

INFORME 2025

Los nuevos contornos de la desinformación



Los nuevos contornos de la desinformación

Resumen

Este estudio tiene como objetivo actualizar el catálogo de formatos utilizados para la desinformación y los discursos de odio durante la campaña electoral de las elecciones europeas del 9 de junio de 2024 y entender cómo la inteligencia artificial es aplicada en técnicas de manipulación digital. La identificación de formatos se hizo a partir de la escucha social realizada desde ocho perfiles de redes sociales creados en Instagram, Facebook, X y TikTok a los que se dotó de una personalidad propia, con perfiles ideológicos o neutros y distintos campos de interés y experiencias de usuarios. El trabajo de observación permitió constatar la generalización de las recomendaciones polarizadoras por parte de las redes sociales, así como la emergencia y evolución de nuevos formatos de desinformación, a menudo alimentados con inteligencia artificial (IA) cada vez más sofisticados y extendidos.

Palabras clave

Desinformación, elecciones, redes sociales, inteligencia artificial, polarización, política.

Los autores



Carmela Ríos

Periodista.

<https://www.linkedin.com/in/carmela-r%C3%ADos-calvo-1103a12b/?originalSubdomain=es>



Rafael Rubio

Catedrático de Derecho Constitucional, Universidad Complutense.

<https://www.linkedin.com/in/rafa-rubio-18616b1/?originalSubdomain=es>



José Suarez de Lezo

Director del Laboratorio de Periodismo

Fundación Luca de Tena

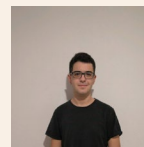
<https://www.linkedin.com/in/jos%C3%A9-suarez-de-lezo-b550507/?originalSubdomain=es>



Luis Jurado Castillo

Asistente de investigación.

<https://www.linkedin.com/in/luisjuradocastillo/?originalSubdomain=es>



Enrique Aparicio Carracedo

Asistente de investigación.

<https://www.linkedin.com/in/enrique-aparicio/?originalSubdomain=es>





Introducción

Las campañas electorales constituyen, como señala el catedrático Rafael Rubio “el punto central de la vida política, la clave de bóveda sobre la que se sostiene el sistema democrático, al proveer al legítimo detentador del poder de una forma de selección periódica de los gobernantes, y a estos de la legitimidad de origen para desempeñar su labor.” (1) El funcionamiento de esta herramienta esencial para la vida de las democracias vive una amenaza creciente desde 2010, fecha a partir de la cual las técnicas de manipulación psicológicas desarrolladas a partir del uso del big data y con las redes sociales como campo principal de acción (2) han facilitado que los procesos electorales de todo el mundo se vean alterados por la acción de actores domésticos e internacionales cuyo objetivo es sembrar la duda sobre la credibilidad de las instituciones de la democracia y las garantías de sus sistemas de votación

“Los ingenieros del caos”, en denominación del ensayista “Giuliano D’Empoli” (3) han perfeccionado desde hace décadas un eficaz sistema de comunicación digital, destinado a alterar la percepción de la realidad de amplios sectores de la población, que se asienta en las vulnerabilidades de las sociedades actuales. Como señala D’Empoli “En Europa, como en otros lugares, las mentiras están en boga porque se funden en un relato político que capta los miedos y las aspiraciones de una parte creciente del electorado (...)” “El carnaval contemporáneo se nutre de dos ingredientes que no tienen nada de irracional: la ira de algunos ámbitos de la clase trabajadora, que se alimenta de motivos sociales y económicos reales, y una máquina de comunicación imponente”

El impacto de las campañas de manipulación dentro de las redes sociales en periodo electoral ha podido ser observado en su justa medida desde 2017, fecha a partir de la cual investigadores universitarios, expertos en verificación y equipos especializados dentro de grandes medios de comunicación pudieron realizar recogidas significativas de material desinformativo y dispusieron de herramientas para el análisis de grandes flujos de comunicación y de comprobación multimedia. (4)

Desde las elecciones presidenciales en Estados Unidos de 2016 los procesos electorales han sufrido una profunda transformación silenciosa. De un lado las operaciones de desinformación desarrolladas en entornos digitales no operan bajo ningún condicionante temporal de tiempo, es decir, puede desarrollarse de forma continuada, como si las campañas electorales no tuvieran principio ni fin. Los actores de la desinformación necesitan tiempo para que sus campañas “filtren” en el corazón de las sociedades a través de las cuentas de las redes sociales por lo que la misma concepción de campaña electoral como un periodo definido en el tiempo por la autoridad nacional se ha visto ampliamente superado.

La segunda transformación significativa tiene que ver con la mutación narrativa que ha presentado la desinformación en la última década dada la aparición de nuevas funcionalidades en las redes sociales y la emergencia de las técnicas de inteligencia artificial que han colocado la desinformación en otro nivel de sofisticación y extensión geográfica. El número de formatos para la desinformación y la polarización política se ha multiplicado, más bien, han terminado incorporándose a los que ya existían. Podríamos hablar de la era de desinformación amontonada. Vemos en 2024 bulos o montajes cuyo contenido o narrativa se repite desde hace varios años y que constituyen un patrón repetido en procesos electorales en distintos lugares del planeta,





como es el caso de los referidos a la poca seguridad de los procesos de votos por correo y las supuestas irregularidades en colegios electorales o en las mecánicas de recuento electoral. Se trata de lo que podíamos denominar “bulos comodín” o recurrentes que aparecen durante la campaña y suelen desaparecer una vez celebrados tras la celebración de los comicios.

La historia reciente, sin embargo, nos recuerda que, movidas con efectividad y desde cuentas con influencias social, estas afirmaciones falsas pueden acarrear consecuencias de una singular gravedad como sucedió en Washington en enero de 2021, cuando centenares de estadounidenses asaltaron el Capitolio convencidos de un “pucherazo” electoral en favor del candidato Joe Biden.

El trabajo que nos ocupa tiene como objetivo identificar nuevas formas de desinformación atendiendo a los dos factores anteriormente mencionados: la prolongación en el tiempo de la acción política fuera del radar oficial y la coexistencia de formatos antiguos y novedosos. Establecimos un periodo de observación y escucha digital que se inició el 1 de marzo de marzo de 2024, tres meses antes la celebración de las elecciones europeas y que se prolongó hasta dos días después de las elecciones.

El objetivo con esta ampliación del campo temporal era detectar formatos y actores de la “campaña permanente” antes de que la conversación electoral se viera alimentada con otros actores, bien oficiales o que las plataformas digitales se mostraran más reactivas a la hora de eliminar la desinformación, empujadas por la presión de los reguladores nacionales, las autoridades europeas o los medios de comunicación y expertos en verificación, instituciones que se han mostrado, durante la campaña europea, atentas y activas a la hora de revelar y denunciar campañas de desinformación. Durante el estudio pudimos, efectivamente, observar una evolución en la incidencia y los formatos a medida que avanzaba la campaña europea.



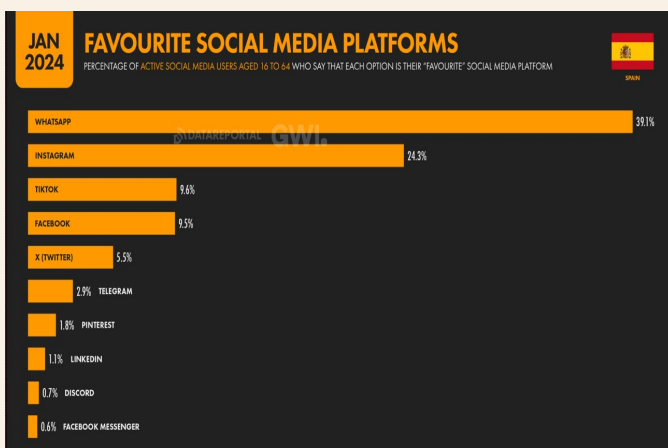
01

Metodología



Una vez definido el periodo de observación, creamos para este trabajo los puestos de observación que nos permitiría hacer una escucha rica y representativa en la que viéramos reproducida la experiencia de usuario de un grupo de ciudadanos de edad, procedencias, categoría socioeconómica e intereses muy diferentes. Centramos nuestra actividad en las cuatro redes sociales más utilizada por los ciudadanos españoles, según el último informe de Hootsuite sobre consumo digital en España (5). Se trata de Instagram, utilizada por el 24,3% de los usuarios españoles de redes sociales, TikTok (9,6%) Facebook (9,5%) y X (5,5%)

Para ello decidimos crear ocho perfiles en distintas redes sociales, TikTok, Instagram, Facebook y X (Twitter) Imaginamos un ramillete de trazos biográficos y profesionales para establecer para cada uno de ellos, una identidad a partir de la cual definir el perfil y, con él, una experiencia de usuario. Decidimos que cada uno de nuestros ciudadanos tuviera una “conexión” vital o familiar con algunos países de la Unión Europea para abrir los contenidos a cuentas y publicaciones que no fueran sólo españolas, dado que nos encontrábamos en unos comicios supranacionales. A fin de vincular los perfiles con los temas centrales de la campaña europea, identidad nacional, inmigración, crisis climática, soberanía, salud electoral o energía, cada uno de los perfiles mostró con su actividad un interés particular por alguno de estos temas.



Acordamos asimismo que la actividad de estas cuentas tendría como único objetivo “entrenar” al algoritmo de recomendación de cada red social, es decir, ofrecer rasgos suficientes de nuestra identidad digital para recibir contenidos personalizados. Para ellos limitamos al máximo las interacciones con otros usuarios y basamos nuestro diálogo en las redes sociales en mensajes de aprobación y algunas publicaciones simples.

Hemos reproducido un método de trabajo al que han recurrido medios e investigadores de todo el mundo con el objeto de reproducir la experiencia de usuario de distintas categorías de ciudadanos. Así, en 2023, la fundación 5Rights y el organismo investigador Revealing Reality pusieron en marcha el proyecto “Pathways” (6) Su objetivo era investigar cómo el diseño de las plataformas digitales afecta la experiencia de los niños en redes sociales como Instagram. Para esto, se crearon perfiles falsos, o "avatares", basados en comportamientos reales de niños y adolescentes, para simular su experiencia sin poner en riesgo a ningún menor real.





Cada perfil de nuestro objeto de estudio fue distinguido con una identidad gráfica ficticia, que construimos a partir de una fotografía generada con la herramienta "This person doesn't exist", una aplicación de inteligencia de generación de rostros humanos en función de unos parámetros y criterios establecidos.

Identidad y características de los perfiles del estudio

Manu. Es el único usuario menor de edad de nuestro estudio. Resultaba del máximo interés para nuestro estudio reproducir la experiencia de un usuario de un joven menor de edad que, como millones de adolescentes, falsean su edad real a la hora de registrarse en una red social, en este caso TikTok, y utilizan su actividad digital para el ocio. Su vinculación con Europa estaba asegurada con un familiar en Francia. Ideológicamente Manu es un perfil neutro, sin preferencias políticas, con más interés por el deporte, la música, las fiestas de su Sevilla natal o las cuentas de humor con las que pasar un buen rato con su teléfono móvil.

Manu es, junto a David, el perfil más joven del estudio, lo que nos parecía de todo interés dado la existencia de investigaciones que señalan la especial vulnerabilidad de los usuarios más jóvenes respecto a los contenidos polarizadores o la desinformación. En este sentido uno de los de estudios más recientes, publicados en julio de 2024 por el think tank alemán Interface y la organización AI Forensics, documenta cómo los algoritmos de recomendación de TikTok condujeron a los jóvenes electores alemanes hacia contenidos del partido de extrema derecha AfD de forma preferente cuando estos buscaban información de partidos políticos. (7) La experiencia de usuario de Manu en TikTok se reveló útil para documentar experiencias similares.



La vida de Manu

Manu tiene 15 años y estudia 4º de la ESO en un instituto de Sevilla. Pronto tendrá que elegir si estudia un carrera o accede a la FP. Vive junto a sus padres, que mantienen una pequeña explotación agrícola. Un hermano del padre de Manu trabaja y vive en Toulouse (Francia)



Manu es un chaval extrovertido y le encanta Tik Tok como a todos sus amigos. Sigue cuentas de música, influencers y sigue cuentas relacionadas con la situación del agro que tanto preocupa en casa. Como chapurrea el francés, gracias a las estancias con su tío, sigue a usuarios franceses.




David. Situamos en Fuenlabrada, en la comunidad de Madrid, el segundo perfil joven de la serie. Se trata de David, un estudiante de informática, aficionado al fútbol, sin grandes inquietudes políticas, pero sí interesado por la cuestión de los flujos migratorios en Europa y de las supuestas entradas masivas de inmigrantes a través de las costas italianas de las que alerta un familiar residente en aquel país. David mantiene perfiles en TikTok e Instagram.

La vida de David

David tiene 18 años y vive con su madre en Fuenlabrada (Madrid) Su padre es italiano y volvió a Catania cuando la pareja se divorció. El hombre rehizo su vida y es padre de una niña de 10 años. David estudia un grado superior de EP en Informática.





David es un loco del Real Madrid y del trap. En Tik Tok e Instagram sigue a músicos, periodistas deportivos. Su padre, con el que pasa los veranos, está muy preocupado por la inmigración masiva en Italia así que ha empezado a seguir varias cuentas que hablan de este problema.



Aitana. De padre francés y madre española, Aitana es la usuaria binacional “pura” de nuestro estudio. Quisimos introducir un perfil capaz de recibir y comprender información en dos idiomas y más interesado por la actualidad que los perfiles jóvenes, un tipo de contenido que busca activamente en las redes. Ideológicamente, Aitana se sitúa en opciones progresistas, pero, sobre todo, muestra un interés abierto por las grandes cuestiones medioambientales que le preocupan especialmente.

La vida de Aitana

Aitana tiene 25 años y es Barcelona acaba de terminar Bellas Artes y estudia un máster de Interiorismo. De padre catalán y madre francesa, Aitana es bilingüe. Disfruta de una vida cómoda y estable. Es una apasionada del yoga y del arte abstracto.







Aitana es una persona muy concienciada con los grandes desafíos medioambientales que vive su generación, desde la escasez de agua a la crisis climática. Por eso sigue en redes sociales cuentas relacionadas con sostenibilidad, además de “influencers” de cultura y arte, además de medios franceses y españoles.

Pedro. Madrileño, economista de profesión, Pedro representa la centralidad en la política y la moderación como norma. Consideramos del máximo interés introducir en el estudio un perfil de redes sociales cuya experiencia de usuario le convirtiera, al menos sobre el papel, en incompatible con entornos polarizadores o publicaciones más radicales. Con perfiles en X e Instagram, Pedro usa las redes sociales para seguir a periodistas y medios de referencias y estar al día en temas de gastronomía, una de sus grandes pasiones.

La vida de Pedro

Pedro tiene 37 años, es economista, vive en Madrid y trabaja en una importante consultora de Casado, estudió parte de su carrera en Alemania, país del que admira muchas cosas y cuya situación política sigue de cerca.







Pedro huye de radicalidades, le gusta el centro político y es un europeísta convencido aunque tiene algunas dudas vista la situación en Alemania. En sus cuentas de Instagram y Twitter sigue a medios y periodistas españoles y alemanes y muchas cuentas de gastronomía porque le encanta cocinar.

Mar. Se trata, junto a Mercedes, del perfil ideológicamente más definido de todo el estudio. Una votante de EH Bildu, con una fuerte convicción soberanista y comprometida con causas sociales. Introdujimos en este perfil una variante específica: su preocupación por la inseguridad (19) que, parece, protagonizan inmigrantes en sus barrios. Queríamos ver cómo respondía a esta inquietud el algoritmo de recomendación de Facebook y X, las dos redes elegidas por Mar para su vida digital.

La vida de Mar

Mar es de Bilbao tiene 46 años y es profesora de Educación Especial en un colegio de Baracaldo. Le encanta su trabajo. Se divorció hace seis años. Le encanta comer y practicar senderismo todos los fines de semana. Tiene familia en San Juan de Luz, a los que visita con frecuencia.

Mar vota a EH Bildu desde hace tiempo. Es miembro del Comité de Empresa del centro. Este año está preocupada porque la delincuencia no deja de aumentar. Tiene varias amigas a las que le han robado el móvil en la calle y los autores parecen ser siempre inmigrantes. ¿Qué hacer?

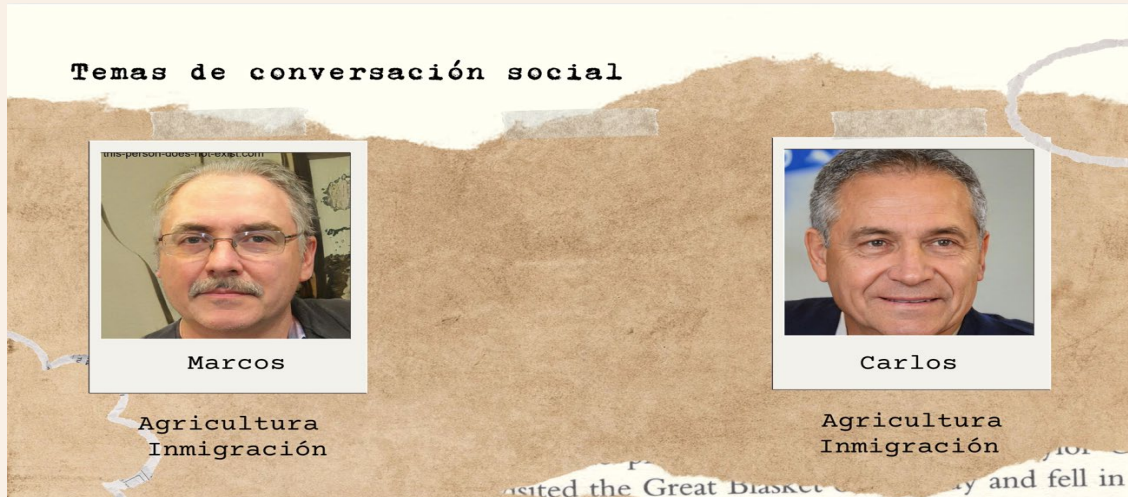
Carlos. Buscamos para Carlos el perfil de un empresario almeriense, propietario de un próspero negocio en Almería. De una orientación ideológica conservadora y moderada. buscábamos con este perfil recrear una experiencia de usuario de un ciudadano poco dado a los excesos pero preocupado por la evolución de la crisis climática e interesado por la cuestión migratoria.

La vida de Carlos

Carlos tiene 55 años, está casado con dos hijos. Goza de una buena situación económica. Es propietario de uno de los primeros y mejor situados supermercados de la provincia de Almería, donde nació y reside desde siempre.

Carlos, uno de cuyos hijos está en Italia con una beca Erasmus, está preocupado por el impacto del clima en el campo y la gestión de la inmigración, que es tan importante para una región como la suya con una alta demanda de mano de obra. Está en Twitter y Facebook

Marcos. Una restricción temporal impuesta por X al perfil de Carlos durante algunas semanas nos llevó a replicar el mismo perfil y la misma identidad digital en otra cuenta de X. De esta forma, nuestro empresario de Almería siguió llamándose Carlos en Facebook, pero cambió su nombre en X por el de Marcos. Para él también generamos un nuevo rostro con la herramienta This Person Doesn't Exist.



Mercedes. La veterana del grupo es también la que ideológicamente presenta un color más definido. Preocupada por la unidad nacional y los temas de inmigración, Mercedes es una votante de las opciones de la derecha radical, entorno del que busca contenidos dentro de las redes sociales para conectar con otros descontentos como ella. Mercedes tiene otros centros de interés como la copla y las celebraciones religiosas de su ciudad.





Categorización de la desinformación para el estudio

Una vez establecido los perfiles y definido nuestro margen de actuación como gestores de perfiles de mera observación ideológica, iniciamos a principios de marzo la tarea de recogida de información de cada perfil y, sobre todo, de la detección de cualquier material desinformativo o polarizador. En función de la información analizada se han podido definir estos tipos de desinformación o contenidos polarizadores.

1. **Bulos genéricos** - Bulo o desinformación que no tiene mayor objetivo aparente que el de desinformar, sin causar mayores daños añadidos.
2. **Contenido de odio** Hace referencia a un texto, fotografías, vídeos u otro formato de carácter ofensivo dirigido a un grupo o individuo y que se basa en características inherentes (como son la raza, la religión o el género)
3. **Estafa financiera** - Publicaciones cuyo objetivo es captar dinero de los usuarios con la promesa de grandes rendimientos a través de plataformas financieras que no existen.
4. **Clickbait** - Contenido destinado a generar consultas, visualizaciones y acciones de compartir generalmente con fines económico.
5. **Contenido radical** - Publicaciones en redes sociales de opciones ideológicas extremas y/o que no corresponden al perfil ideológico del usuario que las recibe.
6. **Falso medio** - Páginas web que reproducen la estética de un medio de comunicación sin respetar sus normas deontológicas.
7. **Estafa económica** - Publicaciones que persiguen atraer el dinero de los usuarios bajo la promesa de un ingreso de dinero fácil.
8. **Bulo medioambiental** Material que, en las redes sociales, busca negar el cambio climático o generar controversia o visualizaciones.
9. **Bulo electoral**- Son aquellas publicaciones falsas en torno a la mecánica electoral y que tienen como objetivo lanzar dudas sobre la limpieza de las instituciones encargadas de organizar las elecciones.
10. **Contenido con IA** - Publicaciones, ya sea texto, vídeo, fotografía o ilustraciones generadas con inteligencia artificial.
11. **Contenido prorruso** - Hace referencias al material de apoyo al régimen ruso o publicaciones de procedencia rusa cuyo objetivo es la crítica de las naciones occidentales y sus instituciones.





Se decidió establecer dos áreas o espacios de observación que nos permitieran englobar los contenidos con más de un contenido desinformativo.

1. **Área del bulo primario.** El que contiene la desinformación principal.
2. **Área del bulo secundario.** Alude a la técnica con la que fue realizado el bulo primario o especifica un segundo tipo de desinformación concentrado en el mismo material. Por ejemplo: contenido radical realizado con IA, o estafa financiera desde un falso medio.

Siguiendo este método, entendemos que la información significativa y relevante aparece en la primera área, bulo primario mientras que la segunda, bulo secundario, resulta de utilidad para completar la información sobre la desinformación principal.

Herramientas de análisis y verificación usada en el estudio

Con objeto de afinar el análisis de la desinformación incorporamos a nuestro trabajo un catálogo de aplicaciones, programas, y herramientas de verificación de información en diferentes soportes, así como el acceso a los repositorios de los anuncios en distintas redes sociales.

Las herramientas usadas por el equipo fueron, principalmente, las siguientes:

1. Biblioteca de anuncios de Meta. (Facebook e Instagram)
2. Twitter Ads Repository
3. TikTok Ads Center.
4. Google Images.
5. Google Trends
6. Google Fact Check Explorer
7. Election24 Check
8. Yandex Pictures.
9. Wayback Machine.
10. AI Detector
11. The Markup
12. TrueMedia.org
13. Trends24.in
14. Telegram Analytics.

De todo el caudal de publicaciones recibidas por todos los perfiles se decidió analizar todos aquellos contenidos que presentaban desinformación y mensajes polarizadores y que permitieron definir una pauta concreta. A partir de estas pautas este estudio define e identifica tipo de formatos de desinformación. El total de posts seleccionados y analizados bajo estos parámetros fue de 168 publicaciones en el periodo de tiempo establecido.





02

Resultados

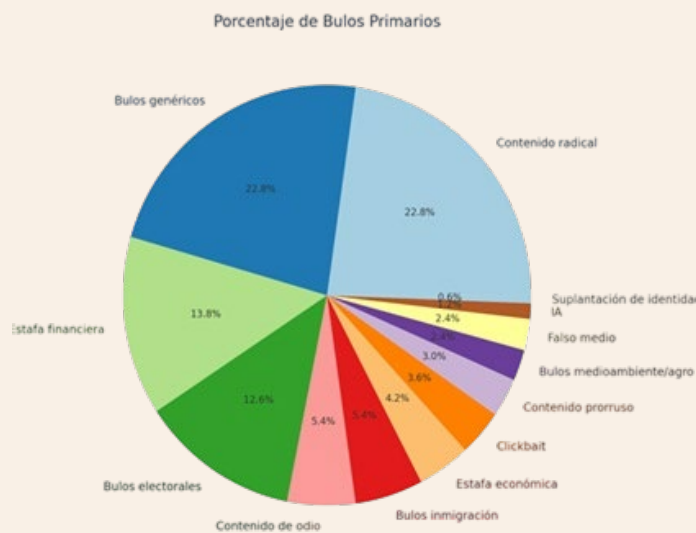




De todo el caudal de publicaciones recibidas por todos los perfiles se decidió analizar todos aquellos contenidos que presentaban desinformación y mensajes polarizadores permitían definir e identificar distintos patrones de desinformación o publicaciones polarizadoras que habían circulado por Facebook, Instagram, Twitter y TikTok durante la precampaña y la campaña europea. Del material recogido entre los meses de marzo a junio de 2024 realizamos una selección de los 168 contenidos. Los perfiles sociales de nuestras “personas” digitales arrojaron distintas experiencias de usuarios y aportaron material adaptado por el algoritmo pero que, con frecuencia, presentaba denominadores comunes.

Tipos de desinformación

Siguiendo el criterio de categorías que establecimos para este estudio, constatamos que dos categorías se reparten el grueso del material desinformativo que nuestros perfiles recibieron. Los bulos genéricos suponen el 22,8% del material mientras que lo que denominados contenidos radicales, aquellos que se sitúan en una posición más extrema de la que algorítmicamente corresponde al perfil, constituyen también el 22,8% del material analizado.



Cabe destacar el alto número de bulos relacionados con estafas financieras algo que podría deberse a la coincidencia del estudio con el fuerte impacto que ha tenido en Europa una campaña de estafa financiera con criptomonedas lanzada desde Rusia para varias redes sociales y en varios países. Las primeras referencias de esta organización aparecieron durante el mes el pasado mes de enero cuando una empresa británica de marketing detectó una primera campaña fraudulenta con la imagen del entonces primer ministro británico, Richi Sunak como señuelo. (7) Desde entonces diversas organizaciones y diarios como Le Monde (8) han destapado la mecánica de estas redes criminales que tienen un impacto significativo en las redes sociales y que, como vimos en el estudio, impactaron en cuatro de los ocho perfiles de estudio: Carlos (3), David (2) Mercedes (1) y Pedro (1).

La desinformación de carácter electoral también ha ocupado un espacio importante de del material recibido por las identidades digitales objeto de este estudio, algo que no resulta extraño dado que el periodo de





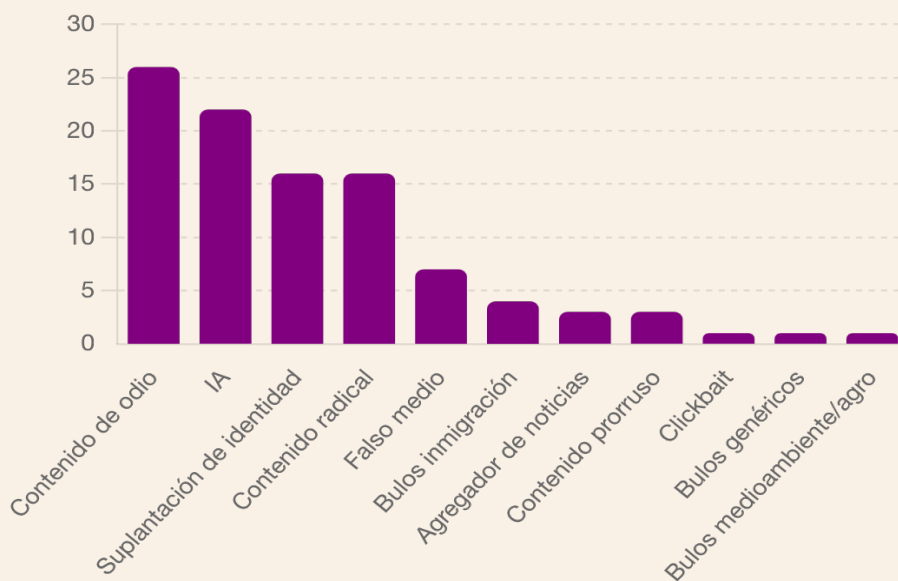
observación se estableció durante los meses previos a la campaña de las elecciones europeas. Sin embargo, las convocatorias anticipadas de los comicios autonómicos en el País Vasco, el 21 de abril y el Cataluña el 20 de mayo, hicieron que la conversación digital, orgánica o inferida, durante este periodo fuera aún más intensa. Los bulos electorales ocupan el 12,6% del material analizado y han impactado sobre seis de los perfiles: Marcos (6 contenidos) Pedro (6) Aitana (4) Carlos (3) Mercedes (2) y David (19).

A menudo, concomitantes, los contenidos de odio y la desinformación sobre la actualidad de los inmigrantes representan casi el 12% del material estudiado y ha incidido con mayor intensidad en los perfiles masculinos: en el caso de los contenidos de odio, David encabeza el listado de receptores, con 5 contenidos, seguido de Pedro con 2 contenidos. Marcos, el empresario de Almería, encabeza la recepción de bulos sobre inmigración (6) mientras David repite en los primeros puestos también en esta categoría (2).

En este sentido es importante resaltar que, dado que los perfiles tenían una conexión con algún país de la Unión Europea, el flujo de la desinformación sobre inmigración también fue internacional. Los contenidos analizados se refieren a supuestas noticias en España, pero también fuera de España como las presuntas avalanchas de inmigrantes en la isla italiana de Lampedusa, como sucedió en el caso de David, como veremos en detalle posteriormente.

Incidencia de bulos secundarios

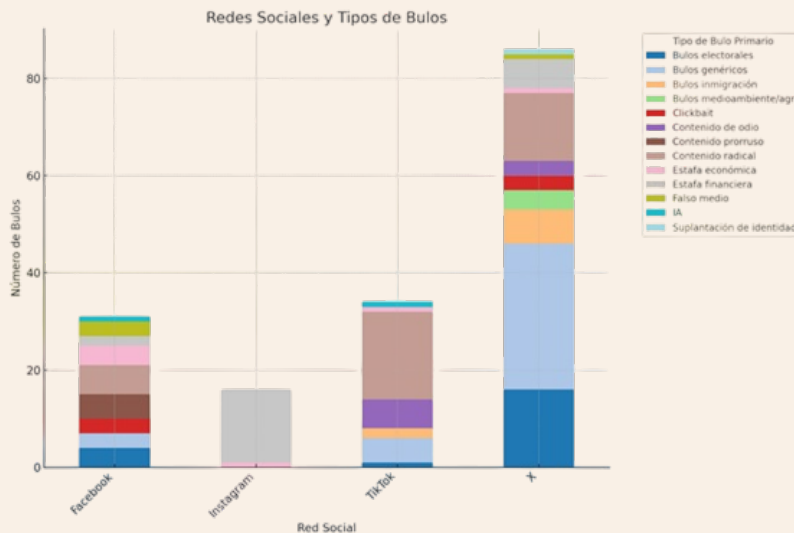
Si bien, la categoría de bulo secundario tiene como objetivo principal poder recoger la información relativa a la creación de la desinformación o la coexistencia de dos narrativas desinformadoras, ha resultado de gran utilidad para identificar la emergencia de la inteligencia artificial (9) como la tecnología base con la que se fabrica una parte cada vez más significativa de los bulos o los contenidos polarizantes. Comprobamos también como el recurso a las técnicas de suplantación de identidad gana terreno a la hora de idear acciones fraudulentas.





Incidencia por redes sociales

Dos redes sociales, X y TikTok, coparon el grueso de la desinformación recibida por las cuentas objetos de estudio%. Las desinformaciones procedentes de X o Twitter supusieron el 51,50% del material mientras TikTok acaparó el 20,6%. Las redes de Meta cierran la lista: en Facebook se registró el 18,56% de las publicaciones analizadas mientras Instagram sólo aportó el 9,58 de los contenidos del estudio.



Estos resultados resultan del todo coherentes con la realidad en materia de comunicación electoral en redes sociales, que hemos visto evolucionar a lo largo de 2024. Desde la llegada de Elon Musk a la presidencia de X en abril de 2022 esta red social ha visto drásticamente reducidos sus equipos de moderación de contenidos que resultaban indispensables para, al menos, tratar de contener el caudal de desinformación y odio que registra la red en todo el mundo. Estos mecanismos han sido sustituidos por otros, como las denominadas “notas de la comunidad” en la que son los propios usuarios lo que se encargan de verificar los contenidos de otros. Ello, unido a la decisión de Elon Musk de “liberalizar” los contenidos de X lo más posible, nos lleva a la situación actual en la que la desinformación es poco controlada.

En este punto resulta pertinente recordar que la Unión Europea mantiene abierto un procedimiento contra X desde el pasado mes de diciembre por incumplimiento de sus obligaciones previstas en la Ley de Servicios Digitales en lo que se refiere a la difusión de contenidos ilícitos y la lucha contra la manipulación de la información en la plataforma. El pasado 12 de julio, la Comisión Europea reforzó su acusación por el masivo uso fraudulento que, dentro de X, se hace de las cuentas verificadas previo pago con un distintivo azul, convertidas en otro elemento con los que los actores de la desinformación pueden engañar a los usuarios, como recuerdan las autoridades europeas.

Además de X, TikTok es el segundo vector de desinformación (9) según los datos de este estudio. El material analizado permite constatar cómo las poderosas capacidades narrativas y viralizadoras de TikTok, con sus canciones tendencias y sus “challenges” son aprovechadas para propagar contenido desinformativo y radical. Es de resaltar que este material se mimetiza perfectamente entre contenidos “blancos” ya que usa formatos creativos y el tono, a menudo humorístico que permite que los contenidos filtren en categorías de usuarios jóvenes como Manu o David, más inclinados a seguir cuentas de ocio, deporte o hiperlocales.



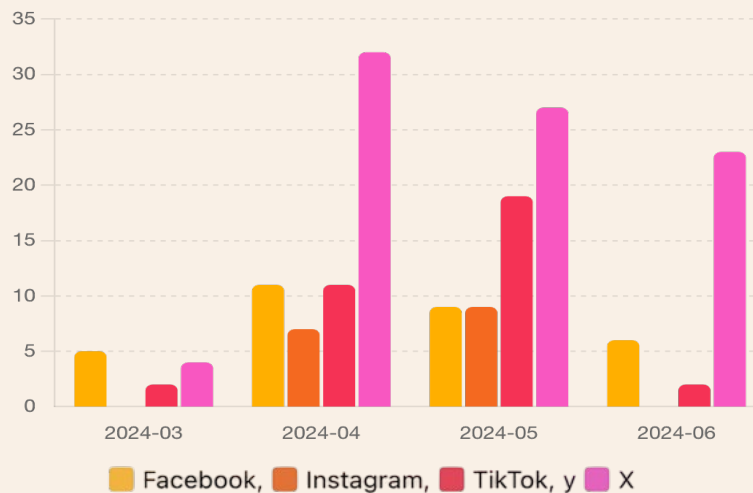


Conviene recordar que, en los últimos años, TikTok ha implementado nuevas políticas para proteger la integridad electoral. Dirigen a los usuarios hacia información autorizada y coloca etiquetas con enlaces en las publicaciones relacionadas con las elecciones. También ha prohibido los anuncios políticos. Sin embargo, todos estos esfuerzos no parecen suficientes para contener el caudal de desinformación que circula desde cuentas anónimas. El algoritmo parece, además recomendar, dentro de los procesos electorales, los temas que resultan más controvertidos, es decir, que generan una mayor viralización, como sugiere uno de los estudios del centro de estudios del algoritmo CrossOver, con sede en Finlandia. (12)

En tercer lugar de la clasificación de redes con desinformación, Facebook aloja una novedosa variedad de formatos de bulos y contenidos polarizantes: desde las publicaciones dentro de grupos hiperlocales, como veremos más adelante, a las ilustraciones con IA. Cabe destacar que los actores desinformación recurren con frecuencia a la difusión pagada gracias a la poderosa herramienta publicitaria que Meta ponen a disposición de los anunciantes. La red de estafas financieras a las que aludíamos anteriormente, ha recurrido masivamente a este recurso de Facebook para extender la difusión de sus publicaciones fraudulentas.

Evolución temporal en la incidencia de la desinformación

El periodo de observación objeto de nuestro estudio arrojó diferencias notables entre los distintos meses respecto al caudal de la información recibido. En el caso del mes de junio resulta lógica que el material sea inferior en cantidad ya que el trabajo de observación se detuvo con la celebración de las elecciones europeas, el día 9 de junio. Sin embargo, percibimos diferencias significativas en los tres meses anteriores, con un mes de marzo poca actividad y los meses de marzo y abril con volúmenes más alto y ligeras oscilaciones en el comportamiento de cada cuenta y de cada red social en cada mes.



Los datos muestran claramente como la prevalencia de X y TikTok como escenarios de la desinformación es constante, así como el papel más moderado que juegan redes como Facebook e Instagram, red esta última donde la mayor parte de la desinformación analizada se refiere a estafas. Dada la presión recibida por los responsables de Meta para que realizaran vigilancia más estrecha sobre la red en periodo electoral, como señalaban medios de información como la cadena France 24, (es inferible que la movilización de efectivo y esfuerzos para la moderación de contenidos tuviera como consecuencia una disminución en el número de desinformación política.





Manu Torres, humor y polarización en bandeja

La experiencia de usuario de un menor de edad

¿Puede quedar un usuario de TikTok menor de edad al margen de la conversación sobre unas elecciones europeas si no da muestras de interés por ella ni por la política en general? ¿Qué sucede cuando un adolescente busca en TikTok información sobre la actualidad? La identidad digital de Manu fue creada con la intención de responder a estas preguntas. Por esta razón se diseñó el perfil y su experiencia de usuario de la forma más neutra posible: un menor de edad sevillano que, como otros muchos jóvenes de su edad no tiene grandes inquietudes ideológicas y cuya principal preocupación era, en este caso, la situación del negocio familiar, una pequeña explotación agrícola. La elección de esta “vulnerabilidad” estaba justificada por el creciente malestar de los agricultores europeos por los rigores de la normativa europea y la reducción de los márgenes de beneficio de sus cosechas.

Manu es, por tanto, un perfil ideológicamente neutro con intereses que se ven reflejados en las cuentas que siguen: futbolistas, clubes como el Betis FC o el Sevilla FC, así como cantantes andaluces y músicos de rap, reggaetón o flamenco como @OmarMontes, @luis cortés_sbr o @_theyounggypsy e “influencers” como @MartaDíaz o @IbaiLlanos. Manu incorporó a su cuenta de TikTok numerosas cuentas de humor y parodia como @gabimuozruiz o así como perfiles como @Semantasantadesevilla o numerosas hermandades de Sevilla ya que vive con mucha pasión. Como medio de actualidad, Manu ha optado por seguir a @ac2ality y @101tvsevilla. Dado su interés por el campo, el joven empezó a seguir a @elcampoesnuestro o @g_marfer, una joven ganadera y agricultora granadina que comparte con sus seguidores escenas de su actividad cotidiana.

Evolución en los contenidos sugeridos para Manu por el algoritmo de recomendación de TikTok

Durante las primeras semanas del estudio, en el mes de marzo, los contenidos que Manu recibía en su feed respetaban en general la gama de preferencias que revelaban las cuentas que había decidido seguir inicialmente. Hay que tener en cuenta que la Semana Santa de Sevilla tuvo lugar entre el 21 y el 28 de marzo, por lo que existía una gran circulación de contenidos relacionados con esta temática que llegaban masivamente a la cuenta de Manu como material sugerido. TikTok sugirió además nuevas cuentas de humor y de usuarios relevantes de la comunidad andaluza.

Contenidos políticos recomendados para un perfil neutro

Sin embargo, los contenidos políticos se presentaron en el feed. El primero llegó el 23 de marzo: un vídeo del partido Vox con un fragmento de un discurso de su presidente, Santiago Abascal realizado durante un mitin de la campaña electoral para las elecciones gallegas que se habían celebrado el 18 de febrero de 2024. Nos llamó la atención que el algoritmo no hubiera servido como prioritario el contenido de una fuerza política con implantación andaluza o, incluso, algún contenido del entorno ideológico de Vox de esta región cuando la procedencia del usuario parecía estar pesando en TikTok a la hora de elaborar recomendaciones para Manu. El 24 de marzo de 2024 el algoritmo sugirió un segundo contenido político, esta vez un vídeo subido el día 2 de febrero desde la cuenta @yoconsantiagoabascal con el fragmento de una entrevista política al líder de la formación española Vox.





Durante el mes de abril Manu recibió un importante caudal de información relacionado con la Feria de Sevilla, que se celebró este año del 14 al 20 de abril que dominó claramente el feed de TikTok y las sugerencias. Sin embargo, emergieron de nuevos algunos contenidos políticos desarrollados con una narrativa humorísticas. El día 7 de abril el algoritmo sugirió en la pestaña “Para ti” el vídeo en el que unos amigos regalaban a su “amigo rojo”, una tarta de cumpleaños en las que aparece pintada una ilustración de Pedro Sánchez vestido de payaso y el mensaje “Sánchez, vete ya” impreso en la tarta.



(Captura pantalla TikTok @bygarcia)





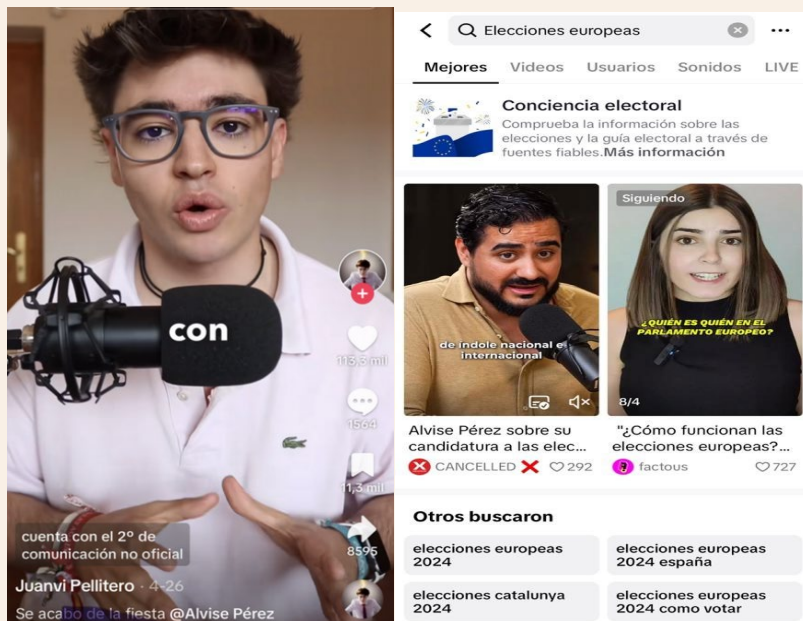
El jueves 25 de abril, tras el anuncio de Pedro Sánchez sobre tomarse un tiempo de reflexión, Manu pudo informarse por una aportación de la cuenta de @infovlogger, de la @cadenacope y de @informativost5, ambas sugeridas por el algoritmo. Posteriormente llegó un montaje de vídeo realizado con inteligencia artificial, servido desde la cuenta de @ac2ality en la que se veía a Pedro Sánchez bailando.

Este será el primero de una serie de vídeos de cuentas servidas desde la pestaña “Para ti” con videos generados con IA de Pedro Sánchez en diversas posiciones y bailes. Las publicaciones de la cuenta @xx_elangel, especializada en vídeos políticos con IA llegaron entonces de forma recurrente.

Información electoral

Una vez que se fue acercando la campaña de cara a los comicios europeos de junio los contenidos de ocio de los ámbitos de interés de Manu comenzaron a mezclarse con las publicaciones política de varios “influencers” de la derecha radical que Manu no seguía. Es el caso de @JuanviPellitero, con 89,400 seguidores en TikTok, de quien Manu recibió el vídeo “El show de Sánchez, sólo en cines” o de @cakeminuesa con una publicación que recogía el testimonio de un ciudadano de El Hierro respecto a las consecuencias de la información ilegal. Hay que descartar que durante todo este tipo la cuenta de Manu no se vio impactada por contenidos desinformativos puros, como estábamos comprobando que sucedía en los perfiles de las identidades de más edad.

En la recta final de las elecciones europeas, a mitad de mayo decidimos “resetear”, es decir, borrar el historial de recomendaciones de la cuenta de TikTok de Manu para que las sugerencias comenzarán a llegar desde una experiencia de usuario completamente nueva que no se viera contaminada. Queríamos, con esta acción, cerciorarnos de si existía la posibilidad de que otras sugerencias políticas, fuera del entorno Vox, pudieran llegar a la cuenta de Manu sugeridas por el algoritmo. Asimismo, decidimos que Manu realizara algunas búsquedas relacionadas con las elecciones europeas.



Captura pantalla TikTok @juanvipellitero Contenidos sugeridos por TikTok tras las búsquedas “elecciones europeas”





Como resultado de las acciones posteriores Manu recibió dos contenidos sugeridos sobre información electoral procedentes de las cuentas @factous_es y @eu_genias, ambos perfiles de divulgación sobre el funcionamiento de las instituciones europeas. Sin embargo, no apareció publicación de ningún partido político o candidato a excepción del contenido de la cuenta @cancelledpodcast con una entrevista al candidato de Se Acabó la Fiesta, Alvisé Pérez. El 30 de mayo, el algoritmo recomienda otro contenido de Alvisé Pérez, este desde la cuenta @Tiki_tube sobre varias convocatorias de actos en Barcelona y Valencia.



Esta fue la tónica general en el comportamiento de la cuenta de Manu hasta la celebración de elecciones que deparó un resultado muy satisfactorio para la formación de Alvisé Pérez, cuya actividad digital tuvo, con toda seguridad, una incidencia mayor en la conexión con un nuevo y joven electorado. La formación de Pérez, Se acabó la fiesta, obtuvo tres escaños en el Parlamento Europeo al obtener 800.763 votos, es decir, el 4,59%.





David, el votante novato

David es el perfil más joven que tenemos en nuestro estudio que sea mayor de edad. Tiene apenas 18 años y acaba de obtener el vital derecho al voto y a decidir sobre los asuntos políticos que le incumben. Como la mayoría de los jóvenes de su edad, muestra un alto grado de desafección por la política, pero, paralelamente, muestra interés en participar en las elecciones, como así lo señalan los informes realizados en sus redes sociales se comportará como cualquier otro adolescente: para seguir a sus amigos y cuentas de su interés.

Para su experiencia de usuario, David utiliza Instagram para estar al tanto de las novedades de las cuentas que le interesan mientras que TikTok lo utilizará, como cualquier otro joven, para descubrir una infinidad de vídeos que le recomiende el algoritmo como puro entretenimiento.

A David le interesa mucho el fútbol, por lo que seguirá cuentas como @realmadrid o @elchiringuitotv. También le interesa lo relacionado con su entorno en el que vive, Fuenlabrada, y sigue a cuentas como @soydefuenla o @ayuntamientodefuenlabrada. El trap y la música urbana le encanta por lo que sigue a cuentas de artistas importantes como @asaprocky y @thekidkeo.

Aunque David no tiene una opinión política formada, dado que su padre vive en Catania y le comenta el problema que tienen allí con la inmigración irregular, sigue a cuentas locales de la zona como @lasicilia.it y @visit.catania. Además, seguirá a medios nacionales italianos como @corriere, @larepubblica y @ilmessaggero. Esta inquietud por la inmigración en Italia es la vulnerabilidad de nuestro joven madrileño y queremos comprobar si la desinformación puede colarse por este motivo a través de sus redes sociales.

Durante las semanas previas a la campaña electoral europea, se interesará más sobre cómo funciona la Unión Europea, expectante y ciertamente ilusionado por poder votar por primera vez, quizás lo más característico de alcanzar la mayoría de edad y convertirse en ciudadano de pleno derecho. Por esta razón empezará a seguir cuentas como @parlamentoeuropeo y @eucouncil, y a un medio europeo como @arte.tv_es.

Nuestro personaje no tiene demasiada experiencia contrastando información, pero mantiene un alto consumo de contenidos en las redes sociales, entre ellos contenidos informativos. No le atrae, sin embargo, la información más convencional que podría encontrar en la oferta de un periódico impreso o en los noticiarios de televisión. Es el patrón mayoritario para jóvenes de su edad. Como indica el informe Reuters 2024 el 85% de los jóvenes españoles (de 18 a 24 años) utilizan redes sociales como su principal fuente de noticias. Instagram y TikTok son las plataformas más populares, con un 70% y 60% de los jóvenes utilizándolas para acceder a noticias, respectivamente. David, es por tanto, potencialmente más vulnerable a la hora de caer en la trampa de titulares y publicaciones engañosas. Su corta experiencia no le hace mostrado aún la necesidad de contrastar la información que recibe por sus canales digitales.

Definido el perfil de nuestro joven David y su futura experiencia de usuario, analizaremos el contenido que le llegará a sus redes sociales de forma orgánica y, en ocasiones, también como búsqueda específica sobre un tema en concreto.

Evolución de la desinformación recibida

Como señalamos anteriormente, no quisimos dotar de identidad ideológica alguna a este perfil sino generar con él la experiencia de un usuario neutral, alejado de polémica, contenidos radicales o de una intensa





conversación política. Por ello no interactuamos en ningún momento con contenidos político ni quisimos seguir información de ninguna cuenta ideologizada. La observación de este perfil nos permitió constatar que, con una gestión neutra, era mucho más sencillo conseguir la conversación deseada desde Instagram. Esta red respetó más nuestra selección inicial.

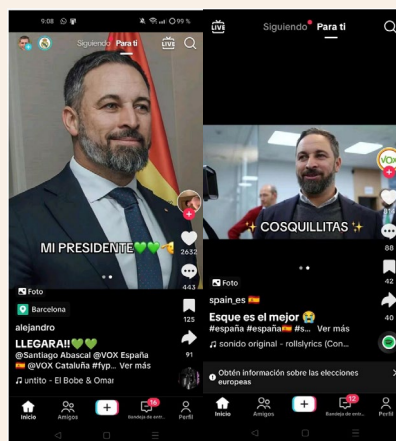
Antes de empezar a analizar el contenido recibido hay que mencionar especialmente que el 92%, ha llegado a través de TikTok frente a Instagram. Como conclusión sacamos que Instagram se aleja por completo del plano de conversación política y se centra más en las interacciones personales y en el contenido multimedia que suben.

La principal conclusión de nuestro análisis es que más de la mitad de la desinformación recibida por David corresponde con contenido radical y contenido de odio, normalmente con publicaciones correspondientes a propaganda de extrema derecha y sus discursos de odio. Ya hay estudios que advierten de que estos discursos han encontrado una fácil interacción en las redes sociales entre los jóvenes por lo que no sorprende la enorme cantidad de estos discursos que nos hemos encontrado en las redes de nuestro joven David. Ya que la mayoría de las publicaciones de desinformación recibidas han llegado a través de TikTok, a continuación, nos referiremos a ellas presumiendo que son de TikTok a no ser que se especifique que son de Instagram.

Contenido radical

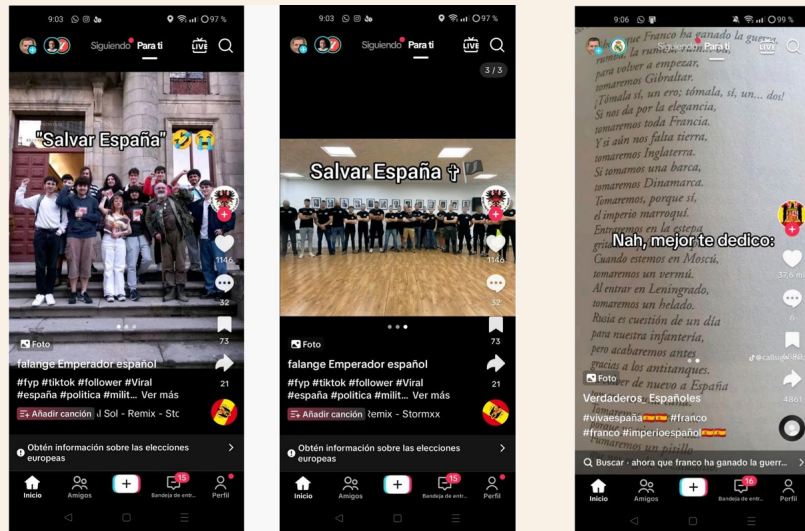
David ha sido objeto de numerosos bulos o publicaciones de contenido político radical. Este tipo de desinformación incluye mensajes extremos que buscan polarizar a la audiencia y fomentar sentimientos de animosidad hacia ciertos grupos. Estos bulos pueden influir en la percepción de David sobre temas sensibles, exacerbando sus preocupaciones y generando una visión distorsionada de la realidad. La mayor cantidad de estas publicaciones que nos hemos encontrados eran propaganda del grupo político Vox, algo que sorprende ya que David no se interesa por la política aún ni ha interactuado con ningún contenido propagandístico.

Las sugerencias de opciones políticas fuera del espectro de los partidos moderados no se detuvieron en Vox. El algoritmo sirvió numerosas publicaciones afines a la ideología franquista o que defienden el legado del dictador español, como supuestas comparaciones de la situación económica de España durante el franquismo o la democracia. Otras sugerencias recupera viejas imágenes con canciones actuales o, al contrario, combina secuencias de vídeo con letras de himnos como el “Cara al sol”, que vive un renacimiento gracias a su difusión por las redes sociales.



Captura pantalla TikTok. @alejandrop Captura pantalla @spain_es



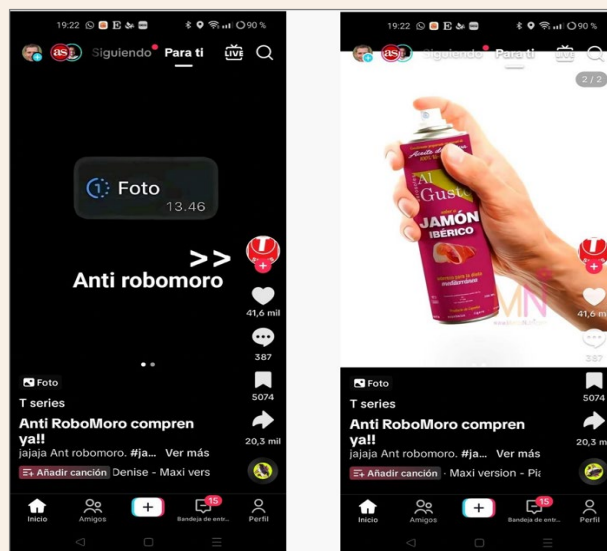


Captura pantalla TikTok

Captura TikTok @verdaderos_españoles

Contenido de odio

Como contenido de odio principal, nos hemos encontrado con publicaciones islamófobas, con imágenes de “espray de jamón” para prevenir los robos o capturas de pantalla de la red social X donde un usuario se pregunta cómo sería vivir en España sin musulmanes, haciendo similitud a vivir en Hogwarts y eliminar a Voldemort con una varita de jamón. Mención especial a la dinámica de este tipo de contenido en TikTok que muchas veces es una réplica de una publicación de la red social X que es republicada en forma de captura de pantalla y con una canción de fondo.



Bulos genéricos y electorales

También ha recibido bulos genéricos y electorales. Los bulos genéricos pueden alimentar una desconfianza general en la información que encuentra en línea, mientras que los bulos electorales pueden distorsionar su





comprensión de procesos democráticos y decisiones políticas. Los bulos electorales llegaron sobre todo en junio, coincidiendo con las fechas en las que se produjeron las votaciones en la Unión Europea, apareciendo mucho en las publicaciones recomendadas contenido del candidato de Se Acabó la Fiesta Alvisé Pérez, colándose en el ideario de nuestro joven David. Nos encontramos por ejemplo bulos que menciona este candidato a las elecciones europeas sobre cuánto cobra un europarlamentario, dando cifras sobredimensionadas y falsas, como comprobamos al realizar la verificación de información con las herramientas señaladas.

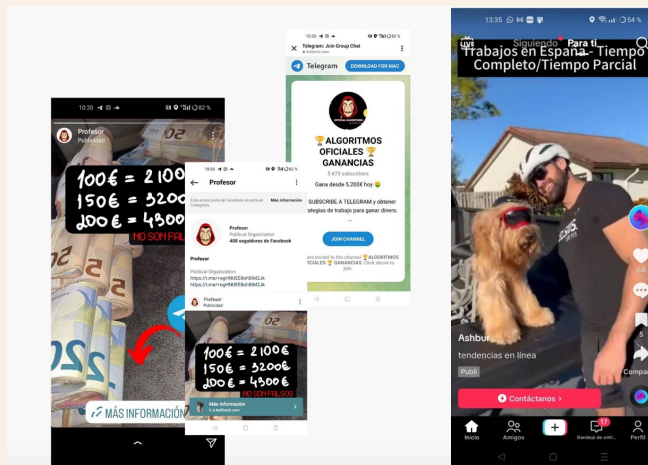


Captura pantalla TikTok @nuvoclips

Como bulos generales nos encontramos, por ejemplo, capturas a supuestos titulares falsos como que los legisladores de Texas van a castigar con pena de muerte a las personas gestantes que quieran abortar, o más repeticiones de X en forma de captura de pantalla donde aseguran que el Tribunal Constitucional ha establecido como derecho constitucional la okupación, con una supuesta captura de la sentencia donde no habla de tal cosa.

Estafas económicas y financieras

Algunos bulos recibidos por David también incluyen estafas económicas, que trataban de engañarle bajo la promesa de obtener beneficios financieros. El feed de David reflejó publicaciones en forma de publicidad que conducían a un contacto de WhatsApp donde se ofrecían trabajo fácil y rápido desde casa mediante la realización de tareas organizadas desde la red Telegram.



Captura pantalla Instagram @profesor

Captura pantalla TikTok @ashbur

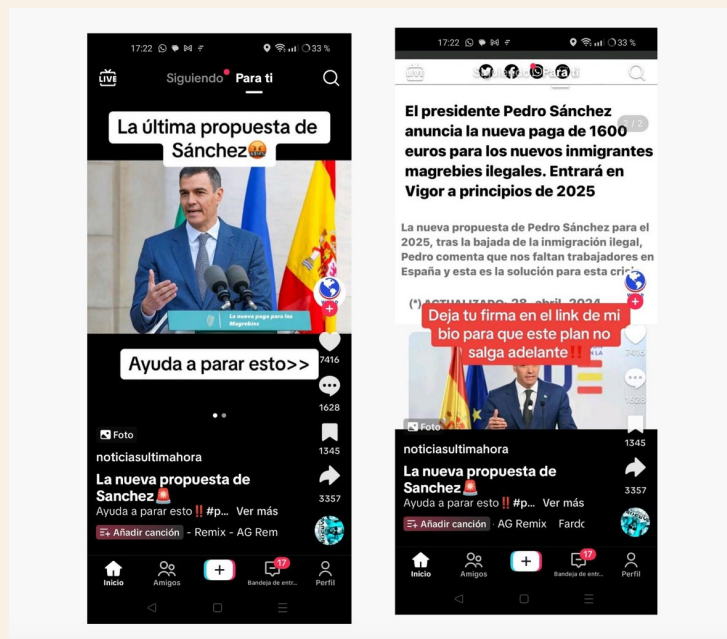




En Instagram, las únicas dos publicaciones recibidas por David desde esta red social también resultaron intentos de estafa a nuestro joven David con publicidad pagada y que le ha entrado en forma de historias. Una de ellas se presentaba con el reclamo de ganar grandes cantidades través de Telegram con “El Profesor”, una creatividad inspirada en la serie de televisión La Casa de Papel con la cual enriquecerse a través de fallos en las casas de apuestas online, La segunda es otra historia con difusión de pago que promociona un supuesto perfil de un partido político alemán, al que suplanta su identidad. que animaba a invertir dinero en una página falsa.

Bulos sobre Inmigración

Como se ha señalado en el estudio, la vulnerabilidad ideológica de David era la inmigración irregular, sobre todo de Italia. Entre las publicaciones recibidas se detectó la de publicaciones un “falso medio” en el que aseguraban que España había llegado a un acuerdo con Marruecos para no devolver a los inmigrantes que llegasen de forma ilegal y que tendrían todos los derechos. Asimismo, el algoritmo sugirió un contenido que presenta la captura de un titular, de un supuesto medio, en la que aseguraba que Pedro Sánchez pagará 1600€ a todo magrebí que entre a España a partir de 2025.



Captura pantalla TikTok- Noticias Ultima Hora.

Respecto a la inmigración italiana, interés de David, no ha aparecido en los perfiles de David en forma de desinformación como se esperaba al principio de esta monitorización.

Inteligencia Artificial y suplantación de identidad

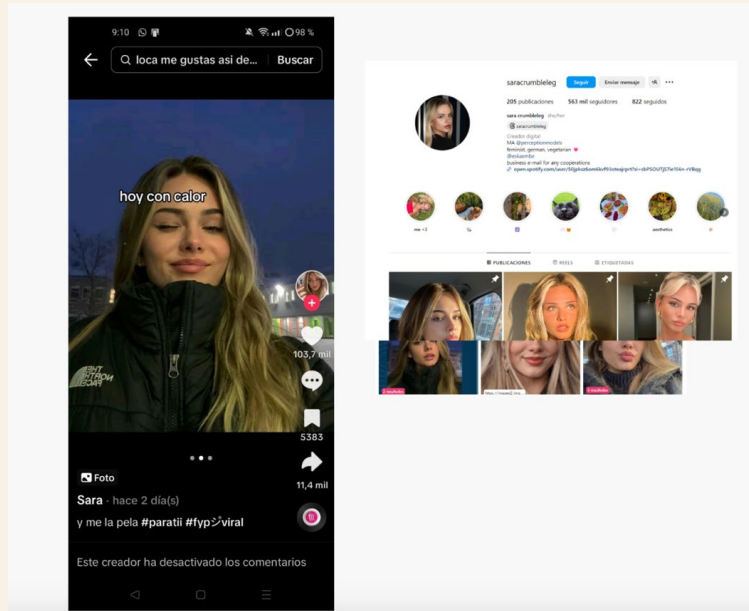
David ha sido expuesto a desinformación relacionada con inteligencia artificial y suplantación de identidad. Esto se ha manifestado en forma de “deep fakes” con vídeos editados con inteligencia artificial y mediante el uso páginas web que calcan el diseño de páginas legítimas, sobre todo de medios de comunicación.

Se han detectado principalmente vídeos de Cristina Pedroche editados con Inteligencia Artificial ofreciendo una entrevista y hablando de la importancia de invertir dinero en una determinada página web con la que ella habría ganado mucho dinero. También David ha recibido desde la cuenta del medio “El Español” un vídeo en





formato de noticia en el que unos supuestos hackers iluminan la puerta de Brandeburgo, en Berlín, con la bandera de la URSS, si bien se trata de un montaje con IA. Otra de las suplantaciones de identidad que hemos detectado ha sido la utilización de fotos de una influencer alemana para hacerla pasar por simpatizante de Vox.



Captura de pantalla TikTok @privadaysecret

Cuenta Instagram modelo @saracrumbleleg

El algoritmo de TikTok y su tendencia a la radicalización

Durante el análisis del perfil de David en TikTok sorprende cómo el algoritmo de TikTok recomienda vídeos que pueden ser de interés para el usuario y que esas recomendaciones acaban radicalizándose hacia la derecha. David tiene un perfil neutro y los intereses que tiene en TikTok y con los que interacciona son los detallados en su identidad digital. Sin embargo, sorprende cómo, a medida que se utiliza su cuenta en la plataforma, aparecen cada vez más contenidos radicales y de extrema derecha. En ningún momento se ha interactuado con esos vídeos y, sin embargo, el algoritmo de la plataforma insistía mostrarle más contenido de este estilo. Incluso habiendo reestablecido el algoritmo de la cuenta, TikTok continuó ofreciendo contenido radicalizado hacia la derecha, aunque en menor medida a partir de este restablecimiento.





Aitana, una usuaria avezada

Aitana es también uno de nuestros perfiles más jóvenes que analizamos. Tiene 25 años y, al contrario que David, Aitana ya tiene una opinión política un poco más formada y empieza a ser menos manipulable. Hemos elegido este tipo de perfil por tener a un tipo de joven más maduro con unos intereses más definidos para constatar si la desinformación llegaba por un lado u otro y de qué tipo y forma lo hacía.

Aitana utilizará por un lado Instagram para seguir a cuentas inspiradoras que le interesen, sobre todo de carácter más artístico, pero también a cuentas que hablen sobre el mundo que le rodea en formato audiovisual. En X, antes Twitter, Aitana está más interesada en informarse sobre el mundo que le rodea y está más activa en conversaciones políticas. Sigue a cuentas más informativas y politizadas, y utilizará esta cuenta para mostrar apoyo a sus preocupaciones sociales y leer otras opiniones.

A Aitana le gusta mucho el arte y sigue a cuentas e influencers relacionados como @jr, @banksy o @studiolafureliasson. Es una apasionada del yoga por lo que seguirá además a cuentas relacionadas como @xuanlanyoga o @somayogabcn.

También como es catalana tiene un cierto sentimiento nacionalista y sigue a cuentas de partidos políticos como @esquerrarepublicana o @juntspercat así como a políticos catalanes como @carlespuigdemont o @adacolau. Al seguir a cuentas tanto de la izquierda como de la derecha catalana independentista, estamos ubicando a Aitana en un centro catalanista marcado por el sentimiento nacionalista e independentista, aunque no estará especialmente interesada en el independentismo durante el uso de sus redes sociales. También le interesa lo que ocurre en su comunidad, por lo que seguirá a cuentas institucionales del Govern y del Ayuntamiento de Barcelona como @gencat, @agentsrurals y @barcelona_cat.

Le hemos querido dar una vulnerabilidad importante a su perfil, ya que Aitana está profundamente preocupada por el medio ambiente, el cambio climático y los enormes desafíos a los que se enfrenta su generación. Con este especial interés queremos comprobar cómo se cuela, o no, la desinformación a través de esta preocupación. Para ello, Aitana sigue a numerosas cuentas relacionadas con este tema: institucionales y gubernamentales como @agenda2030 o @unclimatechange, ONGs como @nature_org o @greenpeace_esp, a activistas como @gretathunberg y medios sobre medioambiente como @_efeverde.

Aitana tiene raíces francesas y entiende francés perfectamente, por lo que sigue a medios franceses para estar enterada de lo que sucede en su segundo país como @france24 o @brutoficiel. Al añadir cuentas francesas, queremos comprobar si este perfil puede recibir desinformación proveniente de ese país.

Por último, debido a su compromiso con el medio ambiente y sus retos, Aitana está decidida a votar en las elecciones europeas. Su perfil es claramente europeísta y está convencida de que las siguientes elecciones son cruciales para continuar y mejorar las políticas verdes que se impulsan desde la Unión Europea y que afectan a todos los países miembros. Por ello, Aitana va a seguir a muchas cuentas institucionales europeas, entre las que destacamos a @lifeprogramme, @europeanparliament y @ursulavonderleyen. También seguirá a medios europeos como @arte.tv_es. Con este interés queremos situar a Aitana en el blanco de las desinformaciones que puedan ocurrir durante la campaña electoral.





Evolución de la desinformación recibida

Aitana, no tiene una ideología muy definida pero sí cuáles son sus preocupaciones, y apoya a aquellas personas que las defienden como ella. Interactúa positivamente con cualquier contenido que se encuentre relacionado con el medio ambiente y la lucha climática. Sin embargo, trataremos de interactuar lo menos posible con el contenido puramente político que no contengan medidas medioambientales. Con esto intentaremos que el algoritmo de las redes sociales que utiliza polarice su experiencia de usuario hacia un lado u otro.

En el análisis de la desinformación recibida por Aitana vemos que el grueso corresponde a lo que hemos denominado bulos genéricos que representa casi un 50% del total. El segundo tipo de desinformación más numeroso ha sido el de bulos electorales, con una especial incidencia sobre todo durante el mes de junio en plena campaña y durante las elecciones. El resto de desinformación de distinto tipo ha sido un poco más residual, aunque destaca también el uso del clickbait.

Hacemos especial mención a que la desinformación recibida ha sido únicamente a través de X, Instagram no sirvió en esta cuenta ningún tipo de contenido desinformador ya que el uso de esa red social ha quedado relegado a contenido audiovisual y, creemos que estos son debido a que la preferencia que Instagram se centra a los temas de ocio.

El caudal de contenido recibidos durante el mes de marzo no resultó representativo. El pico de desinformación se produce durante el mes de abril si bien la incidencia de desinformación se redujo drásticamente durante el mes de mayo. Inferimos que ha existido un esfuerzo por parte de las redes sociales de luchar contra la desinformación en un momento previo a las elecciones europeas, durante mayo, y a esto se debe también la importante reducción de desinformación recibida, un hecho que recogieron medios europeos como la cadena de noticias pública France 24 (17)

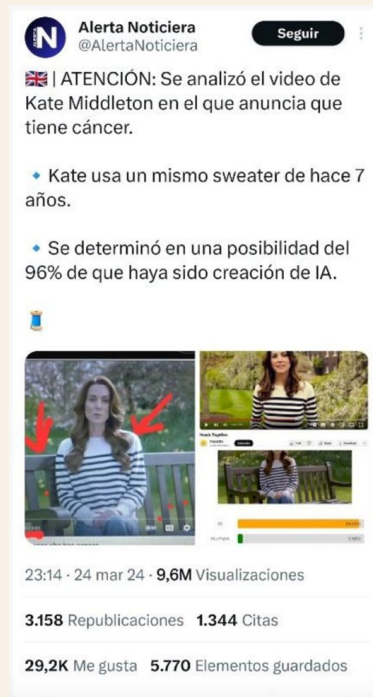
Bulos genéricos

En primer lugar, tenemos a los bulos genéricos a los que Aitana fue repetidamente expuesta a lo largo de estos meses. Sin duda este ha sido el tipo de desinformación más encontrado en su perfil y representa el 50% del total. Por bulos genéricos hemos querido definir cualquier bulo o desinformación que no tiene mayor objetivo que el de simple desinformar, sin causar mayores daños añadidos.

La mayoría de estos bulos genéricos encontrados de forma orgánica en el timeline de Aitana eran fácilmente identificables ya que añadían una nota de la comunidad informando de que ese tweet era desinformación. Las notas de la comunidad son una nueva herramienta introducida en X desde diciembre de 2022 y que se han empezado a utilizar masivamente recientemente y que, durante algunos meses, han resultado de ayuda para identificar desinformación especialmente durante estos meses electorales.

Entre los bulos recibidos nos encontramos con un agregador de noticias afirmando que un vídeo con el que la princesa de Gales, Kate Middleton, anunciaba su enfermedad el pasado mes de marzo, era falso y que había sido creado por inteligencia artificial.





Captura pantalla X @AlertaNoticiera

El algoritmo sirvió asimismo el vídeo que mostraba a un numeroso grupo de musulmanes celebrando el fin del periodo del Ramadán celebrando el fin del Ramadán, si bien el titular aseguraba que la multitud se encontraba celebrando un ataque iraní sobre Israel que tuvo lugar una semana más tarde, en el mes de abril de este mismo año.



Captura pantalla X @EuricSanti

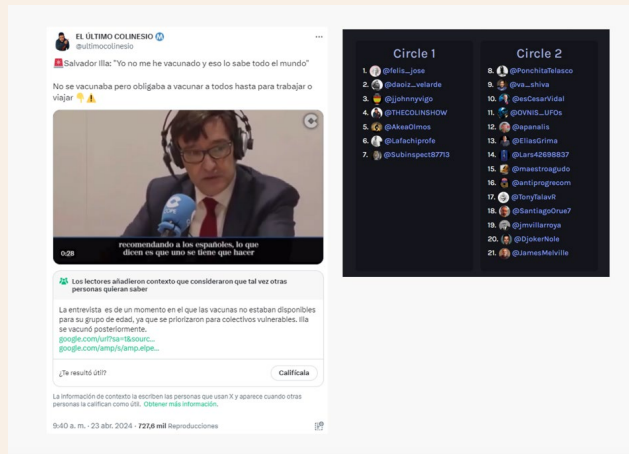


Captura pantalla Verificación @malditobulo





En tercer lugar, o un clip de vídeo de una entrevista al político catalán Salvador Illa que, sacado de contexto, afirma que no se ha vacunado, cuando realmente en la entrevista dice que no se ha vacunado del COVID antes que los de su grupo de edad.

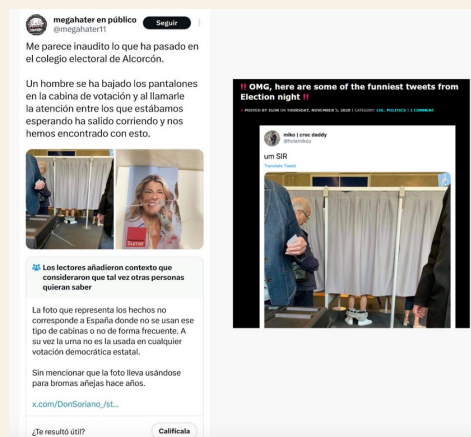


Captura pantalla X @Elultimocolinesio

Bulos electorales

Aitana tenía un claro interés por los temas que atañen a la Unión Europea desde un principio por lo que esperábamos que los bulos electorales apareciesen en algún momento. Sin embargo, además de reducirse a la mitad la desinformación recibida el mes previo a las elecciones, no se registró ningún bulo electoral hasta el mes de junio. Sin embargo, el momento en el que más bulos electorales alcanzan a nuestra joven catalana es el propio día de las elecciones y al día siguiente de estas. Podríamos deducir que, en ese sentido, ha jugado un papel importante el esfuerzo desplegado por las plataformas de X y de Instagram en su luchar contra la desinformación Una acción reforzada con la introducción de centros de información electoral oficial en las propias redes sociales.

Dentro de los bulos electorales recibidos se encuentran publicaciones en las que se afirma que un votante se ha masturbado dentro de un cubículo de votación con una foto de Yolanda Díaz, cuando la foto pertenece a una situación que se registró en otra localización y momento. Las referencias cronológicas y de espacio son fácilmente verificables con herramienta de búsqueda inversa de imágenes.



Captura pantalla X @megahater11





La cuenta de Aitana también recibió impactos de algunas publicaciones donde aseguraban que se produciría un pucherazo electoral ya que habría un retraso en la comunicación de los resultados de 3 horas, cuando simplemente había que esperar a que todos los países de la UE terminasen de votar para publicar los resultados.



Captura pantalla X @betaniatv y @benemeritoverde

Un último ejemplo de bulo electoral lo constituye una publicación en la que aseguran que, si se introduce una papeleta de Vox y del PP juntas, el voto valdría doble, algo claramente falso.

Clickbait

El tercer tipo de desinformación más numerosa que ha recibido el perfil de Aitana ha sido el de contenidos de clickbait, diseñados para captar su atención. Estos contenidos se fabricaron, en varias ocasiones, a partir de noticias reales de medios contrastados si bien eran sacadas de contexto o para las que se elegían titulares llamativos a la hora de presentarlas en las redes sociales.

Cabe resaltar que toda la desinformación recibida por Aitana fue a través de su cuenta de X ya que en Instagram no se ha encontrado ningún tipo de desinformación de la que se está analizando. El único contenido que aparecía en su cuenta estaba relacionado con sus cuentas de interés y estaba muy adaptado a sus gustos y cuentas que sigue.

Es cierto que Instagram se utiliza para conversaciones sociales más centradas en el estilo de vida o para estar al corriente de la vida de amigos o famosos. Sin embargo, no es una plataforma tan útil a la hora de compartir desinformación. Hay que mencionar también que Instagram decidió añadir una opción predeterminada por defecto que limitaba el contenido político que una cuenta podía recibir, si bien ningún usuario había sido avisado de ello. La opción puede ser desactivada lo que podría explicar también por qué no se recibió esta desinformación de forma orgánica.





Pedro, un centrista a la caza de buena información

La identidad digital de Pedro Garea corresponde a la de un economista de mediana edad que trabaja en Madrid como consultor. Este perfil está orientado a una ideología de centro, buscando el consenso y la moderación dentro de un clima político convulso. Del mismo modo, tiene un vínculo con Alemania por lo que está interesado en la situación política tanto nacional, como internacional. Por ello, nos encontramos ante un perfil poco vulnerable, el cual sabe discernir, al menos en teoría, en la gran mayoría de casos entre lo que es real de lo que es irreal o manipulado.

El reto con este perfil se halla en que la desinformación consiga incidir dentro de sus principales preocupaciones: la soberanía, el fraude electoral y Europa. De este modo, la metodología irá ligada a fomentar las áreas de vulnerabilidad de Pedro para lograr que aparezca la desinformación en el perfil, tanto de X (Twitter), como de Instagram.

Para lograr el objetivo anterior se optó por seguir a cuentas de medios de comunicación españoles, @el_pais o @elmundoes, como alemanes, @derspiegel Süddeutsche Zeitung @SZ, Frankfurter Allgemeine @faznet, así como de cuentas de divulgación política y relativas a resultados electorales como @electo_mania, @politico o @ElectsWorld En el catálogo de cuentas se encuentran partidos y líderes políticos nacionales, europeos y grupos parlamentarios europeos como y europeos como @CiudadanosCs, @ppopular, @PSOE, @EPPGroup, @TheProgressives, @NunezFeijoo, @sanchezcastejon o @jorde_canyas.

El interés de Pedro por la economía se refleja en su "timeline". Sigue a como algunas instituciones financieras, como el Banco Central Europeo @ecb, Funcas @FUNCASES el Ministerio de Economía Gobierno y Empresa @_minecogob. Le gusta la cocina y ha incorporado @albertochicote, @karguiñano o Eva Arguiñano @evaarguinano.

Si bien el material hallado en X e Instagram difiere, existe un marco común. Ambas cuentas no recibieron flujos de desinformación electoral de forma explícita y/o de forma orgánica. En cambio, aparecieron otros formatos que se repitieron en otros de los perfiles del estudio.

Las estafas económicas

El perfil de Instagram registra una predominancia de las estafas económicas relativas a la compraventa de criptomonedas o de plataformas de inversión, las cuales utilizaban el supuesto patrocinio de personas de la esfera pública, como Isabel Díaz Ayuso, el Rey Felipe VI o la reina Letizia, a la que achacaban una falsa afirmación.

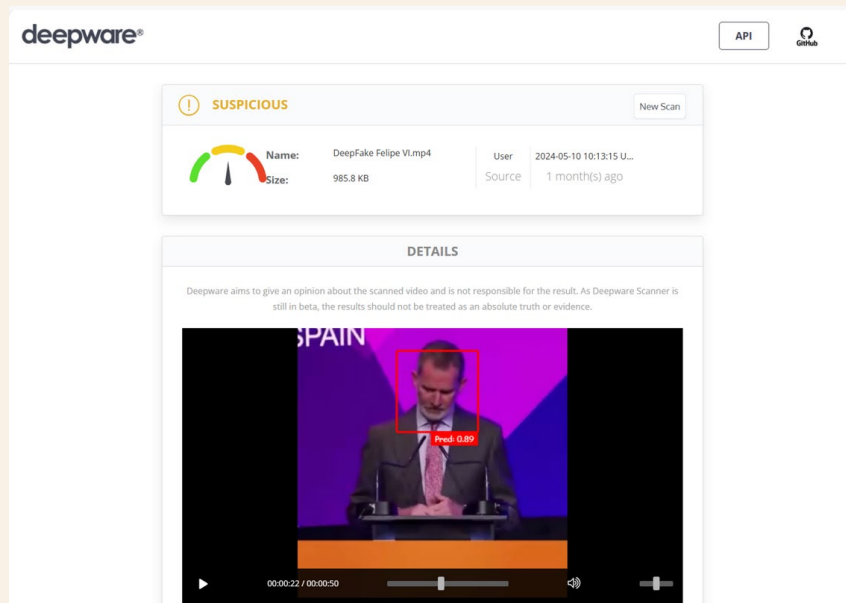




“Me arrepiento de mi comportamiento”, como señuelo de uno de los contenidos alojados en una falsa página de RTVE. Dentro de estas estafas, cabe destacar el uso de la IA y la manipulación tanto del vídeo como del audio de las entrevistas y contenidos audiovisuales.



Captura pantalla web suplantada de RTVE



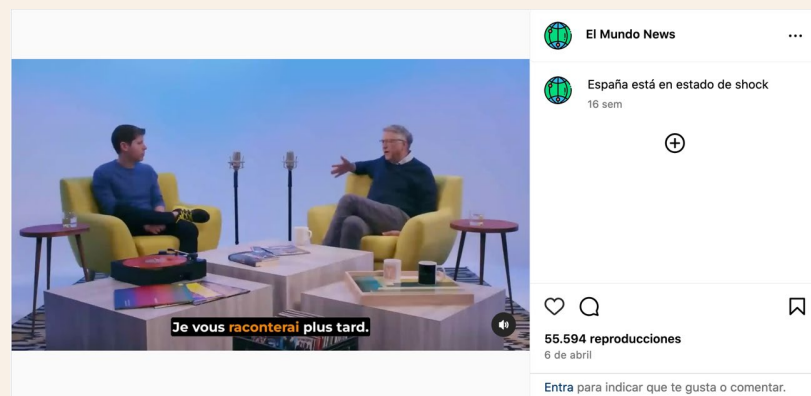
Captura pantalla Verificación Deepware

Estas publicaciones siguen un mismo patrón: utilizan un vídeo manipulado o un titular pretencioso para atraer la atención del usuario, el cual hace clic en el enlace patrocinado de la publicación que lleva a una página que reproduce la estética de un medio de comunicación o de una empresa privada. Dicha página web incluye una invitación para que el usuario deje sus datos personales. Un ejemplo de estas publicaciones constituye un vídeo publicado por @signals8533 en el que aparece Bill Gates promocionando una página de inversión en criptomonedas.





Captura pantalla Instagram @ElMundoNews



Otros ejemplos lo constituyen el publicado por @groovytek donde aparece una supuesta noticia sobre Isabel Díaz Ayuso que, al hacer clic en el enlace, lleva a una página que suplanta la identidad de RTVE con el fin de que se dejen los datos personales para comenzar la estafa.



Por otro lado, en el segundo, publicado por @foodg4811, aparece un vídeo manipulado de Elon Musk alojado también en la página web a la que lleva el enlace promocionado, persistiendo la desinformación.

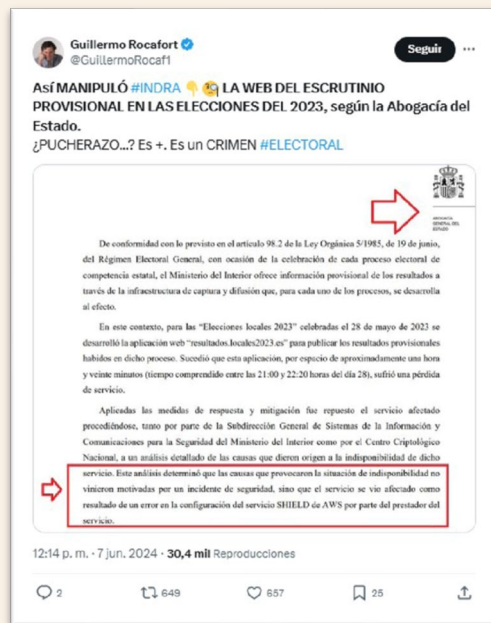




Residualmente, cabe destacar diversas suplantaciones de identidad que, utilizaban el mismo procedimiento que las anteriores desinformaciones mencionadas, pero reutilizaban el contenido real para dar una sensación de seguridad, confianza y calidad.

Desinformación electoral

Por otro lado, dentro del perfil de X (Twitter) de Pedro, ha registrado una menor irrupción de la desinformación de forma orgánica. La desinformación electoral o relacionada con las preocupaciones del perfil ha sido encontrada gracias a las búsquedas de palabras clave como pueden ser hashtags de las Elecciones Europeas o de temas relacionados con la economía y que están presentes en la conversación social, como “alquiler”. De esta manera, la desinformación ha estado de forma más presente durante la última semana de mayo hasta las Elecciones Europeas, lo que demuestra varios aspectos que serán detallados en las conclusiones.



Captura pantalla X. @GuillermoRocafort

Frente a las recurrentes estafas económicas que aparecían en Instagram, dentro del perfil de X (Twitter) únicamente se ha registrado un caso un caso mientras que encontramos una predominancia de bulos relacionados con los comicios, realizados con Inteligencia Artificial o contenidos descontextualizados. que son fácilmente verificables mediante una búsqueda en Google, la comprobación dentro de las Notas de la Comunidad o en algún portal de verificación de información, como es el caso de Elections24Check.

El procedimiento típico de la desinformación encontrada en X (Twitter) es mediante la publicación de una foto o un vídeo con un titular llamativo, logrando captar la atención del usuario. En este caso no hay cesión de datos personales, pero el mero hecho de que éste entienda que la información falsa o tergiversada que está viendo es veraz es una alarma dentro de la irrupción de la desinformación.





Contenido de odio

Algunos ejemplos en X (Twitter) son vídeos como el publicado por @MrDavidTweets en el que aparece una multitud supuestamente entonando la frase “Alemania para los alemanes”. Sin embargo, una nota de la comunidad desmiente este hecho y asegura que el audio del vídeo ha sido manipulado.



En otro tuit encontramos un titular falso de @Puntual24H que afirma que el programa gubernamental “Verano joven” subvenciona al cien por cien los viajes de los MENAS, si bien los menores de edad no pueden acceder a él.

El último caso es un vídeo manipulado con inteligencia artificial publicado por @MartosDiego en el que aparece una persona musulmana pidiendo el voto para el PSOE.



Captura pantalla X @MartosDiego



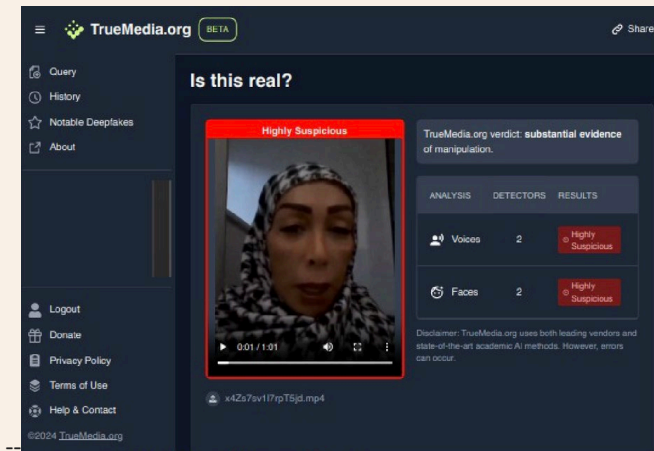


Imagen Verificación vídeo con TrueMedia.Org

Tras la observación de la cuenta de Pedro podemos concluir que las diferencias existentes entre el material desinformativo en función de cada red social y los formatos utilizados. Vemos cómo incluso la experiencia de usuario desde una misma identidad digital difiere profundamente en función de las redes sociales: mientras que en Instagram son estafas económicas, en X (Twitter) son relacionadas con las elecciones o con temas del día a día. Asimismo, en Instagram, la desinformación, que es de carácter financiero, viene dada en forma de anuncios que aparecen de forma orgánica, mientras que en X (Twitter) se realiza por medio de publicaciones de distintas cuentas.



Captura pantalla Instagram @FDP Rheinessen-Vorderpfalz

Por otro lado, es evidente la influencia tanto las legislaciones de la Unión Europea como por medio de iniciativas propias de las redes sociales. Ejemplo de esto es la opción de Instagram de limitar el contenido político dentro de las recomendaciones de la aplicación, algo que impide que una parte de la desinformación política entre dentro del perfil.

Por todo lo anterior, dentro de los contenidos de las redes sociales, hay una parte de desinformación que, en este caso de estudio, ha incidido con mayor fuerza en las últimas semanas de la investigación y abarcando temas económicos en Instagram, y temas electorales y sociales en X (Twitter).





El viaje de Carlos (y Marcos) hacia la polarización

La identidad digital de Marcos Rodríguez en X (Twitter) y la de Carlos Massa (Facebook) correspondía, como señalamos anteriormente, a la de un señor almeriense de 55 años y propietario de un supermercado. Este perfil está orientado a una ideología de centroderecha, propia del Partido Popular, buscando lo tradicional y lo conservador, pero abierto a avances. De la misma manera, tiene un vínculo con Italia, y se muestra interesado por la situación social del país. Por tanto, nos encontramos ante un perfil que tiene una vulnerabilidad alta debido a la edad y a la forma de vida tranquila, que no está en constante alerta ante los posibles intentos desinformativos.

El reto con esta identidad digital está en que la desinformación consiga incidir dentro de sus principales preocupaciones: la agricultura y el medioambiente, la inmigración y la situación laboral. De este modo, el estudio irá orientada a fomentar las áreas de vulnerabilidad de Carlos y Marcos para lograr que la desinformación se haga visible en los perfiles de Facebook y Twitter, respectivamente.

Para lograr el objetivo anterior, se optó por realizar las siguientes acciones, divididas en dos grandes bloques: acciones relativas al seguimiento de cuentas y acciones relativas a la interacción del perfil con el contenido y la realización de búsquedas.

Carlos (y Marcos) siguieron cuentas de diversos medios de comunicación de Andalucía, como @lavozdealmeria, @DiarioDALmeria, @EPAndalucia o canalsur, así como medios de comunicación italianos como el Corriere della Sera @Corriere, @sole24ore o @fattoquotidiano. Carlos confía además en la información que proporcionan cuentas oficiales: nacionales, como las de la Junta de Andalucía @AndaluciaJunta], el Ayuntamiento de Almería, @aytoalm, @AlmeriaJunta, el Ministerio de Transición Ecológica y Reto Demográfico, @mitecogob, o el Ministerio de Agricultura, Pesca y Alimentación, @mapagob. La conexión italiana justificaba el seguimiento de cuentas como la del Senado Italiano, @SenatoStampa, @Palazzo_Chigi o @minGiustizia

Las preferencias políticas de Carlos justifican su seguimiento de cuentas del Partido Popular como las de @JuanMa Moreno. Seguimiento de cuentas de representantes políticos del Partido Popular nacional y andaluz como @JuanMa_Moreno, @CarmenCrespoPP o @NunezFeijoo así como políticos conservadores europeos @Antonio_Tajani o Ursula Von Der Leyen @vonderleyen. También de cuentas europeas como el Partido de los Conservadores y Reformistas Europeos @ECR, el @EEPGroup, ECR Group, @ECRGroup. Carlos ha incorporado cuentas relativas al medio ambiente como @AgriculturAnd, @MedioAmbAND, o @ideagro o hiperlocales como @Culturalmeria o @AlmeriaInfo37, @Senderosalmeria o @almerialovers

Por último, cabe destacar la importancia de herramientas de verificaciones de información como Elections24Check o de análisis de tendencias, como es el caso de Google Trends, las cuales han permitido la búsqueda de los temas de conversación que, o bien no aparecían de forma orgánica, o daban nuevos subtemas de interés relacionados con las preocupaciones del perfil.

Aunque se deban analizar por separado los resultados de los perfiles de X (Twitter) y Facebook, independientemente de que sean la misma identidad digital, hay varias conclusiones generales que son de especial interés.



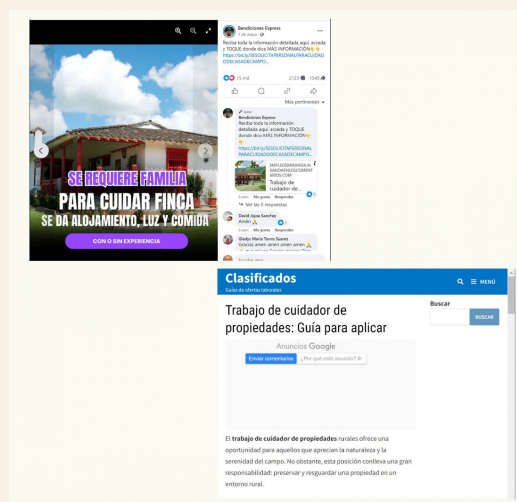


En primer lugar, cabe destacar la falta de desinformación orgánica y/o de forma explícita dentro de los dos perfiles, especialmente en Facebook, siendo necesarias búsquedas para lograr encontrar temas, especialmente el electoral. Por otro lado, la cantidad y el tipo de desinformación encontrada en ambas plataformas es considerablemente diferente, incluso con una misma identidad. Mientras que en Facebook destacan las estafas económicas y los contenidos genéricos, en X (Twitter) hay contenidos de odio y contenidos radicales ligados a desinformaciones sobre la inmigración, desinformaciones sobre medioambiente y la salud y sobre temas alimenticios.

Comenzando por Facebook destaca la aparición y recomendación de grupos de Facebook donde algunos usuarios publican supuestas ofertas de empleo bajo reclamo de grandes salarios sin concretar las condiciones. Los autores remiten a un formulario externa o una conversación privada con algunos de los responsables. Alguno de estos grupos, como Trabajos y Empleos en Campo Agrícola, congregan a más de 40.000 seguidores.



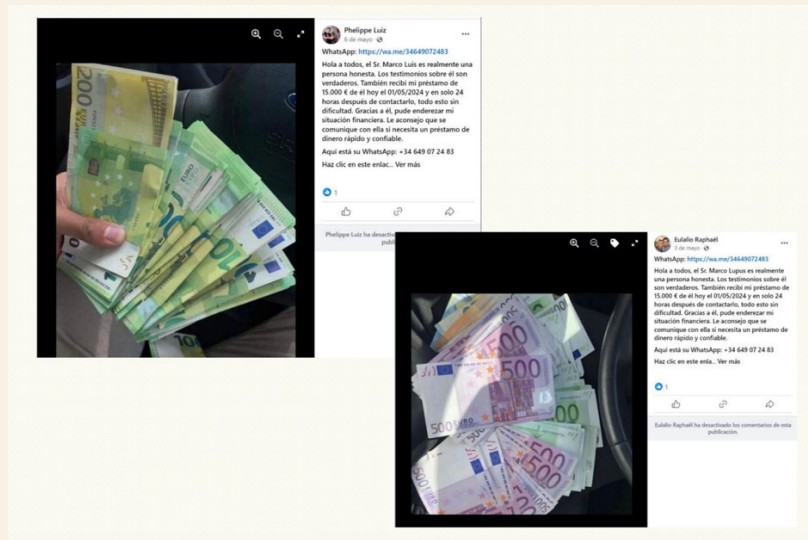
(Captura pantalla grupo Facebook “Trabajos y Empleos en Campo Agrícola/Agricultura”)



(Captura Página “Bendiciones Expres”)

Asimismo, es relevante destacar que aparecieron estafas económicas que utilizan suplantaciones de identidad a través de publicaciones en grupos de Facebook.



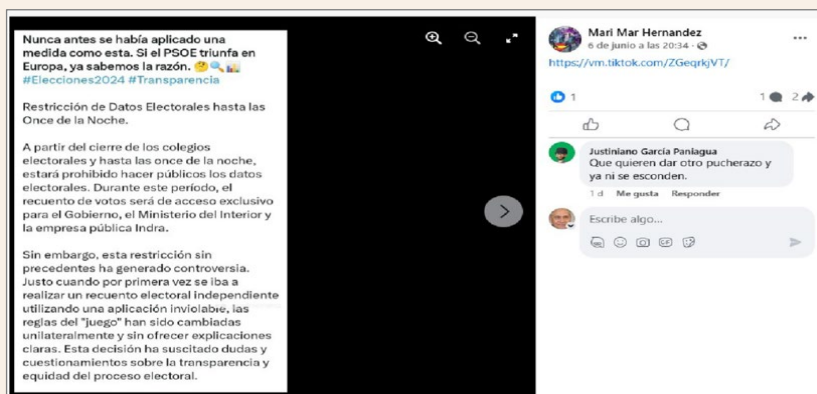


(Captura pantalla Instagram Philippe Luiz y Eulio Raphaël)

Estas estafas siguen el siguiente patrón: a) publica un contenido en un grupo de Facebook con un perfil robado en el que se aboga por préstamos o empleo, atrayendo al usuario, y b) se apela al usuario a que contacte, o por mensaje privado, o por WhatsApp a través de un enlace, moviendo la conversación al terreno personal.

Ejemplos de desinformación los encontramos en casos como la publicación de Eulio Raphaël en el que promociona los servicios de un supuesto prestamista Marco Lupus para la concesión de créditos. Sin embargo, al inspeccionar el contenido, se ve que es una suplantación de identidad, ya que aparece el mismo post en diferentes grupos.

El segundo caso corresponde a una desinformación cruzada, ya que enlaza a un post de X (Twitter). En este, publicado por Un Ciudadano Ejemplar, se vierten contenidos de odio a Susana Gisbert Grifo por aplicar una determinada condena contra varios comentarios en contra de la inmigración ilegal, pero al buscar la información no aparece ningún resultado más. Por último, publicado por Mari Mar Hernández, cabe destacar un post con desinformación cruzada que da difusión a un bulo sobre un apagón informativo en relación con las elecciones europeas.



En relación con X (Twitter), la mayoría de desinformación que ha llegado al perfil ha sido bulos genéricos como el que asegura que el Tribunal Constitucional ha dado validez legal a la ocupación de viviendas.





Captura pantalla cuenta @SrLiberal

No obstante, cabe resaltar la radicalización del perfil, siendo una identidad de carácter moderado y conservador. De esta manera, ha habido recomendaciones de cuentas que promovían el contenido de odio hacia los inmigrantes de forma constante, así como contenido en contra del cambio climático. Estas cuentas actuaban como falsos medios compartiendo vídeos de inmigrantes realizando actos ilegales o vandálicos en la vía pública, contribuyendo a la percepción negativa de este colectivo. Es el caso del contenido recomendado procedente de Europe Invasion, que reúne a más de 435.000 seguidores con sus publicaciones de corte racista que fustigan la presencia de inmigrantes en territorio europeo.



Captura pantalla @EuropeInvasionn

Siguiendo esta línea, los principales temas de la desinformación que han llegado a la cuenta son la inmigración, el cambio climático, la alimentación y algunas publicaciones sobre el COVID-19. En este sentido cabe destacar las publicaciones de una activa cuenta en X, @dejanirsilveir cuyas publicaciones han sido recomendadas en varias ocasiones por el algoritmo y que centra sus actividades sobre temas medioambientales. Un ejemplo es la falta atribución a un Premio Nobel de Física, Ivan Giaever, de la afirmación de que “el calentamiento global ES UNA ESTAFA”





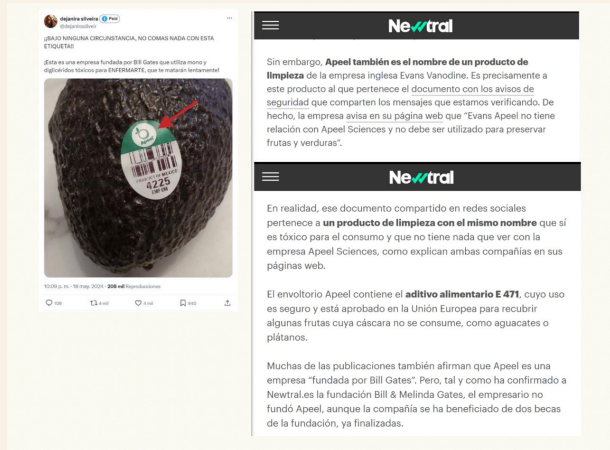
Algunos ejemplos de desinformación en la cuenta de X (Twitter) de Marcos Rodríguez son los casos El primero de ellos, publicado por @PapiTrumpo, corresponde a una fotografía manipulada por Inteligencia Artificial entre Elon Musk y Trump que entierran a los que denominan “Fake News Media”



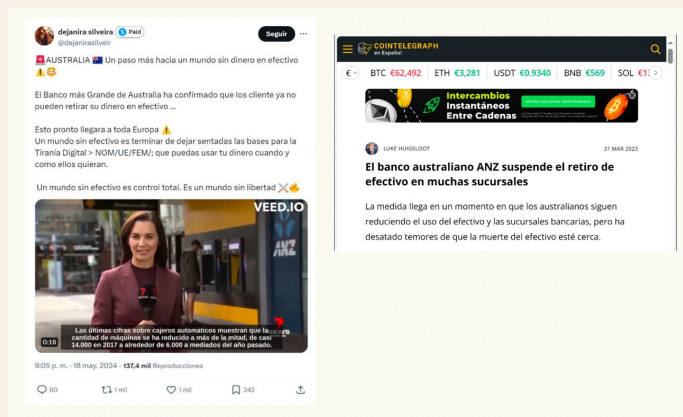
Captura pantalla Twitter @papitrumpo

Especialmente activa en contenidos falsos se muestra el perfil @dejanirasilveir, es sobre una empresa que emplea productos no aptos para la salud en la conservación de sus frutas. Sin embargo, una búsqueda de la desinformación demuestra que confunden deliberadamente el nombre de la empresa (dos empresas comparten una parte del nombre, pero omiten la que es diferente). Asimismo, atribuyen la empresa a Bill Gates cuando esta información no es real.





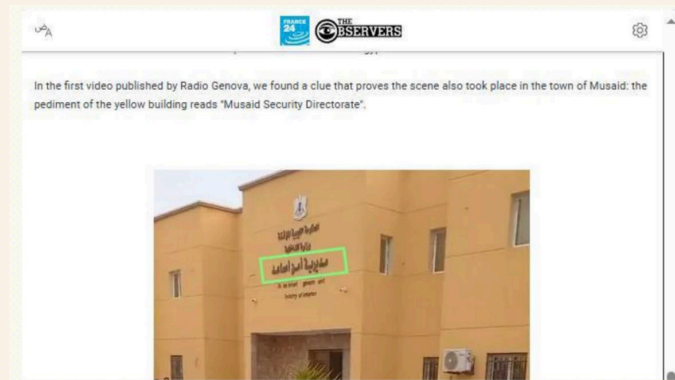
En otro registro la misma cuenta ofrece la noticia de la suspensión de las operaciones económicas en efectivo en territorio australiano.



Captura pantalla @dejaniresilveir

Un último ejemplo de desinformación lo constituye el recibido desde el perfil @orgullonaciona1. En él se afirma que ha llegado a Lampedusa una oleada de 1.200 inmigrantes. El vídeo no corresponder a Lampedusa (Italia), sino Libia.





Captura pantalla verificación imagen del equipo “The Observers” (France24)

Estos contenidos han sido verificados con búsquedas en Google, las notas de la comunidad en X (Twitter) o portales de verificación de información, como Elections24Check, una herramienta que reúne las verificaciones realizadas por organismos y medios certificados por el ICFN (International Fact-Checking Network) sobre publicaciones y material gráfico relacionado con el proceso electoral europeo. Su consulta resultó especialmente valiosa en la recta final del trabajo |





Mar en el planeta IA

Con el perfil de Mar quisimos comprobar cómo impactaba la desinformación en una cuenta con una preferencia ideológica muy clara, en este caso, EH Bildu, aunque sus convicciones políticas no fueron los únicos marcadores que utilizamos para establecer su identidad digital en Facebook y X, las dos redes sociales en las que abrimos perfiles para ella. Mar es una profesora de educación especial de mediana edad que ejerce su profesión en un centro de Barakaldo, en el País Vasco, de talante sociable que ama la montaña y el senderismo, un deporte que practica con frecuencia por su tierra o por la cercana Francia, donde viven algunos familiares. Mar utiliza esencialmente su cuenta de Twitter como un canal de información general y política. Entre los perfiles que sigue se encuentran medios de comunicación del territorio, como @eitbnoticias, @elcorreo_com, @deia_eus, @naiz_info, @garanet, algunas antenas locales de Barakaldo como @BarakaldoDigital o servicios de información autonómicos como @euskalmet, @112_SOSDeiak, @Biz_Basozainak. Como simpatizante de la izquierda abertzale, Mar sigue perfiles como los de @ehbildu, @mertxeaizpurua, @arnaldoOtegi, entre otros. Nuestro objetivo era “colorear” el perfil de una forma muy clara frente a un periodo electoral en Europa que se reveló más intenso de lo esperado. El 26 de febrero de 2024 el lehendakari Urkullu firmó el decreto de disolución del Parlamento vasco y convocó elecciones autonómicas que se celebraron el día 21 de abril. Por tanto, la experiencia de usuario de Mar se vio alterada por una campaña electoral dentro de su propia comunidad autónoma durante el periodo de observación que habíamos establecido para este estudio.

Definimos para Mar una especial preocupación por la inseguridad ciudadana, en concreto los robos de teléfonos móviles que sufren personas del entorno de Mar. Se trata de una preocupación trasladada desde la realidad de unos de los problemas que preocupan en el País Vasco. Según los datos oficiales, el territorio batió en 2023 el récord de criminalidad con más de 105.800 infracciones penales. Los robos con intimidación y violencia experimentaron un fuerte incremento en ese año. Mar no es ajena a esta realidad como también es consciente que muchos de estos robos están protagonizados por jóvenes extranjeros, como señalan las estadísticas. Queríamos comprobar cómo Mar vivía digitalmente esta situación en la que, desde una posición de absoluto rechazo al racismo, se encontraba confrontada a una situación que podía “criminalizar” a un colectivo inmigrante. Sin embargo, estos temas estuvieron ausentes de las recomendaciones que tanto X como Facebook realizaron a nuestro perfil de estudio durante todo el periodo de observación.

Nos llamó la atención también que, desde el principio, el algoritmo definió de forma muy clara la identidad autonómica de Mar y, en el caso de X, dejó fuera de su cuenta publicaciones de alcance nacional. También marcó su experiencia de usuario su afición por el fútbol y, específicamente, por su equipo el Athletic de Bilbao que, durante el periodo objeto de estudio, protagonizó una brillante trayectoria en la Copa del Rey de fútbol que culminó en la victoria del equipo en la final, celebrada en Sevilla el 6 de abril de 2024.

La movilización de los aficionados alcanzó una magnitud significativa y ocupó la conversación durante las semanas en que el equipo fue superando las eliminatorias hasta llegar a la final.





El éxito del equipo se convirtió en un factor de cohesión social que las redes sociales alimentaron viralizando todo tipo de contenidos emocionales que reflejaban el orgullo de pertenencia a una comunidad futbolística. Mar no fue ajena a esta realidad.

Los contenidos de apoyo al Athletic aparecieron en X y sobre todo en Facebook, cuenta en la que Mar sigue grupos como “Quejas de Barakaldo” algunas páginas sobre naturaleza y senderismo como Senderismo Euskadi o páginas de cocina como las de Carlos Arguiñano y Eva Arguiñano. mantenía contacto con algunos grupos de ciudadanos de Barakaldo. Mar también sigue en Facebook algunos medios franceses como las antenas regionales de la cadena de televisión France 3, France 3 Euskal Herria y France 3 Occitanie o las páginas del medio SudOuest Landes, a ser estas las más cercanas a su entorno geográfico.

Durante el periodo de observación los contenidos desinformativos o polarizadores que Mar recibió en X y Facebook se revelaron muy diferentes. Podríamos distinguirlos con la siguiente clasificación.

1. Estafas financieras. El perfil de Mar no quedó al margen de la oleada de publicaciones en X con las que, desde operaciones de suplantación de identidad de medios y celebridades, se intentaba que los usuarios dejaran sus datos personales para probar una aplicación de inversiones con criptomonedas con las que se obtenían, aseguraban, rentabilidades aseguradas. El 12 de abril llegó a la cuenta de X de Mar una primera publicación fraudulenta con la presentadora Cristina Pedroche como reclamo.





Una consulta al repositorio de anuncios de X, donde están reflejadas las publicaciones por cuya difusión que alguien ha pagado. Se identificó a partir del 22 de abril nuevas estafas, en este caso con el magnate estadounidense Elon Musk, como celebridad señuelo.



“Una oportunidad única para los residentes españoles. Inversión inteligente” Mar siguió recibiendo estos anuncios falsos en su cuenta durante todo abril y con menor intensidad desde finales de mayo hasta la celebración de las elecciones europeas.

El 12 de abril, Mar fue tentada por una nueva estafa, esta vez sugerida por el algoritmo de Meta con una operativa distinta a la anterior. Meta recomendó a Mar la publicación que realizaba un conciudadano de Barakaldo, Alberto López, dentro de un grupo de Facebook denominado Barakaldo. Alberto decía sentirse “asombrado” ante lo sucedido a otra ciudadana de Barakaldo “Un giro inesperado del destino llevó a Sofía de Barakaldo de la desesperación (al ser despedida de su trabajo en McDonalds) a la redención con 52.980 euros, después de que un cliente arruinase su vida”.

La publicación llevaba a una página web de inversiones que, comprobamos, pertenecía a una IP registrada en Rusia. Alberto López repitió el anuncio con diferentes enunciados durante todo el mes.



Captura de pantalla Grupo Facebook “Barakaldo”

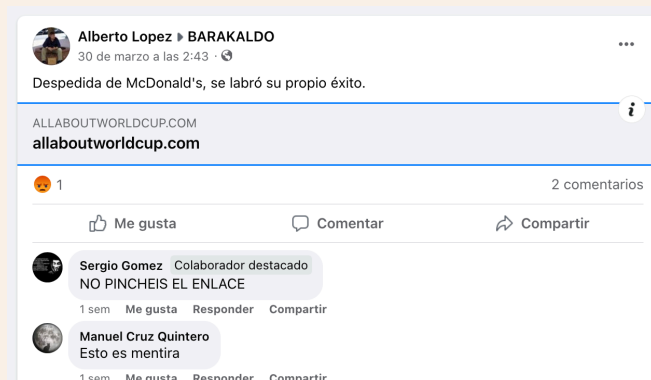




“De las patatas fritas a la fortuna. La trayectoria de una antigua alumna de MacDonald’s es realmente inspiradora” o “¡Es alucinante! No puedo creer que alguien despedido de un restaurante de comida rápida pueda llegar tan alto”.



El hecho de que la referencia estuviera integrada en una página local de Barakaldo y firmada por un ciudadano, al que probablemente, “hackearon” la cuenta de Facebook, introducía un factor adicional de peligro, especialmente al contacto con colectivos más vulnerables a la desinformación, como pueden ser personas de media o avanzada edad que consultan con frecuencia los contenidos aportados por sus vecinos en esos grupos de Facebook de carácter hiperlocal. Afortunadamente, algunos usuarios alertaron de la estafa dentro del hilo de comentarios.



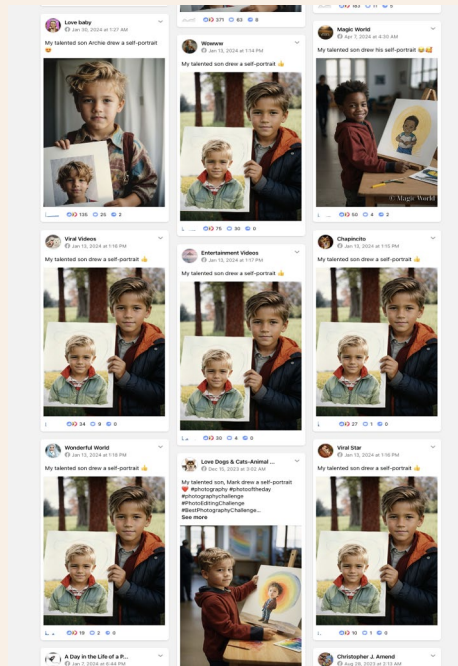
Falsos medios y páginas de viralización creados con IA

De la mano de la cuenta de Facebook de Mar vimos emerger desde principios del mes de mayo otra modalidad de desinformación alimentada con inteligencia artificial en todos los procesos de creación y difusión. El 6 de mayo Facebook sugirió a Mar un contenido de la página “Bonitos animales” que reproducía dos fotografías de un bonito y alegre pastor alemán.





Ambas imágenes presentaban un gran nivel de realismo así que decidimos que Mar aprobara la publicación con un “me gusta”. Comprobamos que habían sido realizadas con IA, si bien la aprobación de Mar ayudó a que llegaran recomendaciones de páginas similares como “Wowwww” en la que el usuario comparte la foto de su propio hijo “Mi talentoso hijo dibujó un autorretrato”, en las que aparece la imagen de un niño de 12 con años que sostiene un cuadro con su propio autorretrato.



Tras este, llegan a la cuenta de Mar otros contenidos similares sugeridos por páginas como Baby World o Simply Beautiful de imágenes realistas de animales, niños o jóvenes que realizan proezas artísticas que despiertan la admiración, y las ganas de compartir, de usuarios que, como Mar, pueden pensar que esos contenidos son reales.





Este es un fenómeno creciente de desinformación que los expertos Renee di Resta y Josh A Goldstein, del Observatorio de Internet de la Universidad de Standford han diseccionado a partir de un estudio de cien páginas de Facebook similares a las que fueron presentadas a Mar por el algoritmo de recomendación.



“Cabra pigmea” generada con IA para la cuenta @simplybeautiful

Como señalan los autores, esta tecnología está sirviendo a distintos objetivos de desinformación o actividad maliciosa: crecimiento de páginas, spam o estafas.

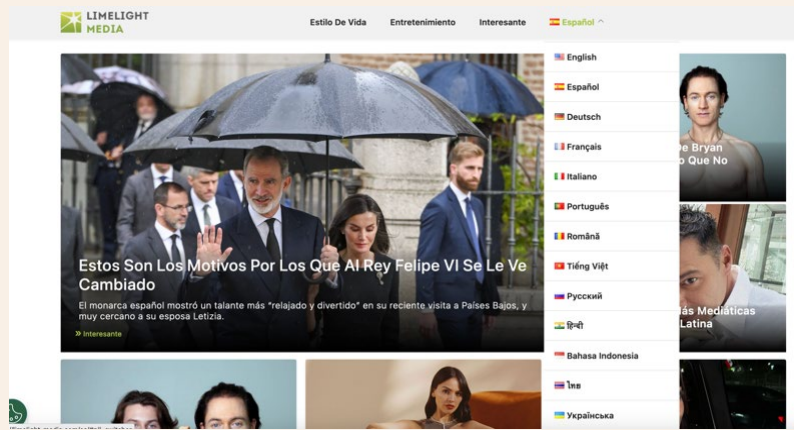
Los falsos medios alojados en el primer comentario de una publicación.

La consulta de los contenidos generados por IA que Mar empezó a recibir en su cuenta de forma cada vez más intensa permitió la identificación de un patrón común a muchas de estas publicaciones: se convertían en señuelos desde las que se invitaba a los usuarios a acceder a un enlace con más información que, sistemáticamente, estaba alojado en el primer comentario del post y no en la publicación principal.





Así el 22 de mayo, Meta sugiere a Mar una publicación de la página “La planeta increíble” en la que, junto a la imagen de un apuesto cantante y una niña” aparece el texto “La superestrella le pide a una niña pequeña que cante ‘You Raise me up’. Segundos después la niña derribó la casa Mira el vídeo en los comentarios abajo” En el primer comentario se aloja un enlace web que conduce a una granja de contenidos con titulares como “Vergüenza real. Las diez fotos más desafortunadas de Letizia”. Mar consulta este contenido y en la siguiente página aparece un nuevo medio, “Limelight”, con versiones en trece idiomas diferentes. La versión española está encabezada por una supuesta noticia “Estos son los motivos por los que al rey Felipe VI se le ve tan cambiado.” La página web que aloja este pseudomedio muestra anuncios de diferentes marcas nacionales e internacionales servida a través de los sistemas de publicidad programática.



Captura de pantalla Facebook del falso medio “Limelight Media”

Este patrón, contenido IA con una cascada de enlaces desde el primer comentario de la publicación inicial, se repetirá en la cuenta de Mar hasta el final del tiempo de observación.

Las fotos con IA serán el contenido predominante de todos los sugeridos en la cuenta de Mar hasta acabar por dominar la experiencia de usuario de forma clara.

Esta operativa es el exponente de unos de los negocios más lucrativos de la industria de la desinformación: los ingresos por publicidad digital que llegan a pseudo medios de información sin que, en muchas ocasiones, las marcas sean conscientes de que sus anuncios están siendo colocados en entornos de desinformación y contenidos “clickbait” poco recomendables. Desarrollaremos en el capítulo de conclusiones esta modalidad de desinformación que constituye una amenaza más a la industria de los medios de comunicación reales y comprometidos con la verdad.





Los inesperados nuevos amigos de Mercedes

Decidimos reproducir, en el caso de Mercedes, la forma de estar en las redes sociales de una persona de la tercera edad con intereses variados y una adscripción ideológica clara al partido político Vox, en el espectro ideológico de la derecha radical. Asignamos para ella la red Facebook, la tercera red más usada en España (69,7% de los usuarios de redes sociales según el estudio We Are Social de Hootsuite y, junto a Instagram y Whatsapp, la preferida por la población de más de 55 años como indican los datos del Estudio Anual de Redes Sociales que publica IAB Spain.

Imaginamos a Mercedes como una señora de Valladolid, jovial y activa, muy unida a sus hijos, interesada por la actualidad y aficionada a la copla. En la cuenta de Facebook de Mercedes se cruzan las páginas de admiradores de Antonio Molina como Recordando a Antonio Molina con otras relacionadas con la Semana Santa de su ciudad, como el grupo público Semana Santa, que congrega en Facebook a 8.934 miembros. Mercedes sigue también varias páginas de cocina como Cocinera y Madre donde se congregan más de 270.000 seguidores. Otras páginas como Acuérdate, con 834.000 seguidores, ofrecen recuerdos y vestigios de la vida de los españoles desde 1960.

Como se indicó anteriormente, asignamos a esta identidad digital una tendencia conservadora, la de una señora preocupada por los desafíos a los que España se enfrenta y, sobre todo, preocupada por la unidad de España y la conservación de valores éticos e identitarios que Mercedes considera hay que salvaguardar para que España no pierda la esencia de los que ha sido históricamente. Dimos forma a estas inquietudes siguiendo a perfiles identificados con Vox, como el del presidente del partido, pero también gracias a grupos de Facebook donde se veían representadas distintas sensibilidades de la derecha como España, Luz de Trento, Mi querida España, Orgulloso de ser español, España y punto, Club de fans Santiago Abascal Patriotas por España, Isabel Díaz Ayuso, la futura Presidenta de España, Es decir, tratamos de integrar a Mercedes de pleno en una conversación con una gran vitalidad dentro de los canales de Facebook.

Paralelamente establecimos para Mercedes una conexión internacional con Francia a través de uno de sus hijos, residente en Estrasburgo. Gracias a esta vinculación el perfil fue diseñado con un interés en Francia, su política y, en el caso de Mercedes, las actividades del partido que en Francia se sitúa en el mismo espectro político de Vox como es Rassemblement National, el partido de Marine Le Pen. Por esta razón Mercedes comenzó a seguir las publicaciones del grupo Amis Patriotes de Marine Le Pen que resultó un caudal activo de publicaciones como veremos posteriormente.

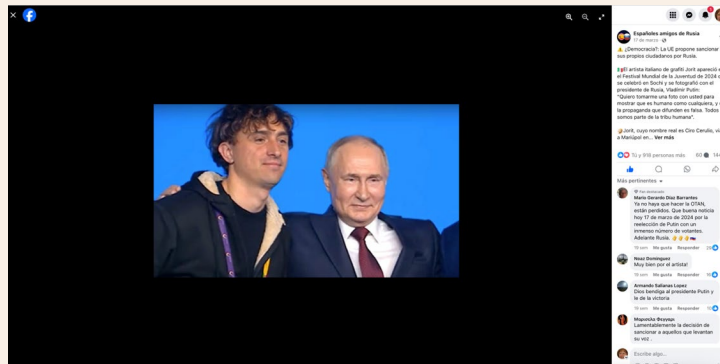
La irrupción de los contenidos prorrusos

Desde el principio de nuestro periodo de observación el algoritmo de recomendación de Facebook empezó a recomendar a Mercedes páginas relacionadas con Rusia o su presidente Vladimir Putin. Así el 18 de marzo llegó al feed la publicación del grupo Españoles amigos de Rusia titulada: “¿Democracia? La UE propone sancionar a sus propios ciudadanos en Rusia. La publicación iba acompañada de una foto de Putin posando con un supuesto grafitero italiano que afirmaba “Quiero tomarme una foto con usted y demostrar que la propaganda que difunden es falsa”





Capturas pantalla Grupo Facebook “Españoles amigos de Rusia”



Los contenidos antieuropeos desde páginas rusas se repitieron durante todo el periodo de observación. (24) El 3 de abril, el algoritmo integró en el feed de Mercedes varias publicaciones de la página De Russie en français, que alternaba fotografías aparentemente reales del metro de Moscú con memes insultantes hacia el presidente francés, Emmanuel Macron y que cuestionan la sexualidad de su esposa Brigitte. El meme reproduce la foto de Macron que habla por teléfono a Putin, en la parte inferior de la fotografía: “¡Hola, Vlad tenemos que hablar de hombre a hombre” “Entonces pásame a Brigitte”.

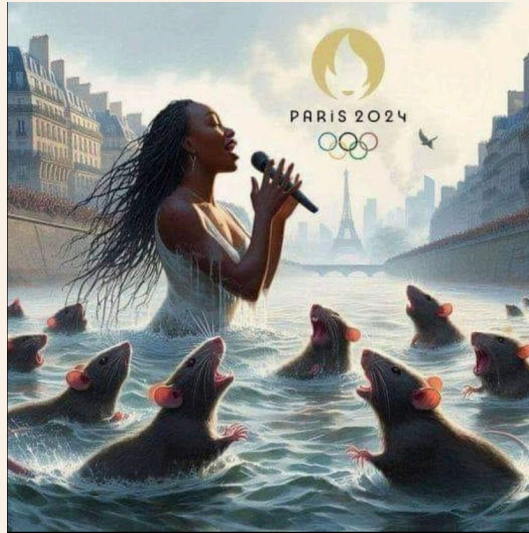


(Captura pantalla grupo Soutien a la Russie)

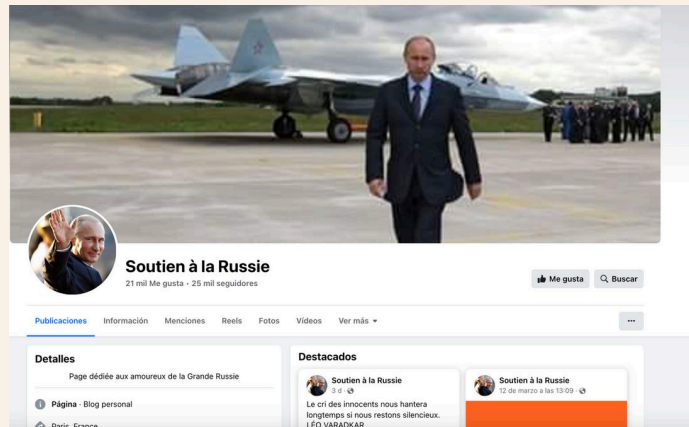




Poco más tarde otro contenido anti Francia llega desde la misma cuenta. Es una ilustración generada por inteligencia que representa a una mujer de raza negra que canta desde las aguas del Sena rodeada de ratas, en una evocación de uno de los momentos de la ceremonia de inauguración de los Juegos Olímpicos.



El algoritmo incorpora un tercer grupo prorruso a la cuenta de Mercedes, en este caso, Soutien a la Russie (Apoyo a Rusia) encabezada por una gran foto de Vladimir Putin.



(Captura pantalla grupo Facebook “Soutien á la Russie”)

Desde esta cuenta aparecen, como en el caso anteriores, publicaciones muy diversas, desde las fotografías de bonitas calles de Moscú a publicaciones insultantes hacia Francia y sus políticos. Por ejemplo, esta página denomina a la alcaldesa de París, Anne Hidalgo, “la emperatriz de la mediocridad y de las ratas” tras declarar Hidalgo, presuntamente, que los atletas rusos y bielorrusos “no son bienvenidos” a los Juegos Olímpicos.





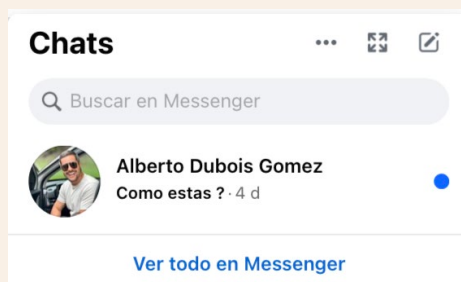
Captura pantalla "Soutien a la Russie"

Asimismo, encontramos en esta cuenta numerosas publicaciones que tratan de dañar la imagen del presidente francés, Emmanuel Macron, particularmente atacado por las campañas de desinformación durante la campaña electoral para las elecciones europeas. Macron es un consumidor de drogas, de una sexualidad muy libre.



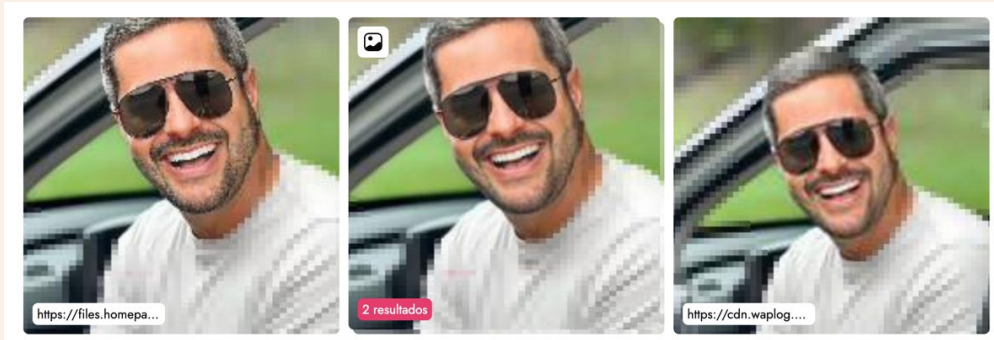
Suplantación de identidad

Mercedes ha recibido algunas solicitudes de amistad por parte de hombres atractivos y más jóvenes que ella. Un patrón que se repite de forma generalizada en las redes sociales y que a veces son el origen de estafas económicas a señoras de cierta edad (27) que creyeron estar iniciando con estos hombres relaciones de amistad o íntimas en la distancia. Los medios de comunicación se hacen eco regularmente de estos casos.



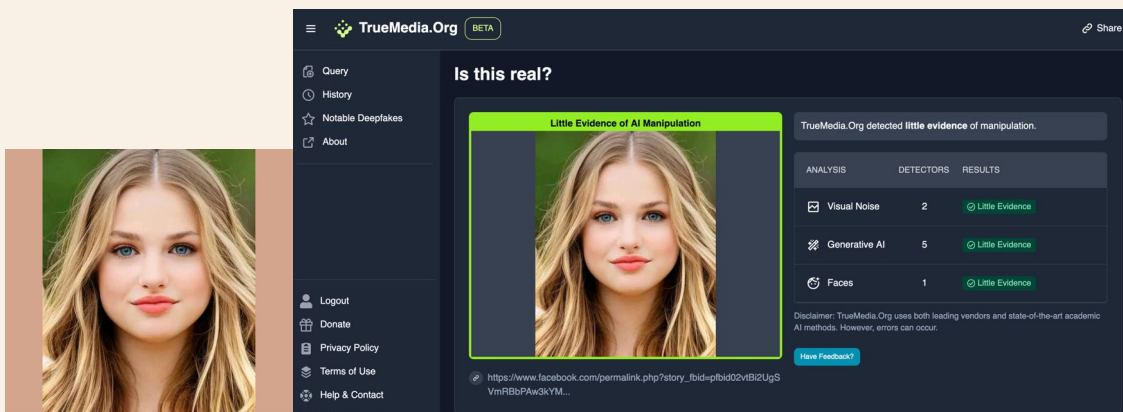


El 11 de marzo, el usuario de Facebook Albert Dubois envió una solicitud de amistad a la señora Mercedes. La foto de su perfil mostraba a un hombre de unos cincuenta años, abundante pelo canoso, gafas de sol y una gran sonrisa que posaba en el interior de un coche. Albert Dubois dirigió un mensaje privado en español a Mercedes. Recurrimos a la herramienta de IA de reconocimiento facial PimEyes para verificar la imagen con la que Dubois se presentaba en su foto de perfil. El mismo hombre aparecía con identidades diferentes en otros canales digitales.



Falsos medios de IA

“Intenta no jadear cuando veas estas fotos escandalosas de Letizia”, asegura el titular de uno de los falsos medios generados con IA al que Mercedes accedió desde una página de contenidos con IA con una operativa similar a las que vimos en el perfil de Mar, es decir, ilustraciones atractivas que son utilizadas como señuelos para que el usuario de Facebook entre a falsos medios que usan titulares llamativos para generar tráfico dentro de su entramado de páginas.



En el caso de Mercedes, algunos de los señuelos utilizado durante el periodo de ejercicio fueron ilustraciones y fotografías falsas de la princesa Leonor que habían sido generadas con IA. Discurso antieuropeo

Al caudal de contenidos críticos con líderes europeos como Macron que emanaban de los contenidos prorrusos recomendados a Mercedes por el algoritmo de Meta, se unieron algunas aportaciones desde el grupo “Amis Patriotes de Marine Le Pen”. La fotografía de un perro defecando sobre la bandera de la Unión Europea apareció a mediados de mayo como contenido destacado en el menú de publicaciones del Facebook de Mercedes. Toda una declaración de principios.







03

Conclusiones



El estudio que hemos realizado sobre la desinformación en redes sociales durante el periodo de las elecciones europeas de 2024 arroja luces sobre un fenómeno creciente y complejo, que se articula en torno a nuevas formas de manipulación digital que amenazan los cimientos de nuestras democracias. A lo largo de esta investigación, que se basó en la observación de ocho perfiles ficticios con diferentes características ideológicas y demográficas, se han identificado patrones de desinformación que no solo se sostienen en técnicas conocidas, sino que se sofistican gracias a la inteligencia artificial (IA) y la personalización algorítmica.

1. La desinformación como fenómeno continuo y adaptativo

Una de las principales conclusiones que emergen de este estudio es la constatación de que la desinformación ya no es un fenómeno puntual vinculado exclusivamente a momentos de alta actividad política, como las campañas electorales. En cambio, se ha transformado en un flujo continuo, con actores que operan fuera del radar institucional y que, gracias a la persistencia de las plataformas digitales, pueden mantener sus narrativas durante largos periodos. Esto supone un desafío mayor para las instituciones democráticas, que tradicionalmente han estructurado sus esfuerzos de control y verificación en torno a calendarios electorales.

La capacidad de estos actores para adaptar sus mensajes y técnicas a diferentes perfiles, utilizando IA para generar contenido personalizado y dirigido, incrementa la eficacia de sus campañas. Esto se observa claramente en la interacción de los perfiles que creamos, donde aquellos con intereses específicos o vulnerabilidades claras, como la preocupación por la inmigración o el medio ambiente, fueron objeto de mensajes desinformativos especialmente diseñados para resonar con esas preocupaciones.

2. La inteligencia artificial y la sofisticación de los bulos

Otro punto clave es el papel cada vez más relevante de la inteligencia artificial en la generación de contenidos desinformativos. Este estudio ha documentado la proliferación de imágenes, vídeos y textos manipulados mediante IA, lo que no solo amplifica el alcance de la desinformación, sino que dificulta enormemente su detección y refutación. Los deepfakes y las suplantaciones de identidad digital se han convertido en herramientas habituales para los desinformadores, quienes ahora cuentan con tecnologías que les permiten recrear con gran realismo figuras públicas o fabricar hechos completamente falsos.

Este avance tecnológico supone un reto para los mecanismos de verificación tradicionales, que se ven obligados a evolucionar rápidamente para no quedar obsoletos. Además, la IA facilita la creación de lo que podríamos llamar "bulos comodín", narrativas falsas que se replican con ligeras variaciones en distintos contextos geográficos y temporales, aprovechando el desconocimiento o la confusión de los usuarios sobre temas complejos o polémicos.

3. La polarización como estrategia y resultado

La polarización emerge en este estudio como una estrategia central en las campañas de desinformación, pero también como un resultado preocupante de estas prácticas. Los perfiles que analizamos recibieron de manera sistemática contenidos que exacerbaban divisiones sociales y políticas, presentando realidades simplificadas y antagonistas que buscan movilizar a los usuarios hacia posiciones extremas.

Una de las conclusiones más inquietantes de nuestro estudio es la aparente imposibilidad de mantener una postura moderada o apolítica en las redes sociales. A pesar de que los perfiles creados para esta investigación





no expresaban en su mayoría preferencias ideológicas claras, el algoritmo de las plataformas insistió en exponerlos a contenidos polarizadores. Esta tendencia parece reflejar un sesgo estructural dentro de los sistemas de recomendación, diseñados para maximizar la interacción del usuario, lo que a menudo significa priorizar contenidos que generen fuertes reacciones emocionales, como la indignación o el miedo.

Este fenómeno no es nuevo y ha sido documentado en estudios anteriores. Un informe del Center for Democracy & Technology (2020) concluyó que los algoritmos de recomendación en plataformas como YouTube, Facebook y Twitter tienden a promover contenidos más extremistas o polarizantes, incluso cuando el usuario no ha mostrado interés previo en tales temas. Otro estudio, realizado por el Pew Research Center en 2021, encontró que los usuarios que pasaban más tiempo en redes sociales eran más propensos a adoptar posiciones políticas extremas, sugiriendo que la exposición continua a contenidos polarizadores refuerza actitudes radicales.

Lo que nuestro estudio confirma es que esta dinámica de radicalización no discrimina; afecta por igual a usuarios ideológicamente indefinidos, arrastrándolos hacia narrativas extremistas a través de la personalización algorítmica. Esta inercia hacia los extremos plantea serias preguntas sobre el impacto de las redes sociales en la moderación del discurso público y la salud de las democracias contemporáneas.

4. La segmentación y personalización de la desinformación

La segmentación de la desinformación es otro aspecto crítico que el estudio pone de relieve. Los perfiles fueron objeto de una personalización extrema en los contenidos desinformativos recibidos, que se ajustaban no solo a sus intereses manifiestos, sino también a sus vulnerabilidades latentes. Esto se hizo evidente en casos como el de David, el joven con interés en la inmigración, que recibió una avalancha de contenidos radicales y xenófobos, o el de Aitana, cuya preocupación por el medio ambiente la convirtió en blanco de bulos sobre políticas climáticas.

Esta personalización no solo incrementa la eficacia de las campañas de desinformación, sino que también plantea preguntas sobre la responsabilidad de las plataformas en la protección de sus usuarios. Las redes sociales, al priorizar la retención de usuarios sobre la calidad de la información, permiten que estos contenidos encuentren su camino hacia las personas más susceptibles de ser influenciadas.

5. Desinformación transnacional y narrativas recurrentes

Por último, el estudio subraya la dimensión transnacional de la desinformación, con narrativas que se replican en diferentes países y contextos, adaptándose ligeramente para encajar en los discursos locales. La cuestión migratoria es un claro ejemplo de esto, con bulos que cruzan fronteras y se adaptan a los miedos y prejuicios de cada sociedad. Estos "bulos comodín" no solo se reproducen, sino que se refuerzan mutuamente, creando una sensación de verdad debido a su omnipresencia.

La recurrencia de estas narrativas plantea un desafío adicional para la verificación y educación mediática, ya que los usuarios se enfrentan a una repetición constante de las mismas mentiras, que terminan por normalizarse o aceptarse como posibles.

Un aspecto crucial dentro de la desinformación transnacional es el papel destacado que ha jugado Rusia en las estrategias globales de manipulación informativa contra Occidente. (24) Desde hace más de una década, Rusia





ha perfeccionado un modelo de desinformación que opera a múltiples niveles y en diversos contextos, adaptándose con notable eficacia a las particularidades de cada país objetivo. Esta estrategia, a menudo denominada como parte de una "guerra híbrida", combina operaciones de ciberseguridad, manipulación mediática y uso de redes sociales para debilitar la cohesión interna de las naciones occidentales, sembrando dudas sobre sus instituciones democráticas.

La maquinaria de desinformación rusa ha sido prolífica en la creación y difusión de narrativas que explotan las divisiones sociales y políticas en Occidente. A través de redes sociales como X (anteriormente Twitter) y Facebook, y con el respaldo de medios estatales como RT y Sputnik, se han amplificado temas sensibles como la inmigración, el nacionalismo, y la soberanía, con el objetivo de polarizar a las sociedades occidentales. Estas narrativas, cuidadosamente diseñadas, a menudo recurren a la fabricación de bulos o la manipulación de hechos, presentando a Rusia como un contrapeso moral y político a un Occidente corrupto y decadente.

En este sentido, la cuestión migratoria ha sido uno de los temas recurrentes explotados por la desinformación rusa, presentando a Europa como una región en decadencia, incapaz de controlar sus fronteras y vulnerable a una supuesta "invasión" de inmigrantes. Estos relatos, aunque falsos o exagerados, han encontrado eco en sectores de la población occidental ya predispuestos a desconfiar de sus gobiernos o instituciones supranacionales como la Unión Europea.

La utilización de inteligencia artificial por parte de actores pro-Kremlin ha añadido una capa adicional de sofisticación a estas campañas, permitiendo la creación de contenido altamente personalizado y creíble, que se difunde rápidamente gracias a las poderosas herramientas de segmentación que ofrecen las plataformas digitales. De este modo, Rusia no solo actúa como un propagador de desinformación, sino como un modelador de las percepciones públicas en Occidente, exacerbando tensiones internas y debilitando la unidad frente a desafíos globales como la seguridad y la estabilidad democrática.

6. El auge de las estafas financieras en redes sociales: un problema generalizado

Uno de los aspectos más alarmantes que ha evidenciado en nuestro estudio es la proliferación de estafas financieras, particularmente aquellas relacionadas con criptomonedas, que han inundado las redes sociales en los últimos años. Estas estafas no solo afectan a usuarios individuales, sino que también socavan la confianza en las plataformas digitales, las cuales parecen incapaces de controlar el flujo de contenido fraudulento que circula a través de sus sistemas.

Durante nuestro análisis, observamos que varios de los perfiles de estudio, sin importar su demografía o intereses específicos, fueron repetidamente expuestos a este tipo de contenido. Las estafas con criptomonedas, en particular, han encontrado un terreno fértil en redes como Facebook e Instagram, donde los estafadores se aprovechan de la falta de conocimientos técnicos de los usuarios y de la creciente popularidad de estas monedas digitales para engañarlos. (16) A menudo, estas estafas se presentan bajo la apariencia de ofertas irresistibles, respaldadas por supuestas celebridades o figuras públicas, cuyas identidades han sido suplantadas mediante sofisticadas técnicas de manipulación digital.

El problema de las estafas financieras en redes sociales no es nuevo, pero ha crecido exponencialmente con el auge de las criptomonedas (21) Un informe de The Independent (2021) señalaba que las denuncias por estafas relacionadas con criptomonedas en redes sociales se habían disparado un 60% en solo un año, reflejando una





tendencia preocupante. Facebook, en particular, ha sido objeto de críticas por su incapacidad para detener la difusión de este tipo de contenido. Según un estudio publicado por la organización Which? en 2022, Facebook fue la plataforma en la que se reportó el mayor número de estafas financieras, a menudo relacionadas con inversiones fraudulentas en criptomonedas.

El modus operandi de estas estafas es cada vez más sofisticado. Los estafadores emplean técnicas de ingeniería social para ganar la confianza de sus víctimas, prometiendo rendimientos exorbitantes y rápidos a cambio de pequeñas inversiones iniciales. Estos fraudes son facilitados por el diseño de las plataformas sociales, que permiten una fácil segmentación y personalización del contenido publicitario.

El uso de inteligencia artificial ha permitido a los estafadores crear contenido visual y textual de gran calidad, que imita a la perfección las características de publicaciones legítimas. Esto no solo dificulta la detección de estas estafas por parte de los usuarios, sino también por parte de los mecanismos de control de las propias plataformas, que a menudo dependen de la denuncia de los usuarios para tomar medidas.

La expansión de las estafas financieras en redes sociales subraya la necesidad urgente de una mayor regulación y de herramientas más efectivas para detectar y eliminar contenido fraudulento. Sin una respuesta contundente, las redes sociales corren el riesgo de convertirse en un campo de minas digitales donde la seguridad financiera de los usuarios está en constante peligro.

7. Los falsos medios y las granjas de contenidos, una seria amenaza a la supervivencia de los medios de información

Como hemos observado profusamente a lo largo de este trabajo de investigación, una gran parte de la desinformación que encontramos en las redes sociales no tiene por objeto modificar la percepción que la opinión pública de una sociedad como la nuestra alberga sobre cuestiones ideológicas o políticas. Más bien, tras sofisticadas operaciones de manipulación de contenidos, se esconde la voluntad de enredar a los usuarios en una impostura de medios de comunicación con un fin meramente económico.

Este fenómeno no es nuevo. Desde 2010, la desinformación sobre temas de salud en las redes sociales ha experimentado un crecimiento notable. Inicialmente, circulaban mitos y remedios caseros sin evidencia científica, pero con la expansión de las plataformas digitales, estos contenidos se multiplicaron y diversificaron. Estudios previos a la pandemia ya advertían sobre la proliferación de información errónea en salud y alertaban sobre la facilidad de publicación y difusión en redes sociales que permitía que se compartieran ampliamente consejos médicos no verificados y teorías sin fundamento. Esta tendencia no solo desinformaba al público, sino que también ponía en riesgo la salud de las personas que seguían recomendaciones inapropiadas.

La pandemia de COVID-19, declarada en 2020, acentuó significativamente este problema. Durante este periodo, se observó un aumento drástico en la difusión de información falsa relacionada con la enfermedad y sus tratamientos. (36) Las redes sociales se inundaron de enlaces a sitios web que, bajo la apariencia de medios legítimos, promovían remedios no comprobados e incluso peligrosos. Estas prácticas no eran fortuitas; el principal objetivo detrás de estos falsos medios era aprovechar el miedo y la incertidumbre para generar ingresos mediante la monetización de visitas y publicidad. Al atraer a usuarios con promesas de soluciones rápidas o milagrosas, estos portales lograban captar una audiencia considerable, incrementando sus ganancias a través de anuncios programáticos o incluso ventas directas de productos de dudosa eficacia





En un contexto como el actual, donde las principales redes sociales como X o Meta (Facebook e Instagram) han eliminado sus programas de verificación de información al sustituirlos por mecanismos de comprobación colaborativa (notas de comunidad) cabe preguntarse si no estamos ante una nueva oleada de desinformación sobre temas de salud, como investigadores en redes sociales y expertos en salud pública.

La inteligencia artificial ha llevado este negocio a otro nivel. En primer lugar esta tecnología permite generar ilustraciones o pretendidas fotografías cuyo contenido apela directamente a las emociones de los usuarios al resultar conmovedoras, tiernas o simplemente chocantes. (35) En un reflejo adquirido por nuestras sociedades en la era del consumo digital, las personas tienden a interactuar con ese tipo de contenidos que son comentados y compartidos. En esta interactividad reside la primera parte del negocio: cuánto más viral, más exposición tendrá la página en una red social. El contenido “señuelo” siempre incluye, como hemos comentado anteriormente, un primer comentario que invita al usuario a ampliar la información en un enlace externo. Se trata de la puerta de acceso a un falso medio o granja de contenidos. En esta fase la inteligencia artificial alimenta todo el proceso, desde el diseño y la creación de una web como la de los contenidos, en su mayoría falsos artículos de prensa sobre personajes conocidos y celebridades. Estas publicaciones suelen abordar temas controvertidos o emocionalmente impactantes, lo que asegura así un alto nivel de visibilidad. Aprovechando esta exposición, los actores malintencionados insertan comentarios que contienen enlaces o referencias a sitios web que, aunque parecen ser fuentes legítimas, son en realidad portales diseñados para engañar al usuario.

Una característica destacada y especialmente preocupante de estos sitios es que frecuentemente contienen espacios publicitarios que generan ingresos económicos. Lo alarmante es que muchas veces las empresas reales pagan por estos anuncios a través de sistemas de publicidad programática sin conocer el contexto en el que aparecerán sus marcas. (37) Esto no solo proporciona financiamiento a los falsos medios, sino que también les otorga una legitimidad aparente ante los ojos de los usuarios. Además, estas páginas pueden recopilar información personal de los visitantes, aumentando los riesgos para la privacidad y seguridad de los datos.

A fin de conocer en profundidad la operativa de las redes de falsos medios ha resultado de gran valor documental el trabajo de investigación realizado por un equipo de la cadena France 2 y plasmado en el documental. “Fake News, la machine à fric” (Noticias falsas, la máquina de dinero) (34) que indaga sobre el lucrativo negocio de la desinformación en línea. El reportaje explica cómo sitios web que difunden noticias falsas generan ingresos a través de publicidad de grandes marcas.

La disección deja el negocio al descubierto: los sitios de noticias falsas, como los encontrados durante este trabajo de investigación, se financian principalmente a través de la publicidad en línea. Las marcas pagan para que sus anuncios aparezcan en estos sitios, a menudo sin ser conscientes de que están financiando la desinformación. Los sitios de noticias falsas utilizan redes publicitarias, como Google, para mostrar anuncios a sus visitantes. Estas redes obtienen una comisión por cada anuncio que se muestra, lo que genera un flujo constante de ingresos para los sitios de noticias falsas. Los responsables de estas páginas web fraudulentas consiguen en ocasiones, además, eludir los sistemas de filtrados de las marcas cambiando el nombre y la URL de los falsos medios que, como los aparecidos en nuestra investigación, aparecen cuando un usuario cliquea en el primer comentario de una publicación de Facebook.

El trabajo de investigación realizado en Francia pone de relieve además una grave disfunción: aunque muchas marcas tienen políticas de control para evitar aparecer en sitios de desinformación, a menudo sus anuncios terminan apareciendo allí debido a errores en la configuración de las campañas publicitarias por parte de las agencias de publicidad. Las marcas, a menudo sin saberlo, se convierten en fuente de ingresos para estos sitios.





Los resultados de una investigación realizada por **NewsGuard y comScore** (38) concluyen que la industria de la **desinformación ingresa 2.600 millones de dólares cada año procedentes de la publicidad programática**, incluyendo millones de ingresos que respaldan sitios web con declaraciones falsas relacionadas con salud, mitos sobre las vacunas, propaganda partidista o desinformación electoral.

El auge de los falsos medios o granjas de contenidos no solo desinforma al público, sino que también socava el modelo económico de los medios de comunicación legítimos. Al capturar una porción significativa del tráfico digital y los ingresos publicitarios, estas granjas de contenido privan a los medios auténticos de recursos esenciales para mantener la calidad y el alcance de su labor informativa. Según estudios recientes, el desvío de ingresos publicitarios hacia sitios fraudulentos afecta especialmente a los medios locales y regionales, que dependen en gran medida de la publicidad para financiar sus operaciones.

Existe actualmente una dificultad añadida para determinar el alcance exacto de este fenómeno de desinformación generada con IA: la imposibilidad de su observación general. Si tenemos en cuenta el importante papel que las redes sociales desempeñan en la creación de su modelo de negocio, toda desaparición iniciativa de transparencia por parte de las plataformas digitales es una mala noticia. En este sentido es importante recordar que, en el caso de Meta (Facebook e Instagram) el pasado verano se decidió el cierre de Crowdtangle, una herramienta interna de acceso público con la que expertos en desinformación y en redes sociales disfrutaron de un acceso global a la actividad de ambas redes en el mundo (39) Ello permitió un conocimiento de las dinámicas de comunicación en las redes imprescindibles para su observación y comprensión. Con una herramienta como Crowdtangle, hubiera sido posible categorizar y vigilar, con mayor precisión y método, las páginas de contenidos generados con inteligencia artificial, lo que nos hubiera permitido tener una idea más precisa de la entidad de este negocio fraudulento en alza.

Estas perversas dinámicas en torno a la falsedad como negocio crean una competencia desleal, donde los medios reales deben enfrentar no solo la pérdida de ingresos, sino también una reducción en su capacidad de llegar a su audiencia. A medida que disminuyen los recursos disponibles, se hace más difícil para las redacciones invertir en periodistas, investigaciones y tecnología, lo que a su vez debilita su posición frente a los desafíos del ecosistema digital. Además, la proliferación de tanta desinformación deteriora la confianza del público en los medios de comunicación en general, alienta el escepticismo hacia las fuentes legítimas.

El fenómeno de los falsos medios en las redes sociales plantea un desafío significativo para la integridad del ecosistema informativo y la sostenibilidad económica de los medios legítimos. El abordaje de este problema requiere un esfuerzo coordinado entre las plataformas digitales, los anunciantes y los propios medios de comunicación. Las redes sociales deben implementar medidas más rigurosas para identificar y eliminar comentarios que redirijan a falsos medios, mientras que los anunciantes deben adoptar políticas estrictas para garantizar que sus recursos no financien desinformación. Por su parte, los medios legítimos pueden desempeñar un papel activo al educar al público sobre la importancia de verificar las fuentes y al colaborar en iniciativas para combatir la propagación de contenido fraudulento.

Solo mediante un enfoque integrado y estratégico será posible mitigar el impacto de esta amenaza y garantizar un acceso a la información veraz y de calidad.





Conclusión general

En conjunto, este estudio revela un panorama preocupante donde la desinformación se ha convertido en un componente permanente y sofisticado de la conversación digital. La combinación de inteligencia artificial, personalización algorítmica y la capacidad de operar en múltiples contextos geográficos y culturales, hacen de la desinformación un fenómeno difícil de erradicar.

Las conclusiones aquí presentadas deben servir como base para futuras acciones tanto a nivel de regulación como de educación mediática, con el fin de fortalecer la resiliencia de nuestras sociedades frente a estas amenazas. Solo a través de un esfuerzo coordinado y sostenido podremos mitigar el impacto de la desinformación y preservar la integridad de democracias de las que los medios de comunicación constituyen un actor esencial.





BIBLIOGRAFÍA

- (1) Rubio, R. (2023, May 25). Anatomía de un bulo de fraude electoral. Iustel. Retrieved August 26, 2024, from https://www.iustel.com/diario_del_derecho/noticia.asp?ref_iustel=1233669
- (2) Más allá de Cambridge Analytica. Los académicos que extraen datos de Facebook” Sheera Frenkel. (2018) Richtel, M., & Rosenberg, M. (2018, May 9). FACEBOOK PONE RESTRICCIONES A ACADÉMICOS TRAS LA CONTROVERSA DE CAMBRIDGE ANALYTICA. The New York Times. Retrieved August 26, 2024, from <https://www.nytimes.com/es/2018/05/09/espanol/facebook-academicos-cambridge-analytica.html>
- (3) De Empoli, G. (2019). LOS INGENIEROS DEL CAOS. Oberon. <https://www.oberonlibros.com/libro/los-ingenieros-del-caos-9788441542198/>
- (4) Maldita.es. (2023, May 26). *Guía de desinformación electoral: Bulos y desinformación en las elecciones de España y Argentina*. Maldita.es. <https://maldita.es/malditobulo/20230526/desinformacion-electoral-bulos-elecciones-espana-argentina>
- (5) We Are Social Spain. (2024, January). Digital 2024. *We Are Social*. <https://wearesocial.com/es/blog/2024/01/digital-2024/>
- (6) 5Rights Foundation. (n.d.). PATHWAYS: HOW DIGITAL DESIGN PUTS CHILDREN AT RISK. Retrieved January 28, 2025, from <https://5rightsfoundation.com/resource/pathways-how-digital-design-puts-children-at-risk/>
- (7) Beard, M. (2024, January 13). Over 100 deep-faked Rishi Sunak ads found on Meta’s platform. *Fenimore Harper Communications*. <https://www.fenimoreharper.com/research/deepfake-political-advertising>
- (8) Le Monde. (2024, March 26). DERRIERE LES FAUSSES INTERVIEWS D’ÉLISE LUCET OU DE JAMEL DEBBOUZE, LES MAFIAS DES ARNAQUES AUX PLACEMENTS. Le Monde. https://www.lemonde.fr/pixels/article/2024/03/26/derriere-les-fausse-interviews-d-elise-lucet-ou-de-jamel-debbouze-les-mafias-des-arnagues-aux-placements_6224203_4408996.html





- (9) Euractiv. (2023, July 23). TikTok fails disinformation test before EU vote, study shows. *Euractiv*. <https://www.euractiv.com/section/elections/news/tiktok-fails-disinformation-test-before-eu-vote-study-shows/>
- (10) Schüler, M., Degeling, M., Romano, S., & Meßmer, K. (2024, June 6). Spanish-language ads amplify scams using audio deepfakes of public figures. *TikTok Audit*. <https://tiktok-audit.com/blog/2024/Search-Suggestions/>
- (11) Schüler, M., Degeling, M., Romano, S., & Meßmer, K. (2024, June 6). Spanish-language ads amplify scams using audio deepfakes of public figures. *TikTok Audit*. <https://tiktok-audit.com/blog/2024/Search-Suggestions/>
- (12) Crossover. (2024, July). The TikTok effect. *Crossover*. <https://crossover.social/wp-content/uploads/2024/07/TheTiktokEffect.pdf>
- (13) Foundation for European Progressive Studies. (2024). *Youth support for democracy in Spain*. <https://feps-europe.eu>
- (14) Newman, N., Fletcher, R., Schulz, A., Andi, S., Robertson, C. T., & Nielsen, R. K. (2024). *Reuters Institute Digital News Report 2024*. Reuters Institute for the Study of Journalism. <https://www.digitalnewsreport.org/>
- (15) Instituto de la Juventud de España. (2021). *El extremismo de derecha: Un estudio sobre la percepción de los jóvenes*. https://www.injuve.es/sites/default/files/adjuntos/2021/05/estudio_injuve_el_extremismo_de_derecha.pdf
- (16) Ruiz, A. (2022, noviembre 10). *Los jóvenes y la desinformación recibida en las redes sociales*. InformaUVA. <https://informauva.com/los-jovenes-y-la-desinformacion-recibida-en-las-redes-sociales/>
- (17) France 24. (2024, April 30). LA UE PRESIONA A INSTAGRAM Y FACEBOOK POR RIESGOS DE DESINFORMACIÓN ELECTORAL ANTES DE LAS EUROPEAS. France 24. <https://www.france24.com/es/minuto-a-minuto/20240430-la-ue-presiona-a-instagram-y-facebook-por-riesgos-de-desinformaci%C3%B3n-electoral-antes-de-las-europeas>



- (18) Eguía, L. (2023, julio 20). *Instagram limita por defecto contenido político y usuarios no lo sabían*. Wired <https://es.wired.com/articulos/instagram-limita-por-defecto-contenido-politico-y-usuarios-no-lo-sabian>
- (19) Ertzaintza. (n.d.). *Estadísticas delictivas: Memoria*. Gobierno Vasco. Retrieved August 26, 2024, from <https://www.ertzaintza.euskadi.eus/lfr/web/ertzaintza/estadisticas-delictivas>
- (20) Stanford Internet Observatory. (2024, August 7). AI SPAM ACCOUNTS BUILD FOLLOWERS WITH 'FOLLOW TRAINS' AND ENGAGEMENT BAIT. Stanford University. <https://cyber.fsi.stanford.edu/io/news/ai-spam-accounts-build-followers>
- (21) Himelein-Wachowiak, M., Giorgi, S., Devoto, A., Rahman, M., Ungar, L., Schwartz, H. A., Epstein, D. H., Leggio, L., & Curtis, B. (2021). Bots and misinformation: Quantifying the impact of COVID-19 on online discourse. *Journal of Medical Internet Research*, 23(5), e26933. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10956171/>
- (22) Salas, J. (2024, March 1). ¿Engañan las noticias falsas sobre todo a los más mayores? *Solidaridad Intergeneracional*. <https://solidaridadintergeneracional.es/wp/enganan-las-noticias-falsas-sobre-todo-a-los-mas-mayores/>
- (23) IAB Spain. (2024). Estudio de redes sociales 2024. *IAB Spain*. <https://iabspain.es/estudio/estudio-de-redes-sociales-2024/>
- (24) Bengani, P., & Jackson, J. (2024, July 6). Russian disinformation networks ramp up attacks on European elections. *The Bureau of Investigative Journalism*. <https://www.thebureauinvestigates.com/stories/2024-07-06/russian-disinformation-networks-ramp-up-attacks-on-european-elections/>
- (25) Weiss, M., Dobrokhotov, R., & Grozev, C. (2024, July 4). "Morality and ethics should play no part": Leaks reveal how Russia's foreign intelligence agency runs disinformation campaigns in the West. *The Insider*. <https://theins.ru/en/politics/272870>
- (26) Suárez Pérez, D. (2024, June 6). Spanish-language ads amplify scams using audio deepfakes of public figures. *DFRLab*. <https://dfrlab.org/2024/06/06/spanish-language-ads-amplify-scams-using-audio-deepfakes-of-public-figures/>



- (27) Lacort, J. (2024, January 27). Facebook se ha convertido en un coladero para las estafas de amor: el crimen de Morata es solo el ejemplo extremo. *Xataka*. <https://www.xataka.com/seguridad/facebook-se-ha-convertido-coladero-para-estafas-amor-crimen-morata-solo-ejemplo-extremo>
- (28) Maldita.es. (2024, June 11). La respuesta de las plataformas a la desinformación sobre las Elecciones Europeas 2024. *Maldita.es*. <https://maldita.es/nosotros/20240611/plataformas-desinformacion-respuesta-elecciones-europeas/>
- (29) Beard, M. (2024, January 13). Over 100 deep-faked Rishi Sunak ads found on Meta's platform. *Fenimore Harper Communications*. <https://www.fenimoreharper.com/research/deepfake-political-advertising>
- (30) Qurium Media Foundation. (2024, March 26). Into the crypt of cloned media. *Qurium Media Foundation*. <https://www.qurium.org/alerts/into-the-crypt-of-cloned-media/>
- (31) Le Monde. (2024, March 26). DERRIERE LES FAUSSES INTERVIEWS D'ÉLISE LUCET OU DE JAMEL DEBBOUZE, LES MAFIAS DES ARNAQUES AUX PLACEMENTS. *Le Monde*. https://www.lemonde.fr/pixels/article/2024/03/26/derriere-les-fausse-interviews-d-elise-lucet-ou-de-jamel-debbouze-les-mafias-des-arnaques-aux-placements_6224203_4408996.html
- (32) European Commission. (2023, July 13). Commission adopts new measures to protect citizens from online disinformation. *European Commission*. https://ec.europa.eu/commission/presscorner/detail/en/ip_23_6709
- (33) Disinfo.eu. (2024, February 5). TikTok factsheet. *Disinfo.eu*. https://www.disinfo.eu/wp-content/uploads/2024/02/20240205_TikTok_factsheet.pdf
- (34) .France 2. (2020). COMMENT LES FAKE NEWS MENENT A DES ARNAQUES EN LIGNE. *France 2*. Recuperado de <https://www.youtube.com/watch?v=IRMvfuS8-uc>





- (35) DiResta, R. (2024). SPAMMERS, SCAMMERS, AND AI-GENERATED IMAGES: EXPLOITING FACEBOOK'S ALGORITHMS. Harvard Kennedy School Misinformation Review. Recuperado de https://misinforeview.hks.harvard.edu/wp-content/uploads/2024/08/diresta_spammers_scammers_ai_images_facebook_20240815.pdf
- (36) ABC News. (2025, 13 de enero). COULD META ENDING FACT-CHECKING LEAD TO RISE IN HEALTH MISINFORMATION? ABC News. Recuperado de <https://abcnews.go.com/Health/meta-ending-fact-checking-lead-rise-health-misinformation/story?id=117480644>
- (37) Petit, Q. (2024, diciembre 16). LOS ANUNCIANTES ANTE LA DESINFORMACIÓN: “LAS COMPAÑÍAS TIENEN QUE SABER DÓNDE PONEN SU DINERO”. EL PAÍS. <https://elpais.com/comunicacion/2024-12-16/los-anunciantes-ante-la-desinformacion-las-companias-tienen-que-saber-donde-ponen-su-dinero.html>
- (38) NewsGuard. (n.d.). BRANDS SEND BILLIONS TO MISINFORMATION WEBSITES: NEWSGUARD-COMSCORE REPORT. NewsGuard. <https://www.newsguardtech.com/special-reports/brands-send-billions-to-misinformation-websites-newsguard-comscore-report/>
- (39) KGI at Georgetown University. (n.d.). THE CROWDTANGLE FUNERAL. Georgetown University. Retrieved January 28, 2025, from <https://kgi.georgetown.edu/research-and-commentary/the-crowdtangle-funeral/>





www.fundaciónlucadetena.org