emergentes*, 21(1). <https://doi.org/10.7195/ri14.v21i1.1890>

Resumen

Este trabajo busca determinar la aplicación de estrategias de *Astroturfing* en Twitter, a nivel español, durante el periodo de pandemia a causa de la covid-19, en la primavera del año 2020. Se aplica análisis estadístico, análisis de redes y técnicas de *machine learning*, en 32.527 mensajes publicados a partir del decreto de estado de alarma en España (14 de marzo de 2020) hasta finales de mayo del mismo año, asociados ocho etiquetas que abordan temas vinculados a contenidos desinformativos identificados por dos de los principales proyectos de *fact-checking* (Maldito Bulo y Newtral). Los datos permiten observar la participación de usuarios (no *bots*), que ejercen un rol de influenciadores pese a que cuentan con un perfil promedio o alejado a ser considerados como personalidad pública. Se aprecia la aplicación del *Astroturfing* como estrategia de comunicación empleada para posicionar temas en las redes sociales a través de la distribución, amplificación y la inundación de contenidos desinformativos. El escenario permite comprobar la presencia de un escenario comunicativo digital que favorecería un marco difícil de detección, desde estrategias como la estudiada, orientadas a romper el efecto campana y filtro de burbuja de redes sociales. Todo con el fin de posicionar temas a nivel de opinión pública.

Palabras clave: contenidos desinformativos; bulos; *Astroturfing*; comunicación política; medios digitales; redes sociales

Abstract

The paper seeks to determine the application of Astroturfing strategies on Twitter, at the Spanish level, during the period of the pandemic caused by covid-19 in the spring of 2020. Statistical analysis, network analysis and machine learning techniques, around 32,527 messages, published from the state of alarm decree in Spain (March 14, 2020) until the end of May of the same year, associated with eight tags that address issues related to disinformative content identified by two of the main fact-checking projects (Maldito Bulo and Newtral). Data allows us to observe the participation of users (not bots), who play the role of influencers despite having an average profile or a profile that is far from being considered a public personality. The application of Astroturfing can be seen as a communication strategy used to position issues on social networks through the distribution, amplification and flooding of disinformative content. The scenario allows us to verify the presence of a digital communication scenario that would favour a framework difficult to detect, from strategies such as the one studied, aimed at breaking the echo chamber and filter bubble of social networks. All with the aim of positioning issues at the level of public opinion.

Keywords: disinformative content; hoaxes; Astroturfing; political communication; digital media; social networks

Resumo

Este documento procura determinar a aplicação de estratégias de Astroturfing no Twitter, a nível espanhol, durante o período pandémico devido à covid-19, na Primavera de 2020. Análise estatística, análise de rede e técnicas de aprendizagem de máquinas são aplicadas a 32.527 mensagens publicadas desde o decreto do estado de alarme em Espanha (14 de Março de 2020) até ao final de Maio de 2020, associadas a oito etiquetas que abordam tópicos relacionados com conteúdos desinformativos identificados por dois dos principais projectos de verificação de factos (Maldito Buló e Newtral). Os dados permitem-nos observar a participação dos utilizadores (não dos bots), que desempenham o papel de influenciadores apesar de terem um perfil médio ou um perfil que está longe de ser considerado uma personalidade pública. A aplicação de Astroturfing pode ser vista como uma estratégia de comunicação utilizada para posicionar questões sobre redes sociais através da distribuição, amplificação e inundação de conteúdo desinformativo. O cenário permite-nos verificar a presença de um cenário de comunicação digital que favorecería um quadro difícil de detectar, a partir de estratégias como a que foi estudada, visando quebrar o efeito sino e filtrar a bolha das redes sociais. Tudo com o objectivo de posicionar as questões ao nível da opinião pública.

Palavras-chave: conteúdo desinformativo; hoaxes; Astroturfing; comunicação política; meios de comunicação digitais; redes sociais

1. Introducción

El creciente avance del fenómeno de transmisión de información falsa desde las redes sociales ha generado un aumento exponencial de estudios académicos alrededor de los tres grandes conceptos que suelen emplearse para el abordaje de este fenómeno (Boididou et al., 2018), sin una clara demarcación aparente o debate sobre las diferencias que estos pueden tener entre sí: información falsa, desinformación y contenido desinformativo (Said-Hung et al., 2021). Definir estos términos no es tarea sencilla dado que las similitudes, diferencias y yuxtaposiciones entre ellos son muy sutiles y no existen suficientes mecanismos para identificarlos (Estrada-Cuzcano et al., 2020). Las *fake news* consideradas estrategias de desinformación intencional con objetivos políticos y sociales (Magallón, 2019) han empezado a difundirse rápidamente gracias a los medios digitales sustituyendo los tradicionales bulos. Estos últimos definidos como “todos aquellos contenidos falsos que alcanzan difusión pública, fabricados intencionadamente por múltiples motivos, que pueden ir desde la simple broma o parodia, hasta la controversia ideológica, pasando por el fraude económico” (Salaverría et al., 2020, p. 4).

El estudio de la difusión de los contenidos desinformativos, no solo se ha centrado en entender el fenómeno desde la influencia que estas generan en procesos electorales y en la opinión pública (Van-der-Linden, 2017; Zerback & Töpfl, 2021), o la incidencia que traen consigo las redes sociales en lo que se conoce como efecto de cámara de resonancia, al amplificar ideas y creencias, desde la transmisión y repetición constante de mensajes (no necesariamente veraces) hasta lograr la censura de visiones diferenciadas (Flaxman et al., 2016; Guess et al., 2018). También, se ha centrado en el efecto de filtro de burbuja

promovido desde la sobreabundancia de información en los escenarios digitales actuales (Flaxman et al., 2016).

Los estudios alrededor de la producción y difusión de noticias falsas en Internet (*fake news* en inglés) definidas como “publicaciones virales basadas en cuentas ficticias y hechas para parecer noticias reales” (Tandoc et al., 2018, p. 2), han girado en los últimos años en la construcción de modelos que ayuden a la detección e identificación de las estrategias de difusión de este tipo de contenidos, a través de técnicas computacionales, análisis de redes y semióticas (Zheng et al., 2017; Howard et al., 2017; Zhao et al., 2020).

Uno de los aspectos que han centrado el estudio de la desinformación es la presencia y/o uso de *bots* o los “*bots* sociales” (Ferrara et al., 2016; Allem, & Ferrara, 2018; Luceri et al., 2019), definidos como “cuentas automatizadas de las redes sociales operadas automáticamente por actores maliciosos con el objetivo de manipular la opinión pública” (Gallwitz & Kreil, 2021, p. 1). Sin embargo, la propagación de contenidos desinformativos se da de diferentes formas (Kucharski, 2016), no ha de limitarse al estudio de *bots* ni a la búsqueda de cuentas individuales con un comportamiento “robótico” pronunciado (Keller et al., 2019).

Estudios realizados por autores como Zheng et al. (2017), Howard et al. (2017) o Zhao et al. (2020) apuntan a diferentes estrategias de propagación, los cuales varían significativamente respecto con los contenidos veraces. Mensajes que, no necesariamente provienen de *bots* sino de usuarios orgánicos (reales), coordinados alrededor de un único propósito: lograr posicionar un tema a nivel social desde las redes sociales, alrededor de un determinado acontecimiento y durante un período de tiempo. Usuarios que Zheng et al. (2017) han dado en denominar “*Internet water army*” o “cuerpos sibilinos de élite”, al estar al servicio de un grupo político o económico, interesado en “inundar” las redes sociales e internet con contenidos intencionados, que intentan sacar máximo provecho de la acción comunicativa colectiva, propia de los escenarios digitales actuales.

Como bien apuntan Guess et al. (2018), la estrategia aplicada por movimientos políticos desde las redes sociales se lleva a cabo con el respaldo de páginas web creadas, para dar sentido de realidad a los diferentes argumentos a los mensajes creados y re-distribuidos (Zhao et al., 2020), o distribuidos, amplificados e inundados por unos usuarios (soldados) alejados del perfil *influencers* (muchos seguidores y una alta actividad). Una estrategia donde estos usuarios (micro- y nano-*influencers*, usuarios con hasta 100.000 y 10.000 seguidores o menos, respectivamente) aparentan no estar organizados ni relacionados entre sí, pero que actúan coordinadamente, evitando, desde el “anonimato aparente”, los celos que pueden traer consigo el liderazgo o relevancia que pueden tener actores políticos y sociales claramente identificables dada su relevancia a nivel social (Howard et al., 2017; Ong et al., 2019). Una estrategia, que se ha dado en llamar *Astrourfing*, empleada en política, relaciones públicas y publicidad desde hace décadas (Sorensen et al., 2017; Keller et al., 2019), y que, desde el auge digital actual, ha cobrado un creciente uso e interés para la divulgación de contenido desinformativo (Elmas et al., 2021). Ello, desde

el uso de técnicas de *marketing* como la denominada *Thunderclap*, donde se busca que personas, consideradas normales e identificables por el resto de la población como sus iguales, introducen dentro de sus respectivos ámbitos un mensaje determinado (Sorensen et al., 2017). Bajo esta técnica (*Thunderclap*), los usuarios que lideran (alfa) no siguen la cuenta del cliente (*influencers*) para evitar su detección. Además, los alfas cuentan con un equipo de usuarios (no necesariamente *bots*, denominados beta), quienes actúan de forma itinerante y con varias cuentas en redes sociales (en nuestro caso Twitter), y se encargan de: 1) contestar las diferentes interacciones generadas por los mensajes difundidos; y 2) pasar el contenido desinformativo transmitido por los usuarios líderes (alfa) a usuarios *influencers* (ej. periodistas y medios), con el propósito de generar tendencia (*trending topics*). Labores que sirven al monitoreo de la tasa de efectividad de la campaña llevada a cabo por ambos perfiles (alfas y betas), como bien señalan autores como Pérez (2020).

Para estudiar el *Astroturfing* no es suficiente centrarse en la búsqueda de los *bots* dado que no todas las cuentas lo son. Keller et al. (2019) recomiendan una estrategia de identificación basada en patrones de coordinación, argumentando que "un comportamiento similar entre un grupo de cuentas gestionadas por personas es una señal más fuerte de una campaña de desinformación, que un comportamiento individual "similar al de un *bot*" (p. 2). Los retos científicos actuales consisten en identificar patrones en las campañas y los ataques más que en el comportamiento de actores individuales. Algo, que requiere observaciones longitudinales, así como datos de múltiples medios sociales y plataformas en línea para el análisis de la dimensión espacial (Grimme et al., 2018).

El estudio del *Astroturfing* parte de tres enfoques (tabla 1), que ayudan a entenderlo y adecuar su aplicación desde las redes sociales: 1) desde la importancia de las publicaciones en internet para crear y co-crear valor compartido entre los usuarios (Sorensen et al., 2017); 2) desde características comunes que autores como Keller et al. (2019) han detectado en diferentes campañas de *Astroturfing* en distintas partes del mundo; y 3) desde tipologías más "efímeras" de *Astroturfing online* (Elmas et al., 2021) basadas en *Trending Topic* (TT), y en la capacidad de este tipo de estrategias en el condicionamiento de la opinión pública.

Tabla 1. *Conceptualizaciones teóricas de Astroturfing*

Autores	Conceptualizaciones teóricas	Rasgos característicos de la teoría
Sorensen et al. (2017)	Co-creación de valor mediante publicaciones en redes sociales	Las publicaciones en redes sociales ofrecen oportunidades para la co-creación de valor asociado. Al transmitir información a otros usuarios, se comparte, se aprende y genera compromiso por parte de los consumidores, a través de un <i>click</i> (un <i>retweet</i>) en lugar de la narración textual de un partidario.
Keller et al. (2019)	<i>Astroturfing</i> político como estrategia de coordinación <i>online</i>	En las redes sociales, el <i>Astroturfing</i> político adopta la forma de una campaña coordinada y organizada, en la que los participantes se hacen pasar por "usuarios ordinarios", que publican opiniones favorables a los instigadores de dicha campaña; distrayendo a los demás con noticias negativas, llaman la atención sobre cuestiones políticas divisivas o atacan a los oponentes y/o a los críticos. Este proceso se conoce como <i>co-tweeting</i> , es decir, la publicación de un mismo mensaje por cuentas dentro de una ventana de tiempo muy corta (e.j. unos minutos). Este tipo de coordinación es el que mejor distingue al <i>Astroturfing</i> , ya que cuentas ordinarias rara vez publican un parecido mensaje original al mismo tiempo.
Elmas et al. (2021)	<i>Astroturfing</i> efímero y los Trending Topic (TT) efímeros en Twitter	Una palabra clave o un tema elegido se promueven de manera artificial, mediante una actividad coordinada y no auténtica. Este tipo de acciones son muy frecuentes en los <i>Trending Topic</i> de Twitter. El <i>Astroturfing</i> efímero supera a otros enfoques en tres aspectos clave: (i) permite el uso de cuentas activas y comprometidas, como fuentes de interacciones falsas; (ii) evade la detección por parte de los usuarios, la plataforma y estudios académicos; y (iii) evita que los usuarios denuncien la actividad maliciosa, al verse los clasificadores de <i>spam</i> tradicionales incapaces de prevenir futuros ataques.

Fuente: Elaboración propia.

Este fenómeno enlazaría con la teoría sociológica de Granovetter (1973) de fortaleza de los vínculos débiles, donde los contactos lejanos al grupo central tendrían una gran influencia con el aporte de nuevas ideas sobre el grupo. De esta forma, los mensajes de usuarios alejados de la centralidad de la red serían los más indicados para introducir nuevos mensajes y hacerlos llegar más allá (Ribera, 2014).

2. Materiales y métodos

El objetivo de este trabajo es determinar el uso de estrategias de *Astroturfing online* político a nivel español, a través del estudio de mensajes vinculados a temas o etiquetas que han sido tendencias en Twitter, durante el periodo de pandemia por covid-19, en la primavera del 2020. Para ello se pretende:

- Establecer cómo se comportan las cuentas o usuarios que intervienen en los mensajes analizados.
- Identificar los rasgos que caracterizan a las cuentas difusoras de contenidos desinformativos.
- Determinar cómo es la evolución temporal de diseminación.
- Estimar la probabilidad de presencia de *bots* en la viralización.

La principal hipótesis que se espera confirmar o no en este trabajo, es la siguiente:

H1: Se está aplicando alrededor del debate político español en Twitter la estrategia *Astroturfing* para la diseminación de contenido desinformativo, en los términos expuestos por autores como Sorensen et al. (2017), Keller et al. (2019) y Elmas et al. (2021).

El análisis se centrará en detección de *Astroturfing* en función de su comportamiento, estructura y contenido (Chen et al., 2021). Los datos analizados se recogen a través de la API de Twitter y programación en R, mediante librería RTweet (Kearney, 2019), sobre distintas campañas. El proceso de recogida se realizó de manera geolocalizada, paralelamente a la toma global, mediante las coordenadas de longitud, latitud y radio de influencia de la zona de interés. Dicha técnica, aunque presenta una fiabilidad mundial del 77,84% y 88,15% en Europa (Van-der-Veen et al., 2015), permite detectar posibles interferencias geográficas. Los países escogidos para búsqueda son los identificados por Bradshaw et al. (2021), como lugares de emisión de desinformación, así como diversas capitales españolas.

El trabajo emplea un enfoque cuantitativo, a través del análisis estadístico, de redes y técnicas de *machine learning*, que ayudan a analizar los datos generados alrededor de las etiquetas y mensajes vinculados a estos en Twitter, desde la publicación del Decreto de estado de alarma en España, en marzo de 2020, hasta finales de mayo.

La selección de etiquetas tenidas en cuenta como caso de estudio, se hizo teniendo en consideración: 1) la selección de contenidos desinformativos recogidos en Maldito Bulo y/o Newtral a partir de temas políticos; y 2) etiquetas que se convertían en tendencias a través de mensajes con un claro perfil político muy partisano.

Las etiquetas recogidas para el análisis fueron: "Pedro Duque Denia", "Pedro Sanchez Huete", "respiradores Granada Madrid", "Muerte al Rey", "#SilenciaelPaís", "#Cacerolada21h", "#Pedroelsepulturero" y "#Elgobiernotemiente". El análisis se hizo a

partir de la recolección hecha, durante los primeros momentos de difusión de los mensajes vinculados. Así, se logra identificar el patrón de los mensajes difundidos y relacionados, así como la aplicación o no de estrategias de *Astrourfing* en ellas, desde una aparente espontaneidad de opiniones y contenidos por parte de perfiles no destacados.

Del total de etiquetas preseleccionadas, el análisis se centró en tres: “Pedro Duque Denia”, “#SilenciaelPaís” y “Muerte al Rey”, por implicar diferentes situaciones de interés (a diferencia del resto de etiquetas), alrededor del objetivo general planteado: 1) campaña sin éxito sin inundación posterior de mensajes en un tiempo muy específico, tras ser rápidamente identificada por otros usuarios como bulo (desmentido en maldita.es: <https://bit.ly/3Fr1pjr>); 2) con una inundación, a partir de un cierto éxito de campaña; y 3) con una doble inundación en el tiempo, por éxito de alcance de la campaña.

La muestra final analizada estuvo integrada por un total de 32.527 mensajes, donde se aplicó:

- Análisis de redes, a través del programa Gephi (versión 0.9.2), empleado para conocer las conexiones existentes entre usuarios, a través de su representación a través de puntos, tamaño y grosor de línea. Se usó el algoritmo Open Ord para separar los diferentes grupos (Martin et al., 2011) y conseguir una mejor visualización, además de realizarse análisis de modularidad (Blondel et al., 2008) así como la centralidad de cercanía (*closenesscentrality*), que mide la posición de cada nodo con respecto a otros y ofrece la idea de quiénes están conectados o actúan de forma individual. Así, una conexión débil, cercana al cero, daría cuenta de nodos periféricos y un valor próximo al uno de estar muy cercano a otros (Hansen et al., 2020).
- Análisis de comportamiento *bot*, por medio del paquete en R *tweetbotornot* de Kearney (2018) que, según su autor, presenta un 93,53% de aciertos en clasificación de *bots* y un 95,32% para no *bots* (media de 93,8% aciertos). Este algoritmo, aunque presenta discrepancias con otros métodos, está considerado entre los mejores y más empleados en ciencias sociales (Martini et al., 2021).
- Análisis de centralidad de cercanía o *closenesscentrality*, entre nodos de una red (cuentas de Twitter), lo que ayuda a identificar a los *influencers* de una comunidad (Pozzi et al., 2017).
- Análisis estadístico, a través del software R, empleado para observar el comportamiento entre las variables y el momento temporal.

4. Resultados

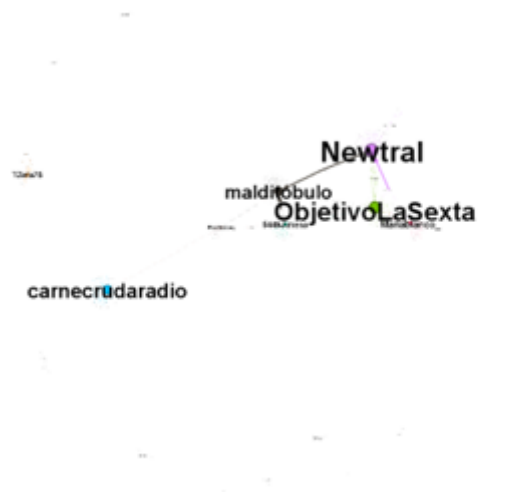
4.1. Campaña: Pedro Duque en Denia

El estudio del tema asociado al ingreso del Ministro de Ciencia de España (Pedro Duque), en un hospital de Denia (Valencia), ofrece un total de 370 mensajes. Solo se consideraron para estudio todos los mensajes que contasen con las palabras “Pedro Duque” o “ministro” y la localidad de “Denia”. Los dos primeros tuits analizados (ID 1241509359722598406 e ID 1241644721975549952) proceden del 21 y 22 de marzo de 2020 a las 23:37 y 08:35 (horarios GMT), respectivamente, haciendo mención de que el ministro ha sido ingresado. A partir de las 10:48 horas del 22/03/2020 comienzan a enviarse tuits en el periodo de distribución, todos con el mismo mensaje (aunque distintos textos, como: “¿Alguien puede verificar esta noticia? ¿Qué hace Buzz Light Year en Denia? ¿No está prohibido el desplazamiento de personas? ¿@astro_duque, donde estás? Pedro Duque ingresa en el Hospital de Denia”) y apoyándose de una información publicada ese día, en el portal espanaesvoz.es. Durante el proceso de viralización de estos mensajes se inicia el periodo de amplificación, con solicitudes de corroboración del hecho a otras webs (ej. “@okdiario Pedro Duque (el ministro astronauta) ingresado en el hospital de Denia... se había ido de fin de semana a Jávea!!! Atajo de imbéciles irresponsables. es cierta esta noticia? Investigad”), pero también otros que enlazan este tema con otras etiquetas vinculadas (ej. #PedroElSepulturero).

Los mensajes continúan hasta que el portal Newtral (dedicado al *fact-checking*), publica el desmentido a las 22:24, y el propio ministro aparece en un canal de televisión en abierto español (La Sexta). A partir de dicho momento, la distribución del contenido desinformativo casi desaparece, y el portal que difundió este hecho emite un comunicado reconociendo la falsedad de este suceso, mientras elimina toda alusión en su web.

El análisis de redes, figura 1, muestra 330 nodos de usuarios y la existencia de cuentas que se erigen como influyentes (según el tamaño del nodo y el tamaño del nombre de la cuenta), cuyos mensajes son seguidos por muchos otros. En dicha figura solo se aprecian cuentas que desmienten el contenido desinformativo, procedentes de medios de comunicación y periodistas, re-difundidos por sus seguidores, pero no se encuentran en el grafo mostrado difusores de contenidos desinformativos. El análisis de *cluster* detecta 62 comunidades de usuarios, con un determinado número de usuarios que actuaron divulgando la desinformación al inicio, pero que no se circunscriben a ningún grupo importante, al contrario de los *influencers* que aparecen posteriormente en el desmentido, tales como @Newtral, @ObjetivoLaSexta, @malditobulo y @carnecrudaradio. Por lo tanto, existe un número importante de mensajes que no aparecen como importantes, ni destacan en un análisis de redes.

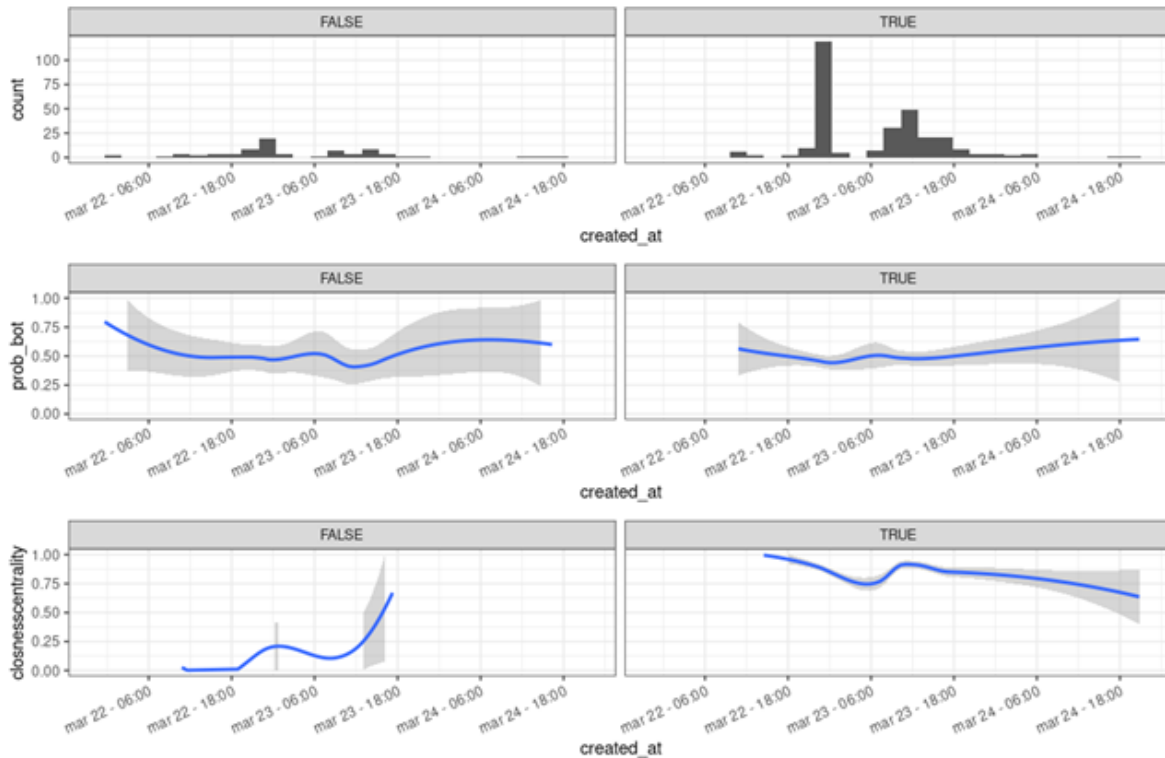
Figura 1 . Análisis de redes sobre la desinformación del Ministro Pedro Duque en Denia



Fuente: Elaborado por los autores.

El estudio de los datos mostrados en la figura 2, donde se representa la probabilidad de ser un *bot* (Kearney, 2018), muestra la presencia de 330 cuentas, quienes difundieron un total de 370 tuits (la mayoría sin ser retuiteados). Los datos mostrados permiten observar que todas las cuentas que difunden el contenido desinformativo en los primeros minutos presentan un valor de cercanía media alrededor del cero. Por lo tanto, no existe relación alguna con otros usuarios y están alejados de cualquier red. Las cuentas que difunden, pese a presentar distintas probabilidades de ser un *bot* tienen un grado medio del 50% de probabilidad de serlo, con un valor al inicio superior al 75% en los primeros mensajes.

Figura 2 . Número de tuits, probabilidad *bot* media y centralidad de cercanía media en contenidos desinformativos sobre Pedro Duque en Denia



Fuente: Elaboración propia. Nota: Los tuits directos son marcados como “False” y los retuits (RT) como “True”.

La tabla 2 muestra cuentas creadas con más de cinco años en Twitter (\bar{x} =2013 y Me =2012), pero con una gran actividad, ya que la media y la mediana supera las varias decenas de miles de tuits emitidos (Me =8 tuits y establecidos casi 4 favoritos por cuenta de manera diaria). La red de seguidores, amigos y contactos primarios de este tipo usuarios detectados, no son muy extensas, pero tampoco escasas, situándose en una Me = 557 seguidores y siguiendo a 770 cuentas. La capacidad de contacto de este tipo de usuarios identificados, unido a un valor de centralidad de cercanía de valor casi cero, sugiere que no se están enviando mensajes a un mismo conjunto de contactos, sino a diferentes grupos de usuarios que toman contacto con los emisores de estos contenidos desinformativos. Esto podría entrar a corresponderse con la técnica del *co-tweeting* de Keller et al. (2019) citado anteriormente, donde cuentas sin conexión con otras en pocos minutos difunden un mismo tema (distribución), mensajes que mencionan hacia determinados medios o usuarios importantes para esperar su difusión, como Luis del Pino (@ldpsincomplejos), OkDiario (@okdiario) o Caso Aislado (@CasoAislado_Es) en el proceso de amplificación. Al ser la desinformación cortada rápidamente no se observa proceso de inundación posterior, sino de contestación contra la desinformación.

El análisis del perfil de las cuentas difusoras permite ver que hacen uso de descripciones orientadas a ser identificados por otros miembros de Twitter como “iguales” o usuarios con un “perfil promedio”: amas de casa, mozos de almacén, periodistas, biólogos, abogados, autónomos, entre otros. También se aprecia que, en las descripciones empleadas hace referencia a valores ideológicos (ej. pro-familia, antifeministas, anticomunistas, heterosexuales, patriotas y liberales) y nacionales (ser español), que sirven para establecer conexión con otros usuarios afines.

Tabla 2. Datos sobre las cuentas que difunden en contenidos desinformativos sobre Pedro Duque en Denia

	Seguidores	Siguiendo	Tuits escritos	Favoritos	Fecha creación
Mínimo	1	0	43	20	13/12/2007
Primer cuartil	243	370,2	8.152	1.903	21/05/2011
Mediana	557	770	23.572	11.325	11/08/2012
Media	1.470	1.439,6	49.478	27.591	06/08/2013
Tercer cuartil	2.027	1.968,8	50.236	32.868	10/04/2015
Máximo	15.773	10.975	511.989	342.166	25/02/2020

Fuente: Elaboración propia, a partir de 162 tuits que difunden la desinformación.

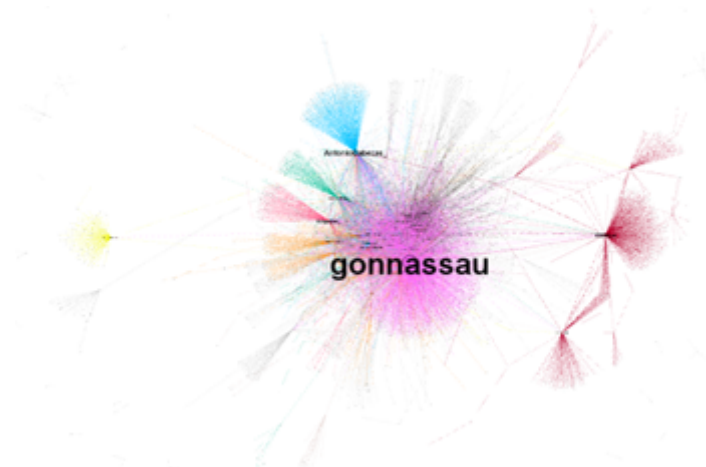
4.2. Campaña: #SilenciaElPaís

La etiqueta #SilenciaElPaís, se inició el 16 de abril de 2020 a las 14:20 horas (horario GMT). A partir de la creación del primer *tweet* comienza la difusión de mensajes, tanto directos como reenvíos (RT), sobre el tema. Un proceso de diseminación en el que se aprecia la oposición de un colectivo de cuentas con mensajes afines a una orientación ideológica de izquierda que no logran parar la difusión inicial. Algo que se hace desde el liderazgo del usuario @03690jul, que actúa de influenciador opositor dentro de esta etiqueta con un 10,72% del tráfico generado. También, aparece otro usuario (@mgonzalezelpais) asume un rol de influenciador y opositor a los contenidos desinformativos transmitidos alrededor de la etiqueta #SilenciaElPaís, quien genera el 2,56% del tráfico, cuyo perfil le identifica como “Periodista de Política Exterior y Defensa de EL PAÍS. Informo de Vox, Vox no quiere que informe de él. ¿Qué debo hacer?”.

Alrededor de la etiqueta analizada se generaron 15.977 tuits, de los cuales 2.771 fueron directos y 13.209 RT. La figura 3 muestra los *influencers* en cada uno de los nodos identificados, en el que destaca la cuenta @gonnasau, con un 36,03% del tráfico de RT, seguido de otras como @AntonioCabezas_, @ElPais_Today o @ElAguijon_, entre otras, las cuales se convierten en nuevos influenciadores a lo largo del tiempo sobre el

grupo principal. En el análisis de los mensajes vinculados, no se aprecia que existan casi conexiones entre grupos a favor de los contenidos difundidos frente sus opositores.

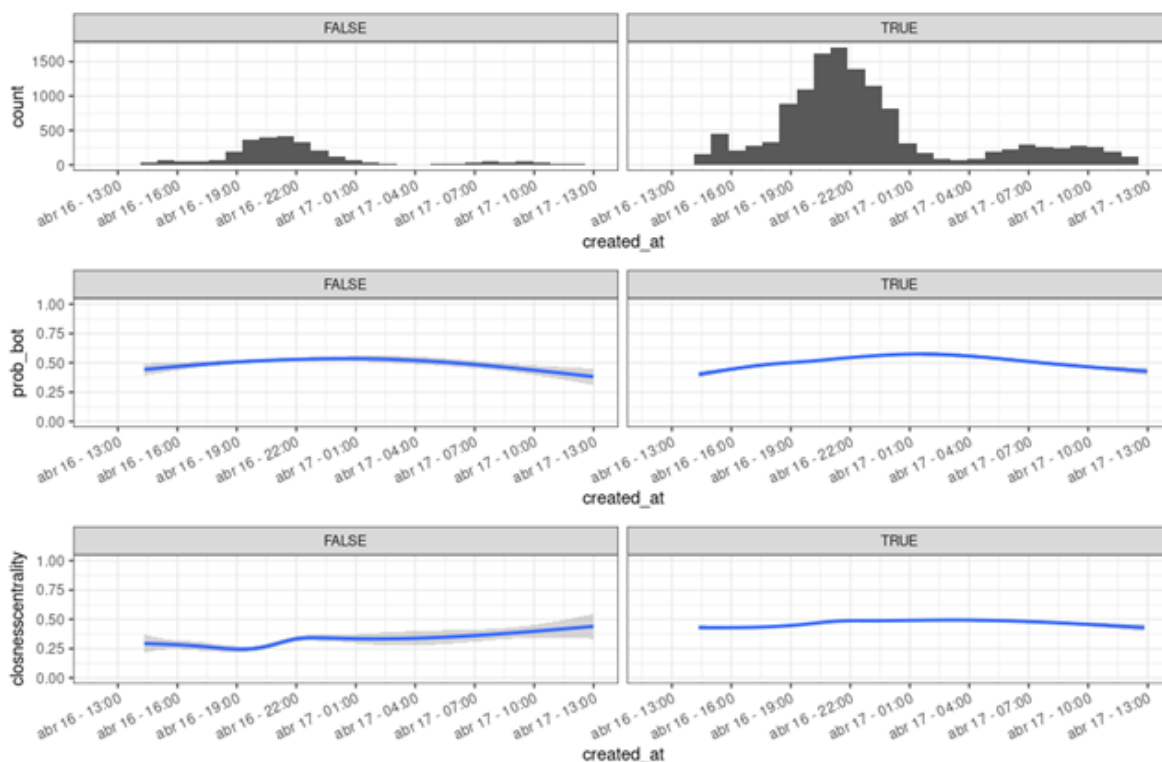
Figura 3. Análisis de redes sobre la campaña #SilenciaElPaís



Fuente: Elaborado por los autores.

En la figura 4, se aprecia cómo desde el inicio de la difusión de mensajes con el hashtag #SilenciaElPaís se producen en los primeros momentos un número de *tweets* con una escasa centralidad de cercanía, los cuales emiten un mismo mensaje (aunque no sea literal) sin conexión con otras cuentas (distribución ej: “Comencemos esta tarde a las 20:00h. Que circule, por favor. #SilenciaElPaís <https://t.co/ctGYI4VQ06>”, amplificación ej: “@Santi_ABASCAL @javiernegre10 #silenciaelpais <https://t.co/G5hfj2bLnE>”, con llamadas al líder de Vox y al fundador de EDATV), y donde la probabilidad media de ser un *bot* es del 50%. No se aprecian grandes cambios de valores de centralidad de cercanía en las cuentas que emiten mensajes directos, con valores medios por debajo de 0,5 aunque subiendo con el tiempo, mientras que en los que *retweetean* hay valores de centralidad de cercanía algo más elevados. En este caso, puede apreciarse una primera fase de amplificación y distribución desde numerosas cuentas independientes para llegar hacia las 8 de la tarde del 16 de abril de 2020 (las 18:00 GMT) a la fase de inundación, en el que se produce un aumento significativo de *tweets*, con un ligero aumento de la centralidad de cercanía y un aumento de la probabilidad *bot* media por encima del 50%. Por la noche del 16 de abril de 2020, los mensajes vinculados a esta etiqueta (*hashtag*) desaparecen, para volver a primera hora de la mañana del 17 de abril de 2020, con el mismo patrón de funcionamiento, pero con una inundación mucho más escasa.

Figura 4. Número de tuits, probabilidad *bot* media y centralidad de cercanía media en la campaña #SilenciaElPaís



Fuente: Elaboración propia. Nota: Los tuits directos son marcados como "False" y los retuits (RT) como "True".

Al igual que lo observado en el estudio de los contenidos desinformativos sobre Pedro Duque (tabla 2), se pueden encontrar usuarios que siguen a un mayor número de usuarios (*followings*) que seguidores tienen (*followers*), además de destacar la actividad intensiva que tienen en Twitter, al escribir varios miles de *tweets*, llegando una cuenta a superar el millón de mensajes (tabla 3).

Tabla 3. Cuentas que difunden la campaña #SilenciaElPaís

	Seguidores	Siguiendo	Tuits escritos	Favoritos	Fecha creación
Mínimo	0	0	1	0	07/03/2007
Primer cuartil	77	179	1.847	1.349	21/11/2011
Mediana	280	494	8.142	6.469	09/09/2014
Media	1.425	1.152	26.648	21.886	11/02/2015
Tercer cuartil	967	1.306	26.931	22.911	12/09/2015
Máximo	1.315.139	98.157	1.149.572	866.264	17/04/2020

Fuente: Elaboración propia.

Un análisis de las fechas de creación de las cuentas permite ver que 2014 y 2015 fueron los períodos donde las medias de estos perfiles fueron abiertas en *Twitter*, pero también que la mayoría se concentra en períodos de fechas muy concretas y no son repartidas en el tiempo (2011, otoño de 2017 y comienzo de 2020).

4.3. Campaña: “Muerte al Rey”

La viralización de contenido, resultado de un vídeo grabado en Alcorcón (Madrid) el 19 de mayo de 2020, durante unos altercados entre distintos manifestantes, donde se oye a una persona con la cara cubierta y con una bandera republicana decir “Muerte al Rey y a sus hijas”, comienza el mismo día a las 19:19 (hora GMT). En los mensajes vinculados a este suceso en *Twitter*, se acusa a grupos de izquierda de querer matar al jefe de Estado y a sus hijas, hecho del que varios medios de comunicación se hacen eco (González, 2020) (ej. distribución “Muerte al rey y a sus hijas” dicen los demócratas estos. No creo que tengan más de 18 años. Me dan pena estas imágenes, la izquierda lleva trabajando desde ZP para volver a la España de la II República. Estas imágenes se parecen bastante. No me gusta. <https://t.co/VSi8r0iUzh>”, amplificación: “@Malumarquez4 @abc_es @policia @guardiacivil Ha dicho “muerte al rey y a sus hijas”).

El estudio de redes del día 19 de mayo hasta el mediodía del día 20 ofrece un total de 16.180 tuits, con 1.970 directos y 14.480 RT, mostrado en la figura 5. El análisis permite comprobar que los mensajes relacionados están muy dispersos entre multitud de pequeños usuarios que ejercen un rol de influenciadores en la viralización de este tema en *Twitter*, sin contar con oposición alguna a sus mensajes.

Figura 5. Análisis de redes sobre la campaña “Muerte al Rey”

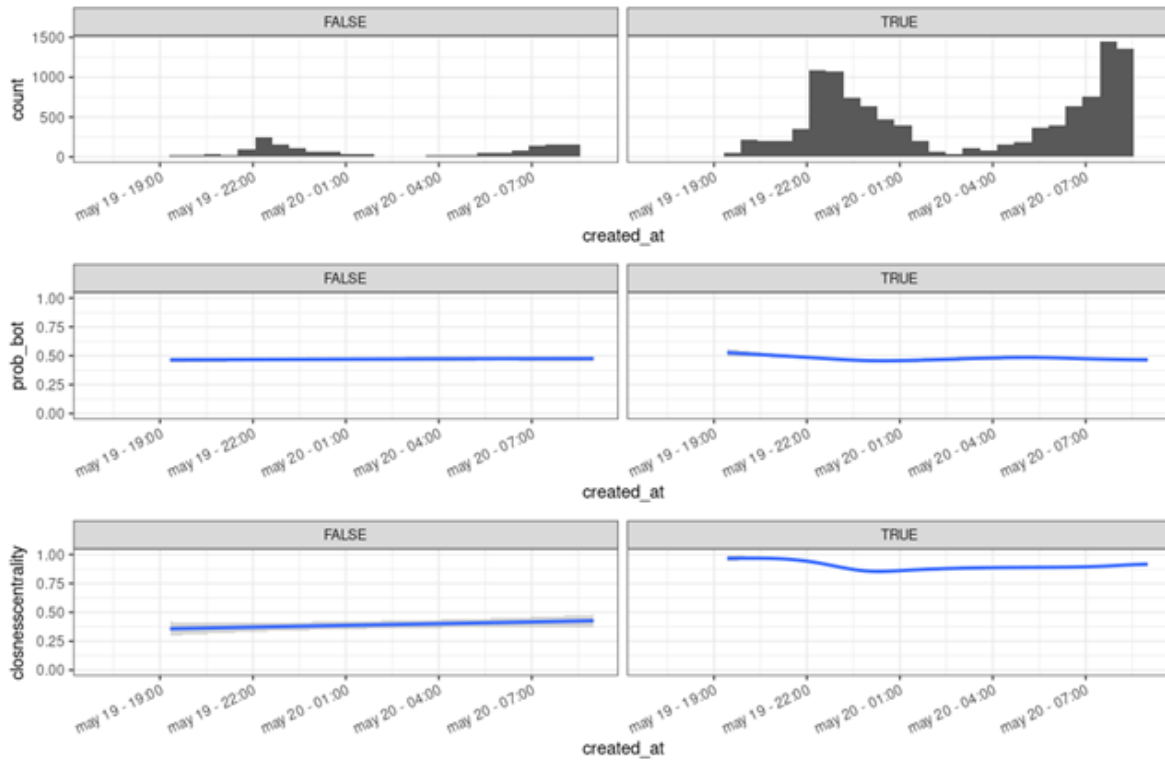


Fuente: Elaboración propia.

La figura 6, permite comprobar cómo en el periodo de tiempo en que se viralizaron contenidos afines al tema, distintas cuentas, de forma directa, publicaron mensajes siguiendo un claro patrón de amplificación y distribución de contenidos divulgados, hasta la llegada de la fase de inundación, que se produce hacia las dos horas del inicio de los primeros mensajes publicados en Twitter. Las cuentas de usuarios que inician la viralización (debate) presentan niveles de centralidad media de cercanía pequeños. La probabilidad *bot* media casi todo el tiempo es del 50%, así como en las cuentas que *retweetean*. La centralidad y proximidad de las cuentas que *retweetean* es muy elevada, cercana a la unidad, con lo que se observa una conexión muy importante entre ellas, lo que demuestra un gran seguimiento partisano, que enlazaría con la fase de *co-retweeting* de Keller et al. (2019).

A diferencia del resto de casos, un rasgo que distingue la estrategia de diseminación de contenidos alrededor del tema de “Muerte al Rey”, es la presencia de una segunda inundación al comienzo de la mañana siguiente (20 de mayo de 2020); además de un ligero aumento en la presencia de *bots* durante este proceso, aunque los tuits con mayor número de marcas de favoritos provienen de cuentas (usuarios) con poca probabilidad de serlo.

Figura 6. Número de tuits, probabilidad *botmedia* y centralidad de cercanía media en la campaña “Muerte al Rey”



Fuente: Elaboración propia. Nota: Los tweets directos son marcados como “False” y los retuits (RT) como “True”.

La Tabla 4 permite observar lo ya indicado en casos previamente abordados en este trabajo (ej. caso Pedro Duque o etiqueta #SilenciaElPaís): usuarios influenciadores más seguidores que seguidos (Me de seguimiento de 367 y 229 seguidores), con una Me de 8.282 tuits escritos, y 7.242 marcados como favoritos. En este caso, lo más llamativo es la presencia de una cuenta que llega a escribir casi un millón y medio de *tweets*, y cuentas creadas el mismo día de la campaña (19 y 20 de mayo de 2020).

Tabla 4. Datos sobre las cuentas que difunden la campaña "Muerte al Rey"

	Seguidores	Siguiendo	Tuits escritos	Favoritos	Fecha creación
Mínimo	0	0	1	0	25/03/2007
Primer cuartil	74,25	151	2156	1.844	03/02/2012
Mediana	229	367	8282	7.242	28/10/2014
Media	1.025,17	909,1	24.788	21.531	22/03/2015
Tercer cuartil	648,75	908,8	25.874	23.024	12/08/2018
Máximo	269.860	50.850	1.459.308	563.868	20/05/2020

Fuente: Elaboración propia.

Los datos recabados muestran una distribución no uniforme en la fecha de creación de los autores de estos contenidos, donde se observan tres momentos en que fueron creados estos usuarios en Twitter (2011, otoño de 2017 y comienzos de 2020).

Si se atiende a la localización se autoproclaman ser los 11.698 usuarios participantes alrededor del tema, se observan cuentas procedentes de todas las comunidades autónomas y muchas ciudades españolas en esta campaña, destacando: España (1.064), Madrid (959), Barcelona (197), Sevilla (173), Valencia (152), Málaga (72), Zaragoza (49). Hay 7.179 cuentas que no ponen en su perfil de dónde son, y apenas se encuentran cuentas procedentes de fuera de España. Por tanto, la procedencia de todos los mensajes, y por ende el debate generado alrededor del tema analizado ("Muerte al Rey") sería nacional, con especial incidencia en Madrid o España en general.

A partir de una captura por sistema de coordenadas de geolocalización a nivel mundial sobre la API de Twitter, se encuentran diferentes resultados: España (10.521), Filipinas (2.688), Estados Unidos (292) y Venezuela (73). El resto de los tuits no pudieron ser rastreados o encontrada su procedencia exacta mediante geolocalización. Los usuarios procedentes de Asturias, Estados Unidos y Venezuela tienen claramente envíos desde diferentes cuentas del mismo mensaje en un lapso corto de tiempo, dentro del periodo de inundación. De esta manera, se observan ciertos mensajes muy difundidos desde cada uno de dichos lugares.

En el caso de los mensajes emitidos desde Filipinas y Estados Unidos se detectan algunos grupos de mensajes alternados emitidos cada pocos segundos en confrontación, donde se encuentran cuentas de perfiles identificados ideológicamente como de izquierda, que consiguen ser seguidos por otras cuentas afines políticamente, intentando así que el impacto de los contenidos disseminados llegue a más gente y a más ámbitos. En ambos

casos, se observa una muy baja probabilidad de presencia de *bot* durante la fase de ampliación y distribución, aunque durante la fase de inundación aumenta la probabilidad de este tipo de usuarios dentro del debate.

5. Conclusiones

Los datos analizados, permiten confirmar la hipótesis principal de este trabajo, al observar una estrategia en *Twitter*, alrededor de los casos estudiados, asentada en la difusión de contenidos desinformativos viralizados a través de la estrategia *Astroturfing* (Elmas, 2019), liderado por usuarios cuyo perfil en las redes sociales no está asociada con aquellos considerados hasta ahora como influenciadores (caracterizados por tener gran cantidad de seguidores y un perfil social destacado o diferenciador). Esto hace cada vez más difícil distinguir al grupo político o de interés promotor de este tipo de debates analizados, salvo por el hecho de tener en común: 1) la crítica, desde diferentes perspectivas, hacia el gobierno nacional; 2) la presencia de usuarios alfas de muy bajo perfil que lideran el debate en una fase inicial de amplificación y distribución alejados del centro de la red (aprovechando la teoría de fortaleza de enlaces débiles de Granovetter (1973)), y usuarios betas, quienes actúan en los términos expuestos por Sorensen et al. (2017), Keller et al. (2019) y Elmas et al. (2021), es decir, contestando y divulgando los contenidos desinformativos divulgados por los usuarios alfas; 3) con una relativa baja presencia probable de *bots*, durante las primeras fases de viralización de estos temas en *Twitter* apoyada por otra claramente que lo son; y 4) el apoyo que pueden hacer de páginas web creadas o contenidos sacados de contexto para dar “credibilidad” al contenido viralizado por estos en la fase inicial de amplificación (Guess et al., 2018; Zhao et al., 2020).

El escenario de diseminación de los contenidos analizado confirmaría la presencia de “soldados”, en los términos expuestos por Zheng et al. (2017), quienes se apoyan de *bots* para reforzar mensajes (primeras etapas de viralización) o inundar el escenario digital donde se difunden los contenidos desinformativos (últimas fases de viralización). Unos usuarios que, no solo, asumen perfiles afines ideológicamente a los contenidos desinformativos diseminados para condicionar la opinión pública desde las redes sociales (en este caso *Twitter*), sino que también adquieren posiciones opositoras (en operación de falsa bandera), con el fin de atraer la atención de un mayor número de usuarios no necesariamente afectos ideológicamente. Todo ello estaría en la línea de los señalado por Pérez-Curiel & Limón (2019), y el uso de mensajes anti-sistemas y el recurrente auto-posicionamiento como *outsiders*, al momento de extender mensajes donde el uso del sentido nacional (españoles) y calificativos contra el hecho de ser de izquierda, derecha, fascistas, marxista o comunista. Una estrategia de la que, a nivel del caso español, hay pocos estudios que centren su atención en la misma, algo muy importante para poder avanzar con nuevos análisis que permitan ahondar en ello, ante el auge de agrupaciones ideológicamente extremistas o populistas, y de la permanente percepción de conflicto y comunicación populista promovida desde las redes sociales. Son aspectos destacados por

autores como Mazzoleni & Bracciale, (2018), donde la apelación de emociones y creencias reducen la capacidad de influencia de los hechos objetivables.

Los datos mostrados en este trabajo también ayudan a confirmar lo expuesto por autores como Howard et al. (2017), es decir, la puesta en marcha del *Astrourfing online* político. Algo que se estaría llevando bajo tres etapas claramente demarcadas: distribución, amplificación e inundación de contenidos desinformativos a nivel de las redes sociales. Todo ello se plantea con el uso de cuentas nano-*influencers*, que participan con diferentes niveles de intensidad o niveles, dependiendo si el contenido desinformativo resulta verificado o no desde los medios de comunicación tradicionales y proyectos periodísticos profesionales orientados a tales propósitos, por ejemplo. Esto favorece un marco más difícil de detección, con estrategias orientadas a romper el efecto campana y filtro de burbuja de redes sociales como *Twitter*, con el fin de posicionar temas a nivel de la opinión pública (Flaxman et al., 2016).

Contribución de los autores

Sergio Arce-García: Conceptualización, Curación de datos, Metodología, Software, Redacción- borrador original, Análisis formal, Investigación y Redacción- revisión y edición. **Elías Said-Hung:** Conceptualización, Redacción-borrador original, Visualización, Análisis formal y Redacción- revisión y edición. **Daria Mottareale-Calvanese:** Redacción- borrador original, Visualización, Conceptualización, Análisis formal y Redacción- revisión y edición. Todos los autores han leído y aceptado la versión publicada del manuscrito. Conflicto de interés: Los autores declaran no tener conflicto de intereses.

Referencias

- Allem, Jon-Patrick; & Ferrara, Emilio (2018). Could social *bots* pose a threat to public health? *American journal of public health*, 108(8), 1005-1006. <https://doi.org/10.2105/AJPH.2018.304512>
- Blondel, Vincent; Guillaume, Jean-Lup; Lambiotte, Renaud; & Lefebvre, Etienne (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10). <https://doi.org/10.1088/1742-5468/2008/10/P10008>
- Boididou, Christina; Middleton, Stuarthe-E.; Jin, Zhiwei; Papadopoulos, Symeon; Dang-Nguyen, Duc-Tien; Boato, Giulia; & Kompatsiaris, Yiannis (2018). Verifying information with multimedia content on Twitter. *Multimedia tools and applications*, 77(12), 15545-15571. <https://doi.org/10.1007/s11042-017-5132-9>
- Bradshaw, Samantha; Bailey, Hanah; & Howard, Philip-N. (2021). *Industrialized disinformation. 2020 Global inventory of organized social media manipulation*. Working Paper 2021.1. Project on Computational Propaganda. <https://cutt.ly/VOgTjO>
- Chen, Tong; Liu, Jiqiang; Wu, Yalun; Tian, Yunzhe; Tong, Endong; Niu, Wenjia, Li, Yike, Xiang, Yingxiao; & Wang, Wei (2021). Survey on Astrourfing Detection and Analysis from an Information

- Technology Perspective. *Security and Communication Networks*, 2021, 3294610. <https://doi.org/10.1155/2021/3294610>
- Elmas, Tugrulcan (2019). *Lateral Astroturfing Attacks on Twitter Trending Topics*. AMLD EPFL. Lausanne. <https://cutt.ly/4yGaj5L>
- Elmas, Tugrulcan; Overdorf, Rebekah; Özkalay, Ahmed-Furkan; & Aberer, Karl (2021). *Ephemeral Astroturfing Attacks: The Case of Fake Twitter Trends*. arXiv preprint arXiv:1910.07783. <https://arxiv.org/abs/1910.07783>.
- Estrada-Cuzcano, Alonso; Alfaro-Mendives, Karen; & Saavedra-Vásquez, Valeria (2020). Desinformación y desinformación, Posverdad y noticias falsas: precisiones conceptuales, diferencias, similitudes y yuxtaposiciones. *Información, cultura y sociedad*, (42), 93-106. <https://doi.org/10.34096/ics.i42.7427>
- Ferrara, Emilio; Varol, Onur; Davis, Clayton; Menczer, Filippo; & Flammini Alessandro (2016). The rise of social bots. *Communications of the ACM*, 59(7), 96–104. <https://doi.org/10.1145/2818717>
- Flaxman, Seth; Goel, Sharad; & Rao, Justin-M. (2016). Filter Bubbles, Echo Chambers, and Online News Consumption. *Public Opinion Quarterly*, 80, 298–320. <https://doi.org/10.1093/poq/nfw006>
- Gallwitz, Florian; & Kreil, Michael (2021). *The Rise and Fall of "Social Bot"*. Research (March 28, 2021). <https://ssrn.com/abstract=3814191>
- González, Fernán (2020, 20 de mayo). Manifestantes de extrema izquierda gritan "¡Muerte al Rey y a sus hijas!". *Ok Diario*. <https://cutt.ly/CyVjTUR>.
- Granovetter, Mark (1973). The strength of weak ties. *American Journal of Sociology*, 78, 1360-1380.
- Grimme, Christian; Assenmacher, Dennis; & Adam, Lena (2018). Changing Perspectives: Is It Sufficient to Detect Social Bots?. In G. Meiselwitz (eds.) *Social Computing and Social Media. User Experience and Behavior* (pp. 445-461). Lecture Notes in Computer Science, vol. 10913. Springer, Cham. https://doi.org/10.1007/978-3-319-91521-0_32
- Guess, Andrew; Nyhan, Brendan; & Reifler, Jason (2018). Selective Exposure to Misinformation: Evidence from the consumption of fake news during the 2016 U.S. Presidential campaign. *European Research Council*. <https://cutt.ly/FOgUe1R>.
- Hansen, Derek-L.; Shneiderman, Ben, Smith, Marc-A.; & Himerlboim, Itai (2020). *Analyzing Social Media Networks with NodeXL: Insights from a Connected World*. Elsevier. <https://doi.org/10.1016/C2018-0-01348-1>
- Howard, Philip-N.; Bolsover, Gillian; Kollanyi, Bence; Bradshaw, Samantha; & Neudert, Lisa-Maria (2017). *Junk News and Bots during the U.S. Election: What Were Michigan Voters Sharing Over Twitter?* Computational Propaganda Project-Oxford Internet Institute, Data Memo, 1. <https://cutt.ly/kRihRoY>
- Kearney, Michael-W. (2018). Tweetbotornot: An R package for classifying Twitter accounts as bot or not. <https://github.com/mkearney/tweetbotornot>
- Kearney, Michael-W. (2019). Rtweet: Collecting y analyzing Twitter data. *Journal of Open Source Software*, 4(42), 1829. <https://doi.org/10.21105/joss.01829>
- Keller, Franziska-B.; Schoch, David; Stier, Sebastian; & Yang, Jung-Hwan (2019). Political Astroturfing on Twitter: How to coordinate a disinformation campaign. *Political Communication*, 37(2), 256-280. <https://doi.org/10.1080/10584609.2019.1661888>
- Kucharski, Adam (2016). Study epidemiology of fake news. *Nature*, 540(525). <https://doi.org/10.1038/540525a>

- Luceri, Luca; Deb, Ashok; Badawy, Adam; & Ferrara, Emilio (2019). *Red bots do it better: Comparative analysis of social bot partisan behavior*. In Companion Proceedings of the 2019 World Wide Web Conference, 1007-1012. <https://arxiv.org/abs/1902.02765>
- Martin, Shawn; Brown, W.-Michael; Klavans, Richard; & Boyack, Kevin-W. (2011). *OpenOrd: An Open-Source Toolbox for Large Graph Layout*. In Proc. SPIE, Visualization and Data Analysis 2011. San Francisco, Estados Unidos. <https://doi.org/10.1117/12.871402>
- Martini, Franziska; Samula, Paul; Keller, Tobias-R., & Klinger, Ulrike (2021). Bot, or not? Comparing three methods for detecting social bots in five political discourses. *Big Data & Society*, 8(2), 1-13. <https://doi.org/10.1177/20539517211033566>
- Mazzoleni, Gianpietro; & Bracciale, Roberta (2018). Socially mediated populism: the communicative strategies of political leaders on Facebook. *Palgrave Communications*, 4(50). <https://doi.org/10.1057/s41599-018-0104-x>
- Magallón, Raúl (2019). *Unfaking News. Cómo combatir la desinformación*. Pirámide.
- Ong, Jonathan-Corpus; Tapsell, Ross; & Curato, Nicole (2019) Tracking Digital Disinformation in the 2019 Philippine Midterm Election. *New Mandala*. <https://cutt.ly/6RhPHt4>
- Pérez-Curiel, Concha; & Limón, Pilar (2019). Political *influencers*. A study of Donald Trump's personal brand on Twitter and its impact on the media and users. *Comunicación y Sociedad*, 32(1), 57-75. <https://doi.org/10.15581/003.32.1.57-75>
- Pérez, Jordi (2020, 21 de may). "Yo fui un bot": las confesiones de un agente dedicado al engaño en Twitter. *El País*. <https://cutt.ly/wRihXGu>
- Pozzi, Federico-Alberto; Fersini, Elisabetta; Messina, Enza; & Liu, Bing (2017). *The aim of Sentiment Analysis*. Elsevier. <https://doi.org/10.1016/C2015-0-01864-0>
- Ribera, Carles-Salom (2014). Estrategia en redes sociales basada en la teoría de los vínculos débiles. *Más poder local*, 19, 23-25. <https://dialnet.unirioja.es/descarga/articulo/4753468.pdf>
- Said-Hung, Elías; Merino-Arribas, Adoración; & Martínez, Javier (2021). Evolución del debate académico en la Web of Science y Scopus sobre unfaking news (2014-2019). *Estudios sobre el Mensaje Periodístico*, 27(3), 961-971. <https://doi.org/10.5209/esmp.71031>
- Salaverría, Ramón; Buslón, Nataly; López-Pan, Fernando; León, Bienvenido; López-Goñi, Ignacio; & Erviti, María-Carmen (2020). Desinformación en tiempos de pandemia: tipología de los bulos sobre la covid-19. *El profesional de la información*, 29(3). <https://doi.org/10.3145/epi.2020.may.15>
- Sorensen, Anne; Andrews, Lynda; & Drennan, Judy (2017). Using social media posts as resources for engaging in value co-creation: The case for social media-based cause brand communities. *Journal of Service Theory and Practice*, 27(4), 898-922. <https://doi.org/10.1108/JSTP-04-2016-0080>
- Tandoc, Edson-C.; Lim, Zheng-Wei; & Ling, Richard (2018). Defining "fake news" A typology of scholarly definitions. *Digital journalism*, 6(2), 137-153. <https://doi.org/10.1080/21670811.2017.1360143>
- Van-der-Linden, Sander; Maibach, Edward; Cook, John; Leiserowitz, Anthony; & Lewandowsky, Stephan (2017). Inoculating Against Misinformation. *Science*, 358(6367), 1141-1142. <https://doi.org/10.1126/science.1257583>
- Van-der-Veen, Han; Hiemstra, Djoerd; Van-den-Broek, Tijs; Ehrenhard, Michel; & Need, Ariana (2015). *Determine the User Country of a Tweet*. *Social and Information Networks*. <https://arxiv.org/abs/1508.02483>

- Zerback, Thomas; & Töpfl, Florian (2021). Forged Examples as Disinformation: The Biasing Effects of Political Astroturfing Comments on Public Opinion Perceptions and How to Prevent Them. *Political Psychology*, 43(3), 399-418. <https://doi.org/10.1111/pops.12767>
- Zhao, Zilong; Zhao, Jichang; Sano, Yukie; Levy, Orr; Takayasu, Hideki; Takayasu, Misako; Li, Daqing; Wu, Junjie; & Havlin, Shlomo (2020). Fake news propagates differently from real news even at early stages of spreading. *EPJ Data Science*, 9(7). <https://doi.org/10.1140/epjds/s13688-020-00224-z>
- Zheng, Haizhong; Xue, Minhui; Hao, Lu; Hao, Shuang; Zhu, Haojin; Liang, Xiaohui; & Ross, Keith (2017). Smoke Screener or Straight Shooter: Detecting Elite Sybil Attack. *Social and Information Networks*. <https://arxiv.org/abs/1709.06916>