Informe C

Desinformación en la era digital

Una amenaza compleja para las democracias

Resumen C 1

Introducción 3

Un marco conceptual en evolución 3
Desinformación y otros desórdenes informativos 3
Narrativas para la desinformación 5
Alcance y relevancia en la era digital 5

Un nuevo contexto socio-informativo 6
Emisor: instigadores y difusores 7
Canales: impacto digital y prevalencia de canales clásicos 7
Contenido 8
Receptor 8

Fenómenos contemporáneos implicados en el auge desinformativo 9
La confianza y el marco democrático 9
Mediación informativa y periodismo 9
Fragmentación social 10
Cognición y vulnerabilidad individual 10
Gobernanza digital y modelo de negocio 12
Tecnologías impulsoras de la desinformación 13

Impacto 16

Combatir la desinformación: agentes y estrategias mitigantes 17
Garantías, detección y neutralización 18
Avances regulatorios 22
Privacidad, seguridad y regulación electoral 25

Una mirada estratégica y participativa hacia el futuro 25

Ideas fuerza 26

Referencias I

Cómo citar este informe:
Oficina de Ciencia y Tecnología del Congreso de los Diputados (Oficina C). Informe C: Desinformación en la era digital. (2023) www.doi.org/10.57952/j3p6-9086

Personal experto e investigador consultado (por orden alfabético)

» **Arteaga, Félix[1].** Investigador Principal, Real Instituto Elcano, España.

» **Arroyo Guardeño, David[1].** Científico Titular, Instituto de Tecnologías Físicas y de la Información "Leonardo Torres Quevedo", Consejo Superior de Investigaciones Científicas, España.

» **Badillo Matos, Ángel[1].** Investigador Principal, Real Instituto Elcano, España. Profesor Titular, Universidad de Salamanca, España.

» **Cano Orón, Lorena[1].** Ayudante doctora, Universitat de Valencia, España.

» **Cardenal Izquierdo, Ana Sofía[1].** Profesora titular, Universitat Oberta de Catalunya (UOC), España.

» **Carrillo, Nereida[1].** Profesora asociada, Universidad Autónoma de Barcelona, España. Cofundadora, asociación de educación mediática Learn to Check.

» **Corredoira y Alfonso, Loreto[1].** Profesora titular, Universidad Complutense, España. Jean Monnet Chair (2020-2023). Co-directora del Observatorio Complutense de la Información y Grupo Dir-Politics.

» **Ecker, Ullrich K.** H. Catedrático, Universidad de Australia Occidental, Australia.

» **García, David[1].** Catedrático, Universidad de Konstanz, Alemania. Miembro del Compelxity Science Hub Vienna, Austria.

» **González Bailón, Sandra.** Catedrática, Universidad de Pensilvania, Estados Unidos.

» **Innerarity, Daniel[1].** Catedrático de Filosofía Política, Ikerbasque, Universidad del País Vasco/Euskal Herriko Unibertsitatea (UPV/EHU). Cátedra Inteligencia Artificial y Democracia, Instituto Universitario Europeo de Florencia, Italia.

» **Jiménez Cruz, Clara.** Chair, European Fact-Checking Standard Network and International Fact-Checking Network, Europa. Directora, Maldita.es, España.

» **Magallón Rosa, Raúl[1].** Profesor Titular, Universidad Carlos III, España.

» **Majó-Vázquez, Silvia[1].** Investigadora Postdoctoral, Instituto Reuters, Universidad de Oxford., Reino Unido. Profesora adjunta, Universidad Libre de Ámsterdam, Países Bajos.

» **Rosso, Paolo[1].** Catedrático, Universitat Politécnica de València, España.

» **Rubio Núñez, Rafael[1].** Catedrático, Universidad Complutense, España. Co-director, Observatorio Complutense de la Información SN-Disorders.

» **Salaverría, Ramón[1].** Catedrático, Universidad de Navarra, España. Miembro, MSI-RES Committee of Experts on Increasing Resilience of Media, Council of Europe.

» **Wagner, Astrid[1].** Científica Titular, Instituto de Filosofía. Consejo Superior de Investigaciones Científica (CSIC), España.

EQUIPO C

Alfonso Cuenca. Letrado de las Cortes Generales. Director de Estudios, Análisis y Publicaciones del Congreso de los Diputados.

Ana Elorza. Coordinadora de la Oficina C en la Fundación Española para la Ciencia y la Tecnología.

Izaskun Lacunza*. Coordinadora de la Oficina C en la Fundación Española para la Ciencia y la Tecnología.

Maite Iriondo de Hond. Técnica de evidencia científica y tecnológica.

Rüdiger Ortiz-Álvarez. Técnico de evidencia científica y tecnológica.

Sofía Otero. Técnica de evidencia científica y tecnológica.

Jose L. Roscales*. Técnico de evidencia científica y tecnológica.

Cristina Fernández-García. Técnica de conexión con la comunidad científica y la sociedad.

*Personas de contacto para este informe

Método de elaboración

Los Informes C son documentos breves sobre los temas seleccionados por la Mesa del Congreso que contextualizan y resumen la evidencia científica disponible para el tema de análisis. Además, recogen las áreas de consenso, disenso, las incógnitas y los debates en curso. Su proceso de elaboración se basa en una exhaustiva revisión bibliográfica que se complementa con entrevistas a personal experto y dos rondas de revisión posterior por su parte. La Oficina C colabora con la Dirección de Documentación, Biblioteca y Archivo del Congreso de los Diputados en este proceso.

Para la redacción del presente informe la Oficina C ha referenciado 492 documentos y consultado a un total de 18 personas expertas en la materia. Se trata de un grupo multidisciplinar en el que el 83 % pertenece a las ciencias sociales relaciones internacionales, ciencias políticas, ciencias de la comunicación, sociología, derecho), y el 17 % pertenece a las ciencias físicas e ingeniería (ciencias de la computación, informática, ingeniería de telecomunicaciones). El 56 % trabaja en centros o instituciones españolas mientras que el 44 % está afiliado al menos a una institución en el extranjero.

La Oficina C es la responsable editorial de este informe.

[1] Especialistas que también han participado en la revisión total o parcial del informe.

Resumen C

El informe en 5 minutos

Relevancia

Internet y el desarrollo digital propician múltiples avances y beneficios económicos y sociales. También un nuevo contexto social e informativo que ha favorecido una amplificación sin precedentes de la desinformación y sus efectos, convirtiéndola en una desatacada amenaza para los sistemas democráticos. Se trata de una cuestión de seguridad nacional que alcanza niveles críticos en situaciones de gran relevancia social, como crisis sanitarias, conflictos bélicos o los procesos electorales. Puede repercutir negativamente sobre bienes públicos como la salud y erosionar los procesos e instituciones democráticos, así como algunos derechos fundamentales, como el de información. A la vez, su gestión es compleja, ya que algunos de estos derechos, como la libertad de expresión, podrían verse restringidos si no se actúa con la cautela y precisión necesarias. La ciudadanía se muestra mayoritariamente preocupada por la cuestión, a la vez que existen síntomas claros de su indefensión.

Este informe ahonda en las causas e impactos del fenómeno, así como en los mecanismos que pueden ayudar a combatirlo.

La desinformación

De acuerdo con la Comisión Europea, la desinformación se refiere a la información verificablemente falsa o engañosa que se crea, presenta y divulga con fines lucrativos o para engañar deliberadamente a la población, y que puede causar un perjuicio público. Puede perseguir lucro económico, tener fines ideológicos y electoralistas o atender a intereses geopolíticos. De hecho, la desinformación es parte del conjunto de acciones propias de las "amenazas híbridas" mediante las que terceros países tratan de explotar las vulnerabilidades de la Unión Europea. Además de instigadores internacionales, existen también instigadores nacionales como grupos de interés por razón ideológica, religiosa, económica u otros, cuyo perjuicio puede ser equiparable.

En la práctica, adopta múltiples formas y no es siempre sencilla de identificar. Bajo amplias narrativas, muchas de alcance internacional, la desinformación construye relatos que sustituyen la verdad por verosimilitud y entremezcla contenidos falsos y veraces. Son flexibles, por lo que se adaptan localmente y a la actualidad informativa para impregnar cualquier tema de relevancia social, incidente o confrontación que pueda emerger. Para tener éxito, no necesariamente necesita generar falsas creencias, es suficiente con provocar confusión, desconfianza, dividir y amplificar sesgos y prejuicios. Persigue así cambios de fondo o estructurales en la esfera pública más que resultados inmediatos en torno a una noticia falsa concreta. Los instigadores se apoyan en técnicas como el uso de la atracción afectiva, de visiones simplistas e incompletas, en la simple repetición o en la utilización de la inteligencia artificial, como en el caso de las ultrafalsificaciones, para aumentar su efectividad. En España, entre los principales temas objeto de desinformación, se encuentran la política y los procesos electorales o algunos desafíos sociales como la migración.

En el foco

Las causas y efectos de la desinformación ahondan en el nuevo contexto informativo, fuertemente mediado por internet, y en el conjunto de factores geopolíticos, económicos y tecnológicos, sociales e individuales, que modulan la relación con la información tanto dentro como fuera de la red. Se trata de un fenómeno multifacético, multifactorial e inmerso en una serie de dinámicas que se solapan y se refuerzan mutuamente, donde no siempre es posible establecer relaciones causa-efecto inequívocas. Aunque la comunidad experta se muestra de acuerdo en los riesgos, la complejidad del fenómeno dificulta un análisis integral del impacto de la desinformación.

Las redes sociales, servicios de mensajería y grandes plataformas digitales han modificado la forma en la que la información fluye y llega a la ciudadanía. Por un lado, propician una explosión informativa de diferentes calidades que dificulta la identificación de contenidos veraces y genera incertidumbre. Por otro, constituyen la principal fuente de información y a la vez difuminan el flujo informativo: cualquiera puede generar contenido, difundirlo y compartirlo. En esta desintermediación, los intérpretes clásicos de la realidad, como medios periodísticos, la televisión o élites políticas, pierden relevancia. A pesar de ello, estos mantienen un papel decisivo en la amplificación de la información falsa o verdadera que circula por internet y, por tanto, de su impacto.

Entre las múltiples dimensiones interrelacionadas, la comunidad experta señala la importancia que la evolución del marco geopolítico tiene en el uso creciente y, de forma cada vez más eficaz, de las operaciones de desinformación como herramienta de desestabilización por parte de algunos países. En clave sistémica, destaca también una disminución de la confianza en las instituciones democráticas. La pérdida de bienestar económico y social, el aumento de la desigualdad o la insatisfacción, entre otros aspectos, pueden actuar como grietas que refuerzan la vulnerabilidad de los estados y sus sociedades ante la desinformación. Los medios de comunicación periodísticos, con estructuras profesionales debilitadas y una confianza ciudadana mermada, pierden eficacia como freno a la amenaza. En el plano más social, el marco de la posverdad, la creciente polarización afectiva y la circulación de teorías conspiranoicas favorecen la fragmentación social y reinterpretan la relación de la sociedad con la falsedad y la veracidad: en su extremo puede encontrarse la negación de la evidencia objetiva y la aceptación de la mentira. A nivel personal, estas realidades convergen con múltiples factores que, como los sesgos cognitivos y algunos factores socioafectivos, pueden predisponer a los individuos a

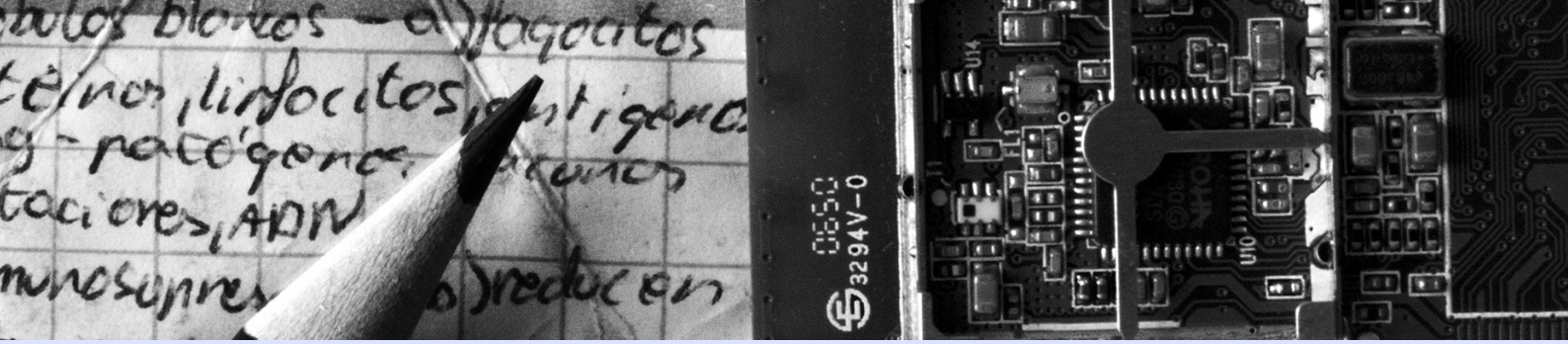

Imagen FOTCIENCIA Fragmentos de memoria ©Gabriel Castilla Cañamero

creer y diseminar la información falsa. En conjunto configuran una fuerte resistencia en las personas para corregir las creencias erróneas o aceptar la falsedad de la información que resulta afín.

Finalmente, la dimensión tecnológica viene definida por la parte del modelo de negocio digital que dificulta la neutralidad y pluralidad de la información que reciben los usuarios. Este modelo persigue captar la atención del usuario para monetizarla con publicidad, para lo que aprovecha los avances de la inteligencia artificial y el análisis masivo de datos que sustentan las plataformas. Destacan, entre ellos, los sistemas algorítmicos de cribado o la publicidad y la propaganda personalizada y dirigida. Se trata de herramientas con gran capacidad para amplificar el impacto de la desinformación. Especialmente, preocupa la falta de transparencia algorítmica, ya que limita la comprensión de su papel al respecto. Además, la falta de conocimiento en torno al flujo de información falsa en las redes de mensajería privada o los rápidos avances en el uso de la inteligencia artificial por ejemplo a través de bots o los *deep-fakes* para desinformar completan el desafío tecnológico.

Horizonte

La comunidad experta recalca que el tratamiento de la desinformación requiere de la combinación y coordinación de múltiples instrumentos y medidas para mitigar sus efectos a corto plazo y para diseñar estrategias que permitan combatirla de forma estructural a largo plazo. Así, apela a la responsabilidad y cooperación de todos los agentes, desde las instituciones y actores políticos e informativos, hasta las grandes plataformas digitales y comerciales, para no explotar la incertidumbre y la desinformación e incorporar, como moderadoras, medidas para frenarla. El objetivo general es la resiliencia y la alfabetización mediática y digital del conjunto de la sociedad. Para lograrlo, las políticas públicas pueden apoyarse en un amplio conjunto de medidas que refuercen el papel de los principales agentes implicados y de la propia ciudadanía, incluyendo la acción regulatoria.

Las instituciones democráticas afrontan el desafío estructural de fomentar un diálogo con la ciudadanía que refuerce la confianza y se adecue al nuevo contexto informativo. Se apunta también al fortalecimiento de la función periodística, potenciando sus capacidades y medios, independencia, transparencia y pluralidad, como medida para mitigar de la desinformación. Además, las agencias de verificación juegan un papel importante y positivo para la sociedad a través de la vigilancia y refutación de la información falsa, papel que además puede ser reforzado por otros actores.

La resiliencia ciudadana a la desinformación puede reforzarse mediante planes de alfabetización mediática. Existen múltiples propuestas al respecto y la psicología avanza en el desarrollo de mecanismos efectivos para neutralizar la desinformación a nivel individual. A nivel social, destaca el fomento de un marco ético que oriente el comportamiento de las personas o cualquier agente relacionado hacia el rechazo de la desinformación y que promueva el rediseño de las arquitecturas propias de redes sociales y plataformas digitales para dificultar el flujo desinformativo.

Las políticas y el marco regulatorio europeos promueven medidas orientadas a defender y reforzar la democracia y consolidar mecanismos que combatan la desinformación de forma sistémica, desde la atribución de responsabilidades a las grandes plataformas digitales o la desmonetización, hasta la extensión de la pluralidad y libertad de medios y la moderación de la contienda electoral online. Recientemente, ha entrado en vigor la Ley europea de servicios digitales que destaca por señalar responsabilidades concretas en torno a la desinformación en el ecosistema digital, entre otros aspectos. De hecho, las grandes redes sociales y plataformas digitales juegan un papel moderador esencial en el flujo desinformativo, por lo que han de considerase aliadas necesarias en la lucha contra la desinformación. Aunque han emprendido múltiples medidas para combatirla, la comunidad experta aun señala importantes desafíos en este aspecto destacando, por ejemplo, la necesidad de una mayor transparencia.

Mientras España avanza con el marco europeo, la comunidad experta hace hincapié en la importancia de consolidar una estrategia nacional que integre las múltiples dimensiones de la desinformación y vertebre el desarrollo de políticas públicas en la materia. Esta incluye, entre otros aspectos, un plan de alfabetización mediática y digital. Al mismo tiempo, son elementos esenciales en todas las actuaciones la cooperación de los sectores público y privado, y con la sociedad civil, la transparencia y rendición de cuentas y la cooperación internacional.

Desinformación en la era digital

Introducción

La desinformación repercute negativamente sobre la salud o la seguridad ciudadana. Erosiona la democracia y algunos derechos fundamentales, que no pueden darse por sentados y han de ser protegidos.

Internet y el avance digital constituyen uno de los pilares del desarrollo económico y social[1]. Supone una gran oportunidad con múltiples beneficios sociales, como la amplificación y transversalización de las posibilidades de las personas para explorar y acceder a la información. Sin embargo, como con otras tecnologías, su uso malicioso o inadecuado puede suponer nuevas amenazas o la amplificación de las que ya existían fuera de línea[2]. Así, aunque la desinformación no es un fenómeno novedoso, sí lo son la extensión y gravedad de los riesgos que conlleva en la era digital.

La información juega un papel central en la sociedad; es la materia prima con la que se construye el conocimiento y la discusión y toma de decisiones democráticas[3–6]. De ser falsa o engañosa, pervierte este proceso, pudiendo llegar a bloquearlo. Así, la desinformación se incluye entre las principales amenazas a los sistemas democráticos, ya que puede afectar negativamente a bienes públicos como la salud, la economía y la seguridad nacional e, incluso erosionar el Estado de derecho y derechos fundamentales o influir sobre los resultados electorales o deslegitimarlos en base a información falsa o engañosa[3,5,7–10]. Si bien el desarrollo digital ha multiplicado el alcance de la información falsa, se trata de un fenómeno complejo con causas sistémicas asentadas en la confianza en las instituciones, la economía, la sociedad y los canales y tecnologías que median la información que se consume[6,11–16].

La desinformación preocupa a la mayoría de la población que, además, muestra signos de indefensión ante la dificultad para reconocer la veracidad de las fuentes o percibir su propia manipulación[17]. Sus efectos son acumulativos y perduran en el tiempo, ya que es muy resistente a la corrección: es difícil aceptar el error en las creencias u opiniones propias. Por otro lado, carece de fronteras y opera también en espacios de gestión privada, como redes sociales, por lo que su gestión requiere la cooperación internacional y público-privada[11,18]. Además, en este desafío, la ciudadanía constituye el escenario central[19]: la mente es el espacio donde se produce la batalla cognitiva de la desinformación. Así, la alfabetización mediática y digital es un paso clave hacia la resiliencia.

La Comisión Europea ha emprendido medidas contra la desinformación, señalando que la democracia no puede darse por sentada; necesita ser cuidada y protegida de forma activa[20]. En España, los poderes públicos reconocen la amenaza que supone para el Estado[21,22].

Un marco conceptual en evolución

Desinformación y otros desórdenes informativos

Aunque la desinformación es un fenómeno concreto que atañe a información falsa o engañosa que persigue causar daño, su presentación formal es diversa y su identificación, causas y efectos son complejos y multidimensionales.

No existe un consenso general en torno a la definición de desinformación, que cuenta con múltiples propuestas para su clasificación[11,23,24]. Una de las clasificaciones más extendidas es la de los desórdenes informativos, que incluye tres conceptos distintos pero relacionados[11,25,26]:

- **Información errónea (*misinformation* en inglés):** Es falsa, se comparte inadvertidamente y sin intención dañina porque alguien ha sido engañado, por creencias genuinas o simple descuido.
- **Información dañina (*malinformation* en inglés):** Puede ser real o falsa, no siempre es verificable y se comparte con el fin de causar daño explícito. Puede incluir opinión, información personal o de otro tipo robada o expuesta sin permiso[25].
- **Desinformación (*disinformation* en inglés):** Es información falsa y persigue causar daño. La Comisión Europea la define como información verificablemente falsa o engañosa que se crea, presenta y divulga con fines lucrativos o para engañar deliberadamente a la población, y que puede causar un perjuicio público[11]. Este perjuicio incluye su capacidad para dañar el debate público y los procesos democráticos, especialmente en el ámbito electoral, además de impactar negativamente sobre bienes públicos como la salud, la economía o la seguridad, entre otros[22,27–30].

En la práctica, la desinformación puede ser difícil de identificar[24,25,31,32]. Por ejemplo, normalmente, no es posible determinar la intencionalidad del emisor: la información falsa puede nacer con una intención dañina, pero su difusión puede ser amplificada por personas que la comparten sin esa intención[5]. De hecho, se aprovecha de la buena intención de la ciudadanía u otros agentes para ser distribuida[33]. Así, la información falsa y errónea tienden a integrarse en el término "fenómenos desinformativos"[34–36]. En esa línea, en español, "desinformación" suele usarse como sinónimo de "información errónea", al margen de la intención del emisor[37]. Desde la perspectiva legal, por otra parte, solo algunas prácticas extremas y vinculadas principalmente a la información dañina se incluyen en el Código Penal[25] (**Cuadro 1**).

La complejidad del fenómeno desinformativo se acentúa cuando se atiende a su naturaleza, causas y efectos, aspectos que aún no se comprenden bien en todo su alcance[4,24,38]. La evidencia señala que el problema se amplifica gracias a los cambios que el desarrollo digital ha supuesto en la forma en la que la información fluye y llega a la ciudadanía[11–14,23,27,39–42], fuertemente definida por las redes sociales[5]. Sin embargo, se trata de un problema multidimensional. A la dimensión tecnológica se suma la de la confianza en la democracia y sus instituciones, la del papel del periodismo como garante democrático, la geopolítica, la económica, la social y la individual o cognitiva[4,5,33,43,44].

Existen tipos extremos de desinformación que pueden tipificarse como delitos incluidos en el Código Penal, pero la mayoría de la desinformación opera dentro del ámbito legal.

Cuadro 1. Trastornos de la información y el Código Penal.

En 2020, a la luz de las graves consecuencias de la desinformación en el contexto de la crisis sanitaria por la COVID-19 y con el objetivo de poder guiar las actuaciones legales[45], la Fiscalía General del Estado identificó varios delitos tipificados en el Código Penal[46] en los que pueden incurrir algunas formas concretas de desinformación[47]:

Delito de odio: En España, los datos señalan una tendencia alcista de este tipo de delitos en los últimos años (1.724 en 2022), siendo el racismo/xenofobia (639 casos y 18 de antigitanismo) y la orientación sexual e identidad de género (466) los más comunes. Les siguen la discriminación por sexo o género y por ideología o religión[48].

Descubrimiento y revelación de secretos: Cuando la desinformación se acompaña de la revelación de datos personales, etc.

Delito contra la integridad moral: En casos graves que afectan a una persona singular.

Desórdenes públicos: Asociado a información falsa sobre atentados terroristas, catástrofes u otros, que causen alarma, situaciones de peligro para la sociedad o requieran la activación de los servicios de emergencia.

Injurias y calumnias: Es injuria la "acción o expresión que lesiona la dignidad de otra persona, menoscabando su fama o atentando contra su propia estimación" y la calumnia es la imputación de un delito "con conocimiento de su falsedad o temerario desprecio hacia la verdad".

Delitos contra la salud pública, estafas e intrusismo: Falsas terapias curativas, falsos métodos de detección de enfermedades, etc.

Delitos contra el mercado y los consumidores: Engloba la aplicación del delito relativo al mercado y los consumidores del código penal, y diversos tipos que castigan la falsedad informativa en el contexto de los mercados y/o los consumidores.

Narrativas para la desinformación

Las narrativas son relatos amplios y flexibles que combinan información veraz y falsa en diferentes grados. Así, aumentan su verosimilitud, se adaptan a vulnerabilidades localmente y refuerzan creencias propias y prejuicios. Son exitosas cuando consiguen engañar, sembrar dudas, confusión, desconfianza o desapego en torno a un tema.

No toda la desinformación es estrictamente falsa, ya que el concepto engloba también la dimensión de engañosa, por lo que no se trata simplemente de las comúnmente denominadas fake news[34] (noticias falsas o sencillamente, falsedades)[4,11,27,34,40,41]. La distorsión de la información puede presentarse en un gradiente diverso de contenidos falsos y manipulados que permite generar narrativas desinformativas amplias y muchas veces, sutiles, lo que dificultan su detección[24,31,34,49–53]. Este gradiente va desde el contenido 100 % falso y elaborado para engañar, hasta distintos grados de alteración, como la modificación del contexto, el establecimiento de falsas conexiones o fuentes, el uso continuado de la parodia o la jajaganda[54], así como imágenes o vídeos manipulados o no relacionados[25,31,55]. Puede incluirse en torno a la propaganda, la alteración sistemática de la información, detalles o fuentes con un fin determinado, mediante, por ejemplo, la supresión o sobrexposición de contenidos, o ser utilizada por teorías de la conspiración[5,7,56,57].

Mediante este amplio conjunto de técnicas, las narrativas suelen entremezclar contenido cierto y falso o manipulado para que la verosimilitud sustituya a la verdad[6,24]. Además, evolucionan en el tiempo, incorporando elementos subjetivos que se adaptan al contexto local y a la actualidad informativa[4]. Por ejemplo, una narrativa global que persiga deslegitimar los procesos electorales interpretará cualquier suceso coetáneo como prueba de su invalidez o manipulación. El Departamento de Seguridad Nacional español destaca algunas narrativas internacionales de origen ruso, como "élites malvadas contra el pueblo", "valores tradicionales amenazados", "soberanía, identidad y valores nacionales amenazados" y señala que se apoyan en el refuerzo de los movimientos políticos y sociales de carácter extremista[54].

Por tanto, cualquier tema de relevancia social, incidente o confrontación son susceptibles de ser explotados por una narrativa desinformativa que fluya a escala internacional[43]. Su éxito no radica tanto en engañar como en sembrar dudas, confusión, desconfianza o desapego en relación a un tema, institución o proceso democrático[10,24,34,50–53], o fomentar la incapacidad de reconocer un conjunto de hechos comúnmente acordado para describir una realidad[58]. De hecho, la información falsa o engañosa rara vez confronta las creencias del receptor, más bien se alinea con sus ideas y sentimientos para activar y reforzar sus propios prejuicios o falsas creencias[59]. De esta forma, se persiguen cambios de carácter estructural o profundos, a largo plazo, en la ciudadanía, más que los resultados inmediatos de una noticia falsa concreta.

Alcance y relevancia en la era digital

Una cuestión de seguridad nacional

El desarrollo digital amplifica el fenómeno y facilita sus objetivos económicos, ideológicos o geopolíticos. Es una cuestión de seguridad nacional que alcanza niveles críticos en situaciones de gran relevancia social.

Internet ha posibilitado el surgimiento de nuevas amenazas a los Estados y la amplificación del alcance, oportunidades y medios de algunas preexistentes, así como de su impacto fuera de línea. Es el caso de la desinformación[11–15], considerada, junto a otras, como parte de las amenazas híbridas[60,61]. Constituye una amenaza sistémica con potencial para desestabilizar los Estados y los procesos democráticos[7], lo que la convierte en una cuestión de seguridad nacional.

Bajo ese marco, confluyen las operaciones de influencia en la información y la injerencia extranjera[4,62]. Las primeras engloban el uso del conjunto diverso de desórdenes informativos y manipulativos de distinta índole y la segunda incluye esfuerzos coercitivos que pueden combinar otras acciones como ciberamenazas, presión financiera o de otro tipo[4,60]. En un

- Noticas falsa o fake-news: El conocido término se ha considerado inadecuado por parte de la comunidad científica y experta principalmente por su falta de precisión (es en sí mismo un oxímoron) y por su asimilación en el discurso político para señalar informaciones o medios que contradicen el punto de vista del emisor. Si se tratara solo de información falsa, el problema podría abordarse simplemente mediante procesos de verificación.
- Jajaganda: Consiste en camuflar la desinformación y la manipulación a través del humor; ridiculizar y humillar a instituciones públicas y políticos, para así socavar la credibilidad y la confiabilidad de un objetivo determinado.
- Amenazas híbridas: Se corresponde con acciones desarrolladas por actores, ya sean estatales o no estatales, que buscan explotar las vulnerabilidades de la Unión Europea (UE) en su propio beneficio, utilizando de manera coordinada una mezcla de medidas (diplomáticas, militares, económicas y tecnológicas), mientras se mantienen por debajo del umbral de la guerra formal.

sentido más amplio de la amenaza, la OTAN acuña el término de guerras cognitivas[63]. A pesar de ello, las campañas de desinformación se encuentran, como otras amenazas estructurales, deliberadamente en una zona gris por debajo del umbral de la guerra y normalmente dentro del espectro de la legalidad[59,64]. En su conjunto, dan forma y potencian las narrativas desinformativas[59]. En consonancia, parte del desafío para frenar la desinformación no está tanto en la detección y neutralización de la información falsa, sino en poder conectar esta con las grandes narrativas y sus objetivos a largo o medio plazo[65,66].

Distintos fines para un mismo medio

No se trata solo de una amenaza exterior. Los objetivos y actores vinculados a la desinformación también se presentan en clave doméstica o nacional, siendo esta la parte que normalmente la población general mejor percibe y más preocupa[17,67]. Las principales motivaciones que existen detrás de las campañas desinformativas pueden ser[7,9,12]:

- **Económicas**: Se derivan de los intereses del instigador y abarcan desde la estafa, la competencia desleal o su uso para reforzar o proteger un sector o actividad económica concreta hasta la capitalización de la desinformación como modelo de negocio[8,9,67,68].
- **Ideológicas**: Pretenden principalmente influir en los resultados de los procesos electorales utilizando desinformación[64,67,69,70]. En España el Departamento de Seguridad Nacional[67] identifica como principales motivos ideológicos desacreditar a los gobiernos, a partidos políticos o candidatos y socavar la confianza ciudadana en la integridad del proceso electoral.
- **Geopolíticas**: Buscan desestabilizar la democracia o provocar una reacción desproporcionada que cuestione el carácter democrático del país o su prestigio internacional, así como generar división e incertidumbre en torno a temas de relevancia social, entre otros[9,13,67]. Se trata así de moldear unos nuevos valores sociales que puedan subvertir los fundamentos de las sociedades democráticas[4].

Muchas veces, estas motivaciones pueden converger y entremezclar actores externos y nacionales, tanto que, en la mayoría de los casos, no es posible identificar ni las primeras ni los segundos[4]. A pesar de ello, existe consenso en que las estrategias de desinformación se multiplican en situaciones de gran relevancia social[24,71,72]. Por ejemplo, el grupo de expertos bajo el marco del Departamento de Seguridad Nacional español destaca[4] algunos estudios sobre su papel en las elecciones presidenciales de EE. UU. en 2016[73] y las francesas en 2017[74], el Brexit[75,76], el conflicto bélico de Ucrania[54] o más recientemente, el conflicto en Oriente Próximo[77,78]. Atención especial requiere la denominada infodemia[79], surgida en torno a la COVID-19, con grandes narrativas engañosas a nivel internacional sobre su origen, prevención o la peligrosidad de las vacunas[53,80–83]. En el caso de España, algunos estudios e informes reflejan el papel de la desinformación en aspectos como la COVID[84–86], los procesos electorales[67,70,87–89], la cuestión territorial y Cataluña[59,90], entre otros[66,70,86,91].

Un nuevo contexto socio-informativo

Las redes sociales, servicios de mensajería y grandes plataformas digitales han modificado la forma en la que la información fluye y llega a la ciudadanía, gracias a una irrupción masiva de nuevos canales y posibilidades que constituyen la base de la desintermediación informativa.

La comunidad experta coincide en que el ecosistema digital, entendido como las grandes redes sociales, plataformas y motores de búsqueda, y la mensajería privada, ha sido decisivo en el acceso globalizado, democratización, aceleración y crecimiento exponencial de la información. En la actualidad, son consideradas su principal medio[11–15], especialmente las redes sociales y de mensajería[24,92], pero la evidencia señala que en ellas la velocidad y el alcance de la información falsa supera a la verdadera, amplificando así la primera[93]. Además, se han multiplicado los contenidos, canales y actores involucrados en cómo se informa la sociedad. Asimismo, se han difuminado los roles clásicos de emisor y receptor fusionándolos como "prosumidores"[12]. Estos aspectos posibilitan el carácter descentralizado, no hay una única fuente sino muchas y normalmente coordinadas, y multidireccional del flujo desinformativo *online*[10]. En este nuevo contexto, se acelera un proceso de desintermediación en el que los intérpretes tradicionales de la realidad, como medios periodísticos, actores políticos u otros, son sustituidos en el ecosistema digital; la propia tecnología que sustenta su funcionamiento y cualquier usuario, real o no, anónimo o reconocible, son prescriptores de contenidos[6,94].

· Infodemia: Concepto acuñado durante la COVID-19 para reflejar el exceso de información sobre el tema.
· Prosumidor: Consumidor de un producto o un servicio, en este caso información y contenidos diversos, que al mismo tiempo participa en la producción del mismo.

Además, el interés general por las noticias, entendido como información periodística, no ha dejado de caer en los últimos años y es muy minoritario entre los internautas, lo que debilita su capacidad para frenar la desinformación[95,96].

Emisor: instigadores y difusores

Pueden ser actores nacionales o extranjeros, o ambos, si sus intereses confluyen. El ecosistema digital dificulta su identificación.

La red amplía las posibilidades para ocultar la identidad, generar confianza y eludir los mecanismos de control destinados a limitar la desinformación[4]. Los emisores pueden ser actores estatales y no estatales, incluyendo los proxies de ambos[4]. Los primeros pueden ser gobiernos y/o sus estructuras asociadas que actúan tanto sobre su población como sobre poblaciones extranjeras. Algunos estudios señalan que alrededor de 81 países usan las redes sociales para la dispersión de propaganda y desinformación apoyándose en grupos cibernéticos o "cyber troops" bien capacitados para ese fin[97]. El volumen de negocio que representa su actividad se estima en torno a los 9 millones de euros (año 2020)[97]. Los actores no estatales incluyen corporaciones, grupos de presión (*lobbies*), agencias de marketing y grupos de interés con base ideológica, como los partidos políticos, o de otro tipo como religiosa, étnica, entre otros[4,7,70,71,98]. Los intereses de actores estatales y no estatales pueden converger. Los agentes que se encarguen de la distribución pueden estar relacionados con los instigadores o no, y pueden actuar de forma consciente o inconsciente, por lo que cualquier ciudadano, organización o institución es susceptible de "emitir" y recibir desinformación[99].

Es importante destacar que el diverso conjunto de actores señalado opera normalmente de manera coordinada, algo facilitado por el ecosistema digital, lo que resalta el carácter sistémico que tiene la amenaza[59].

Canales: impacto digital y prevalencia de canales clásicos

Redes sociales, servicios de mensajería y motores de búsqueda han pasado a ser los principales mediadores de la información y favorecen el flujo desinformativo. Las élites mediáticas o políticas siguen teniendo un papel central en la amplificación masiva de la información falsa que circula por internet.

En el ecosistema informativo actual, conviven nuevos canales *online* con el resto de las puertas a la información que ya existían previamente *offline* y sus actuales versiones digitales[96].

El espacio digital ha posibilitado la irrupción de miles de medios periodísticos[100] y de otros tipos como las redes sociales públicas y privadas de mensajería para transmitir información, con nuevas formas y formatos, como agregadores de noticias, blogs, podcasts y un largo etcétera. Las redes sociales y grandes plataformas digitales son la principal forma de acceso a la información y las noticias, intermediación que tradicionalmente ejercían medios periodísticos, pero sin adoptar sus principios profesionales[6,43,101]. Son especialmente relevantes en el caso de los más jovenes[95], donde dominan redes como Tik-Tok[96]. Para el conjunto de la población, Facebook y WhatsApp son los medios más habituales para informarse e interactuar con las noticias, siendo el segundo el preferido para compartirlas[96]. Se conforma, así, una explosión de posibilidades informativas que es, a la vez, una liberación y una saturación[6]:

- El ecosistema digital favorece una sobreinformación que abarca todo tipo de calidades informativas dentro y fuera de la práctica periodística, mezcla opinión con información e impone inmediatez. Así, potencia el acceso a todo tipo de contenidos, pero no la capacidad de comprenderlos o usarlos como conocimiento[43,102,103].
- Cualquier información puede tener apariencia periodística o de estar respaldada por grupos, instituciones o personas expertas o influyentes[24].
- Resulta barato y rápido inyectar información *online*. De hecho, las campañas de desinformación pueden sustentarse en la creación de una red de medios propios[18].
- Tiene un coste reputacional bajo para los instigadores, por la dificultad de atribución y por la falta de conciencia crítica sobre la mentira, o por una mayor aceptación social de la misma[6–8,71].

· Proxies: Puede tratarse de agencias, organizaciones etc. No se establece un vínculo directo con los instigadores públicamente, pero están vinculados, financiados y controlados por ellos de manera encubierta.
· Grupos cibernéticos: Se trata de grupos o individuos que cuentan con el respaldo económico o de otro tipo de gobiernos, partidos políticos u otras entidades con el encargo de manipular la opinión pública en línea. Su objetivo principal es difundir propaganda y desinformación con el fin de influir en la percepción de eventos, cuestiones políticas, sociales o cualquier otro tema de interés. El concepto anglosajón cyber troop hace referencia clara al concepto de guerra híbrida u operaciones de influencia.
· Sobreinformación: La sobreinformación se refiere al exceso de información disponible o recibida, tanto cierta como falsa, sobre un tema lo que puede resultar en una saturación cognitiva o dificultades para procesar y asimilar la información de manera efectiva.

Junto a estos cambios, parte de la comunidad experta establece y destaca una fuerte relación entre la denominada crisis de desintermediación informativa y el flujo de desinformación[6,24,59,104–106]. En España, la mitad de la población recibe información curada algorítmicamente en vez de por una decisión editorial o atendiendo a criterios profesionales de corte periodístico o institucional que puedan definir la calidad informativa. Además, en algunas redes sociales la fuente de noticias prioritaria no es necesariamente el periodismo profesional[95,96].

Sin embargo, a pesar de los profundos cambios señalados, las élites tradicionales, entendidas como medios periodísticos masivos, partidos políticos o instituciones relevantes, siguen teniendo un papel destacado[44,70,98,107–109]. El impacto y alcance de la información falsa aumenta exponencialmente cuando esta es recogida y diseminada por ellos. De hecho, de forma generalizada, la televisión sigue de cerca a los medios digitales como canal predilecto para la información. Se refleja así la responsabilidad e importancia que aún tienen este tipo de canales y actores[70,96,110].

Contenido

La desinformación se basa en técnicas generales como la atracción afectiva o repetición para aumentar su influencia. En España, entre los principales temas objeto de desinformación, se pueden encontrar la política o desafíos sociales como la migración.

Cualquier tema puede ser objeto de información falsa, pero las tendencias varían con el tiempo y entre países, y están condicionadas por el idioma y aspectos culturales locales[66,111]. Por esta razón, los flujos desinformativos son más comunes entre sociedades con un mismo idioma[112,113]. Para mejorar la aceptación de la información falsa se utilizan mensajes basados en la atracción afectiva, que fomentan una respuesta emocional. Pueden incluir componentes visuales, mientras que su simpleza, aparente solidez narrativa y la mera repetición constante aumentan su credibilidad[25,114–116].

El Observatorio de Medios Digitales de España y Portugal, impulsado por la Comisión Europea y vinculado al Observatorio de Medios Digitales Europeo, en línea con los resultados de algunos estudios e informes recientes, agrupa los contenidos de las noticias falsas en España alrededor de los principales temas que ocupan: políticos y electorales, salud, medioambientales, migración, género, personajes famosos, seguridad y sexualidad[66,70,88,89,117,118]. Concretamente, durante el primer tercio de 2023 predominaron los de contenido político y electoral y los vinculados al cambio climático, con narrativas sobre fraude electoral y el clima vinculado a la restricción de derechos[89]. Por su parte, la desinformación sobre salud y ciencia repuntó durante la pandemia por COVID-19 y prevalece, incluso aumenta, desde entonces[34,86,119]. Recientemente, la UE ha destinado algo más de un millón de euros para ahondar en la comprensión de la desinformación en torno a algunos de estos temas[120].

Receptor

Cualquier individuo o institución puede ser objeto de la desinformación a través de campañas desarrolladas de forma personalizada al receptor y los objetivos que persiga. Además, el empoderamiento informativo del individuo conlleva beneficios, pero también puede actuar debilitando su capacidad de hacer frente a la desinformación.

Cualquier persona puede ser receptora y difusora, accidental o no, de desinformación. Aun así, se considera que existen algunos grupos especialmente vulnerables, como personas mayores o población en situación de exclusión social[19,121]. En España, el 53 % de la población, frente al 37 % para la UE, considera estar expuesta a noticias falsas diariamente[17]. No obstante, la distribución no es aleatoria y está cargada de intencionaldiad[6,122]. Cualquier institución u organización pública o privada, así como colectivos sociales o individuos, son objeto de campañas diseñadas *ad-hoc* para alcanzar sus objetivos, apelando en muchas ocasiones a las emociones[18]. A ello se suma la avalancha e inmediatez informativa que, si bien fomentan el empoderamiento o autosuficiencia informativa y rompen el flujo unidireccional clásico (cualquiera es capaz de acceder a "toda" la información), también suscitan importantes desafíos. En el nuevo contexto de la desintermediación, existen dificultades para diferenciar información y opinión, nivel de veracidad, establecer confianza en las fuentes, etc., por lo que los instintos, emociones o sesgos personales y sociales gobiernan fácilmente la relación con la información[6]. Así, en su conjunto, puede fomentar la predisposición de las personas a creer en la desinformación.

· Curación algorítmica: Las redes sociales y las plataformas digitales y motores de búsqueda nos ofrecen contenidos que son seleccionados, o filtrados, y ordenados en función de los criterios de un determinado algoritmo destinado a personalizarlos para mantener la atención del usuario.
· Fuente de noticias prioritaria: A nivel nacional, los profesionales del periodismo y sus marcas informativas son la fuente de noticias más popular en Facebook (46 %), YouTube (44 %), Instagram (42 %) y Twitter (57 %). En el otro extremo, la sociedad en general (44 %) y medios alternativos (35%) dominan como fuentes informativas en TikTok.

Fenómenos contemporáneos implicados en el auge desinformativo

La desintermediación informativa se produce en un contexto social en el que un amplio grupo de factores fuertemente interrelacionados facilitan la presencia de la desinformación y su impacto. La complejidad y el carácter dinámico del fenómeno dificultan la clara distinción de sus causas y efectos.

La crisis de desintermediación interacciona con las múltiples dimensiones que definen la realidad institucional, política, económica, social o individual que dan forma al fenómeno desinformativo[12,16,33,123]. De esta manera, la desinformación se configura como un fenómeno multifacético, multifactorial e inmerso en una serie de dinámicas que se solapan y se refuerzan mutuamente. Algunos autores destacan la digitalización de la esfera pública a través del filtrado algorítmico, el debilitamiento de las estructuras profesionales informativas, la polarización, la autocratización, la pérdida de confianza o la creciente conspiranoia, entre otros, y todo esto ante retos cada vez más complejos[12,33]. Su compleja relación puede situar a algunos de estos factores a la vez entre las causas y los efectos de la desinformación, en función del enfoque o fuentes consultadas[7,8,12].

La confianza y el marco democrático

La desafección democrática como telón de fondo

La confianza en las instituciones democráticas, vinculada al bienestar ciudadano, refuerza la resiliencia de los estados y su sociedad ante la desinformación.

Algunos estudios refieren una crisis global y profunda que afecta tanto a la cantidad como la calidad de las democracias. Esta se relaciona con el incremento de la desafección y desconfianza hacia las instituciones democráticas y otros garantes de las mismas, como los medios de comunicación periodísticos[6,59,124–126]. Se trata de una vulnerabilidad que debilita la capacidad de las instituciones para combatir las narrativas desinformativas lo que incrementa la susceptibilidad ciudadana y sus efectos[26,127], a la vez que constituye una limitación en la repuesta[20,127]. Aunque no es posible establecer unas causas concretas e inequívocas, la comunidad experta señala que estas tienen un componente estructural, principalmente relacionado con el efecto y respuesta ante las sucesivas crisis económicas y el aumento de la desigualdad. También socioafectivo[5,6,58,59,115,124,128], que agrupa por ejemplo sentimientos colectivos o individuales de agravio personal, desencanto, desinterés general por el sistema democrático, las instituciones y la política, o la insatisfacción y la búsqueda de sentido e identidad en el contexto globalizado. Todo ello favorece la aparición de personas dispuestas a creer ya sea por razones sociales o psicológicas. Por ejemplo, durante la COVID-19, se puso de manifiesto que la confianza en las instituciones era un aspecto clave para mitigar la influencia de la infodemia y la fatiga derivada de la situación excepcional de crisis[129–131].

Geopolítica y marco internacional

Existe un uso creciente de la desinformación como herramienta para la desestabilización de terceros países en el marco geopolítico.

La confianza entre los Estados también juega un papel importante. En la última década, la comunidad experta pone de relieve una creciente toma de conciencia y utilización de los denominados poderes blandos (*soft-power*) en el contexto internacional. Se asocian con el uso de la cultura y la comunicación, incluyendo la desinformación, como una herramienta más en las operaciones de los Estados y polos de influencia geopolítica[42,132]. En la misma dirección, el avance y consolidación de las denominadas guerras híbridas, que incluyen la desinformación, configuran un marco proclive y propicio para esta amenaza[133,134]. Estas tensiones internacionales se han agudizado especialmente durante la COVID-19[135] y, más recientemente, tras la invasión rusa de Ucrania[4,62,136–138]. Entre los actores más reconocidos, la UE[139–141] y múltiples informes[7,9,71,142] apuntan a Rusia, concretamente a la Agencia de Investigación de Internet Rusa, y en menor medida a otros países, como China[71,142,143].

Mediación informativa y periodismo

La crisis económica, la precarización del sector y la falta de confianza ciudadana minan la capacidad del periodismo para luchar contra la desinformación y su papel como institución garante de la democracia.

Distintos factores han debilitado el papel de los medios periodísticos como principales mediadores de la información y salvaguarda frente a la desinfomración[6,43,49]. La comunidad experta y los datos en España señalan una crisis financiera y profesional en el sector que descapitaliza y precariza las redacciones, así como un problema de confianza[4,24,39,43,49,95,144,145]. La falta de recursos, la sobreinformación y la inmediatez, junto a la competencia por la atención, la publicidad y el posicionamiento en el ecosistema digital[43,95,146,147], dificultan el rigor informativo[148,149].

Este contexto propicia la aparición de un amplio abanico de calidades informativas marcado entre otros aspectos por: estándares de rigor variable, incluyendo los denominados pseudomedios[150], un aumento del sensacionalismo para obtener atención (*click-bait*)[151,152], el uso de las redes como fuente[39,43,150,153–155] o la aparición de informantes[156]. Todo ello, junto a la publicación por error, falta de medios o intencionadamente de desinformación[43,49] se relaciona con la crisis de confianza que sufre el sector[43,127,157,158] y una mayor dependencia económica respecto a los poderes públicos y audiencias fieles[39,159,160]. En conjunto, estos aspectos contribuyen a minar la confianza y pueden incrementar, en contra del código periodístico[161], la falta de independencia política o empresarial o la percepción de la misma[39,159,160] y la polarización de medios y audencias[95,162,163]. En el caso de España, parte del personal experto ha relacionado esta percepción con falta de trasparencia y arbitrariedad[163–165] en el uso de la publicidad institucional en los medios[166].

Fragmentación social

La fragmentación del debate público facilita la influencia de la desinformación. La falta de cultura deliberativa, social o política, pueden potenciar estos fenómenos a través de la polarización afectiva o el marco de la posverdad.

El crecimiento de fenómenos como la posverdad[12,167–169], la conspiranoia[57,170,171] y la polarización afectiva e ideológica[172], que pueden relacionarse con la cultura política y deliberativa de los Estados, incide en la fragmentación progresiva del espacio público y la capacidad de penetración de la desinformación[6–8,12,31,33,49,70,170]. Estos fenómenos han modificado la forma en que la sociedad se relaciona con la falsedad, pero también con la veracidad. Por ello, parte del personal experto señala una crisis o alteración epistemológica o, incluso, ontológica de la verdad[6,12,98,173,174]. Por un lado, favorecen el despliegue y aceptación de mensajes sesgados o falsos apelando directamente a las emociones, sentimientos y creencias. Por otro, especialmente en el caso de las teorías de la conspiración, alteran la relación de la sociedad con la evidencia objetiva y el propio conocimiento científico, pudiendo de facto rechazarlos. En conjunto, obstaculizan el pensamiento racional y facilitan la aceptación de narrativas desinformativas[175,176], lo que dificulta y puede llegar a bloquear el debate social y amplificar la desconfianza en las instituciones[4,33,49]. España es considerado uno de los países más polarizados a nivel afectivo y social, no tanto político, de la Unión Europea[24]: incluso temas sin carga ideológica son afectados hoy en día por este prisma[12,177]. No obstante, no todos los contenidos emocionales o polarizantes hacen uso de la desinformación.

La polarización afectiva ha aumentado de forma prácticamente generalizada. La evidencia sobre las causas es aún escasa, especialmente fuera de EE. UU., con resultados no necesariamente convergentes entre países o estudios[12,101,178,179]. Entre estas causas, destacan la cultura política, marcada por la polarización de las élites, la proliferación de redes sociales y medios de comunicación hiperpartidistas, o la falta de cultura deliberativa[12,24,101,178–180], junto con los factores de orden estructural vinculados a la desafección antes señalados. Esta situación converge con aspectos socioafectivos derivados del sentimiento de insatisfacción social o desafección democrática, que favorecen la difusión de desinformación[33,115]: para avivar la indignación generalizada[181,182], incitar al caos o apelar al denominado deseo de "ver el mundo arder"[183].

Cognición y vulnerabilidad individual

La psicología juega un papel importante en por qué se comparte o se cree en la desinformación, pero existen otras causas variables que pueden cambiar entre temáticas informativas.

Las personas juegan un papel central en la amplificación de la desinformación, pues son sus principales receptores y difusores[15,93]. Por ello, entender los factores y los atajos mentales o emociones que modulan la decisión sobre qué es verdadero y falso y cómo se forman creencias al respecto[115,184], es un aspecto de gran importancia[115,185,186]. No obstante, el conjunto de factores que delimitan la susceptibilidad, aceptación y potencial diseminador de la información falsa es muy diverso y no es posible establecer un patrón generalizado[15,86,115,187]. Su influencia puede variar en función de las personas, del contexto personal o social, de un país a otro o, incluso, entre temáticas diferentes[86,111,115,123].

- Pseudomedios: Conjunto de medios digitales surgidos con el fin de explotar el modelo publicitario basado en la economía de la atención del ecosistema digital, sin atender a criterios periodísticos de calidad, éticos o de otro tipo.
- Posverdad. Se trata de una percepción de la realidad que está ligada o denota circunstancias en las que los hechos objetivos son menos influyentes en la formación de la opinión pública que las apelaciones a la emoción y la creencia personal, lo que dificulta la percepción de la veracidad. La posverdad no es solo mentira. Es una distorsión de la verdad cargada, sobre todo, de intencionalidad.
- Teorías de la conspiración: Promueven ideas simplistas y autojustificadas independientes de los hechos reales o más probables en torno a cuestiones complejas. Afirman que ciertos eventos o situaciones son manipulados en secreto por poderosas fuerzas con intenciones negativas. Suelen tener un conjunto de elementos en común bien identificados.
- Polarización afectiva: Se refiere a la distancia emocional entre el afecto que despiertan quienes simpatizan con nuestras mismas ideas políticas en contraposición con el rechazo hacia quienes tienen opiniones distintas.
- Polarización ideológica: Se refiere al grado en el que la población tiene creencias divergentes que son consistentemente conservadoras o progresistas, en relación a un amplio conjunto de temas de relevancia social.

Factores demográficos

No existe un consenso claro sobre la influencia de estos factores.

Los resultados basados en factores demográficos son contradictorios[15,122]. El aumento de la vulnerabilidad con la edad es un factor bien confirmado en EE. UU. pero en Europa los datos no son concluyentes[5,188]. En el caso de los menores, preocupa especialmente su exposición a través de plataformas basadas en contenidos audiovisuales, aunque falta información al respecto para entender el alcance del problema[189]. En la misma línea, la relación general del nivel educativo, económico o el género con la tendencia a creer o compartir desinformación varía en función del estudio[15].

Predisposición cognitiva para creer y compartir

Existen sesgos cognitivos que predisponen a los individuos a creer en la información falsa y a diseminarla.

Múltiples factores cognitivos se relacionan con la tendencia a creer información falsa[15,115,190]:

- El nivel de pensamiento crítico[115,191]. Se puede vincular con la teoría del proceso dual, confiar en la intuición en vez de deliberar[115,192], y con los sesgos propios del razonamiento motivado que, como el de confirmación, favorecen el refuerzo de ideas y sesgos propios.
- El efecto de la verdad ilusoria[115,186]. Esencialmente predispone a creer en información que resulta familiar[193], aunque sea por repetición o consistencia con recuerdos previos[194], y en mensajes sencillos, favorecidos por la avaricia cognitiva y la heurística, aunque sean falsos.
- Fallos cognitivos. Están relacionados principalmente con la forma en que se perciben las fuentes y el conocimiento o predisposición que se tiene respecto a un tema y respecto a la propia desinformación[115,195]. Incluyen la falta de atención o precisión sobre la información que se recibe[185] o la excesiva confianza en una fuente[196,197].

En cuanto a por qué se comparte la información falsa, destaca, cuando se hace de forma inintencionada, la buena voluntad de difundir algo importante[198], así como el impulso derivado del simple hábito de compartir información en redes[199], o de la falta de interés o atención[185,200]. Cuando se comparte a conciencia, la evidencia señala principalmente la búsqueda de interés personal, como señalar la pertenencia a un grupo[201] o alcanzar notoriedad[202], o como mecanismo para cubrir otras necesidades psicológicas vinculadas al malestar social[181–183].

Contexto socioafectivo

La desinformación puede explotar aspectos derivados del encaje de los individuos en la sociedad y el estado afectivo y emocional derivado del mismo.

Los aspectos afectivos amplifican el poder de persuasión de la información[203]. Sirven para generar falsas creencias y difundirlas al explotar un componente emocional o moral insertado en la información[181,182,204]. Es algo típico del sensacionalismo, que se aprovecha del propio estado emocional de las personas[115,205] o intenta inducir uno explotando emociones como el miedo o la inseguridad. Algunos estudios asocian rasgos de la personalidad[15,185], hábitos y creencias[187,206] o la ideología[200,207] con la vulnerabilidad y tendencia a interactuar con la desinformación. Además, destacan la fuerte influencia de medios periodísticos masivos, *influencers*, élites políticas y comunidad experta sobre el alcance de la desinformación y la percepción pública[115,208].

Perdurabilidad y efecto continuado

Existe amplia evidencia científica sobre la resistencia de las personas a corregir las creencias erróneas o aceptar la falsedad de la información que les es afín.

Las personas muestran una gran resistencia a aceptar la corrección de la desinformación y modificar así falsas creencias. Estas pueden perdurar mucho tiempo, con independencia de la habilidad cognitiva del individuo o de haber sido desmentidas. Es lo que se conoce como efecto de influencia continuada[184,209]. De forma más extrema, la corrección continuada puede provocar un efecto rebote que refuerce el aferramiento a falsas creencias, en vez de corregirlas (fenómeno denominado *backfire* en inglés)[210,211]. Si bien parece un efecto adverso que se ha sobredimensionado, la resistencia a la corrección y el efecto rebote son especialmente

· Teoría del proceso dual: Enfrenta el pensamiento analítico, basado en el razonamiento lógico, y el intuitivo, basado en las emociones y sentimientos, aunque evidencias recientes cuestionan su relacionan con la desinformación y destacan su complejidad. La avaricia cognitiva y la heurística favorecen los mensajes sencillos cuya comprensión no requiere un alto procesamiento.
· Razonamiento motivado: Validamos lo que está en línea con nuestra ideología y visión del mundo. El sesgo de confirmación o la disonancia cognitiva refuerzan esta tendencia.
· Avaricia cognitiva: hace referencia a un tipo de sesgo que inclina la percepción hacia la información que confirma las propias creencias, reduciendo así el esfuerzo mental de procesarla

acusados cuando la refutación entronca con creencias o valores que conforman la visión identitaria de uno mismo[115].

Gobernanza digital y modelo de negocio

El modelo de negocio digital basado en la atención responde a intereses económicos que pueden favorecer la falta de privacidad y la exposición y circulación de desinformación, sin atender a los efectos negativos que estas tienen sobre la ciudadanía.

Los aspectos geopolíticos, la confianza institucional y los sesgos sociales y psicológicos que operan fuera de internet se pueden reflejar en ella. La comunidad experta pone de relieve que no se trata de una red universal homogénea, anónima y neutral. Si bien las tecnologías que la sustentan sí son universales, los estándares con los que opera varían entre países o plataformas[212,213].

Existe una parte de la economía digital que se relaciona con la desinformación. Concretamente, la que se basa en un modelo de negocio estructural de múltiples plataformas digitales que monetiza la atención de los usuarios mostrándoles publicidad[5]. En esta relación, el poder de intervención de los últimos es además muy limitado[5]. El valor de un contenido está en su capacidad de captar la atención[13,151]; y el de la atención, reside en su utilización para modificar el comportamiento para, por ejemplo, comprar un producto determinado[214]. Así, la venta de espacios publicitarios online suele ser la principal fuente de ingresos de las grandes redes sociales y muchos sitios web[13]. Este modelo favorece la distribución de contenidos desinformativos capaces de captar la atención por encima de los veraces[13,101,215,216], mediante atracción afectiva y mensajes característicos que aumentan su aceptación[151], como el sensacionalismo[151,217,218]. Esto resultan altamente rentable, si se monetiza con publicidad[7,11,27,68,214].

Proliferan así las empresas que básicamente generan contenidos falsos para rentabilizarlos de esta manera[8,219]. Algunos datos señalan que en torno a 200 millones de dólares americanos en publicidad termina en dominios señalados por desinformación a nivel global (datos de 2019)[220], 76 millones si nos centramos en los datos referidos a la UE (datos de 2020)[221]. Además, las estimaciones globales sobre la actividad propagandística y desinformativa de los grupos cibernéticos asociados a Estados asciende a otros 10 millones (datos de 2020)[97].

Los datos de los usuarios son otro elemento de valor en la monetización: son la base de una nueva economía[222]. Estos se recopilan y analizan con el objetivo de poder dirigir de forma personalizada aquellos contenidos que más captan la atención y la publicidad que puede ser más efectiva para cada usuario[7]. Los datos personales y el comportamiento de los usuarios pueden revelarse de forma deliberada y consentida, como ocurre en muchas redes sociales o derivarse de las tecnologías que monitorizan nuestro comportamiento. También pueden ser inferidos de la información e interacciones con otros usuarios, "*amigos*" bajo el marco de muchas redes sociales o, incluso, *offline* si se considera el internet de las cosas[13,223,224]. Asimismo, pueden directamente ser comprados o vendidos con un fin[225]. Las informaciones sensibles como las referentes a la ideología política o religión, protegidas bajo el Reglamento General de Protección de Datos (RGPD), son fácilmente deducibles del análisis de estos conjuntos de datos sin necesidad de vulnerar la RGPD[13,226–228]. Por todo ello, se considera que las redes sociales son un riesgo para la protección de los datos y la privacidad de la ciudadanía[5].

La ciudadanía financia con sus datos servicios "gratuitos", como las propias redes sociales, u otras tipologías freemium, sin que exista una percepción clara de sus inconvenientes y posibles externalidades[214]: desinformación y manipulación, daños psicológicos y adicciones, pérdida de privacidad, etc.[5,214,229]. Las arquitecturas propias de redes sociales y plataformas digitales lo favorecen: suelen impulsar configuraciones de privacidad laxas para aumentar la participación del usuario y también pueden basarse en configuraciones, elegidas o no, que limiten la percepción del individuo respecto a la información disponible, el clima de opinión u otros múltiples aspectos[5].

- Monitorización online del comportamiento: El comportamiento general en línea es rastreado continuamente a través de tecnologías como cookies, scripts e imágenes de seguimiento, anuncios publicitarios, código CSS/HTML e incluso terceras partes que intervienen en la plataformas y sistema que usamos mediante interfaces de programación de aplicaciones (API) u otras técnicas.
- Freemium: Término que se utiliza comúnmente en el ámbito de los servicios digitales y aplicaciones para describir un modelo de negocio en el que se ofrece una versión básica de un producto de forma gratuita, pero se proporcionan características adicionales o premium por un costo adicional. Es muy común en múltiples aplicaciones móviles, servicios digitales etc.
- Arquitectura: En el contexto de sitios web y redes sociales se refiere a cómo se diseñan y presentan las opciones y configuraciones disponibles para los usuarios. Estas arquitecturas son estrategias de diseño que influyen en las decisiones que los usuarios toman al interactuar en línea.

Este modelo de negocio explota los datos personales para que las tecnologías que lo gobiernan, algoritmos y técnicas publicitarias personalizadas, muestren a los usuarios el mundo que, según sus cálculos, deberían querer ver[6]. Lo hace, como señalan algunos expertos, atendiendo principalmente al beneficio privado, pudiendo así limitar la rendición de cuentas públicas y las soluciones que se pueden lograr sin intervención regulatoria[5,230]. De igual manera, su maquinaria puede ponerse a disposición de terceros con fines distintos a los puramente comerciales, pudiendo consistir en propaganda ideológica, aspecto especialmente crítico bajo el prisma electoral[8,231].

Tecnologías impulsoras de la desinformación

Se basan en el uso de un conjunto de tecnologías asociadas al análisis masivo de datos y a la inteligencia artificial[13]. Están en constante evolución, contribuyen al papel de internet como mediador de la información en redes sociales y grandes plataformas de búsqueda o navegadores, y sustentan la economía de la atención. Por ello, tienen un papel importante en torno a la desinformación[232], contribuyendo a su generación y difusión. Entre otras[13], pueden destacarse las siguientes:

Algoritmos: cámaras de resonancia y burbujas de filtro

El curado algorítmico de la información cuestiona la neutralidad y pluralidad de la información que reciben los usuarios. Puede favorecer la dispersión de desinformación en la búsqueda de la atención de los usuarios y generar espacios propicios para la misma.

El orden en el que aparece la información que se muestra a cada usuario en redes sociales y en los buscadores viene adaptado por algoritmos de curación y recomendación. Estos algoritmos pueden poner en cuestión el acceso a la información y la deliberación de forma neutral y plural en internet[233,234], y esta mediación, preocupa a más de la mitad de la población española[96]. Son propios de cada red social y buscador y están en constante desarrollo, lo que señala que la premisa de que las redes sociales son representativas de la opinión pública puede considerarse falsa[6]. De este modo, internet está fuertemente controlado por algoritmos corporativos privados diseñados para maximizar ganancias capturando la atención, sin atender necesariamente a posibles efectos psicológicos o sociales[230]. De hecho, la falta de transparencia habitual de estos algoritmos es un factor que limita la lucha contra la desinformación por parte de usuarios o personal experto al dificultar[169,178,235–238]: dificulta la identificación de sesgos, la comprensión de su influencia en el comportamiento social e individual, o el desarrollo de mecanismos de detección y prevención.

Los algoritmos basados en la atención pueden favorecer la exposición a la desinformación por ofrecer contenidos de naturaleza impactante, de corte sensacionalista o tendentes a la radicalización y el extremismo[13,239–241]. Este modo de oferta de la información se ha relacionado con las "burbujas de filtro" y "cámaras de resonancia", conceptos distintos pero que en común apuntan a la falta de exposición a opiniones diferentes a la propia, a la creación de silos de verdades autorreferenciales, que incrementan la polarización y bloquean el debate público[6,13,179,233,241–244]. No obstante, las redes sociales y plataformas dan también voz a comunidades marginadas y desfavorecidas[5]. Además, existe un debate científico activo sobre estos fenómenos y sus efectos[178,241,245,246].

Por un lado, la evidencia emergente señala que las cámaras de resonancia son menos comunes de lo que se viene asumiendo[178,247,248] y la proporción de la población que llega a las mismas es minoritaria[92,207,233] y altamente polarizada con anterioridad[178,247]. No obstante, pueden tener un efecto significativo sobre el debate público[178]. Por otro, aunque el efecto del filtrado algorítmico sobre la calidad de la dieta informativa depende mucho de la dieta con la que se compare, la evidencia empírica más reciente señala que el filtrado algorítmico no necesariamente limita la dieta informativa y su calidad[178] y cuestiona algunos de sus efectos negativos, como la polarización[245,246,249] y la radicalización[239,240,246,250] entre otros. Incluso, hay estudios que apuntan a que podría reducir la exposición a desinformación[245]. Aunque estos

· Burbujas de filtro: Fenómeno por el que una persona se encuentra expuesta principalmente a información y perspectivas que refuerzan sus creencias y puntos de vista preexistentes, debido a algoritmos de recomendación y personalización en plataformas digitales y redes sociales. Ocurre sin acción voluntaria por parte del usuario.
· Cámaras de resonancia: Se trata de un entorno o plataforma donde las ideas, mensajes o conceptos encuentran una audiencia receptiva y amplifican su impacto. Requieren de un papel activo del usuario que demanda activamente determinados contenidos por diferentes razones, incluidas las ideológicas, y donde puede circular más libremente la desinformación.

resultados refuerzan la idea de otros factores distintos del tecnológico como motor de la polarización[179,246], también pueden reflejar cambios recientes en los algoritmos de diversas plataformas destinados al aumento de la moderación[239]. Sin embargo, la falta de transparencia algorítmica dificulta enormemente la confirmación de estos estudios[169,178,235–238].

Perfiles engañosos: bots, ciborgs, troles y grupos falsos

Las cuentas y grupos engañosos tienen relevancia en la difusión de información falsa en el ecosistema digital. También, pueden causar daño reputacional a terceros, disparar la polarización afectiva o conseguir colarse en la agenda informativa colocando desinformación en ella o al ser usados como ejemplo de lo que ocurre en la sociedad.

En las plataformas sociales abundan perfiles y grupos falsos cuya función principal es la dispersión, casi siempre automatizada y a gran escala, de desinformación[13]. Estos perfiles pueden ser sistemas totalmente automatizados basados en inteligencia artificial (bots), cuentas semiautomatizadas con intervención humana (ciborgs) o usuarios enteramente humanos (trol)[13,41]. Pueden contar con el respaldo económico del actor que orquesta las campañas de desinformación[41] y muchas veces conforman grupos falsos o híbridos (bots y humanos interactuando)[8].

A pesar de los esfuerzos de la plataformas y redes sociales por limitar su presencia, el uso malicioso de bots y perfiles falsos para difundir desinformación se extiende y explota económicamente bajo un modelo de servicio por demanda[251]. Su número y actividad aumenta exponencialmente antes de los procesos electorales[13,252,253]. No obstante, aunque existe cierto disenso[31,93,254,255], las cuentas verificadas (personas reconocibles por así decirlo) parecen tener un papel más relevante en la diseminación que las artificiales y falsas[31,93,255]. Por otra parte, el avance de las técnicas de inteligencia artificial hace cada vez más difícil distinguir los bots de los usuarios humanos[13,31,256,257].

Además de como diseminadores, las cuentas falsas también actúan inflando la popularidad o silenciando a usuarios y contenidos[8], lo que influye sobre el filtrado algorítmico[258–262], colando desinformación en la agenda informativa o política[261,263] y promoviendo corrientes de opinión artificiales aparentemente espontáneas (*astroturfing*)[258,264,265]. Estas cuentas también favorecen la crispación y polarización afectiva mediante mensajes de falsa bandera que enfatizan de forma extrema y en ocasiones ridiculizan una posición ideológica, social o científica para generar una repuesta y un clima de opinión reaccionario[266,267]. En un país del tamaño de España, el control sobre unos pocos cientos de cuentas[8] es suficiente para tener un impacto significativo en una red social como X (antes conocida como Twitter)[13].

Análisis de datos y microsegmentación.

La microsegmentación permite dirigir a nivel individuo o en pequeños grupos la publicidad, pero también la propaganda ideológica, lo que puede socavar el debate público. Además, facilita la influencia sobre los procesos electorales en base a información falsa.

La microsegmentación es una práctica publicitaria común y útil que permite dirigir contenidos a grupos específicos de usuarios basándose en sus características, sentimientos, etc., y personalizando los mensajes que reciben[13]. Algunos trabajos recientes señalan su potencial para alcanzar un nuevo nivel de utlrapersonalización (nanosegmentación) en redes como Facebook, con campañas dirigidas a nivel de individuo[268]. El uso con fines desinformativos de la microsegmentación puede afectar negativamente a la población ya que permite el control y la automatización de la recopilación de datos y la selección de canales y contenidos para impactar en los usuarios[13,184,269–271]. Puede apoyarse en múltiples herramientas, algunas ofrecidas a nivel comercial incluso por las propias plataformas digitales, entre las que destacan la prospección dinámica[13,272,273], la publicidad programática o la psicografía[35]. En Europa, la microsegmentación está limitada por el RGPD[274], aunque esto es considerado insuficiente por algunos expertos[228,275], y los datos señalan un rechazo público sobre esta práctica[5].

- Astroturfing: Es una estrategia de manipulación en línea que implica la creación o promoción de una falsa impresión de apoyo u oposición pública a una causa, idea, producto, individuo o política. Individuos o grupos intentan hacer que sus mensajes o acciones parezcan orgánicos y espontáneos cuando, en realidad, están siendo orquestados o financiados de manera oculta.
- Prospección dinámica: Es una técnica publicitaria que implica la adaptación automatizada en tiempo real de los anuncios y mensajes publicitarios en función de la información y el comportamiento, atencional y emocional, del usuario. Permite generar miles de versiones de un anuncio con el fin de atraer de forma customizada la atención. Se basa en la recopilación de datos sobre el usuario, como sus preferencias, historial de navegación o ubicación geográfica, para personalizar el contenido publicitario que se le muestra de manera individualizada.
- Publicidad programática: Es una forma automatizada de comprar y vender espacio publicitario en línea, que utiliza algoritmos y tecnología para facilitar y optimizar el proceso y los objetivos publicitarios.
- Psicografía: Se refiere al estudio y la clasificación de las características, actitudes, valores, intereses y comportamientos de un grupo de personas o de una audiencia específica.

Un aspecto que preocupa especialmente es la utilización de esta tecnología para la difusión de contenidos ideológicos o políticos[228,276,277,277,278], sobre todo, en la contienda electoral. Puede favorecer la polarización emocional y dificultar la comprensión de las posturas generales de los partidos o invisibilizar o reforzar de forma personalizada sólo determinadas posturas. De esta forma puede socavar el debate público y democrático y potencialmente, manipular el voto[5,13,225,226,275,277]. Algunos expertos sugieren la necesidad de regulaciones electorales adaptadas a estas prácticas, así como más investigación empírica en el contexto europeo para comprender quién las utiliza y sus efectos, con el fin de limitar su posible influencia en el derecho de libertad de voto[7,225,228,275,277,279]. La UE aborda en la actualidad la regulación de la segmentación de la publicidad política[280].

Encriptación y mensajería privada

La mensajería privada es uno de los principales vehículos de información falsa, pero la encriptación, necesaria desde la perspectiva de la privacidad, dificulta su identificación y mitigación.

Las plataformas de mensajería privada son el canal preferente para compartir información y noticias[96]. Incluyen servicios internos de mensajería de redes sociales y aplicaciones desarrolladas exclusivamente para tal fin, como WhatsApp, Telegram, Signal, Skype, etc. Entre las distintas causas que las hacen propicias a la desinformación, destaca la mayor confianza con los interlocutores y la opacidad que procura el encriptado en la mayor parte de ellas[13]. Por un lado, esta última característica es clave para la privacidad y mejorar la ciberseguridad de las comunicaciones, en consonancia con las políticas europeas y nacionales, como destacan muchos expertos (nota Ciberseguridad[2]). Por otro, la encriptación impide identificar la información falsa y los actores asociados, etc., incluso, por parte de las compañías que proveen los servicios, lo que además limita su responsabilidad al respecto[13]. La comunidad experta señala la importancia de afrontar el reto que supone mejorar la comprensión sobre cómo fluye la desinformación en estos espacios[281–283].

La inteligencia artificial generativa

La inteligencia artificial es capaz de generar textos, imágenes o videos falsos indistinguibles muchas veces de contenidos veraces, pudiendo incluso afectar a algunos derechos fundamentales. Se considera un frente que puede requerir avances regulatorios.

Las distintas técnicas que engloba la inteligencia artificial generativa destacan por su capacidad para producir contenidos falsos y manipulados de cada vez mayor calidad en imágenes, audios, vídeos y textos. Aunque estas tecnologías ofrecen grandes oportunidades e impactos positivos[284], también suponen importantes riesgos como la posibilidad de vulnerar distintos derechos fundamentales[285–287]. Existe un amplio gradiente de calidad en la manipulación, desde niveles obvios como el de los memes y la jajaganda, que también pueden desinformar[13], hasta los modelos masivos del lenguaje, las ultrafalsificaciones y clonadores de voz (**Cuadro 2**). Mediante su uso se pueden producir contenidos falsos prácticamente indistinguible de los verdaderos, tanto por la inteligencia humana como por la artificial. Constituyen la muerte del *ver para creer*[6]. Por ello, representan un importante desafío en la desinformación[288].

Se trata de un frente de la inteligencia artificial que puede requerir regulaciones adicionales en el corto-medio plazo. A pesar de los esfuerzos, las tecnologías de generación superan a las de detección tanto en el caso del texto[294,310–312] como en las ultrafalsificaciones[288,299]. Estas técnicas evolucionan rápidamente, dificultando la detección temprana y la formulación de políticas públicas preventivas[285]. Para paliar este problema, se han desarrollado materiales con el fin de facilitar su detección por parte de la ciudadanía[313] y para guiar los esfuerzos regulatorios[286,292,314–317] más allá del marco genérico de la inteligencia artificial[318,319], tanto a nivel global[320,321] como en la UE[285,319].

Los modelos de lenguaje masivos como los usados por chat GPT o las ultrafalsificaciones son algunos de los avances tecnológicos con mayor potencial para llevar la desinformación a un nuevo nivel no conocido hasta ahora.

Cuadro 2. La evolución tecnológica como herramienta para generar contenido falso.

Los modelos de leguaje masivos, entre los que destacan modelos como el conocido GPT (en inglés, *generative pre-trained transformer*)[289], permite analizar y generar texto: resumir contenidos, traducirlos, responder a preguntas complejas etc. La comunidad experta señala que el texto generado por máquinas puede llevar la desinformación a niveles no conocidos hasta ahora[290–293], especialmente en aspectos como la salud o la cuestión climática[292–295]:

- Permiten disfrazar contenido falso como auténtico imitando estilos y formatos de fuentes confiables y, así, evadir su detección[31,288,296].
- Amplifican al integrarse en webs falsas, bots, o influenciando el filtrado algorítmico mediante citación cruzada de contenidos generados artificialmente[13,288].
- La falta de control exhaustivo sobre las fuentes que se usan para que los modelos aprendan y puedan responder facilita la generación de contenidos erróneos, imprecisos, falsos o sesgados[288,297].

El término ultrafalsificación (*deepfake*)[298] se define como medios audiovisuales manipulados o sintéticos producidos mediante técnicas de inteligencia artificial, que parecen auténticos y que muestran a personas que parecen decir o hacer algo que nunca han dicho o hecho[285]. Las posibilidades para sembrar dudas sobre cualquier cosa que se ve o escucha son casi infinitas y su accesibilidad general es cada vez mayor[288,298–302], poniendo en cuestión el conocido *ver para creer*[234]. Además de permitir una desinformación a la carta más eficaz, en la actualidad existen múltiples ejemplos de usos dañinos de esta tecnología[303,304]. Entre ellos destaca su uso para generar contenido pornográfico no consentido[305], incluyendo el caso de menores[306]. Se trata de un aspecto que afecta especialmente a las mujeres y que implica, además de la difamación otras formas de violencia como la intimidación o la extorsión[285,307].

El área de investigación sobre generación de contenidos, y también sobre su detección, más activa es la manipulación de caras en imágenes y vídeos[13,288,299,308].

Otros espacios digitales que pueden ser relevantes son el uso de la realidad virtual y aumentada, que incluye el denominado metaverso[13,309]. Destacan, asimismo, los asistentes o dispositivo inteligentes, las tecnologías autónomas distribuidas (Ver nota Ciberseguridad[2]), los juegos multimedia y las denominadas narrativas transmedia[13].

Impacto

Son necesarios avances que permitan una mejor comprensión del impacto de la desinformación, especialmente a nivel social. Por ahora, la evidencia disponible la relaciona con importantes efectos sobre los individuos que pueden desencadenar desórdenes democráticos.

Existe un amplio consenso entre el personal experto sobre el impacto potencial y el peligro que representa la desinformación, lo que justifica el gran número de estudios al respecto. Sin embargo, fortalecer el nivel de evidencia empírica que permita establecer causalidad[7,9,24,29,39,322] constituye aun un importante desafío[24,101,115,185,323].

La mayor parte de la evidencia causal se centra en efectos a corto-medio plazo sobre el individuo, incluyendo efectos sobre su comportamiento en línea y su respuesta emocional[93,324–329], así como consecuencias sobre la salud en el caso de la infodemia[15,84,86,324]. En el plano individual de la toma de decisiones y a nivel social, existen estudios que atribuyen efectos negativos a la desinformación y, sobre todo, al uso de las redes sociales[322], sin poder establecer siempre causalidad. Estos incluyen el desarrollo de conductas peligrosas, actitudes hostiles e incluso crímenes de odio[5,203,330], daño en la confianza en las instituciones[44,208,331,332], un aumento de la polarización[331,333,334] y alteración del voto mediante el uso de información falsa o engañosa[331,334,335].

· Narrativa transmedia: Se trata de contar historias cuyo universo narrativo se desarrolla a través de múltiples medios y plataformas de comunicación como redes sociales, sitios web, videojuegos, podcast, un programa de televisión, un libro etc. Cada medio o plataforma aporta una parte única de la historia y se utiliza para enriquecer la experiencia global del espectador o lector.
· Causalidad: Persigue poder establecer una variable o factor como causa que explica el comportamiento de otra que se puede denominar efecto. En investigación el método referencia para establecer este tipo de relación son los experimentos controlados aleatorizados, muy comunes en el ámbito clínico. Implica la asignación aleatoria de los participantes a un grupo control (no sometido a intervención) y otro experimental, lo que dificulta su aplicación a la dimensión social de la desinformación.

Diversos factores dificultan el avance del conocimiento sobre el impacto. Por un lado, es metodológicamente complejo establecer la relación entre las actitudes y comportamientos que determinan la exposición a la desinformación y su posterior traducción en un cambio o refuerzo de actitudes o toma de decisiones a nivel individual. Aún más desafiante es establecer el siguiente salto, su traducción en efectos a nivel social[24,101,115,175,185,324,336]. Por otra parte, las consecuencias de la desinformación son acumulativas en el tiempo por lo que, además, pueden verse alterados por factores como cambios algorítmicos o de otro tipo. Sin embargo, no existen datos longitudinales ni transparencia suficiente para abordar esta cuestión. Por ello, es difícil evaluar los efectos a largo plazo. Esto junto a otros factores, se relaciona con la fragmentación existente sobre el conocimiento del impacto, que impide una visión sistémica, debido a diversos aspectos que dificultan su propia delimitación conceptual[24]:

- El alcance del impacto y sus causas son multinivel lo que abarca cada eslabón de la cadena, desde el individuo a lo social[324].
- Se pueden diferenciar áreas de impacto distintas y muchas veces difíciles de conectar: psicológica, financiera y social[285].
- Existe una baja representatividad y comparabilidad entre los estudios derivada de los múltiples canales y factores individuales[324–326,337,338], geográficos o lingüísticos que influyen[31,257,257,327].
- Los medios periodísticos, como prensa, televisión o radio, y en general medios *offline* están escasamente representados en los estudios de impacto[24].
- Las lagunas del conocimiento sobre si las redes sociales son un motor o una amenaza para los sistemas democráticos se relacionan con las dificultades y desacuerdos en la comprensión de los impactos de la desinformación[5,71,94,101,322].

El hecho de que, hasta la fecha, la evidencia sólida de la influencia causal que la desinformación y otras dinámicas tienen en las actitudes y comportamientos *offline* sea muy limitada no significa que no exista o que no requiera una respuesta, como de hecho señala la evidencia existente[101,245,339]. Para avanzar, es importante reforzar la evidencia observacional y relacional en el ecosistema desinformativo y poder validar las relaciones detectadas con experimentos en entornos controlados que permitan atribuir causalidad[327,340,341]. La comunidad experta destaca la necesidad de un abordaje multidisciplinar a gran escala[101], lo que incluye el desarrollo y uso de indicadores armonizados y comparables[24]. Para ello, conviene reforzar la colaboración, transparencia y el acceso a los datos de redes sociales y plataformas digitales así como de medios periodísticos, lo que puede requerir políticas públicas e incentivos[24,101,340–343].

Combatir la desinformación: agentes y estrategias mitigantes

El diseño de estrategias para luchar contra la desinformación requiere de la combinación de múltiples instrumentos que puede inspirarse en los principios democráticos de igualdad, representación y participación.

La comunidad experta señala la importancia de coordinar múltiples intervenciones en las estrategias de respuesta a la desinformación[5]. Además, la respuesta debe combinar respuestas a corto plazo, orientadas a los efectos inmediatos de la desinformación, y estructurales o a largo plazo, que persigan elevar la resiliencia de la ciudadanía y los sistemas democráticos y sus instituciones[10,11]. Unas no deben ser sustitutivas de otras, ya que mayormente actúan de forma complementaria, aunque también existen sinergias negativas[99]. En conjunto, las medidas intervienen simultáneamente en los niveles de prevención, mitigación o a nivel sistémico. Además, engloban un variado conjunto de actores públicos y privados que pueden aplicarlas, donde destaca el papel de moderación de las plataformas digitales. Este conjunto ha de contemplar también a la propia ciudadanía. La UNESCO[99] agrupa las estrategias de acuerdo al papel de estos actores, pero existen propuestas que los clasifican en función del efecto que persiguen, como el Centro Común de Investigación de la UE, que utiliza tres principios democráticos[5].

- **Igualdad**: Agrupa estrategias orientadas a reducir asimetrías derivadas de la acumulación o falta de conocimiento, información o datos, poder o responsabilización de algunos actores frente al resto[5,6]. Se basan principalmente en fortalecer instrumentos e instituciones garantes, a través de la identificación y neutralización, del fomento de la resiliencia social e individual y de reforzar el frente regulatorio.

- **Representación**: Agrupa medidas centradas en proteger los procesos electorales y los datos, privacidad y libre elección de los ciudadanos.
- **Participación**: Persiguen asentar el papel de la ciudadanía, la investigación y los actores públicos y privados en base a una mirada hacia el futuro.

Garantías, detección y neutralización

Instituciones garantes: responsabilidad como primera línea preventiva

Reforzar la confianza en las instituciones democráticas y fortalecer la mediación periodística potenciando sus capacidades, medios, independencia, transparencia y pluralidad, constituyen medidas de fondo necesarias para mitigar los efectos de la desinformación.

Las democracias y sus instituciones pueden responder y responsabilizarse de la desafección ciudadana y los mecanismos que, como la desinformación, pueden minarlas y aumentar la vulnerabilidad de la población[6,7,33,344,345]. Se enfrentan al desafío estructural de generar un nuevo diálogo con los ciudadanos, adaptado al contexto propio de la desintermediación digital, que gestione la incertidumbre asociada y que genere nuevos vínculos de confianza. Por un lado, la comunicación institucional y el desarrollo de políticas públicas diseñadas desde la comprensión de los mecanismos desinformativos pueden ser beneficiosas para que las instituciones mitiguen sus efectos[7,59,345–347]. En este sentido, la comunicación estratégica juega un papel muy importante para poder generar narrativas que neutralicen aquellas propias de la desinformación[347,348], lo que requiere de vigilancia y anticipación[349]. Por otro lado, los actores políticos o mediáticos pueden también combatir sus efectos, evitando explotar las vulnerabilidades asociadas a la fragmentación social, la polarización o la erosión de la confianza democrática[33,44,70,98,125,127,175].

Además, existe un amplio consenso en la comunidad experta y las entidades europeas sobre el papel central que el periodismo y los profesionales de la información pueden ejercer como freno estructural a la desinformación[11,29,99,350]. Sin embargo, son escasos los proyectos y estudios científicos sobre los medios de comunicación como barrera contra la desinformación[24,148,351] más allá de las agencias de verificación[99,352–354]. De hecho, aunque existe abundante información sobre el comportamiento y la dieta informativa en redes sociales y plataformas digitales, la información *offline* o respecto a medios clásicos, o en ámbitos privados como la mensajería instantánea, es muy limitada.

Aunque no existe un consenso claro entre las múltiples propuestas para reforzar el sector periodístico[39,352,355–357], la discusión se centra en fortalecer sus capacidades y medios, independencia, transparencia y la pluralidad, así como delimitar la responsabilidad respecto a la desinformación y el papel de la publicidad y la relación con las redes sociales[33,99,352–354]. También se proponen nuevas formas de comunicar que favorezcan la conexión con la ciudadanía y refuercen una visión constructiva y proactiva ante las dificultades globales[358–361]. Todo ello sin alterar la libertad de expresión. Para acercarse a estos objetivos estratégicos, se trata de avanzar con el desarrollo de instituciones y marcos regulatorios, como el Observatorio Digital de Medios (EDMO)[362,363] o la Ley Europea de Libertad de Medios Digitales[354]. La propuesta de esta ley persigue proteger el pluralismo y la independencia con un conjunto de normas que abordan desde la financiación estable de medios y la transparencia, hasta mecanismos de protección contra injerencias políticas o editoriales y desequilibrios derivados de la concentración de medios[364]. La comunidad experta señala su relevancia en el largo-medio plazo, aunque algunos aspectos han despertado ciertos disensos entre los actores relacionados[365].

Vigilancia y verificación

Las agencias de verificación juegan un papel importante y positivo en la lucha contra la desinformación. Este papel puede ser extensible a otros actores y no está exento de desafíos.

Los hechos también son vulnerables, ya que la verdad no es lo mismo que la objetividad y la exactitud[6]. Las agencias de verificación o *fact-checkers* evalúan la exactitud de la información para detectar y corregir aquella que es falsa. La manera de corregirla se basa, principalmente, en exponer los hechos en torno a las inexactitudes[356,366–369] y proporcionar información precisa cuando hay evidencia[370–372] (conocido como *debunking*), por un lado, y proporcionar contexto o la información disponible, cuando el contenido no es verificable, por otro. La neutralización de la información falsa puede también consistir en evaluar y exponer la plausibilidad de las fuentes y su credibilidad[115]. Además, las agencias amplifican su alcance mediante la colaboración con redes sociales y medios de comunicación y ahondan en la

· Comunicación estratégica: Es en un enfoque especializado para distribuir y recibir información. Consiste en comunicar el mensaje más adecuado, a través de los canales correctos, a las personas adecuadas, en el momento preciso y utilizando la retroalimentación de este proceso para mantenerse enfocado en los objetivos que se persiguen.

alfabetización mediática, la prevención anticipada (*prebunking*) contra la desinformación e incluso la promoción de políticas públicas[368,373–375].

La evidencia científica señala que el desmentido de la información falsa tiene efectos positivos a nivel social[5], es efectiva para combatir la desinformación y, en la mayoría de los casos, es preferible a la inacción[184,376–383]. No obstante, no es un mecanismo infalible y existen consideraciones que pueden afectar a su efectividad entre las que destacan las derivadas de:

- **Aceptación y alcance de la refutación**: Presenta dos importantes desafíos referidos al *qué* y *cómo*[379]. Dada la sobreinformación y el constante incremento de las ultrafalsificaciones, no es posible verificar toda la información. Así, un aspecto crítico es el uso de criterios metodológicos claros y transparentes que expliquen qué contenidos se priorizan para ser verificados[19,377,379,384]. La comunidad experta destaca que estos criterios incluyan la fiabilidad de fuentes, viralidad y potencialidad dañina[379]. Por otro lado, la corrección debe atender a estándares de calidad que potencien su aceptación y alcance y minimicen el rechazo a la misma[184,210,379,385–388].
- **Creencias personales y el escepticismo sobre las agencias[184,376–378]:** Los estándares y políticas que garantizan la confianza, como la neutralidad y transparencia metodológica, financiera, política, etc., están definidos internacionalmente y se certifican[389–391]. Además, la actividad de las agencias de verificación puede verse favorecida y amplificada mediante mecanismos orientados a la inclusión y escucha social[392], la participación y el refuerzo de la confianza de los usuarios[377,393–395], así como con el uso de inteligencia artificial[396–399].

La vigilancia y la refutación atienden también a la dimensión estratégica en la lucha contra la desinformación. Por un lado, no son un tipo de intervención exclusiva de verificadores; otros agentes del periodismo, comunidad experta y científica, instituciones, etc. pueden participar de las mismas para, de forma conjunta, prevenir a la ciudadanía ante operaciones hostiles de influencia en la información[11]. Por otro, la vigilancia constituye la base que posibilita cierto nivel de anticipación necesario para la comunicación estratégica u otro tipo de medidas estructurales asociadas al refuerzo de la resiliencia ciudadana[11].

Automatización: la inteligencia artificial como aliada

La inteligencia artificial tiene el potencial de amplificar la capacidad de detección de la desinformación y sus emisores, conectarla con narrativas generales y amplificar el alcance de refutación.

Aunque la inteligencia artificial puede favorecer la desinformación (**Cuadro 2**), existen estudios que evalúan el uso de técnicas de aprendizaje automático para combatirla[31,288]. La evidencia científica demuestra su potencial para detectar y caracterizar la desinformación, así como a los emisores en redes sociales, en función de características propias del mensaje o de su contexto[31,288,400,401]. Esto incluye además del texto, el análisis de imágenes o vídeos que lo acompañan, la forma en que se comparten o las emociones que despiertan entre otros aspectos[31,257,400]. Además de desenmascarar y caracterizar actualmente bots, troles o la propia desinformación, pueden ser clave para facilitar la labor de los verificadores y detectar interconexiones entre información y grandes narrativas a nivel internacional[396–399]. También puede ser de ayuda para amplificar el alcance de la información corregida[257].

En general, aunque existen aplicaciones desarrolladas, estas técnicas aún requieren de avances para su aplicación generalizada por parte de ciudadanos, administraciones o verificadores[288,399,402,403]. Los desafíos son, entre otros, la fragmentación de los datos disponibles, los posibles sesgos, la falta de transparencia y explicabilidad algorítmica, el posible impacto sobre la privacidad y cuestiones éticas[257,404].

Resiliencia y capacitación social

Puesto que la batalla contra la información falsa se da en la mente de las personas es necesario dotar a la ciudadanía de competencias y mecanismos que refuercen su capacidad de identificarla y reduzcan su influencia.

La mitad de la población española no confía en su capacidad para identificar la información falsa[17]. Existe un consenso generalizado en la comunidad experta sobre el papel central que ocupa la sensibilización y capacitación ciudadana para reducir el impacto de la desinformación[4,21,405,406,406]. Por un lado, la educación es la base para generar pensamiento crítico[19]. Por otro, la novedad y la sutileza del fenómeno de la desinformación impiden una percepción adecuada del riesgo y, con ello, una sensibilización que facilite las intervenciones y políticas públicas al respecto[6].

· Explicabilidad algorítmica: Se refiere a la importancia de poder entender y explicar las decisiones apoyadas por la inteligencia artificial que tienen un impacto en la vida de las personas.

Existe una amplia gama de intervenciones que pueden agruparse en función de cuándo se aplican:

- Previas a la exposición:
  - A largo plazo o de carácter estructural como la **alfabetización mediática y digital**, las **intervenciones tecnocognitivas** o la búsqueda de un **marco ético**
  - Dirigidas al desarrollo de competencias concretas o tipos de desinformación específica denominadas comúnmente por su término anglosajón ***prebunking***.
- Tras la exposición: Persiguen modificar la creencia falsa mediante la refutación y se agrupan bajo el concepto de ***debunking***.

Alfabetización mediática y digital

Una mejor comprensión de las formas y múltiples dimensiones que conforman la desinformación puede preparar a las generaciones venideras y presentes contra la misma.

Puede combatir la desinformación a nivel estructural al incorporar competencias y habilidades técnicas, cognitivas, sociales, cívicas, éticas y creativas en la ciudadanía para desenvolverse de una manera más crítica en los medios actuales, a la hora de producir contenido, comunicar y de entender la información que se recibe[19,407–411]. Fomenta una percepción crítica de las distintas dimensiones que conforman el desafío digital e informativo. Aunque se han propuesto un gran un número de intervenciones, la tarea es compleja[19,407,408,410] y pocos estudios evalúan su eficacia y aplicabilidad en diferentes contextos y grupos demográficos[408,410,412–416]. Parte de la evidencia reciente señala que aprender a reconocer los indicios concretos de la desinformación[271,410] o destacar recurrentemente la importancia de la atención y la exactitud ante la información[417,418] son estrategias efectivas. Además, deben evitarse actuaciones que generen un aumento del escepticismo generalizado ante todo tipo de información[419,420].

Desde una perspectiva sistémica, los expertos en España abogan por la educación mediática en todos los niveles educativos pero también orientada a docentes, profesionales de la comunicación y grupos vulnerables[19,121]. Sin embargo, la inclusión actual de estas competencias en el currículo formativo puede ser insuficiente[19,421–425]. El profesorado no siempre cuenta con la información o los recursos para abordarlo y, respecto al alumnado, por ejemplo, solo un tercio de los estudiantes de secundaria distingue correctamente opinión de información[425]. Existen múltiples propuestas y guías orientadas al sistema educativo reglado[19,406,409,410,413,426–428], así como campañas, guías y enfoques de aprendizaje informal[19,413,429–434]. Consolidar un enfoque sistémico y coordinado, con objetivos a largo plazo, y centrado en los contextos de redes y temas que tratan los más jóvenes, puede reforzar la resiliencia individual y la colectiva. Se trata, en definitiva de asentar un marco social propicio para el análisis y discusión de este tipo de desafíos. Actualmente, la Ley General 13/2022 de Comunicación Audiovisual[435] refleja la alfabetización mediática de forma algo marginal desde el punto de vista de algunos expertos que destacan que existe margen para ampliar acciones, planes y políticas públicas a este respecto[70].

Respuesta ética y norma social

Consiste en fomentar un marco ético que oriente el comportamiento de las personas, instituciones y sus agentes, enfocado al rechazo social de la desinformación.

Persiguen moldear un marco social compartido que, en base a la comprensión del riesgo que supone la amenaza, apele a un conjunto de normas y comportamientos éticos al respecto que se ubique por encima de las respuestas reactivas y en el largo plazo[12,33,99]. La alfabetización mediática e informativa pueden contribuir en esta dirección. También, el fomento, desde las instituciones y actores asociados, de acciones que fortalezcan el debate democrático evitando su fragmentación. Estas intervenciones están alineadas con las normas internacionales sobre derechos humanos y códigos éticos sobre el comportamiento ante la desinformación y en las redes sociales[99]. No solo se trata de identificar la información falsa o engañosa. Una actitud ciudadana y social que la denuncie y rechace puede elevar la mitigación a un plano colectivo[33]. A nivel del individuo, existe evidencia sobre cómo el daño reputacional puede actuar como freno a la hora de compartir desinformación y amplificar la atención y precisión sobre los contenidos[115].

Intervenciones tecnocognitivas y mecanismos de confianza

El flujo desinformativo puede frenarse con el rediseño de las arquitecturas propias de redes sociales y plataformas digitales, así como con el desarrollo y aplicación de mecanismos que permitan valorar la veracidad de la información.

En base al conocimiento sobre economía conductual, comunicación y ciencias de la computación existe evidencia sobre cómo rediseñar la forma en que se utilizan las plataformas o sus arquitecturas para generar un impulso (conocido por el anglicismo *nudging)* que reduzca la tendencia a compartir o interactuar con la información falsa[5,173,379,436,437]. La implementación de mecanismos que dificultan el reenvío automático y favorecen la reflexión son un ejemplo[379,438]. Los mecanismos de confianza son herramientas y métodos que ayudan a los usuarios a entender mejor la información que ven y sus fuentes, permitiéndoles ajustar su nivel de confianza[99]. Pueden incluir el etiquetado, como el caso de *"reenviado muchas veces"* de algunas redes de mensajería privadas, el acceso a indicadores de confiabilidad sobre un contenido, su calidad, la de su fuente etc. o información para su contextualización. Otra opción es el uso de aplicaciones, webs y plataformas para rastrear el origen de la información y versiones de la misma, lo que puede incluir el uso de inteligencia artificial[99,438–440].

Este tipo de intervenciones son difícilmente asumibles por las plataformas digitales si no existe presión pública o gubernamental[379].

Refutación e intervenciones conductuales

Refutación anticipada (*prebunking*)

Incluyen mecanismos para preparar a la ciudadanía de manera anticipada frente a la llegada de la información falsa y para poder refutarla cuando se entra en contacto con la misma.

Las intervenciones más simples van desde presentar información factualmente correcta con fines preventivos hasta las advertencias genéricas sobre desinformación antes de que esta se difunda[115]. Las más sofisticadas y eficaces, de acuerdo con la evidencia científica[376,441–446], pueden basarse en[379]:

- Un tema concreto sobre el que se muestran y explican contenidos desinformativos a la vez que se refutan, antes de estar expuesto a una campaña real[376]. Es lo que la comunidad experta denomina comúnmente como inoculación.
- Mecanismos de razonamiento lógico aplicables a cualquier tema para comprender las técnicas engañosas y aumentar la resistencia ante las mismas. Se basan en cinco técnicas principales a identificar: falsos expertos, falacias lógicas[447], expectativas imposibles, selección sesgada de evidencias y teorías conspiranoicas[448].

Así se fortalece la capacidad de identificar la información falsa o engañosa durante campañas reales. La inoculación requiere un profundo conocimiento estratégico sobre la desinformación específica a abordar; cómo se presentará, cuándo, etc.[184,376,449]. Su escalado generalizado a la población y multiplicidad de canales, temas etc. es compleja y requiere más investigación[184]. Sin embargo, las prácticas basadas en razonamiento lógico son más versátiles y existen avances prometedores, sobre todo mediante el uso de juegos, que son transferibles a distintos contextos, como el educativo o las redes sociales[376,449–454].

Refutar en diferido: resistencia psicológica a la corrección

Aunque la refutación de la información falsa reduce el nivel de engaño y concepciones erróneas, su efectividad a nivel social y ante todo tipo de desinformación es limitado[379]. No obstante, la comunidad experta destaca su utilidad frente a la inacción[115,379].

La corrección de información falsa debe atender a criterios concretos para maximizar su impacto positivo e influencia a largo plazo[115,385,455] y reducir la resistencia a la refutación y el efecto rebote. La comunidad experta[379,456,457] propone prácticas basadas en la repetición, la empatía, el uso de explicaciones alternativas, de fuentes de confianza generalizada y la aplicación selectiva en el momento adecuado.

· Falacias lógicas: Argumentos que parecen válidos, pero no los son. Incluyen el argumento ad hominem, ad ignorantiam, ad populum, la falacia de autoridad, la del hombre de paja, la anecdótica entre otras o técnicas como la elusión de la carga de prueba, la pendiente resbaladiza y su contrario o la falsa dicotomía.
· Expectativas imposibles: Consiste en el uso de expectativas u objetivos inalcanzables para desacreditar o neutralizar una información. Por ejemplo, "Las pruebas de PCR para el coronavirus no son 100% precisas, por lo que no deberíamos molestarnos en administrarlas".

Avances regulatorios

La regulación en torno a la desinformación es compleja, ya que puede colisionar con algunos derechos fundamentales, carece de un mandato político homogéneo y existen dificultades para dirimir la responsabilidad de redes sociales y plataformas digitales o su autoría y atribución a los instigadores.

Cualquiera de las intervenciones propuestas, como las conductuales, corren el riesgo de ser instrumentalizada por los propios canales explotados por la desinformación. Las plataformas digitales u otros actores pueden trasladar la responsabilidad de la detección y propagación de la desinformación y otras externalidades de la economía de la atención, a los usuarios u otros agentes en vez de ser asumirlas como propias[5]. La intervención regulatoria puede reforzar las intervenciones descritas y la responsabildiad[5,115].

Desafíos de la acción regulatoria

Buena parte de las acciones gubernamentales e intergubernamentales se han orientado en el eje de la seguridad nacional e institucionalización de la misión de detectar, denunciar y actuar contra las campañas de desinformación extranjeras[458]. Así lo indica la OTAN[459], la acción exterior y de contrainteligencia[349] y el Centro de Excelencia para la Lucha Contra las Amenazas Híbridas, en el caso de la UE [460], y acciones a nivel estatal[18,22]. Existe otro eje orientado al aumento de la resiliencia social y la responsabilización de los canales respecto a la desinformación[458], sobre el que se han acometido también estrategias y regulaciones[5,169] en clave nacional y a nivel europeo.

Sin embargo, se trata de una cuestión altamente compleja, ya que a las dificultades conceptuales y del conocimiento sobre la materia se unen otros desafíos transversales como:

Colisión de derechos fundamentales: Deben prevalecer la libertad de expresión[29,99,353,461,462] e información[463] y la protección de la propia democracia y sus valores[4,5,7]. Conviene por tanto, evitar la criminalización o clasificación de la información bajo el pretexto de falsedad, lo que podría socavar la democracia y dar poder discrecional a los estados[461].

Neutralidad, transparencia y mandato político: Los Estados carecen de la potestad de decidir qué es información verdadera o falsa y no siempre son neutrales[4,461]. La transparencia y los mecanismos de control en el desarrollo normativo[4], como la colaboración directa con la sociedad civil y el sector privado[4,346,362,363], así como la cooperación a nivel internacional, pueden ampliar la legitimación de las acciones[4]. Estas no son incompatibles con acciones gubernamentales[4,464]. Cada una presenta sus propios riesgos y ventajas, por lo que puede ser conveniente la utilización de ambas aproximaciones[4]. Además, el mandato político no es homogéneo, por ejemplo, a nivel de la UE[142]. Algunos países o instituciones pueden no estar igualmente comprometidos en abordar este problema, lo que puede ser otro desafío[142].

Responsabilidad: Por un lado, la comunidad experta destaca el avance en los medios técnicos y humanos para posibilitar la atribución a los instigadores. Por otro, apunta la necesidad de discutir la responsabilidad de las plataformas digitales en relación con los efectos de su actividad[7,33,169]. Delimitar qué es un medio de comunicación afecta a sus responsabilidades sociales, éticas y legales. Algunos autores señalan que la denominación de "tecnologías" de las plataformas permite eludir su responsabilidad transnacional como mediador en la información e, incluso, en torno a la moderación electoral[465]. Por ejemplo, Google indica que en la primera mitad de 2023 mostró 20.441 anuncios políticos en la UE, lo que reportó 4.5 millones de euros de beneficio, mientras que rechazó 141.823 anuncios políticos por no pasar los procesos de verificación de identidad[466]. En esta línea, distintos avances normativos recientes a escala europea aluden directamente al papel moderador de las plataformas digitales.

Por todo ello, la comunidad experta señala la importancia de evitar la regulación orientada a los contenidos y centrarse en mecanismos que contrarresten el fenómeno ya que, en buena medida puede ser considerado como una consecuencia de la propia libertad[4,37,467].

El marco europeo

La regulación europea se enfoca hacia un marco co-regulatorio y destaca la importancia de detectar y contrarrestar la información falsa, la coordinación y acción conjunta, la movilización del sector privado y la resiliencia social.

El marco europeo no ha dejado de avanzar en los últimos años (**Cuadro 3**). Aun así, la efectividad de los múltiples instrumentos que integran el modelo europeo ha sido cuestionada[37,468–470], lo que puede relacionarse con su evolución desde propuestas centradas en la autoregulación de la grandes plataformas digitales y redes sociales, como el Código de Buenas Prácticas, a intervenciones co-regulatorias como la reciente Ley de Servicios Digitales (**Cuadro 3**).

Cuadro 3. Hitos regulatorios y herramientas en el marco europeo.

(2015) **El Servicio Europeo de Acción Exterior** (SEAE) estableció la división de comunicación estratégica (**Stratcom**), con cuatro grupos de trabajo en la actualidad para monitorizar y desacreditar la desinformación internacional que afecta a la Unión Europea, coordinar la respuesta de la UE y colaborar con socios internacionales[349].

(2018) Se publica el enfoque europeo para combatir la desinformación[28] y el informe del **grupo de expertos** internacional que asesora a la UE[11]. Desembocan en el **Plan de Acción Europeo Contra la Desinformación**[29] y el **Código de Buenas Prácticas** voluntario para el sector privado (plataformas digitales, redes sociales y sector publicitario)[471].

(2019) **Sistema de Alerta Rápida** (RAS, por sus siglas en inglés)[473] para el intercambio de información y la coordinación de una respuesta temprana.

(2020) Múltiples intervenciones:

- **La comunicación** «La lucha contra la desinformación acerca de la COVID-19: contrastando los datos»[474], que examina las actuaciones ante la infodemia.
- **El Plan de Acción para la Democracia Europea**[20], que aborda pautas para revisar y mejorar el código de buenas prácticas y para fortalecer la acción exterior europea y la atribución en la materia[32].
- Tras una serie de **evaluaciones** que señalan un impacto insuficiente y una baja aplicación del Código de Bunas Prácticas en 2020[469] y de la estrategia en su conjunto[468] en 2021, en el año 2022 se refuerza el código[68,475].
- Se crean el **Observatorio Europeo de Medios Digitales** (EDMO)[362] que reúne a verificadores. investigadores académicos, plataformas digitales y redes sociales, medios periodísticos y profesionales de la alfabetización mediática. Existen observatorios específicos de carácter regional que en el caso español y portugués es denominado IBERIFIER[363]. Su misión conjunta es mejorar el conocimiento sobre la desinformación en Europa y consolidar avances que permitan implantar políticas públicas efectivas.

La Ley de Servicios Digitales es la norma más reciente sobre desinformación. Trata de depurar la responsabilidad de las grandes plataformas digitales en torno a la desinformación entre otros aspectos.

(2023) La **Ley de Servicios Digitales** persigue ser la puerta a la definición de la responsabilidad de las plataformas digitales y consolidar legalmente buena parte del código de buenas prácticas[472,476]. **Ley de Mercados Digitales** se propone garantizar un sector digital competitivo y justo, permitiendo que las empresas digitales innovadoras crezcan y asegurando la seguridad de los usuarios en línea[477].

La injerencia extranjera sobre los procesos electorales europeos es de gran relevancia para la UE, lo que se ha traducido en recomendaciones por parte de la Comisión a los Estados miembros[478]. En el año 2020, se crea en el Parlamento europeo de la comisión especial de injerencias extranjeras (ING 1)[479] recientemente renovada (ING 2)[480].

Otras medidas son la propuesta de **Reglamento sobre la Transparencia y la Segmentación de la Publicidad política**[280] y **Ley Europea de Libertad de los Medios de Comunicación**[344]. Esta última pretende respaldar y proteger el pluralismo y la independencia de los medios de comunicación en la UE[354]. La propuesta del Reglamento denominado como **Ley de Inteligencia Artificial** señala también aspectos relacionados, como la obligatoriedad de identificar ultrafalsificaciones creadas con inteligencia artifical[319].

La acción europea se vertebra en torno a cuatro líneas principales: aumentar las capacidades de detectar, analizar y exponer la desinformación, crear mecanismos para la coordinación y acción conjunta, movilizar el sector privado y, por último, la concienciación y resiliencia social[405]. Algunas de las medidas que vertebran las propuestas son:

- Desmonetizar la desinformación[68,469,471].
- Regular la propaganda política y técnicas como la microsegmentación[4,7,9,280].
- Transparencia algorítmica y prácticas tolerables[319,472].
- Identificación de las ultrafalsificaciones (*deepfake*) como tales[285,319].
- Desarrollo de herramientas, normas y sistemas de incentivos para el acceso y fortalecimiento del avance científico y los procesos de verificación en el espacio digital[24,68].
- Alfabetización mediática y participación ciudadana (mecanismos de denuncia de contenidos y su etiquetado, diseño más seguro de la arquitectura de los servicios, etc.).

En Europa, existen a nivel estatal ejemplos[4,7] de marcos co-regulatorios, así como regulaciones clásicas que responsabilizan, incluyendo sanciones económicas, a las redes sociales del contenido falso y su eliminación[7]. Es el caso de Francia (parcialmente revisada por el Conseil Constitutionnel)[481] o Alemania[482], entre otros. Sin embargo, parte de la comunidad experta señala que este último tipo de aproximaciones pueden tener efectos negativos sobre la libertad de expresión o conllevar graves sobrecargas judiciales[7,9].

España: una apuesta por la cooperación

Como en el caso de otras amenazas híbridas o la ciberseguridad, España avanza en la consolidación de un marco de cooperación que aúne la visión de los sectores público, privado y sociedad civil para encauzar el desarrollo normativo y las intervenciones frente a la desinformación.

La comunidad experta destaca que el marco regulatorio español y las políticas públicas en vigor parten del marco europeo[70] (**Cuadro 3**) y así conviene que siga siendo[4]. En línea también con el consenso experto, España trata de consolidar la cooperación público-privada y de la sociedad civil en la lucha contra la desinformación[18,22,54] (**Cuadro 4**).

Actualmente, la transposición de la Ley de Servicios Digitales requiere de la identificación de un organismo garante que persiga y vigile su cumplimento. Los mecanismos de atribución a medios periodísticos respecto a los contenidos (por ejemplo, el Artículos 30 del Código Penal)[46] pueden inspirar el principio de responsabilidad subsidiaria y escalonada de las plataformas digitales como mediadores informativos.

Cuadro 4. El recorrido español en la lucha contra la desinformación.

- (2019) La estrategia nacional de ciberseguridad reconoce el peligro de la desinformación[483], materia que queda asignada a los órganos y organismos que forman parte del Sistema de Seguridad Nacional (DSN, Departamento de Seguridad Nacional). El Centro Nacional de Inteligencia (CNI) realiza el seguimiento de los agentes vinculados a las campañas de desinformación y, a nivel interior, también intervienen las Fuerzas y Cuerpos de Seguridad del Estado dentro de sus competencias[484]. El Centro Criptológico Nacional (CCN), dependiente del CNI, crea una unidad sobre desinformación.
- (2020) El DSN:
  - Publica el Procedimiento de actuación contra la desinformación[464], que cuenta con el respaldo de la Comisión europea.
  - Crea el grupo de expertos de la sociedad civil para que, junto a representantes de la Administración pública, lleven a cabo un esfuerzo conjunto de análisis sobre la amenaza y las posibles estrategias para combatirla desde el plano social, informativo, tecnológico y regulatorio.
- (2021) La Estrategia de Seguridad Nacional incluye como principal actualización los riesgos derivados de las campañas de desinformación y los retos para su gestión[22], si bien las amenazas concretas se vienen recogiendo en los informes anuales de seguridad nacional desde la COVID-19[54,484,485].
- (2022) Se crea el Foro Nacional contra las campañas de desinformación donde con fines consultivos cooperan sociedad civil, sector privado y administración en torno a nueve grupos de trabajo que abordan todas las dimensiones del problema[346].
- (2023) Tanto el grupo de expertos como el Foro y las Administraciones involucradas coinciden en la necesidad de una estrategia nacional contra la desinformación y trabajan en su desarrollo[18].

El sector privado, un aliado necesario

Las plataformas digitales han emprendido múltiples medidas para mitigar la desinformación, pero se han considerado insuficientes. El marco regulatorio trata de delimitar su responsabilidad en el tema, a la vez que debe contemplarlas como aliadas necesarias a través de formas de co-regulación.

Las grandes plataformas digitales y las redes sociales han incorporado progresivamente diversos tipos de mecanismos e iniciativas destinadas a su moderación[19,24,59,67,99]. Además, bajo el marco del Código de Buenas Prácticas en materia de desinformación (**Cuadro 3**), estas publican periódicamente[465] los avances en su implantación[466]. Existe evidencia que señala la importancia y utilidad de incorporar los mecanismos descritos de verificación, corrección, confianza, etc. en el ecosistema digital aunque algún estudio también señala aspectos negativos al respecto[115,486].

A pesar de las múltiples iniciativas, la comunidad experta señala la necesidad de una mayor transparencia y acceso a los datos por parte de la comunidad investigadora para impulsar el conocimiento sobre desinformación y poder fortalecer las políticas públicas[5,24,101,342]. Se trata de un actor llamado a jugar un papel fundamental en la lucha contra esta amenaza, cuya cooperación y colaboración son necesarias[18,67,475].

Privacidad, seguridad y regulación electoral

Existen múltiples iniciativas en el marco internacional que apuntan hacia la necesidad de afinar la regulación respecto a la contienda electoral *online* colocando en el centro la privacidad e independencia de voto de la ciudadanía.

La publicidad política durante los períodos electorales está fuertemente regulada en la Unión Europea, tanto en los sectores de radiodifusión como de prensa. Sin embargo, las redes sociales, en su mayoría, no están cubiertas por estas medidas[5]. La comunidad experta vincula la protección de los procesos electorales, asimilándolos como una infraestructura crítica, con la implementación de avances en ciberseguridad, la regulación de las campañas *online* y la protección de los datos y la privacidad de la ciudadanía para evitar posibles abusos derivados de la propaganda customizada u otras amenazas[2,5,7,10,228].

En materia electoral, existen recomendaciones de actuación por parte de la UE (**Cuadro 3**) que España ha seguido, como la creación de la Red de Coordinación para la Seguridad en Procesos Electorales[487]. El grupo de expertos formado en torno al Departamento de Seguridad Nacional ha hecho propuestas concretas y de carácter estratégico para reforzar la resiliencia democrática ante la desinformación en largo plazo[67]. Concretamente, la comunidad experta destaca la necesidad de emprender la reforma de la Ley Orgánica 5/1985 del Régimen Electoral General para evitar las injerencias y garantizar un marco en el que se minimice el impacto de la desinformación[10,67]. Países como Canadá[488,489], EE. UU.[490] o Nueva Zelanda[491] y también del entorno[492], como Francia[493], Irlanda[494] o Reino Unido[495] han incluido reformas de sus reglamentos electorales o trabajan en ellas con el objetivo, entre otros, de mejorar la transparencia sobre el contenido y distribución de la propaganda política en línea y fuera de ella, o reforzar la comunicación institucional, sobre lo que también existe una propuesta europea[280]. Estas iniciativas pueden guiar la actuación de España en ese ámbito[67].

Una mirada estratégica y participativa hacia el futuro

El desarrollo del campo de conocimiento multidisciplinar de la desinformación, así como la consolidación de una estrategia nacional que integre sus múltiples dimensiones de forma estratégica son aspectos claves para frenar el avance de esta amenaza.

Para abordar la complejidad de la desinformación, la comunidad experta destaca la importancia de un enfoque sistémico[59,405] que pueda cristalizar en una Estrategia Nacional que identifique debilidades, principios y objetivos para combatirla[18,22]. Este enfoque debe unir la acción exterior de los Estados con el fortalecimiento de la seguridad y la resiliencia social, conectando a los múltiples actores involucrados y colocando en el centro la defensa de los valores democráticos[5]. Requiere de la responsabilización individual y el fortalecimiento de la confianza entre los actores, ya sean medios de comunicación, ciudadanos, instituciones públicas o agentes políticos[5,18,33].

Estos objetivos requieren abordar el desarrollo del área de conocimiento de la desinformación y mapear a sus actores, un ámbito necesariamente multidisciplinar[18]. Es importante avanzar en el desarrollo de estudios a gran escala y la evaluación combinada de múltiples intervenciones de mitigación para diseñar estrategias efectivas, así como el uso de la prospección para la planificación y visión a largo plazo[5,24,101,115]. La profesionalización en torno a la materia puede también favorecer su integración y representatividad en el marco de las instituciones y políticas públicas.

De igual manera, la cooperación ha de trascender al ámbito privado y llegar al conjunto de la sociedad civil, en línea con el Foro Nacional (**Cuadro 4**), así como al ámbito internacional[18,70]. Dada la sensibilidad del fenómeno, la transparencia en las acciones de todos los actores, públicos y privados, es clave, así como la rendición de cuentas con la ciudadanía[5]. Además, teniendo en cuenta su rápida y constante evolución conviene atender a herramientas adaptables, dinámicas y actualizadas[70]. Esta puede beneficiarse del desarrollo de nuevas plataformas para la deliberación pública y la alfabetización mediática; no solo se trata de aumentar la resiliencia contra la información falsa, también de fortalecer un marco social que genere un sustrato propicio para enraizar las estrategias destinadas a la defensa democrática, como es el caso de la lucha contra la desinformación[5].

Ideas fuerza

- La era digital permite una amplificación sin precedentes de la desinformación y otros desórdenes informativos que supone un importante riesgo para las democracias.
- El tratamiento de la desinformación es un reto porque debe proteger a la ciudadanía y ampliar sus derechos sin restringir otros como la libertad de expresión o el derecho a la información veraz.
- El éxito de una campaña desinformativa no necesariamente recae en generar falsas creencias, sino en generar confusión, desconfianza, dividir y amplificar sesgos y prejuicios. Para ello, los instigadores suelen explotar componentes afectivos y sustituir verosimilitud por veracidad. Persigue cambios estructurales en la percepción ciudadanía más que efectos concretos a corto plazo.
- La desinformación en la era digital se ve favorecida por un entorno en el que la intermediación clásica y el flujo informativo se difuminan: cualquiera puede generar contenido, difundirlo y compartirlo. Esto genera una explosión informativa de diferentes calidades que dificulta la identificación de contenidos veraces generando incertidumbre.
- La desinformación se explica en un contexto sociopolítico en el que la crisis de confianza democrática, la situación geopolítica, los factores sociales y psicológicos y el propio modelo de negocio digital, apoyado en tecnologías poco transparentes y en constante evolución, tienen un papel muy relevante.
- Aunque existe consenso en torno a los riesgos y la necesidad de poner en marcha mecanismos para combatirlo, la complejidad del fenómeno dificulta un análisis integral del impacto de la desinformación.
- Se apela a la responsabilidad y cooperación de todos los agentes (políticos, informativos, comerciales) para no explotar la incertidumbre y la desinformación.
- Las instituciones democráticas y sus garantes deben fomentar un diálogo con la ciudadanía que refuerce la confianza y se adecúe al nuevo contexto informativo.
- Las medidas para combatir la desinformación tienen como objetivo final la resiliencia y la alfabetización mediática y digital del conjunto de la sociedad.
- El marco europeo promueve medidas orientadas a defender y reforzar la democracia ante la desinformación y consolidar mecanismos que la combatan de forma sistémica, desde la atribución de responsabilidades o su desmonetización hasta la extensión de la pluralidad y libertad de medios y la moderación de la contienda electoral online.
- Los nuevos desarrollos en inteligencia artificial pueden suponer un antes y un después para la desinformación. Aunque intensifican el alcance y peligro que supone también ofrecen nuevas oportunidades para detectarla y combatirla.

Referencias:

1. Nai Fovino, I. *et al.* Cybersecurity, our digital anchor. EUR 30276 EN, Publications Office of the European Union. Luxemburgo. (2020) www.doi.org/10.2760/352218.
2. Oficina de Ciencia y Tecnología del Congreso de los Diputados (Oficina C). Informe C: Ciberseguridad. (2022) www.doi.org/10.57952/c8hy-6c31.
3. Marciel, R. Democracia, desinformación y conocimiento político: algunas aclaraciones conceptuales. *Dilemata* 65–82 (2022).
4. Grupo de expertos de la sociedad civil - Departamento de Seguridad Nacional. *Lucha contra las campañas de desinformación en el ámbito de la seguridad nacional. Propuesta de la sociedad civil.* Ministerio de la Presidencia, Relaciones con las Cortes y Memoria Democrática https://www.dsn.gob.es/sites/dsn/files/LibroDesinfoSN.pdf (2022).
5. Lewandowsky, S. *et al.* Technology and Democracy: Understanding the influence of online technologies on political behaviour and decision-making. *JRC Publications Repository* https://publications.jrc.ec.europa.eu/repository/handle/JRC122023 [30/05/2023] www.doi.org/10.2760/709177.
6. Innerarity, D. & Colomina, C. La verdad en las democracias algorítmicas - Truth in algorithmic democracies. *Revista CIDOB d'Afers Internacionals* 11–24 (2020).
7. Bayer, J., Bitiukova, N., Bard, P., Szakács, J., Alemanno, A. & Uszkiewicz, E. Disinformation and Propaganda – Impact on the Functioning of the Rule of Law in the EU and its Member States. *SSRN Electronic Journal* (2019) www.doi.org/10.2139/ssrn.3409279.
8. Panel for the Future of Science and Technology (STOA). *Automated tackling of disinformation: major challenges ahead.* https://data.europa.eu/doi/10.2861/368879 (2019).
9. Colomina, C., Margalef, H. S. & Youngs, R. *The impact of disinformation on democratic processes and human rights in the world.* (2021).
10. Rubio Núñez, R. La amenaza tecnológica en los procesos electorales: Una respuesta jurídica. *Revista de privacidad y derecho digital* **3**, 109–146 (2018).
11. High Level Expert Group on Fake News and Online Disinformation. *A multi-dimensional approach to disinformation.* https://digital-strategy.ec.europa.eu/en/library/final-report-high-level-expert-group-fake-news-and-online-disinformation (2018).
12. Wagner, A. Retos filosóficos de las sociedades digitales: esbozo de un enfoque sistémico. *Dilemata* **38**, 13–29 (2022) https://www.dilemata.net/revista/index.php/dilemata/article/view/412000497.
13. van Boheemen, P., Munnichs, G. & Dujso, E. *Digital threats to democracy. On new technology and disinformation.* https://www.rathenau.nl/en/digitalisering/digital-threats-democracy (2020).
14. Tennøe, T. & Barland, M. *Elections, technology and political influencing.* Norwegian Board of Technology (NBT) (2019).
15. Bruns, H., Dessart, F. J. & Pantazi, M. *Covid-19 misinformation: Preparing for future crises.* Joint Research Center https://publications.jrc.ec.europa.eu/repository/handle/JRC130111 (2022) www.doi.org/10.2760/41905.
16. Del-Fresno-García, M. Desórdenes informativos: sobreexpuestos e infrainformados en la era de la posverdad. *Profesional de la información / Information Professional* **28**, (2019) www.doi.org/10.3145/epi.2019.may.02.
17. Eurobarometer. *Fake news and disinformation online.* https://europa.eu/eurobarometer/surveys/detail/2183 (2018).
18. Grupo de expertos de la sociedad civil - Departamento de Seguridad Nacional. Principios para una estrategia contra la desinformación. *Lucha contra las campañas de desinformación en el ámbito de la seguridad nacional. Propuesta de la sociedad civil. Ministerio de la Presidencia, Relaciones con las Cortes y Memoria Democrática. https://www.dsn.gob.es/sites/dsn/files/LibroDesinfoSN.pdf* (2022).
19. Grupo de expertos de la sociedad civil - Departamento de Seguridad Nacional. La alfabetización mediática, herramienta clave en la lucha contra la desinformación. *Lucha contra las campañas de desinformación en el ámbito de la seguridad nacional. Propuesta de la sociedad civil. Ministerio de la Presidencia, Relaciones con las Cortes y Memoria Democrática. https://www.dsn.gob.es/sites/dsn/files/LibroDesinfoSN.pdf* (2022).
20. *COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS On the European democracy action plan.* (2020).
21. Centro Criptológico Nacional. Ministerio de Defensa. Gobierno de España. *Buenas Prácticas en la Desinformación en el Ciberespacio.* CCN-CERT BP/13 https://www.ccn-cert.cni.es/es/informes/informes-de-buenas-practicas-bp/3549-ccn-cert-bp-13-desinformacion-en-el-ciberespacio/file?format=html (2021).
22. Departamento de Seguridad Nacional. Estrategia de Seguridad Nacional 2021. https://www.dsn.gob.es/es/documento/estrategia-seguridad-nacional-2021 [18/04/2022].
23. Chong, M. & Choy, M. An Empirically Supported Taxonomy of Misinformation: *Advances in Media, Entertainment, and the Arts* (eds. Dalkir, K. & Katz, R.) 117–138 (IGI Global, 2020). ISBN: 978-1-79982-543-2.
24. Badillo-Matos, A. *et al. Analysis of the Impact of Disinformation on Political, Economic, Social and Security Issues, Governance Models and Good Practices: The cases of Spain and Portugal.* IBERIFIER Report. DOI: www.doi.org/10.15581/026.002 https://iberifier.eu/2023/06/21/report-analysis-impact-disinformation-june-2023/ (2023).
25. Wardle, C. & Derakhshan, H. *Information Disorder. Toward an interdisciplinary framework for research and policymaking.* Council of Europe (2017).
26. Bennett, W. L. & Livingston, S. The disinformation order: Disruptive communication and the decline of democratic institutions. *European Journal of Communication* **33**, 122–139 (2018) www.doi.org/10.1177/0267323118760317.
27. House of Commons Digital, Culture, Media and Sport Committee. UK Paliament. Disinformation and 'fake news'. **Eighth Report of Session 2017–19**, (2019).
28. *COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS Tackling online disinformation: a European Approach.* (2018).
29. European Commission. Action Plan against Disinformation. https://digital-strategy.ec.europa.eu/en/library/action-plan-against-disinformation [05/06/2023].
30. Borges do Nascimento, I. J., Pizarro, A. B., Almeida, J. M., Azzopardi-Muscat, N., Gonçalves, M. A., Björklund, M. & Novillo-Ortiz, D. Infodemics and health misinformation: a systematic review of reviews. *Bulletin of the World Health Organization* **100**, 544–561 (2022) www.doi.org/10.2471/BLT.21.287654.
31. Ruffo, G., Semeraro, A., Giachanou, A. & Rosso, P. Studying fake news spreading, polarisation dynamics, and manipulation by bots: A tale of networks and language. *Computer Science Review* **47**, (2023) www.doi.org/10.1016/j.cosrev.2022.100531.
32. Comisión Europea. Plan de Acción para la Democracia Europea. https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/new-push-european-democracy/european-democracy-action-plan_es [23/06/2023].
33. Wagner, A. Deliberación, polarización y posverdad. Repensar la responsabilidad en la sociedad digital. *Quaderns de Filosofia* **10**, 51–67 (2023) www.doi.org/10.7203/qfia.10.2.26616.
34. Salaverría-Aliaga, R. *Informe del GTM: Entender y combatir la desinformación sobre ciencia y salud.* http://hdl.handle.net/10261/239480 (2021).
35. Zubiaga, A., Aker, A., Bontcheva, K., Liakata, M. & Procter, R. Detection and Resolution of Rumours in Social Media: A Survey. *ACM Computing Surveys* **51**, 32:1-32:36 (2018) www.doi.org/10.1145/3161603.
36. Zubiaga, A., Liakata, M., Procter, R., Hoi, G. W. S. & Tolmie, P. Analysing How People Orient to and Spread Rumours in Social Media by Looking at Conversational Threads. *PLOS ONE* **11**, e0150989 (2016) www.doi.org/10.1371/journal.pone.0150989.
37. Corredoira y Alfonso, L. European Regulatory Responses to Disinformation. Special Attention to Election Campaigns. *Derecom* 5 (2020).
38. *Willemo J., (2019). Trends and Developments in the Malicious Use of Social Media. Riga: NATO Strategic Communications Centre of Excellence.* https://stratcomcoe.org/pdfjs/?file=/publications/download/nato_report_-_trends_and_developments.pdf?zoom=page-fit .
39. IBERIFIER Reports. *The Impact of Disinformation on the Media Industry in Spain and Portugal.* https://iberifier.eu/2023/02/15/iberifier-reports-the-impact-of-disinformation-on-the-media-industry-in-spain-and-portugal/ (2023).
40. Rinehart, A. Fake news. It's complicated. *First Draft* https://firstdraftnews.org/articles/fake-news-complicated/ [22/06/2023].
41. Starbird, K. Disinformation's spread: bots, trolls and all of us. *Nature* **571**, 449–450 (2019).
42. Matos, Á. B. La sociedad de la desinformación: propaganda, «fake news» y la nueva geopolítica de la información. *Real Instituto Elcano* https://www.realinstitutoelcano.org/documento-de-trabajo/la-sociedad-de-la-desinformacion-propaganda-fake-news-y-la-nueva-geopolitica-de-la-informacion/ [30/08/2023].
43. Martens, B., Aguiar, L., GGmez, E. & Mueller-Langer, F. The Digital Transformation of News Media and the Rise of Disinformation and Fake News. *SSRN Electronic Journal* (2018) www.doi.org/10.2139/ssrn.3164170.
44. Ognyanova, K., Lazer, D., Robertson, R. E. & Wilson, C. Misinformation in action: Fake news exposure is linked to lower trust in media, higher trust in government when your side is in power. *Harvard Kennedy School Misinformation Review* (2020) www.doi.org/10.37016/mr-2020-024.
45. Bobadilla, Á. M. & Isidoro, B. del C. M. *Fake News, desinformación y otros desórdenes informativos.* (Editorial Fragua, 2022). ISBN: 978-84-7074-963-6.
46. Ley Orgánica 10/1995, de 23 de noviembre, del Código Penal.
47. Fiscalía General del Estado. *Tratamiento penal de las 'fake news'.* (2020).
48. Secretaría de Estado de Seguridad. Ministerio del Interior. *Informe sobre la evolución de los delitos de odio en España 2021.*

https://www.interior.gob.es/opencms/pdf/archivos-y-documentacion/documentacion-y-publicaciones/publicaciones-descargables/publicaciones-periodicas/informe-sobre-la-evolucion-de-los-delitos-de-odio-en-Espana/Informe_evolucion_delitos_odio_Espana_2021_126200207.pdf (2021).
49. Salaverría, R. & León, B. Misinformation Beyond the Media: 'Fake News' in the Big Data Ecosystem. *Total Journalism* (eds. Vázquez-Herrero, J., Silva-Rodríguez, A., Negreira-Rey, M.-C., Toural-Bran, C. & López-García, X.) vol. 97 109–121 (Springer International Publishing, 2022). ISBN: 978-3-030-88027-9.
50. Ricard, J. & Medeiros, J. Using misinformation as a political weapon: COVID-19 and Bolsonaro in Brazil. *Harvard Kennedy School Misinformation Review* **1**, (2020) www.doi.org/10.37016/mr-2020-013.
51. Pantti, I. K., Mervi. Fake News: The narrative battle over the Ukrainian conflict. *The Future of Journalism: Risks, Threats and Opportunities* (Routledge, 2019). ISBN: 978-0-429-46203-0.
52. Levinger, M. Master Narratives of Disinformation Campaigns. *Journal of International Affairs* **71**, 125–134 (2018).
53. EUvsDisinfo. Actualización del informe especial del SEAE: evaluación de las narrativas y la desinformación en torno a la pandemia de COVID-19 (actualizada del 23 de abril al 18 de mayo). *EU vs Disinfo* https://euvsdisinfo.eu/es/actualizacion-del-informe-especial-del-seae-breve-evaluacion-de-las-narrativas-y-la-desinformacion-en-torno-a-la-pandemia-de-covid-19-actualizada-del-23-de-abril-al-18-de-mayo/ [19/09/2023].
54. Presidencia del Gobierno. *Informe Anual de Seguridad Nacional 2022*. https://www.dsn.gob.es/es/documento/informe-anual-seguridad-nacional-2022 (2023).
55. Tandoc, E. C., Lim, Z. W. & Ling, R. Defining "Fake News". *Digital Journalism* **6**, 137–153 (2018) www.doi.org/10.1080/21670811.2017.1360143.
56. McCright, A. M. & Dunlap, R. E. Combatting misinformation requires recognizing its types and the factors that facilitate its spread and resonance. *Journal of Applied Research in Memory and Cognition* **6**, 389–396 (2017) www.doi.org/10.1016/j.jarmac.2017.09.005.
57. Giachanou, A., Ghanem, B. & Rosso, P. Detection of conspiracy propagators using psycho-linguistic characteristics. *Journal of Information Science* **49**, 3–17 (2023) www.doi.org/10.1177/0165551520985486.
58. *Michlin-Shapir, V. (2021). The Long Decade of Disinformation. Defence Strategic Communications, 9, 17-33. doi: 10.30966/2018.RIGA.9.5.* https://stratcomcoe.org/pdfjs/?file=/publications/download/web_shapir_dsc_vol9-1.pdf?zoom=page-fit .
59. Grupo de expertos de la sociedad civil - Departamento de Seguridad Nacional. La desinformación: una amenaza a la democracia. *Lucha contra las campañas de desinformación en el ámbito de la seguridad nacional. Propuesta de la sociedad civil. Ministerio de la Presidencia, Relaciones con las Cortes y Memoria Democrática. https://www.dsn.gob.es/sites/dsn/files/LibroDesinfoSN.pdf* (2022).
60. *Giannopoulos, G., Smith, H., Theocharidou, M. The landscape of hybrid threats: a conceptual model. The Landscape of Hybrid Threats: A conceptual model.* EUR 30585; JRC123305. ISBN 978-92-76-29819-9, doi:10.2760/44985, https://data.europa.eu/doi/10.2760/44985 (2021).
61. Jungwirth, R. *et al. Hybrid Threats: A Comprehensive Resilience Ecosystem.* Joint Research Centre. ISBN 978-92-76-53293-4, doi:10.2760/867072, JRC129019. https://publications.jrc.ec.europa.eu/repository/handle/JRC129019 (2023) www.doi.org/10.2760/37899.
62. Pamment, J. *The EU's Role in Fighting Disinformation: Crafting A Disinformation Framework*. Carnegie Endowment for International Peace. https://carnegieendowment.org/2020/09/24/eu-s-role-in-fighting-disinformation-crafting-disinformation-framework-pub-82720 (2020).
63. Johns Hopkins University & Imperial College London. Countering cognitive warfare: awareness and resilience. *NATO Review* https://www.nato.int/docu/review/articles/2021/05/20/countering-cognitive-warfare-awareness-and-resilience/index.html [28/09/2023].
64. Bradshaw, S. & Howard, P. The Global Disinformation Order: 2019 Global Inventory of Organised Social Media Manipulation. *Copyright, Fair Use, Scholarly Communication, etc.* (2019).
65. Benedicto Solsona, M. Á. La UE frente a la desinformación de China y Rusia durante la COVID-19. La necesidad de una mayor proactividad narrativa europea a nivel internacional. *Janus.net* **11**, 84–98 www.doi.org/10.26619/1647-7251.DT21.6.
66. Suau, J. & Puertas-Graell, D. Disinformation narratives in Spain: reach, impact and spreading patterns. *Profesional de la información* **32**, (2023) www.doi.org/10.3145/epi.2023.sep.08.
67. Grupo de expertos de la sociedad civil - Departamento de Seguridad Nacional. Propuesta para combatir las campañas de desinformación en proceso electoral. *Lucha contra las campañas de desinformación en el ámbito de la seguridad nacional. Propuesta de la sociedad civil. Ministerio de la Presidencia, Relaciones con las Cortes y Memoria Democrática. https://www.dsn.gob.es/sites/dsn/files/LibroDesinfoSN.pdf.*
68. European Commission. *Strengthened Code of Practice on Disinformation.* https://digital-strategy.ec.europa.eu/en/library/2022-strengthened-code-practice-disinformation (2022).
69. Hameleers, M. Disinformation as a context-bound phenomenon: toward a conceptual clarification integrating actors, intentions and techniques of creation and dissemination. *Communication Theory* **33**, (2023) www.doi.org/10.1093/ct/qtac021.
70. IBERIFIER Reports. *Political and Legal Aspects of Disinformation in Portugal and Spain.* https://iberifier.eu/2023/10/20/iberifier-reports-legal-and-political-aspects-of-disinformation-in-portugal-and-spain-october-2023/ (2023).
71. Jeangène vILMER, J. B., Escorcia, A., Guillaume, M. & Herrera, J. *Information manipulation: A challenge for our democracies. By the Policy Planning Staff (CAPS) of the Ministry for Europe and Foreign Affairs and the Institute for Strategic Research (IRSEM) of the Ministry for the Armed Forces, Paris, August 2018.* (2018).
72. Posetti, J. & Matthews, A. *A short guide to the history of 'fake news' and disinformation.* International Center for Journalists https://www.icfj.org/sites/default/files/2018-07/A%20Short%20Guide%20to%20History%20of%20Fake%20News%20and%20Disinformation_ICFJ%20Final.pdf (2018).
73. Jamieson, K. H. *Cyberwar: How Russian Hackers and Trolls Helped Elect a President: What We Don't, Can't, and Do Know.* (Oxford University Press, 2020). ISBN: 978-0-19-752895-2.
74. Lemke, T. & Habegger, M. W. Foreign Interference and Social Media Networks: A Relational Approach to Studying Contemporary Russian Disinformation. *Journal of Global Security Studies* **7**, ogac004 (2022) www.doi.org/10.1093/jogss/ogac004.
75. Bastos, M. T. & Mercea, D. The Brexit Botnet and User-Generated Hyperpartisan News. *Social Science Computer Review* **37**, 38–54 (2019) www.doi.org/10.1177/0894439317734157.
76. Greene, C. M., Nash, R. A. & Murphy, G. Misremembering Brexit: partisan bias and individual predictors of false memories for fake news stories among Brexit voters. *Memory* **29**, 587–604 (2021) www.doi.org/10.1080/09658211.2021.1923754.
77. Gerken, T. EU gives Meta and TikTok formal Hamas disinformation deadline. *BBC News* https://www.bbc.com/news/technology-67157733 [21/10/2023].
78. Informe EDMO fact-checking network #29 | Iberifier. https://iberifier.eu/2023/11/22/informe-edmo-numero-29/ [23/11/2023].
79. García-Saisó, S. *et al.* Infodemia en tiempos de COVID-19. *Revista Panamericana de Salud Pública* **45**, e89 (2021) www.doi.org/10.26633/RPSP.2021.89.
80. Siwakoti, S., Yadav, K., Thange, I., Bariletto, N., Zanotti, L., Ghoneim, A. & Shapiro, J. N. *Localized Misinformation in a Global Pandemic: Report on COVID-19 Narratives around the World.* Empirical Study of Conflict, Princeton University, pages 1-68, https://esoc.princeton.edu/publications/localized-misinformation-global-pandemic-report-covid-19-narratives-around-world (2021).
81. Zou, S. Mistranslation as disinformation: COVID-19, global imaginaries, and self-serving cosmopolitanism. *The Cultural Politics of COVID-19* (Routledge, 2022). ISBN: 978-1-00-331041-9.
82. Clemente-Suárez, V. J. *et al.* Mis–Dis Information in COVID-19 Health Crisis: A Narrative Review. *International Journal of Environmental Research and Public Health* **19**, 5321 (2022) www.doi.org/10.3390/ijerph19095321.
83. Lee, S. K., Sun, J., Jang, S. & Connelly, S. Misinformation of COVID-19 vaccines and vaccine hesitancy. *Scientific Reports* **12**, 13681 (2022) www.doi.org/10.1038/s41598-022-17430-6.
84. León, B., Martínez-Costa, M.-P., Salaverría, R. & López-Goñi, I. Health and science-related disinformation on COVID-19: A content analysis of hoaxes identified by fact-checkers in Spain. *PLOS ONE* **17**, e0265995 (2022) www.doi.org/10.1371/journal.pone.0265995.
85. Theocharis, Y. *et al.* Does the platform matter? Social media and COVID-19 conspiracy theory beliefs in 17 countries. *New Media and Society* (2021) www.doi.org/10.1177/14614448211045666.
86. IBERIFIER Reports & Fundación Española para la Ciencia y la Tecnología. *Desinformación científica en España.* https://www.fecyt.es/es/publicacion/desinformacion-cientifica-en-espana (2022).
87. Cano Orón, L., Calvo, D., López García, G. & Baviera, T. Disinformation in Facebook Ads in the 2019 Spanish General Election Campaigns. (2021).
88. Paniagua Rojano, F., Seoane Pérez, F. & Magallón Rosa, R. Anatomía del bulo electoral: la desinformación política durante la campaña del 28-A en España. *Revista CIDOB d'Afers Internacionals* **124**, 123–145 (2020) www.doi.org/doi.org/10.24241/rcai.2020.124.1.123.
89. IBERIFIER Reports. *Spain & Portugal fact-checking brief.* Q1 2023 https://iberifier.eu/2023/05/26/fact-checking-brief-q1-2023/ (2023).
90. Aparici, R., García-Marín, D. & Rincón-Manzano, L. Noticias falsas, bulos y trending topics. Anatomía y estrategias de la desinformación en el conflicto catalán. *Profesional de la información* **28**, (2019) www.doi.org/10.3145/epi.2019.may.13.
91. Vicente, A. R. *Disinformation landscape in Spain.* EU DesinfoLab https://www.disinfo.eu/publications/disinformation-landscape-in-spain/ (2023).
92. Rhodes, S. C. Filter Bubbles, Echo Chambers, and Fake News: How Social Media Conditions Individuals to Be Less Critical of Political Misinformation. *Political Communication* **39**, 1–22 (2022) www.doi.org/10.1080/10584609.2021.1910887.
93. Vosoughi, S., Roy, D. & Aral, S. The spread of true and false news online. *Science* **359**, 1146–1151 (2018) www.doi.org/10.1126/science.aap9559.
94. Rubio Núñez, R. Los efectos de la posverdad en la democracia. *Revista de derecho político* 191–228 (2018) www.doi.org/10.5944/rdp.103.2018.23201.
95. Newman, N., Fletcher, R., Eddy, K., Robertson, C. T. & Nielsen, R. K. *Reuters Institute Digital News Report 2023.*

https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2023-06/Digital_News_Report_2023.pdf (2023).
96. Amoedo-Casais, A., Moreno-Moreno, E., Negredo-Bruna, S., Kaufmann-Argueta, J. & Vara-Miguel, A. Digital News Report España 2023. (2023) www.doi.org/10.15581/019.2023.
97. Bradshaw, S., Bailey, H. & Howard, P. N. *Industrialized Disinformation: 2020 Global Inventory of Organized Social Media Manipulation*. Oxford University, UK: Programme on Democracy&Technology https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/12/2021/02/CyberTroop-Report20-Draft9.pdf .
98. Lasser, J., Aroyehun, S. T., Carrella, F., Simchon, A., Garcia, D. & Lewandowsky, S. From alternative conceptions of honesty to alternative facts in communications by US politicians. *Nature Human Behaviour* 1–12 (2023) www.doi.org/10.1038/s41562-023-01691-w.
99. UNESCO. *Balancing Act: Countering Digital Disinformation While Respecting Freedom of Expression*. https://en.unesco.org/publications/balanceact (2020).
100. IBERIFIER. Iberian Digital Media Map. https://map.iberifier.eu/ [28/09/2023].
101. González-Bailón, S. & Lelkes, Y. Do social media undermine social cohesion? A critical review. *Social Issues and Policy Review* **17**, 155–180 (2023) www.doi.org/10.1111/sipr.12091.
102. Bode, L. & Vraga, E. K. In Related News, That was Wrong: The Correction of Misinformation Through Related Stories Functionality in Social Media. *Journal of Communication* **65**, 619–638 (2015) www.doi.org/10.1111/jcom.12166.
103. *Mair D., Smillie L., La Placa G., Schwendinger F., Raykovska M., Pasztor Z., van Bavel R., Understanding our political nature: How to put knowledge and reason at the heart of political decision-making, EUR 29783 EN, Publications Office of the European Union, Luxembourg, 2019, ISBN 978-92-76-08621-5, doi:10.2760/374191, JRC117161.* https://policycommons.net/artifacts/2162869/understanding-our-political-nature/2918428/ (2019).
104. Innerarity, D. & Colomina, C. Introducción: desinformación y poder, la crisis de los intermediarios. *Revista CIDOB d'afers internacionals* 7–10 (2020) www.doi.org/doi.org/10.24241/rcai.2020.124.1.7.
105. Törnberg, P. Echo chambers and viral misinformation: Modeling fake news as complex contagion. *PLOS ONE* **13**, e0203958 (2018) www.doi.org/10.1371/journal.pone.0203958.
106. Torreblanca, J.-ıgnacıo. Social Networks and Democracy: Problems and Dilemmas of Regulating the Digital Ecosystem. *Siyasal: Journal of Political Sciences* **32**, 15–33 (2023) www.doi.org/10.26650/siyasal.2023.32.1252061.
107. Xaudiera, S. & Cardenal, A. S. Ibuprofen narratives in five European countries during the COVID-19 pandemic. *Harvard Kennedy School Misinformation Review* **1**, (2020) www.doi.org/10.37016/mr-2020-029.
108. Benkler, Y., Faris, R. & Roberts, H. The Propaganda Feedback Loop. *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics* (eds. Benkler, Y., Faris, R. & Roberts, H.) 0 (Oxford University Press, 2018). ISBN: 978-0-19-092362-4.
109. Benkler, Y., Faris, R. & Roberts, H. Mainstream Media Failure Modes and Self-Healing in a Propaganda-Rich Environment. *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics* (eds. Benkler, Y., Faris, R. & Roberts, H.) 0 (Oxford University Press, 2018). ISBN: 978-0-19-092362-4.
110. Heiberger, R., Majó-Vázquez, S., Castro Herrero, L., Nielsen, R. K. & Esser, F. Do Not Blame the Media! The Role of Politicians and Parties in Fragmenting Online Political Debate. *International Journal of Press/Politics* **27**, 910–941 (2022) www.doi.org/10.1177/19401612211015122.
111. Rodríguez-Virgili, J., Serrano-Puche, J. & Fernández, C. B. Digital Disinformation and Preventive Actions: Perceptions of Users from Argentina, Chile, and Spain. *Media and Communication* **9**, 323–337 (2021) www.doi.org/10.17645/mac.v9i1.3521.
112. Rodríguez-Pérez, C., Seibt, T., Magallón-Rosa, R., Paniagua-Rojano, F. J. & Chacón-Peinado, S. Purposes, Principles, and Difficulties of Fact-checking in Ibero-America: Journalists' Perceptions. *Journalism Practice* **0**, 1–19 (2022) www.doi.org/10.1080/17512786.2022.2124434.
113. Abonizio, H. Q., de Morais, J. I., Tavares, G. M. & Barbon Junior, S. Language-Independent Fake News Detection: English, Portuguese, and Spanish Mutual Features. *Future Internet* **12**, 87 (2020) www.doi.org/10.3390/fi12050087.
114. Unkelbach, C., Koch, A., Silva, R. R. & Garcia-Marques, T. Truth by Repetition: Explanations and Implications. *Current Directions in Psychological Science* **28**, 247–253 (2019) www.doi.org/10.1177/0963721419827854.
115. Ecker, U. K. H. *et al.* The psychological drivers of misinformation belief and its resistance to correction. *Nature Reviews Psychology* **1**, 13–29 (2022) www.doi.org/10.1038/s44159-021-00006-y.
116. Fazio, L. K. Repetition Increases Perceived Truth Even for Known Falsehoods. *Collabra: Psychology* **6**, 38 (2020) www.doi.org/10.1525/collabra.347.
117. Almansa-Martínez, A., Fernández-Torres, M. J. & Rodríguez-Fernández, L. Desinformación en España un año después de la COVID-19. Análisis de las verificaciones de Newtral y Maldita. *Revista Latina de Comunicación Social* 183–200 (2022) www.doi.org/10.4185/RLCS-2022-1538.
118. Magallón-Rosa, R. The Agenda Below the Radar. Disinformation and Fact-Checking on (IM) Migration. *Migraciones* **52**, 59–87 (2021) www.doi.org/10.14422/mig.i52.y2021.003.
119. Elsevier. 'Fake news', bulos y contenidos en salud: una tendencia con muchos riesgos. *Elsevier Connect* https://www.elsevier.com/es-es/connect/actualidad-sanitaria/fake-news-bulos-y-contenidos-en-salud-una-tendencia-con-muchos-riesgos [23/10/2023].
120. European Commison. €1.2 million for new project to deepen understanding of disinformation on war, elections and gender | Shaping Europe's digital future. https://digital-strategy.ec.europa.eu/en/news/eu12-million-new-project-deepen-understanding-disinformation-war-elections-and-gender [23/11/2023].
121. Moore, R. C. & Hancock, J. T. A digital media literacy intervention for older adults improves resilience to fake news. *Scientific Reports* **12**, 6008 (2022) www.doi.org/10.1038/s41598-022-08437-0.
122. Blanco-Herrero, D., Amores, J. J. & Sánchez-Holgado, P. Citizen Perceptions of Fake News in Spain: Socioeconomic, Demographic, and Ideological Differences. *Publications* **9**, 35 (2021) www.doi.org/10.3390/publications9030035.
123. Serrano Maíllo, M. I. ¿Qué hace que seamos tan vulnerables a la desinformación?: ¿estamos perdidos o aún podemos hacer algo? *Fake News, desinformación y otros desórdenes informativos, 2022, ISBN 9788470749636, págs. 135-149* 135–149 (Fragua, 2022).
124. Repucci, S. & Slipowitz, A. *Democracy under Siege*. Freedom House https://freedomhouse.org/report/freedom-world/2021/democracy-under-siege .
125. Gerschewski, J. Erosion or decay? Conceptualizing causes and mechanisms of democratic regression. *Democratization* **28**, 43–62 (2021) www.doi.org/10.1080/13510347.2020.1826935.
126. A new low for global democracy. *The Economist* https://www.economist.com/graphic-detail/2022/02/09/a-new-low-for-global-democracy [29/09/2023].
127. Humprecht, E. The Role of Trust and Attitudes toward Democracy in the Dissemination of Disinformation—a Comparative Analysis of Six Democracies. *Digital Journalism* **0**, 1–18 (2023) www.doi.org/10.1080/21670811.2023.2200196.
128. The deep roots of polarisation, or on the need to recover the lost narrative. *CaixaBank Research* https://www.caixabankresearch.com/en/economics-markets/public-sector/deep-roots-polarisation-or-need-recover-lost-narrative [23/11/2023].
129. OECD. The territorial impact of COVID-19: Managing the crisis across levels of government. *OECD* https://www.oecd.org/coronavirus/policy-responses/the-territorial-impact-of-covid-19-managing-the-crisis-across-levels-of-government-d3e314e1/ [25/10/2023].
130. Eurofound. *Maintaining trust during the COVID-19 pandemic*. https://data.europa.eu/doi/10.2806/707872 (2022).
131. Roccato, M., Colloca, P., Cavazza, N. & Russo, S. Coping with the COVID-19 pandemic through institutional trust: Rally effects, compensatory control, and emotions. *Social Science Quarterly* **102**, 2360–2367 (2021) www.doi.org/10.1111/ssqu.13002.
132. University of Oxford. *Social media manipulation by political actors now an industrial scale problem prevalent in over 80 countries – annual Oxford report*. https://www.oii.ox.ac.uk/news-events/news/social-media-manipulation-by-political-actors-now-an-industrial-scale-problem-prevalent-in-over-80-countries-annual-oxford-report (2021).
133. Comisión Europea. *Comunicación conjunta al Parlamento Europeo, al Consejo Europeo y al Consejo. Aumentar la resiliencia y desarrollar las capacidades para hacer frente a las amenazas híbridas*. https://eur-lex.europa.eu/legal-content/GA/TXT/?uri=CELEX:52018JC0016 (2018).
134. Comisión Europea. *Comunicación conjunta al Parlamento Europeo y al Consejo. Comunicación conjunta sobre la lucha contra las amenazas híbridas Una respuesta de la Unión Europea*. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52016JC0018 (2016).
135. EUvsDisinfo. EEAS SPECIAL REPORT UPDATE: Short Assessment of Narratives and Disinformation around the COVID-19/Coronavirus Pandemic (Updated 2 – 22 April). *EUvsDisinfo* https://euvsdisinfo.eu/eeas-special-report-update-2-22-april/ [26/06/2023].
136. Council of the EU. 11th package of sanctions on Russia's war of aggression against Ukraine: additional 71 individuals and 33 entities included in the EU's sanctions list and new tools to counter circumvention and information warfare. https://www.consilium.europa.eu/en/press/press-releases/2023/06/23/11th-package-of-sanctions-on-russia-s-war-of-aggression-against-ukraine-additional-71-individuals-and-33-entities-included-in-the-eu-s-sanctions-list-and-new-tools-to-counter-circumvention-and-information-warfare/ [08/09/2023].
137. EUvsDisinfo. UKRAINE. https://euvsdisinfo.eu/ukraine/ [08/09/2023].
138. European Digital Media Observatory. EDMO Task Force on Disinformation on the War in Ukraine. https://edmo.eu/edmo-task-force-on-disinformation-on-the-war-in-ukraine/ [08/09/2023].
139. EUvsDisinfo. 'To Challenge Russia's Ongoing Disinformation Campaigns': Eight Years of EUvsDisinfo. *EUvsDisinfo* https://euvsdisinfo.eu/to-challenge-russias-ongoing-disinformation-campaigns-eight-years-of-euvsdisinfo/ [23/11/2023].
140. Council of the EU. The fight against pro-Kremlin disinformation. https://www.consilium.europa.eu/en/documents-publications/library/library-blog/posts/the-fight-against-pro-kremlin-disinformation/ [23/11/2023].
141. European Commission. *Digital Services Act: application of the risk management framework to Russian disinformation campaigns*. https://data.europa.eu/doi/10.2759/764631 (2023).
142. Pamment, J. The EU's Role in Fighting Disinformation: Crafting A Disinformation Framework. *Carnegie Endowment for International Peace*

https://carnegieendowment.org/2020/09/24/eu-s-role-in-fighting-disinformation-crafting-disinformation-framework-pub-82720 [26/06/2023].
143. European Partnership for Democracy. *Louder than words? Connecting the dots of European democracy support*. https://epd.eu/news-publications/louder-than-words-connecting-the-dots-of-european-democracy-support/ https://epd.eu/news-publications/louder-than-words-connecting-the-dots-of-european-democracy-support/ (2019).
144. Bel Mallén, I. La realidad es un proceso pernicioso. *Fake news, desinformación y otros desordenes informativos. Directoras Ángela Moreno Bobadilla y Beatriz del Carmen Martínez Isidoro.* (Fragua, 2022). ISBN: 978-84-7074-963-6.
145. La independencia de los medios españoles ante los grupos de presión, bajo sospecha | Digital News Report España 2023 (DNR): informe de noticias digitales en español. https://www.digitalnewsreport.es/2022/la-independencia-de-los-medios-espanoles-ante-los-grupos-de-presion-bajo-sospecha/ [18/09/2023].
146. Mont'Alverne, C., Badrinathan, S., Ross Arguedas, A., Toff, B., Fletcher, R. & Kleis Nielsen, R. *The trust gap: how and why news on digital platforms is viewed more sceptically versus news in general*. https://reutersinstitute.politics.ox.ac.uk/trust-gap-how-and-why-news-digital-platforms-viewed-more-sceptically-versus-news-general (2022).
147. Burgess, J. & Hurcombe, E. Digital Journalism as Symptom, Response, and Agent of Change in the Platformed Media Environment. *Digital Journalism* **7**, 359–367 (2019) www.doi.org/10.1080/21670811.2018.1556313.
148. García-Gordillo, M., Palau-Sampio, D. & Rivas-de-Roca, R. Capítulo 3. Pero ¿qué me cuentas? Una revisión del concepto de verdad en el S. XXI. *Espejo de Monografías de Comunicación Social* 61–81 (2023) www.doi.org/10.52495/c3.emcs.11.p98.
149. Gómez Mompart, J. L., Gutiérrez Lozano, J. F. & Palau-Sampio, D. La calidad periodística en España según la percepción de los periodistas. *Estudios sobre el mensaje periodístico* 13–30 (2015).
150. Sampio, D. P. & Carratalá, A. Injecting disinformation into public space: pseudo-media and reality-altering narratives. *Profesional de la información* **31**, (2022) www.doi.org/10.3145/epi.2022.may.12.
151. Palomares, P. J. & Silva, F. G.-F. e. Visibilidad de la información en redes sociales: los algoritmos de Facebook y su influencia en el clickbait. *Caleidoscopio - Revista Semestral de Ciencias Sociales y Humanidades* 173–211 (2019) www.doi.org/10.33064/41crscsh1772.
152. Bravo Araujo, A., Serrano Puche, J. & Novoa Jaso, M. F. Uso del clickbait en los medios nativos digitales españoles. Un análisis de El Confidencial, El Español, Eldiario.es y Ok Diario. *Dígitos: Revista de Comunicación Digital, 7: 185-* (2021) www.doi.org/10.7203/rd.v1i7.184.
153. McGregor, S. C. Social media as public opinion: How journalists use social media to represent public opinion. *Journalism* **20**, 1070–1086 (2019) www.doi.org/10.1177/1464884919845458.
154. Denisova, A. 'Viral journalism', is it a thing? Adapting quality reporting to shifting social media algorithms and wavering audiences. (eds. Morrison, J., Birks, J. & Berry, M.) 271–278 (Routledge, 2021). ISBN: 978-0-367-24822-2.
155. Petre, C. *All the News That's Fit to Click: How Metrics Are Transforming the Work of Journalists*. (Princeton University Press, 2021). ISBN: 978-0-691-17764-9.
156. Corredoira, L. Anonimato, transparencia e identificación de fuentes, informantes y robots en la era del algoritmo. *Derecho Público de la Inteligencia Artificial* (2023).
157. Mancini, P. Comparing Media Systems and the Digital Age. *International Journal of Communication* **14**, 14 (2020).
158. Busquet Durán, J. Sistemas mediáticos comparados: tres modelos de relación entre los medios de comunicación y la política. *REIS: Revista Española de Investigaciones Sociológicas* 165–172 (2010).
159. Newman, N., Fletcher, R., Robertson, C. T., Eddy, K. & Nielsen, R. K. *Reuters Institute Digital News Report 2022*. https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2022-06/Digital_News-Report_2022.pdf (2022).
160. Vara-Miguel, A., Amoedo-Casais, A., Moreno-Moreno, E., Negredo-Bruna, S. & Kaufmann-Argueta, J. *Digital News Report España 2022*. Pamplona: Servicio de Publicaciones de la Universidad de Navarra. DOI: www.doi.org/10.15581/019.2022 (2022).
161. FAPE. Código Deontológico. https://fape.es/home/codigo-deontologico/ [02/11/2023].
162. Democratic Efficacy and the Varieties of Populism in Europe. Partisan Media Erodes Trust in Politics, New Study Claims. https://demos-h2020.eu/en/partisan-media-erodes-trust-in-politics-new-study-claims [18/09/2023].
163. Majó-Vázquez, S. & González-Bailón, S. Polarización en las audiencias de los medios en España. *Center for Economic Policy - EsadeEcPol* https://www.esade.edu/ecpol/es/publicaciones/polarizacion-medios-espana/ [30/09/2023].
164. Rosa, R. M. La transparencia en la publicidad institucional como garantía de pluralismo informativo. *The Conversation* http://theconversation.com/la-transparencia-en-la-publicidad-institucional-como-garantia-de-pluralismo-informativo-149716 [30/09/2023].
165. Rosa, R. M. La publicidad institucional en España. Evolución legislativa, tecnológica y social. (2020) http://dx.doi.org/10.5209/arab.67255.
166. España. *Ley 29/2005, de 29 de diciembre, de Publicidad y Comunicación Institucional. Jefatura del Estado BOE núm. 312, de 30 de diciembre de 2005 Referencia: BOE-A-2005-21524*.
167. Waisbord, S. Truth is What Happens to News. *Journalism Studies* **19**, 1866–1878 (2018) www.doi.org/10.1080/1461670X.2018.1492881.
168. Krasni, J. How to hijack a discourse? Reflections on the concepts of post-truth and fake news. *Humanities and Social Sciences Communications* **7**, 1–10 (2020) www.doi.org/10.1057/s41599-020-0527-z.
169. De Blasio, E. & Selva, D. Who Is Responsible for Disinformation? European Approaches to Social Platforms' Accountability in the Post-Truth Era. *American Behavioral Scientist* **65**, 825–846 (2021) www.doi.org/10.1177/0002764221989784.
170. European Commission. Identifying conspiracy theories. https://commission.europa.eu/strategy-and-policy/coronavirus-response/fighting-disinformation/identifying-conspiracy-theories_en [12/09/2023].
171. Lewandowsky, S. & Cook, J. *The Conspiracy Theory Handbook*. https://digitalcommons.unl.edu/scholcom/246 (2020).
172. Orriols, L. La polarización afectiva en España: bloques ideológicos enfrentados. *ESADE* https://dobetter.esade.edu/es/polarizacion-afectiva [28/09/2023].
173. Lewandowsky, S., Ecker, U. K. H. & Cook, J. Beyond Misinformation: Understanding and Coping with the "Post-Truth" Era. *Journal of Applied Research in Memory and Cognition* **6**, 353–369 (2017) www.doi.org/10.1016/j.jarmac.2017.07.008.
174. Lewandowsky, S. Hannah Arendt and the contemporary social construction of conspiracy theorists. *PsyArXiv* (2020) www.doi.org/10.31234/osf.io/fm8yg.
175. Pantazi, M., Hale, S. & Klein, O. Social and Cognitive Aspects of the Vulnerability to Political Misinformation. *Political Psychology* **42**, 267–304 (2021) www.doi.org/10.1111/pops.12797.
176. Fiedler, K. Metacognitive Myopia: Gullibility as a Major Obstacle in the Way of Rational Behavior. *The Social Psychology of Gullibility* (Routledge, 2019). ISBN: 978-0-429-20378-7.
177. Cazorla, A., Montabes, J. & López-López, P. C. Medios de comunicación, información política y emociones hacia partidos políticos en España. *Revista Española de Ciencia Política* **58**, 83–109 (2022) www.doi.org/10.21308/recp.58.03.
178. Ross Arguedas, A., Robertson, C., Fletcher, R. & Nielsen, R. *Echo chambers, filter bubbles, and polarisation: a literature review*. https://ora.ox.ac.uk/objects/uuid:6e357e97-7b16-450a-a827-a92c93729a08 (2022).
179. Pérez-Escolar, M. & Noguera-Vivo, J. M. *Hate Speech and Polarization in Participatory Society. Routledge Studies in Media, Communication, and Politics*. (2022). ISBN: 978-0-367-62601-3.
180. Torcal, M. *De votantes a hooligans. La polarización política en España - Catarata*. (Catarata). ISBN: 978-84-13-52614-0.
181. Crockett, M. J. Moral outrage in the digital age. *Nature Human Behaviour* **1**, 769–771 (2017) www.doi.org/10.1038/s41562-017-0213-3.
182. Brady, W. J., Wills, J. A., Jost, J. T., Tucker, J. A. & Van Bavel, J. J. Emotion shapes the diffusion of moralized content in social networks. *Proceedings of the National Academy of Sciences* **114**, 7313–7318 (2017) www.doi.org/10.1073/pnas.1618923114.
183. Petersen, M. B., Osmundsen, M. & Arceneaux, K. The "Need for Chaos" and Motivations to Share Hostile Political Rumors. *American Political Science Review* **117**, 1486–1505 (2023) www.doi.org/10.1017/S0003055422001447.
184. van der Linden, S. Misinformation: susceptibility, spread, and interventions to immunize the public. *Nature Medicine* **28**, 460–467 (2022) www.doi.org/10.1038/s41591-022-01713-6.
185. Pennycook, G. & Rand, D. G. The Psychology of Fake News. *Trends in Cognitive Sciences* **25**, 388–402 (2021) www.doi.org/10.1016/j.tics.2021.02.007.
186. Wolters, H., Stricklin, K., Carey, N. & McBride, M. K. *The Psychology of (Dis)information: A Primer on Key Psychological Mechanisms*. https://www.cna.org/reports/2021/10/psychology-of-disinformation-key-psychological-mechanisms (2021).
187. García, P. A. P., Danieli, N. E. & Freire, I. E. M. Cognitive processing of political fake news. Review of experimental studies. *Dixit* **37**, 44–60 www.doi.org/10.22235/d.v37i1.3112.
188. Vidgen, B., Taylor, H., Pantazi, M., Anastasiou, Z., Inkster, B. & Margetts, H. *Understanding vulnerability to online misinformation*. The Alan Turing Institute https://www.turing.ac.uk/news/publications/understanding-vulnerability-online-misinformation (2021).
189. Howard, P. N., Neudert, L.-M., Prakash, N. & Vosloo, S. *Digital misinformation / disinformation and children*. UNICEF https://www.unicef.org/globalinsight/reports/digital-misinformation-disinformation-and-children (2021).
190. Pantazi, M., Kissine, M. & Klein, O. The Power of the Truth Bias: False Information Affects Memory and Judgment Even in the Absence of Distraction. *Social Cognition* **36**, (2018) www.doi.org/10.1521/soco.2018.36.2.167.
191. Prike, T., Arnold, M. M. & Williamson, P. The relationship between anomalistic belief, misperception of chance and the base rate fallacy. *Thinking & Reasoning* **26**, 447–477 (2020) www.doi.org/10.1080/13546783.2019.1653371.
192. Brashier, N. M. & Marsh, E. J. Judging Truth. *Annual Review of Psychology* **71**, 499–515 (2020) www.doi.org/10.1146/annurev-psych-010419-050807.
193. Begg, I. M., Anas, A. & Farinacci, S. Dissociation of processes in belief: Source recollection, statement familiarity, and the illusion of truth. *Journal of Experimental Psychology: General* **121**, 446–458 (1992)

www.doi.org/10.1037/0096-3445.121.4.446.
194. Unkelbach, C. & Rom, S. C. A referential theory of the repetition-induced truth effect. *Cognition* **160**, 110–126 (2017)
www.doi.org/10.1016/j.cognition.2016.12.016.
195. Mitchell, K. J. & Johnson, M. K. Source monitoring 15 years later: What have we learned from fMRI about the neural mechanisms of source memory? *Psychological Bulletin* **135**, 638–677 (2009)
www.doi.org/10.1037/a0015849.
196. Nadarevic, L., Reber, R., Helmecke, A. J. & Köse, D. Perceived truth of statements and simulated social media postings: an experimental investigation of source credibility, repeated exposure, and presentation format. *Cognitive Research: Principles and Implications* **5**, 56 (2020)
www.doi.org/10.1186/s41235-020-00251-4.
197. Briñol, P. & Petty, R. E. Source factors in persuasion: A self-validation approach. *European Review of Social Psychology* **20**, 49–96 (2009)
www.doi.org/10.1080/10463280802643640.
198. Altay, S., de Araujo, E. & Mercier, H. "If This account is True, It is Most Enormously Wonderful": Interestingness-If-True and the Sharing of True and False News. *Digital Journalism* **10**, 373–394 (2022)
www.doi.org/10.1080/21670811.2021.1941163.
199. Ceylan, G., Anderson, I. A. & Wood, W. Sharing of misinformation is habitual, not just lazy or biased. *Proceedings of the National Academy of Sciences* **120**, e2216614120 (2023)
www.doi.org/10.1073/pnas.2216614120.
200. Pennycook, G. & Rand, D. G. Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition* **188**, 39–50 (2019)
www.doi.org/10.1016/j.cognition.2018.06.011.
201. Brady, W. J., Crockett, M. J. & Van Bavel, J. J. The MAD Model of Moral Contagion: The Role of Motivation, Attention, and Design in the Spread of Moralized Content Online. *Perspectives on Psychological Science* **15**, 978–1010 (2020)
www.doi.org/10.1177/1745691620917336.
202. Islam, A. K. M. N., Laato, S., Talukder, S. & Sutinen, E. Misinformation sharing and social media fatigue during COVID-19: An affordance and cognitive load perspective. *Technological Forecasting and Social Change* **159**, 120201 (2020)
www.doi.org/10.1016/j.techfore.2020.120201.
203. Flore, M. Understanding Citizens' Vulnerabilities (II): From Disinformation to Hostile Narratives. *JRC Publications Repository*
https://publications.jrc.ec.europa.eu/repository/handle/JRC118914 [30/05/2023]
www.doi.org/10.2760/271224.
204. Brady, W. J., Gantman, A. P. & Van Bavel, J. J. Attentional capture helps explain why moral and emotional content go viral. *Journal of Experimental Psychology: General* **149**, 746–756 (2020)
www.doi.org/10.1037/xge0000673.
205. Forgas, J. P. Happy Believers and Sad Skeptics? Affective Influences on Gullibility. *Current Directions in Psychological Science* **28**, 306–313 (2019)
www.doi.org/10.1177/0963721419834543.
206. Morosoli, S., Van Aelst, P., Humprecht, E., Staender, A. & Esser, F. Identifying the Drivers Behind the Dissemination of Online Misinformation: A Study on Political Attitudes and Individual Characteristics in the Context of Engaging With Misinformation on Social Media. *American Behavioral Scientist* 00027642221118300 (2022)
www.doi.org/10.1177/00027642221118300.
207. González-Bailón, S. *et al.* Asymmetric ideological segregation in exposure to political news on Facebook. *Science* **381**, 392–398 (2023)
www.doi.org/10.1126/science.ade7138.
208. Guess, A. M., Barberá, P., Munzert, S. & Yang, J. The consequences of online partisan media. *Proceedings of the National Academy of Sciences* **118**, e2013464118 (2021)
www.doi.org/10.1073/pnas.2013464118.
209. McIlhiney, P., Gignac, G. E., Weinborn, M. & Ecker, U. K. H. Sensitivity to misinformation retractions in the continued influence paradigm: Evidence for stability. *Quarterly Journal of Experimental Psychology* **75**, 1259–1271 (2022)
www.doi.org/10.1177/17470218211048986.
210. Prike, T., Blackley, P., Swire-Thompson, B. & Ecker, U. K. H. Examining the replicability of backfire effects after standalone corrections. *Cognitive Research: Principles and Implications* **8**, (2023)
www.doi.org/10.1186/s41235-023-00492-z.
211. Swire-Thompson, B., Miklaucic, N., Wihbey, J. P., Lazer, D. & DeGutis, J. The backfire effect after correcting misinformation is strongly associated with reliability. *Journal of Experimental Psychology. General* **151**, 1655–1665 (2022)
www.doi.org/10.1037/xge0001131.
212. Stoycheff, E., Burgess, G. S. & Martucci, M. C. Online censorship and digital surveillance: the relationship between suppression technologies and democratization across countries. *Information, Communication & Society* **23**, 474–490 (2020)
www.doi.org/10.1080/1369118X.2018.1518472.
213. Ververis, V., Marguel, S. & Fabian, B. Cross-Country Comparison of Internet Censorship: A Literature Review. *Policy & Internet* **12**, 450–473 (2020)
www.doi.org/10.1002/poi3.228.
214. Andrés, M. B. Modelos de negocio basados en datos, publicidad programática, inteligencia artificial y regulación: algunas reflexiones. *IDP. Revista de Internet, Derecho y Política* 1–13 (2022)
www.doi.org/10.7238/idp.v0i36.401947.
215. Ryan, C. D., Schaul, A. J., Butner, R. & Swarthout, J. T. Monetizing disinformation in the attention economy: The case of genetically modified organisms (GMOs). *European Management Journal* **38**, 7–18 (2020)
www.doi.org/10.1016/j.emj.2019.11.002.
216. Bakir, V. & McStay, A. Fake News and The Economy of Emotions. *Digital Journalism* **6**, 154–175 (2018)
www.doi.org/10.1080/21670811.2017.1345645.
217. Mourão, R. R. & Robertson, C. T. Fake News as Discursive Integration: An Analysis of Sites That Publish False, Misleading, Hyperpartisan and Sensational Information. *Journalism Studies* **20**, 2077–2095 (2019)
www.doi.org/10.1080/1461670X.2019.1566871.
218. Staender, A., Humprecht, E., Esser, F., Morosoli, S. & Van Aelst, P. Is Sensationalist Disinformation More Effective? Three Facilitating Factors at the National, Individual, and Situational Level. *Digital Journalism* **10**, 976–996 (2022)
www.doi.org/10.1080/21670811.2021.1966315.
219. Braun, J. A. & Eklund, J. L. Fake News, Real Money: Ad Tech Platforms, Profit-Driven Hoaxes, and the Business of Journalism. *Digital Journalism* **7**, 1–21 (2019)
www.doi.org/10.1080/21670811.2018.1556314.
220. Global disinformation Index. *The Quarter Billion Dollar Question: How is Disinformation Gaming Ad Tech?*
https://www.disinformationindex.org/files/gdi_ad-tech_report_screen_aw16.pdf (2019).
221. Global disinformation Index. *Research Brief: Ad Tech Fuels Disinformation Sites in Europe – The Numbers and Players*.
https://www.disinformationindex.org/research/2020-3-1-research-brief-ad-tech-fuels-disinformation-sites-in-europe-the-numbers-and-players/ (2020).
222. European Commison. A European Strategy for data | Shaping Europe's digital future.
https://digital-strategy.ec.europa.eu/en/policies/strategy-data [10/06/2022].
223. Garcia, D. Privacy beyond the individual. *Nature Human Behaviour* **3**, 112–113 (2019)
www.doi.org/10.1038/s41562-018-0513-2.
224. Bagrow, J. P., Liu, X. & Mitchell, L. Information flow reveals prediction limits in online social activity. *Nature Human Behaviour* **3**, 122–128 (2019)
www.doi.org/10.1038/s41562-018-0510-5.
225. Dommett, K. Data-driven political campaigns in practice: understanding and regulating diverse data-driven campaigns. *Internet Policy Review* **8**, (2019)
www.doi.org/10.14763/2019.4.1432.
226. Zuiderveen Borgesius, F. J. *et al.* Online Political Microtargeting: Promises and Threats for Democracy. *Utrecht Law Review* **14**, 82 (2018)
www.doi.org/10.18352/ulr.420.
227. Gebru, T., Krause, J., Wang, Y., Chen, D., Deng, J., Aiden, E. L. & Fei-Fei, L. Using deep learning and Google Street View to estimate the demographic makeup of neighborhoods across the United States. *Proceedings of the National Academy of Sciences* **114**, 13108–13113 (2017)
www.doi.org/10.1073/pnas.1700035114.
228. García Sanz, R. M. Questioning the protection of European General Data Regulation: identifying key problems. *Derecom* 3 (2019).
229. Center for Humane Technology. Ledger of Harms. https://ledger.humanetech.com/ [11/09/2023].
230. Metzler, H. & Garcia, D. Social Drivers and Algorithmic Mechanisms on Digital Media. *Perspectives on Psychological Science* (2023)
www.doi.org/10.1177/17456916231185057.
231. Crain, M. & Nadler, A. Political Manipulation and Internet Advertising Infrastructure. *Journal of Information Policy* **9**, 370–410 (2019)
www.doi.org/10.5325/jinfopoli.9.2019.0370.
232. Kreps, S. *The role of technology in online misinformation*. Brookings
https://www.brookings.edu/articles/the-role-of-technology-in-online-misinformation/ (2020).
233. Cardenal, A. S., Aguilar-Paredes, C., Galais, C. & Pérez-Montoro, M. Digital Technologies and Selective Exposure: How Choice and Filter Bubbles Shape News Media Exposure. *International Journal of Press/Politics* **24**, 465–486 (2019)
www.doi.org/10.1177/1940161219862988.
234. Innerarity, D. Justicia algorítmica y autodeterminación deliberativa. *Isegoría* e23–e23 (2023)
www.doi.org/10.3989/isegoria.2023.68.23.
235. Shin, D. *et al.* Countering Algorithmic Bias and Disinformation and Effectively Harnessing the Power of AI in Media. *Journalism & Mass Communication Quarterly* **99**, 887–907 (2022)
www.doi.org/10.1177/10776990221129245.
236. Mohseni, S. & Ragan, E. Combating fake news with interpretable news feed algorithms. *arXiv* (2018)
www.doi.org/10.48550/arXiv.1811.12349.
237. Saurwein, F. & Spencer-Smith, C. Combating Disinformation on Social Media: Multilevel Governance and Distributed Accountability in Europe. *Digital Journalism* **8**, 820–841 (2020)
www.doi.org/10.1080/21670811.2020.1765401.
238. Saurwein, F. Emerging structures of control for algorithms on the Internet: Distributed agency - distributed accountability. 196–211 (2019). ISBN: 978-1-351-11578-0.
239. Shaw, A. Social media, extremism, and radicalization. *Science Advances* **9**, eadk2031 (2023)
www.doi.org/10.1126/sciadv.adk2031.
240. Hosseinmardi, H., Ghasemian, A., Clauset, A., Mobius, M., Rothschild, D. M. & Watts, D. J. Examining the consumption of radical content on YouTube. *Proceedings of the National Academy of Sciences* **118**, e2101967118 (2021)
www.doi.org/10.1073/pnas.2101967118.
241. Wolfowicz, M., Weisburd, D. & Hasisi, B. Examining the interactive effects of the filter bubble and the echo chamber on radicalization. *Journal of Experimental Criminology* **19**, 119–141 (2023)
www.doi.org/10.1007/s11292-021-09471-0.
242. O'Callaghan, D., Greene, D., Conway, M., Carthy, J. & Cunningham, P. Down the (White) Rabbit Hole: The Extreme Right and Online Recommender Systems. *Social Science Computer Review* **33**, 459–478 (2015)
www.doi.org/10.1177/0894439314555329.

243. Colleoni, E., Rozza, A. & Arvidsson, A. Echo Chamber or Public Sphere? Predicting Political Orientation and Measuring Political Homophily in Twitter Using Big Data. *Journal of Communication* **64**, 317–332 (2014) www.doi.org/10.1111/jcom.12084.
244. Pariser, E. *The Filter Bubble: What The Internet Is Hiding From You.* (Penguin UK, 2011). ISBN: 978-0-14-196992-3.
245. Garcia, D. Influence of Facebook algorithms on political polarization tested. *Nature* **620**, 39–41 (2023) www.doi.org/10.1038/d41586-023-02325-x.
246. Nyhan, B. *et al.* Like-minded sources on Facebook are prevalent but not polarizing. *Nature* **620**, 137–144 (2023) www.doi.org/10.1038/s41586-023-06297-w.
247. Guess, A., Lyons, B., Nyhan, B. & Reifler, J. *Avoiding the echo chamber about echo chambers: Why selective exposure to like-minded political news is less prevalent than you think*. (2018).
248. Masip, P., Suau, J. & Ruiz-Caballero, C. Incidental exposure to non-like-minded news through social media: opposing voices in echo-chambers' news feeds. *Media and Communication* **8**, 53–62 (2020) www.doi.org/10.17645/mac.v8i4.3146.
249. Guess, A. M. *et al.* How do social media feed algorithms affect attitudes and behavior in an election campaign? *Science* **381**, 398–404 (2023) www.doi.org/10.1126/science.abp9364.
250. Chen, A. Y., Nyhan, B., Reifler, J., Robertson, R. E. & Wilson, C. Subscriptions and external links help drive resentful users to alternative and extremist YouTube channels. *Science Advances* **9**, eadd8080 (2023) www.doi.org/10.1126/sciadv.add8080.
251. Kolomeets, M. & Chechulin, A. Analysis of the Malicious Bots Market. *2021 29th Conference of Open Innovations Association (FRUCT)* 199–205 (2021). www.doi.org/10.23919/FRUCT52173.2021.9435421.
252. Keller, T. R. & Klinger, U. Social Bots in Election Campaigns: Theoretical, Empirical, and Methodological Implications. *Political Communication* **36**, 171–189 (2019) www.doi.org/10.1080/10584609.2018.1526238.
253. Luceri, L., Deb, A., Giordano, S. & Ferrara, E. Evolution of bot and human behavior during elections. *First Monday* (2019) www.doi.org/10.5210/fm.v24i9.10213.
254. Shao, C., Ciampaglia, G. L., Varol, O., Yang, K.-C., Flammini, A. & Menczer, F. The spread of low-credibility content by social bots. *Nature Communications* **9**, 4787 (2018) www.doi.org/10.1038/s41467-018-06930-7.
255. González-Bailón, S. & De Domenico, M. Bots are less central than verified accounts during contentious political events. *Proceedings of the National Academy of Sciences* **118**, e2013443118 (2021) www.doi.org/10.1073/pnas.2013443118.
256. Hussain, M. N., Tokdemir, S., Agarwal, N. & Al-Khateeb, S. Analyzing Disinformation and Crowd Manipulation Tactics on YouTube. *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)* 1092–1095 (2018). www.doi.org/10.1109/ASONAM.2018.8508766.
257. Giachanou, A., Zhang, X., Barrón-Cedeño, A., Koltsova, O. & Rosso, P. Online information disorder: fake news, bots and trolls. *International Journal of Data Science and Analytics* **13**, 265–269 (2022) www.doi.org/10.1007/s41060-022-00325-0.
258. García-Orosa, B. Disinformation, social media, bots, and astroturfing: the fourth wave of digital democracy. *Profesional de la información* **30**, (2021) www.doi.org/10.3145/epi.2021.nov.03.
259. Jachim, P., Sharevski, F. & Pieroni, E. TrollHunter2020: Real-time Detection of Trolling Narratives on Twitter During the 2020 U.S. Elections. *Proceedings of the 2021 ACM Workshop on Security and Privacy Analytics* 55–65 (Association for Computing Machinery, 2021). www.doi.org/10.1145/3445970.3451158.
260. Santini, R. M., Salles, D., Tucci, G., Ferreira, F. & Grael, F. Making up Audience: Media Bots and the Falsification of the Public Sphere. *Communication Studies* **71**, 466–487 (2020) www.doi.org/10.1080/10510974.2020.1735466.
261. Zannettou, S., Caulfield, T., De Cristofaro, E., Sirivianos, M., Stringhini, G. & Blackburn, J. Disinformation Warfare: Understanding State-Sponsored Trolls on Twitter and Their Influence on the Web. *Companion Proceedings of The 2019 World Wide Web Conference* 218–226 (Association for Computing Machinery, 2019). www.doi.org/10.1145/3308560.3316495.
262. Keller, F. B., Schoch, D., Stier, S. & Yang, J. Political Astroturfing on Twitter: How to Coordinate a Disinformation Campaign. *Political Communication* **37**, 256–280 (2020) www.doi.org/10.1080/10584609.2019.1661888.
263. Lukito, J. *et al.* The Wolves in Sheep's Clothing: How Russia's Internet Research Agency Tweets Appeared in U.S. News as Vox Populi. *The International Journal of Press/Politics* **25**, 196–216 (2020) www.doi.org/10.1177/1940161219895215.
264. Schoch, D., Keller, F. B., Stier, S. & Yang, J. Coordination patterns reveal online political astroturfing across the world. *Scientific Reports* **12**, 4572 (2022) www.doi.org/10.1038/s41598-022-08404-9.
265. Chan, J. Online astroturfing: A problem beyond disinformation. *Philosophy & Social Criticism* 01914537221108467 (2022) www.doi.org/10.1177/01914537221108467.
266. Bail, C. A. *et al.* Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences of the United States of America* **115**, 9216–9221 (2018) www.doi.org/10.1073/pnas.1804840115.
267. Van Bavel, J. J., Robertson, C., del Rosario, K., Rasmussen, J. & Rathje, S. Social Media and Morality. *Annual Review of Psychology* www.doi.org/10.31234/osf.io/ywevq.
268. Gonzalez-Cabañas, J., Cuevas, A., Cuevas, R., Lopez-Fernandez, J. & Garcia, D. Unique on Facebook: Formulation and evidence of (nano)targeting individual users with non-PII data. 464–479 (2021). www.doi.org/10.1145/3487552.3487861.
269. Kim, Y. M. *et al.* The Stealth Media? Groups and Targets behind Divisive Issue Campaigns on Facebook. *Political Communication* **35**, 515–541 (2018) www.doi.org/10.1080/10584609.2018.1476425.
270. Heawood, J. Pseudo-public political speech: Democratic implications of the Cambridge Analytica scandal. *Information Polity* **23**, 429–434 (2018) www.doi.org/10.3233/IP-180009.
271. Lorenz-Spreen, P., Lewandowsky, S., Sunstein, C. R. & Hertwig, R. How behavioural sciences can promote truth, autonomy and democratic discourse online. *Nature Human Behaviour* **4**, 1102–1109 (2020) www.doi.org/10.1038/s41562-020-0889-7.
272. Rothwell, J. Perspectives: Find your audience on digital and storytell with data. *Think with Google* https://www.thinkwithgoogle.com/intl/en-apac/marketing-strategies/video/perspectives-find-your-audience-digital-and-storytell-data/ [05/07/2023].
273. Martin, K. *Ethics of Data and Analytics: Concepts and Cases.* (CRC Press, 2022). ISBN: 978-1-00-056626-0.
274. Blasi Casagran, C. & Vermeulen, M. Reflections on the murky legal practices of political micro-targeting from a GDPR perspective. *International Data Privacy Law* **11**, 348–359 (2021) www.doi.org/10.1093/idpl/ipab018.
275. Dobber, T., Ó Fathaigh, R. & Zuiderveen Borgesius, F. J. The regulation of online political micro-targeting in Europe. *Internet Policy Review* **8**, (2019) www.doi.org/10.14763/2019.4.1440.
276. López-López, P. C., Barredo-Ibáñez, D. & Jaráiz-Gulías, E. Research on Digital Political Communication: Electoral Campaigns, Disinformation, and Artificial Intelligence. *Societies* **13**, 126 (2023) www.doi.org/10.3390/soc13050126.
277. Ryabtsev, K. Political Micro-Targeting in Europe: A Panacea for the Citizens' Political Misinformation or the New Evil for Voting Rights. *Groningen Journal of International Law* **8**, 69–89 (2020) www.doi.org/10.21827/GroJIL.8.1.69-89.
278. Gibney, E. The scant science behind Cambridge Analytica's controversial marketing techniques. *Nature* (2018) www.doi.org/10.1038/d41586-018-03880-4.
279. Hinds, J., Williams, E. J. & Joinson, A. N. "It wouldn't happen to me": Privacy concerns and perspectives following the Cambridge Analytica scandal. *International Journal of Human-Computer Studies* **143**, 102498 (2020) www.doi.org/10.1016/j.ijhcs.2020.102498.
280. Comisión Europea. *Propuesta de reglamento del Parlamento Europeo y del Consejo sobre la transparencia y la segmentación de la publicidad política.* (2021).
281. Ozawa, J. V. S., Woolley, S. C., Straubhaar, J., Riedl, M. J., Joseff, K. & Gursky, J. How Disinformation on WhatsApp Went From Campaign Weapon to Governmental Propaganda in Brazil. *Social Media + Society* **9**, 20563051231160630 (2023) www.doi.org/10.1177/20563051231160632.
282. Paris, B. & Pasquetto, I. Hidden Virality and the Everyday Burden of Correcting WhatsApp Mis- and Disinformation. *Governing Everyday Misinformation* (Cambridge University Press, 2024).
283. Chauchard, S. & Garimella, K. What Circulates on Partisan WhatsApp in India? Insights from an Unusual Dataset. *Journal of Quantitative Description: Digital Media* **2**, (2022) www.doi.org/10.51685/jqd.2022.006.
284. European Parliament. Artificial intelligence: threats and opportunities. https://www.europarl.europa.eu/news/en/headlines/society/20200918STO87404/artificial-intelligence-threats-and-opportunities [29/09/2023].
285. European Parliament. Directorate General for Parliamentary Research Services. *Tackling deepfakes in European policy.* (Publications Office, 2021).
286. Westerlund, M. The Emergence of Deepfake Technology: A Review. *Technology Innovation Management Review* **9**, 39–52 (2019) www.doi.org/10.22215/timreview/1282.
287. Guo, D., Chen, H., Wu, R. & Wang, Y. AIGC Challenges and Opportunities Related to Public Safety: A Case Study of ChatGPT. *Journal of Safety Science and Resilience* (2023) www.doi.org/10.1016/j.jnlssr.2023.08.001.
288. Montoro-Montarroso, A. *et al.* Fighting disinformation with artificial intelligence: fundamentals, advances and challenges. *Profesional de la información* **32**, (2023) www.doi.org/10.3145/epi.2023.may.22.
289. Brown, T. B. *et al.* Language Models are Few-Shot Learners. *arXiv:2005.14165 [cs]* (2020) www.doi.org/10.48550/arXiv.2005.14165.
290. Spitale, G., Biller-Andorno, N. & Germani, F. AI model GPT-3 (dis)informs us better than humans. *Science Advances* **9**, eadh1850 (2023) www.doi.org/10.1126/sciadv.adh1850.
291. Fütterer, T., Fischer, C., Alekseeva, A., Chen, X., Tate, T., Warschauer, M. & Gerjets, P. ChatGPT in education: global reactions to AI innovations. *Scientific Reports* **13**, 15310 (2023) www.doi.org/10.1038/s41598-023-42227-6.
292. Meskó, B. & Topol, E. J. The imperative for regulatory oversight of large language models (or generative AI) in healthcare. *npj Digital Medicine* **6**, 1–6 (2023) www.doi.org/10.1038/s41746-023-00873-0.
293. De Angelis, L., Baglivo, F., Arzilli, G., Privitera, G. P., Ferragina, P., Tozzi, A. E. & Rizzo, C. ChatGPT and the rise of large language models: the new AI-driven infodemic threat in public health. *Frontiers in Public Health* **11**, (2023).
294. Sun, Y., He, J., Lei, S., Cui, L. & Lu, C. T. Med-MMHL: A Multi-Modal Dataset for Detecting Human-and LLM-Generated Misinformation in the Medical Domain. *arXiv:2306.08871 [cs.SI]* (2023)

www.doi.org/10.48550/arXiv.2306.08871.
295. Galaz, V., Metzler, H., Daume, S., Olsson, A., Lindström, B. & Marklund, A. *AI could create a perfect storm of climate misinformation*. Stockholm Resilience Centre (Stockholm University) and the Beijer Institute of Ecological Economics (Royal Swedish Academy of Sciences) http://arxiv.org/abs/2306.12807 (2023).
296. Gao, C. A., Howard, F. M., Markov, N. S., Dyer, E. C., Ramesh, S., Luo, Y. & Pearson, A. T. Comparing scientific abstracts generated by ChatGPT to real abstracts with detectors and blinded human reviewers. *npj Digital Medicine* **6**, 1–5 (2023)
www.doi.org/10.1038/s41746-023-00819-6.
297. Wang, B. *et al.* Wang, B., Chen, W., Pei, H., Xie, C., Kang, M., Zhang, C., ... & Li, B. (2023). DecodingTrust: A Comprehensive Assessment of Trustworthiness in GPT Models. arXiv preprint arXiv:2306.11698. *arXiv:2306.11698 [cs.CL]* (2023)
www.doi.org/10.48550/arXiv.2306.11698.
298. Boucher, P. *What if deepfakes made us doubt everything we see and hear?* European Parliamentary Research Service
https://www.europarl.europa.eu/thinktank/en/document/EPRS_ATA(2021)690046 (2021).
299. Masood, M., Nawaz, M., Malik, K. M., Javed, A., Irtaza, A. & Malik, H. Deepfakes generation and detection: state-of-the-art, open challenges, countermeasures, and way forward. *Applied Intelligence* **53**, 3974–4026 (2023)
www.doi.org/10.1007/s10489-022-03766-z.
300. Rafique, R., Gantassi, R., Amin, R., Frnda, J., Mustapha, A. & Alshehri, A. H. Deep fake detection and classification using error-level analysis and deep learning. *Scientific Reports* **13**, 7422 (2023)
www.doi.org/10.1038/s41598-023-34629-3.
301. Greengard, S. Will deepfakes do deep damage? *Communications of the ACM* **63**, 17–19 (2019)
www.doi.org/10.1145/3371409.
302. UC Berkeley School of Information. Hany Farid Explains How AI Voice Cloning Fuels Misinformation.
https://www.ischool.berkeley.edu/news/2023/hany-farid-explains-how-ai-voice-cloning-fuels-misinformation [16/09/2023].
303. Rini, R. & Cohen, L. Deepfakes, Deep Harms. *Journal of Ethics and Social Philosophy* **22**, 143 (2022).
304. Home Security Heroes. 2023 State Of Deepfakes: Realities, Threats, And Impact.
https://www.homesecurityheroes.com/state-of-deepfakes/#appendix [03/11/2023].
305. Mania, K. Legal Protection of Revenge and Deepfake Porn Victims in the European Union: Findings From a Comparative Legal Study. *Trauma, Violence, & Abuse* **25**, 117–129 (2024)
www.doi.org/10.1177/15248380221143772.
306. Ocaña, J. Caso de Almendralejo: ¿por qué los menores crean contenidos falsos de carácter sexual con IA? *EFE Verifica*
https://verifica.efe.com/almendralejo-inteligencia-artiificial-machismo-menores/ [23/11/2023].
307. Universitat Oberta de Catalunya. El 99% del deepfake dirigido contra las mujeres.
https://www.uoc.edu/portal/es/news/actualitat/2023/265-deepfakes-pornograficos-cuando-IA-desnuda-tu-intimidad-vulnera-tus-derechos.html [23/11/2023].
308. Tolosana, R., Vera-Rodriguez, R., Fierrez, J., Morales, A. & Ortega-Garcia, J. Deepfakes and beyond: A Survey of face manipulation and fake detection. *Information Fusion* **64**, 131–148 (2020)
www.doi.org/10.1016/j.inffus.2020.06.014.
309. Dwivedi, Y. K. *et al.* Metaverse beyond the hype: Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management* **66**, 102542 (2022)
www.doi.org/10.1016/j.ijinfomgt.2022.102542.
310. Mitchell, E., Lee, Y., Khazatsky, A., Manning, C. D. & Finn, C. Detectgpt: Zero-shot machine-generated text detection using probability curvature. *arXiv:2301.11305 [cs.CL]* (2023)
www.doi.org/10.48550/arXiv.2301.11305.
311. Sarvazyan, A. M., González, J. Á., Franco-Salvador, M., Rangel, F., Chulvi, B. & Rosso, P. Overview of AuTexTification at IberLEF 2023: Detection and Attribution of Machine-Generated Text in Multiple Domains. *Procesamiento del Lenguaje Natural* **71**, 275–288 (2023).
312. Sarvazyan, A. M., González, J. Á., Rosso, P. & Franco-Salvador, M. Supervised Machine-Generated Text Detectors: Family and Scale Matters. *Experimental IR Meets Multilinguality, Multimodality, and Interaction* (eds. Arampatzis, A. et al.) 121–132 (Springer Nature Switzerland, 2023).
www.doi.org/10.1007/978-3-031-42448-9_11.
313. U.S. GAO. *Science & Tech Spotlight: Deepfakes*.
https://www.gao.gov/products/gao-20-379sp (2020).
314. van der Sloot, B. & Wagensveld, Y. Deepfakes: regulatory challenges for the synthetic society. *Computer Law & Security Review* **46**, 105716 (2022)
www.doi.org/10.1016/j.clsr.2022.105716.
315. Hine, E. & Floridi, L. New deepfake regulations in China are a tool for social stability, but at what cost? *Nature Machine Intelligence* **4**, 608–610 (2022)
www.doi.org/10.1038/s42256-022-00513-4.
316. Khowaja, S. A., Khowaja, P. & Dev, K. ChatGPT Needs SPADE (Sustainability, PrivAcy, Digital divide, and Ethics) Evaluation: A Review. *arXiv:2305.03123 [cs.CY]* (2023)
www.doi.org/10.48550/arXiv.2305.03123.
317. Hacker, P., Engel, A. & Mauer, M. Regulating ChatGPT and other Large Generative AI Models. *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency* 1112–1123 (Association for Computing Machinery, 2023).
www.doi.org/10.1145/3593013.3594067.
318. European Parliament. EU AI Act: first regulation on artificial intelligence.
https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence [29/09/2023].
319. Comisión Europea. *Propuesta de REGLAMENTO DEL PARLAMENTO EUROPEO Y DEL CONSEJO POR EL QUE SE ESTABLECEN NORMAS ARMONIZADAS EN MATERIA DE INTELIGENCIA ARTIFICIAL (LEY DE INTELIGENCIA ARTIFICIAL) Y SE MODIFICAN DETERMINADOS ACTOS LEGISLATIVOS DE LA UNIÓN*. (2021).
320. Lawson, A. A Look at Global Deepfake Regulation Approaches. *RAI Institute*
https://www.responsible.ai/post/a-look-at-global-deepfake-regulation-approaches [17/09/2023].
321. Pavis, M. Rebalancing our regulatory response to Deepfakes with performers' rights. *Convergence* **27**, 974–998 (2021)
www.doi.org/10.1177/13548565211033418.
322. Lorenz-Spreen, P., Oswald, L., Lewandowsky, S. & Hertwig, R. A systematic review of worldwide causal and correlational evidence on digital media and democracy. *Nature Human Behaviour* **7**, 74–101 (2023)
www.doi.org/10.1038/s41562-022-01460-1.
323. Lazer, D. M. J. *et al.* The science of fake news. *Science* **359**, 1094–1096 (2018)
www.doi.org/10.1126/science.aao2998.
324. Colley, T. P., Granelli, F. & Althuis, J. Disinformation's Societal Impact: Britain, COVID and Beyond. *Defence Strategic Communications* **8**, 89–140 (2020)
www.doi.org/10.30966/2018.RIGA.8.3.
325. Bastick, Z. Would you notice if fake news changed your behavior? An experiment on the unconscious effects of disinformation. *Computers in Human Behavior* **116**, 106633 (2021)
www.doi.org/10.1016/j.chb.2020.106633.
326. Bovet, A. & Makse, H. A. Influence of fake news in Twitter during the 2016 US presidential election. *Nature Communications* **10**, (2019)
www.doi.org/10.1038/s41467-018-07761-2.
327. Arcos, R., Gertrudix, M., Arribas, C. & Cardarilli, M. Responses to digital disinformation as part of hybrid threats: a systematic review on the effects of disinformation and the effectiveness of fact-checking/debunking. *Open Research Europe* (2022)
www.doi.org/10.12688/openreseurope.14088.1.
328. Ghanem, B., Rosso, P. & Rangel, F. An Emotional Analysis of False Information in Social Media and News Articles. *ACM Transactions on Internet Technology* **20**, 19:1-19:18 (2020)
www.doi.org/10.1145/3381750.
329. Rosso, P., Ghanem, B. & Giachanou, A. On the Impact of Emotions on the Detection of False Information. *2021 International Symposium on Electrical, Electronics and Information Engineering* 277–282 (Association for Computing Machinery, 2021).
www.doi.org/10.1145/3459104.3459150.
330. Flore, M., Balahur-Dobrescu, A., Podavini, A. & Verile, M. Understanding Citizens' Vulnerabilities to Disinformation and Data-Driven Propaganda. *JRC Publications Repository*
https://publications.jrc.ec.europa.eu/repository/handle/JRC116009 [30/05/2023]
www.doi.org/10.2760/919835.
331. Zimmermann, F. & Kohring, M. Mistrust, Disinforming News, and Vote Choice: A Panel Survey on the Origins and Consequences of Believing Disinformation in the 2017 German Parliamentary Election. *Political Communication* **37**, 215–237 (2020)
www.doi.org/10.1080/10584609.2019.1686095.
332. Sabatini, F. & Sarracino, F. Online Social Networks and Trust. *Social Indicators Research* **142**, 229–260 (2019)
www.doi.org/10.1007/s11205-018-1887-2.
333. Au, C. H., Ho, K. K. W. & Chiu, D. K. W. The Role of Online Misinformation and Fake News in Ideological Polarization: Barriers, Catalysts, and Implications. *Information Systems Frontiers* **24**, 1331–1354 (2022)
www.doi.org/10.1007/s10796-021-10133-9.
334. Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N. & Westwood, S. J. The Origins and Consequences of Affective Polarization in the United States. *Annual Review of Political Science* **22**, 129–146 (2019)
www.doi.org/10.1146/annurev-polisci-051117-073034.
335. Schaub, M. & Morisi, D. Voter mobilisation in the echo chamber: Broadband internet and the rise of populism in Europe. *European Journal of Political Research* **59**, 752–773 (2020)
www.doi.org/10.1111/1475-6765.12373.
336. Tucker, J. *et al.* Social Media, Political Polarization, and Political Disinformation: A Review of the Scientific Literature. *SSRN Electronic Journal* (2018)
www.doi.org/10.2139/ssrn.3144139.
337. Bronstein, M. V., Pennycook, G., Buonomano, L. & Cannon, T. D. Belief in fake news, responsiveness to cognitive conflict, and analytic reasoning engagement. *Thinking & Reasoning* **27**, 510–535 (2021)
www.doi.org/10.1080/13546783.2020.1847190.
338. Dabbous, A., Aoun Barakat, K. & de Quero Navarro, B. Fake news detection and social media trust: a cross-cultural perspective. *Behaviour & Information Technology* **41**, 2953–2972 (2022)
www.doi.org/10.1080/0144929X.2021.1963475.
339. Bak-Coleman, J. B. *et al.* Stewardship of global collective behavior. *Proceedings of the National Academy of Sciences* **118**, e2025764118 (2021)
www.doi.org/10.1073/pnas.2025764118.
340. Lazer, D. M. J. *et al.* Computational social science: Obstacles and opportunities. *Science* **369**, 1060–1062 (2020)
www.doi.org/10.1126/science.aaz8170.
341. Lazer, D., Hargittai, E., Freelon, D., Gonzalez-Bailon, S., Munger, K., Ognyanova, K. & Radford, J. Meaningful measures of human society in the twenty-first century. *Nature* **595**, 189–196 (2021)
www.doi.org/10.1038/s41586-021-03660-7.
342. Green, Y., Gully, A., Roth, Y., Abhishek, R., Tucker, J. A. & Wanless, A. *Evidence-Based Misinformation Interventions: Challenges and Opportunities for Measurement and Collaboration*.
https://carnegieendowment.org/2023/01/09/evidence-based-misinformation-interventions-challenges-and-opportunities-for-measurement-and-collaboration-pub-88661 (2023).

343. Allcott, H., Braghieri, L., Eichmeyer, S. & Gentzkow, M. The Welfare Effects of Social Media. *American Economic Review* **110**, 629–676 (2020) www.doi.org/10.1257/aer.20190658.
344. European Commission. *Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL establishing a common framework for media services in the internal market (European Media Freedom Act) and amending Directive 2010/13/EU.* (2022).
345. Aba-Catoira, A. Los desórdenes informativos en un sistema de comunicación democrático. *Revista Derecho Político* **1**, 119–151 (2020) www.doi.org/10.5944/rdp.109.2020.29056.
346. Departamento de Seguridad Nacional. Constitución del Foro contra las Campañas de Desinformación en el ámbito de la Seguridad Nacional. https://www.dsn.gob.es/es/actualidad/sala-prensa/constituci%C3%B3n-del-foro-contra-campa%C3%B1as-desinformaci%C3%B3n-%C3%A1mbito-seguridad [25/09/2023].
347. European Commission. *RAN C&N Working Group meeting - How to respond to disinformation in public communications from the perspective of frontline practitioners, 27-29 March 2023*. Migration and Home Affairs https://home-affairs.ec.europa.eu/whats-new/publications/ran-cn-working-group-meeting-how-respond-disinformation-public-communications-perspective-frontline_en (2023).
348. Bolt, N. *Strategic communications and disinformation in the early 21st century.* https://cadmus.eui.eu/handle/1814/74494 (2021).
349. European External Action Service. Tackling disinformation: Information on the work of the EEAS Strategic Communication division and its task forces (SG.STRAT.2). https://www.eeas.europa.eu/countering-disinformation/tackling-disinformation-information-work-eeas-strategic-communication_en [08/09/2023].
350. European Digital Media Observatory. EDMO at a Glance. https://edmo.eu/edmo-at-a-glance/ [22/09/2023].
351. Farkas, J. Fake News in Metajournalistic Discourse. *Journalism Studies* **24**, 423–441 (2023) www.doi.org/10.1080/1461670X.2023.2167106.
352. Gov UK. *The Cairncross Review: A sustainable future for journalism.* https://www.gov.uk/government/publications/the-cairncross-review-a-sustainable-future-for-journalism (2020).
353. Judit, B., Irini, K., Olga, B., Bernd, H., Sarah, H. & Katarzyna, L. *The fight against disinformation and the right to freedom of expression.* European Parliament's Committee on Civil Liberties, Justice and Home Affairs. PE 695.445 https://www.europarl.europa.eu/thinktank/en/document/IPOL_STU(2021)695445 (2021).
354. European Media Freedom Act. *European Commission* https://ec.europa.eu/commission/presscorner/detail/en/ip_22_5504 [22/09/2023].
355. Palau-Sampio, D., Carratalá, A., Tarullo, R. & Crisóstomo Flores, P. Reconocimiento de la calidad como prescriptor contra la desinformación. *Comunicar: Revista Científica de Comunicación y Educación* 59–70 (2022) www.doi.org/10.3916/C72-2022-05.
356. Cavaliere, P. From journalistic ethics to fact-checking practices: defining the standards of content governance in the fight against disinformation. *Journal of Media Law* **12**, 133–165 (2020) www.doi.org/10.1080/17577632.2020.1869486.
357. Stanton, L. Trusting News. *Trusting News* https://trustingnews.org/ [11/12/2023].
358. Lough, K. & McIntyre, K. A systematic review of constructive and solutions journalism research. *Journalism* **24**, 1069–1088 (2023) www.doi.org/10.1177/14648849211044559.
359. Aitamurto, T. & Varma, A. The Constructive Role of Journalism. *Journalism Practice* **12**, 695–713 (2018) www.doi.org/10.1080/17512786.2018.1473041.
360. McIntyre, K. Solutions Journalism. *Journalism Practice* **13**, 16–34 (2019) www.doi.org/10.1080/17512786.2017.1409647.
361. Li, Y. Assessing the Role Performance of Solutions Journalism in a Global Pandemic. *Journalism Practice* **17**, 1445–1464 (2023) www.doi.org/10.1080/17512786.2021.1990787.
362. EDMO – United against disinformation. https://edmo.eu/ [07/07/2023].
363. Iberifier | Iberian Digital Media Research and Fact-Checking Hub. https://iberifier.eu/ [25/09/2023].
364. Comisión Europea. Ley Europea de Libertad de los Medios de Comunicación. *European Commission - European Commission* https://ec.europa.eu/commission/presscorner/detail/es/ip_22_5504 [20/11/2023].
365. Cabrera Blázquez, F. J. *The proposal for a European Media Freedom Act*. European Audiovisual Observatory (2022).
366. EFE Verifica. ¿Qué es EFE Verifica? https://verifica.efe.com/que-es-efe-verifica/ [18/10/2023].
367. Maldita.es. Metodología de Maldito Bulo. *Maldita.es — Periodismo para que no te la cuelen* https://maldita.es/metodologia-de-maldito-bulo/ [18/10/2023].
368. Newtral. Metodología y transparencia. *Newtral* https://www.newtral.es/metodologia-transparencia/ [18/10/2023].
369. RTVE. Herramientas de Verificación de Bulos por internet de RTVE. *RTVE.es* https://www.rtve.es/noticias/verificartve/herramientas-de-verificacion/index.shtml [07/07/2023].
370. Barrera, O., Guriev, S., Henry, E. & Zhuravskaya, E. Facts, alternative facts, and fact checking in times of post-truth politics. *Journal of Public Economics* **182**, 104123 (2020) www.doi.org/10.1016/j.jpubeco.2019.104123.
371. Ecker, U. K. H., O'Reilly, Z., Reid, J. S. & Chang, E. P. The effectiveness of short-format refutational fact-checks. *British Journal of Psychology* **111**, 36–54 (2020) www.doi.org/10.1111/bjop.12383.
372. Brashier, N. M., Pennycook, G., Berinsky, A. J. & Rand, D. G. Timing matters when correcting fake news. *Proceedings of the National Academy of Sciences* **118**, e2020043118 (2021) www.doi.org/10.1073/pnas.2020043118.
373. Maldita.es. Qué es pre-bunking y cómo se lucha contra la desinformación antes del desmentido. *Maldita.es — Periodismo para que no te la cuelen* https://maldita.es/nosotros/20230323/prebunking-que-es-antes-desmentido/ [18/09/2023].
374. Rojas Caja, F. El fact checking. Las agencias de verificación de noticias en España. *bie3: Boletín IEEE* 1492–1505 (2020).
375. Maldita.es. Políticas Públicas. https://maldita.es/politicas-publicas-desarrollo-institucional/ [19/09/2023].
376. Roozenbeek, J. & Van der Linden, S. *Inoculation Theory and Misinformation*. NATO Strategic Communications Centre of Excellence https://stratcomcoe.org/publications/inoculation-theory-and-misinformation/217 (2021).
377. Markowitz, D. M., Levine, T. R., Serota, K. B. & Moore, A. D. Cross-checking journalistic fact-checkers: The role of sampling and scaling in interpreting false and misleading statements. *PLOS ONE* **18**, e0289004 (2023) www.doi.org/10.1371/journal.pone.0289004.
378. Bavel, J. J. V. & Pereira, A. The Partisan Brain: An Identity-Based Model of Political Belief. *Trends in Cognitive Sciences* **22**, 213–224 (2018) www.doi.org/10.1016/j.tics.2018.01.004.
379. Vraga, E. K., Ecker, U. K. H., Žeželj, I., Lazić, A. & Azlan, A. A. To Debunk or Not to Debunk? Correcting (Mis)Information. *Managing Infodemics in the 21st Century: Addressing New Public Health Challenges in the Information Ecosystem* 85–98 (2023). ISBN: 978-3-031-27789-4.
380. Porter, E. & Wood, T. J. The global effectiveness of fact-checking: Evidence from simultaneous experiments in Argentina, Nigeria, South Africa, and the United Kingdom. *Proceedings of the National Academy of Sciences* **118**, e2104235118 (2021) www.doi.org/10.1073/pnas.2104235118.
381. Kyriakidou, M., Cushion, S., Hughes, C. & Morani, M. Questioning Fact-Checking in the Fight Against Disinformation: An Audience Perspective. *Journalism Practice* **0**, 1–17 (2022) www.doi.org/10.1080/17512786.2022.2097118.
382. Walter, N., Cohen, J., Holbert, R. L. & Morag, Y. Fact-Checking: A Meta-Analysis of What Works and for Whom. *Political Communication* **37**, 350–375 (2020) www.doi.org/10.1080/10584609.2019.1668894.
383. Porter, E. & Wood, T. J. Political Misinformation and Factual Corrections on the Facebook News Feed: Experimental Evidence. *The Journal of Politics* **84**, 1812–1817 (2022) www.doi.org/10.1086/719271.
384. Salaverría, R., Buslón, N., López-Pan, F., León, B., López-Goñi, I. & Erviti, M.-C. Desinformación en tiempos de pandemia: tipología de los bulos sobre la Covid-19. *Profesional de la información* **29**, (2020) www.doi.org/10.3145/epi.2020.may.15.
385. Carey, J. M., Guess, A. M., Loewen, P. J., Merkley, E., Nyhan, B., Phillips, J. B. & Reifler, J. The ephemeral effects of fact-checks on COVID-19 misperceptions in the United States, Great Britain and Canada. *Nature Human Behaviour* **6**, 236–243 (2022) www.doi.org/10.1038/s41562-021-01278-3.
386. European Media and Information Fund. Boosting Fact-Checking Activities in Europe. https://gulbenkian.pt/emifund/bolsas-lista/boosting-fact-checking-activities-in-europe/ [22/09/2023].
387. McIlhiney, P., Gignac, G. E., Ecker, U. K. H., Kennedy, B. L. & Weinborn, M. Executive function and the continued influence of misinformation: A latent-variable analysis. *PLoS One* **18**, (2023) www.doi.org/10.1371/journal.pone.0283951.
388. Calabrese, C. & Albarracín, D. Bypassing misinformation without confrontation improves policy support as much as correcting it. *Scientific Reports* **13**, 6005 (2023) www.doi.org/10.1038/s41598-023-33299-5.
389. The European Fact-Checking Standards Network. https://eufactcheckingproject.com/ [30/09/2023].
390. IFCN Code of Principles. https://ifcncodeofprinciples.poynter.org/ [07/07/2023].
391. European Digital Media Observatory. Fact-checking EDMO. https://edmo.eu/fact-checking/ [07/07/2023].
392. Allen, J., Arechar, A. A., Pennycook, G. & Rand, D. G. Scaling up fact-checking using the wisdom of crowds. *Science Advances* **7**, eabf4393 (2021) www.doi.org/10.1126/sciadv.abf4393.
393. Li, J. & Chang, X. Combating Misinformation by Sharing the Truth: a Study on the Spread of Fact-Checks on Social Media. *Information Systems Frontiers* 1–15 (2022) www.doi.org/10.1007/s10796-022-10296-z.
394. Vraga, E. K., Bode, L. & Tully, M. Creating News Literacy Messages to Enhance Expert Corrections of Misinformation on Twitter. *Communication Research* **49**, 245–267 (2022) www.doi.org/10.1177/0093650219898094.
395. Nissen, I. A., Walter, J. G., Charquero-Ballester, M., Bechmann, A. & DATALAB Aarhus University. *A method for auditing fact-checking tools and databases*. NORDIS https://www.tjekdet.dk/files/2022-05/AAReport%20task%201.1%20-%20A%20method%20for%20auditing%20fact-checking%20tools%20and%20databases.pdf (2022).

396. Zeng, X., Abumansour, A. S. & Zubiaga, A. Automated fact-checking: A survey. *Language and Linguistics Compass* **15**, e12438 (2021) www.doi.org/10.1111/lnc3.12438.
397. Das, A., Liu, H., Kovatchev, V. & Lease, M. The state of human-centered NLP technology for fact-checking. *Information Processing & Management* **60**, 103219 (2023) www.doi.org/10.1016/j.ipm.2022.103219.
398. Lazarski, E., Al-Khassaweneh, M. & Howard, C. Using NLP for Fact Checking: A Survey. *Designs* **5**, 42 (2021) www.doi.org/10.3390/designs5030042.
399. Ortega, J. Periodismo e inteligencia artificial: los avances de Newtral. https://www.newtral.es/periodismo-inteligencia-artificial-avances-newtral/20220624/ [18/10/2023].
400. Giachanou, A., Zhang, G. & Rosso, P. Multimodal fake news detection with textual, visual and semantic information. *Lecture Notes in Computer Science* **12284 LNAI**, 30–38 (2020) www.doi.org/10.1007/978-3-030-58323-1_3.
401. Zhou, X. & Zafarani, R. A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities. *ACM Computing Surveys* **53**, 109:1-109:40 (2020) www.doi.org/10.1145/3395046.
402. IBERIFIER & Iberian Media Research and Fact-Checking. *Is the 'AI toolbox for disinformation' ready?* https://iberifier.eu/app/uploads/2023/06/202303-IBERIFIER-Report-Is-the-%E2%80%98AI-toolbox-for-disinformation-ready.pdf (2023).
403. Linden, C. G., Dang-Nguyen, D. T., Salas-Gulliksen, C., Khan, S. A., Amelie, M. & Dierickx, L. State of the art in fact-checking technology. *NORDIS - NORdic observatory for digital media and information DISorders* (2022).
404. Oficina de Ciencia y Tecnología del Congreso de los Diputados (Oficina C). Informe C: Inteligencia artificial y salud. (2022) www.doi.org/10.57952/tcsx-b678.
405. *JOINT COMMUNICATION TO THE EUROPEAN PARLIAMENT, THE EUROPEAN COUNCIL, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS Action Plan against Disinformation*. (2018).
406. European Commission. Media Literacy Guidelines | Shaping Europe's digital future. https://digital-strategy.ec.europa.eu/en/library/media-literacy-guidelines [24/09/2023].
407. Buckingham, D. Epilogue: Rethinking digital literacy: Media education in the age of digital capitalism. *Digital Education Review* 230–239 (2020).
408. Buckingham, D. Teaching media in a 'post-truth' age: fake news, media bias and the challenge for media/digital literacy education. *Culture and Education* **31**, 213–231 (2019) www.doi.org/10.1080/11356405.2019.1603814.
409. Council of Europe. Supporting Quality Journalism through Media and Information Literacy. *Freedom of Expression* https://www.coe.int/en/web/freedom-expression/-/supporting-quality-journalism-through-media-and-information-literacy [24/09/2023].
410. Goodman, E. *Media literacy in Europe and the role of EDMO*. https://edmo.eu/wp-content/uploads/2022/02/Media-literacy-in-Europe-and-the-role-of-EDMO-Report-2021.pdf (2021).
411. NAMLE. Media Literacy Defined. https://namle.net/resources/media-literacy-defined/ [30/10/2023].
412. McDougall, J., Zezulkova, M. & van Driel, B. Teaching media literacy in Europe: evidence of effective school practices in primary and secondary education. *NESET* https://nesetweb.eu/en/resources/library/teaching-media-literacy-in-europe-evidence-of-effective-school-practices-in-primary-and-secondary-education/ [24/09/2023].
413. Dumitru, E.-A., Ivan, L. & Loos, E. A Generational Approach to Fight Fake News: In Search of Effective Media Literacy Training and Interventions. *Human Aspects of IT for the Aged Population. Design, Interaction and Technology Acceptance* (eds. Gao, Q. & Zhou, J.) 291–310 (Springer International Publishing, 2022). www.doi.org/10.1007/978-3-031-05581-2_22.
414. Lee, N. M. Fake news, phishing, and fraud: a call for research on digital media literacy education beyond the classroom. *Communication Education* **67**, 460–466 (2018) www.doi.org/10.1080/03634523.2018.1503313.
415. Guess, A. M., Lerner, M., Lyons, B., Montgomery, J. M., Nyhan, B., Reifler, J. & Sircar, N. A digital media literacy intervention increases discernment between mainstream and false news in the United States and India. *Proceedings of the National Academy of Sciences* **117**, 15536–15545 (2020) www.doi.org/10.1073/pnas.1920498117.
416. Jones-Jang, S. M., Mortensen, T. & Liu, J. Does Media Literacy Help Identification of Fake News? Information Literacy Helps, but Other Literacies Don't. *American Behavioral Scientist* **65**, 371–388 (2021) www.doi.org/10.1177/0002764219869406.
417. Pennycook, G., Epstein, Z., Mosleh, M., Arechar, A. A., Eckles, D. & Rand, D. G. Shifting attention to accuracy can reduce misinformation online. *Nature* **592**, 590–595 (2021) www.doi.org/10.1038/s41586-021-03344-2.
418. Pennycook, G., McPhetres, J., Zhang, Y., Lu, J. G. & Rand, D. G. Fighting COVID-19 Misinformation on Social Media: Experimental Evidence for a Scalable Accuracy-Nudge Intervention. *Psychological Science* **31**, 770–780 (2020) www.doi.org/10.1177/0956797620939054.
419. Literat, I., Abdelbagi, A., Law, N. Y., Cheung, M. Y.-Y. & Tang, R. Research note: Likes, sarcasm and politics: Youth responses to a platform-initiated media literacy campaign on social media. *Harvard Kennedy School Misinformation Review* (2021) www.doi.org/10.37016/mr-2020-67.
420. Vraga, E., Tully, M. & Bode, L. Assessing the relative merits of news literacy and corrections in responding to misinformation on Twitter. *New Media & Society* **24**, 2354–2371 (2022) www.doi.org/10.1177/1461444821998691.
421. Herrero-Curiel, E. & La-Rosa, L. Los estudiantes de secundaria y la alfabetización mediática en la era de la desinformación. *Comunicar: Revista Científica de Comunicación y Educación* **73**, 95–106 (2022) www.doi.org/10.3916/C73-2022-08.
422. Herrero Curiel, E. & La- Rosa Barrolleta, L. Cultura, economía y educación: nuevos desafíos en la sociedad digital. *La alfabetización mediática en secundaria: transversalidad y voluntad* 76–95 (Dykinson S.L., 2021). ISBN: 978-84-13-77585-2.
423. Bernabeu Morón, N., Esteban Ruiz, N., Gallego Hernández, L. & Rosales Páez, A. *Alfabetización mediática y competencias básicas. Proyecto Mediascopio Prensa La lectura de la prensa escrita en el aula*. Ministerio de Educación. Instituto de Formación del Profesorado, Investigación e Innovación Educativa (IFIIE) https://sede.educacion.gob.es/publiventa/alfabetizacion-mediatica-y-competencias-basicas-proyecto-mediascopio-prensa-la-lectura-de-la-prensa-escrita-en-el-aula/ensenanza-tecnologias-de-la-informacion-prensa/14484 (2011).
424. Ministerio de Educación, Formación Profesional y Deportes. Desinformación y alfabetización mediática e informacional. https://www.educacionyfp.gob.es/gl/mc/sgctie/comunicacion/blog/2020/octubre2020/alfabetizacion-mediatica.html [30/10/2023].
425. Herrero Curiel, E. & La-Rosa, L. *Estudio de alfabetización mediática en centros de Educación Secundaria Obligatoria*. Ministerio de Educación y Formación Profesional https://sede.educacion.gob.es/publiventa/d/26772/19/1 (2023).
426. European Commison. Reporting on Media Literacy in Europe | Shaping Europe's digital future. https://digital-strategy.ec.europa.eu/en/library/reporting-media-literacy-europe [24/09/2023].
427. Tsourapas, E. Media, Information and Digital Literacy Organisations in Europe. *EAVI* https://eavi.eu/media-information-digital-literacy-organisations-in-europe/ [24/09/2023].
428. Council of Europe. Dealing with propaganda, misinformation and fake news. *Democratic Schools for All* https://www.coe.int/en/web/campaign-free-to-speak-safe-to-learn/dealing-with-propaganda-misinformation-and-fake-news [25/10/2023].
429. European Regulators Group for Audiovisual Media Services. *Improving Media Literacy campaigns on disinformation (ERGA Report)*. https://erga-online.eu/wp-content/uploads/2021/01/ERGA-SG2-Report-2020-Improving-Media-Literacy-campaigns-on-disinformation.pdf .
430. Orosz, G., Paskuj, B., Faragó, L. & Krekó, P. A prosocial fake news intervention with durable effects. *Scientific Reports* **13**, 3958 (2023) www.doi.org/10.1038/s41598-023-30867-7.
431. EU-Citizen.Science. The News Evaluator. https://eu-citizen.science/project/3 [18/10/2023].
432. Maldita.es. Educación: la caja de herramientas de verificación para que no te la cuelen. *Maldita.es — Periodismo para que no te la cuelen* https://maldita.es/malditobulo/20181128/educacion-la-caja-de-herramientas-de-verificacion-para-que-no-te-la-cuelen/ [18/09/2023].
433. CAPCIT. Consell Assessordel Parlament sobre Ciència i Tecnologia. *Desinformació a les xarxes socials: Què és i com identificar-la*. https://www.parlament.cat/document/intrade/267188726 (2022).
434. Arroyo, D. & Degli Esposti, S. *Como protegerme de la desinformación*. Ministerio de Educación y Formación Profesional (2022).
435. Jefatura del Estado. *Ley 13/2022, de 7 de julio, General de Comunicación Audiovisual*. vol. BOE-A-2022-11311 96114–96220 (2022).
436. Konstantinou, L., Caraban, A. & Karapanos, E. Combating Misinformation Through Nudging. *Human-Computer Interaction – INTERACT 2019* (eds. Lamas, D., Loizides, F., Nacke, L., Petrie, H., Winckler, M. & Zaphiris, P.) 630–634 (Springer International Publishing, 2019). www.doi.org/10.1007/978-3-030-29390-1_51.
437. Sartori, R., Tommasi, F., Ceschi, A., Falser, M., Genero, S. & Belotto, S. Enhancing critical thinking skills and media literacy in initial vocational education and training via self-nudging: The contribution of NERDVET project. *Frontiers in Psychology* **13**, (2022) www.doi.org/10.3389/fpsyg.2022.935673.
438. de Freitas Melo, P., Vieira, C. C., Garimella, K., de Melo, P. O. S. V. & Benevenuto, F. Can WhatsApp Counter Misinformation by Limiting Message Forwarding? *Complex Networks and Their Applications VIII* (eds. Cherifi, H., Gaito, S., Mendes, J. F., Moro, E. & Rocha, L. M.) 372–384 (Springer International Publishing, 2020). www.doi.org/10.1007/978-3-030-36687-2_31.
439. Maldita.es. Desinformación en WhatsApp: el chatbot de Maldita.es y el atributo 'Reenviado Frecuentemente' . https://maldita.es/nosotros/20210603/desinformacion-whatsapp-chatbot-frequently-forwarded-reenviado-frecuentemente/ [21/10/2023].
440. Arroyo, D., Degli-Esposti, S., Gómez-Espés, A., Palmero-Muñoz, S. & Pérez-Miguel, L. On the Design of a Misinformation Widget (MsW) Against Cloaked Science. *Network and System Security* (eds. Li, S., Manulis, M. & Miyaji, A.) 385–396 (Springer Nature Switzerland, 2023). www.doi.org/10.1007/978-3-031-39828-5_21.
441. Banas, J. A. & Rains, S. A. A Meta-Analysis of Research on Inoculation Theory. *Communication Monographs* **77**, 281–311 (2010) www.doi.org/10.1080/03637751003758193.

442. Lewandowsky, S. & van der Linden, S. Countering Misinformation and Fake News Through Inoculation and Prebunking. *European Review of Social Psychology* **32**, 348–384 (2021) www.doi.org/10.1080/10463283.2021.1876983.
443. van der Linden, S., Leiserowitz, A., Rosenthal, S. & Maibach, E. Inoculating the Public against Misinformation about Climate Change. *Global Challenges* **1**, 1600008 (2017) www.doi.org/10.1002/gch2.201600008.
444. van der Linden, S., Maibach, E., Cook, J., Leiserowitz, A. & Lewandowsky, S. Inoculating against misinformation. *Science* **358**, 1141–1142 (2017) www.doi.org/10.1126/science.aar4533.
445. Basol, M., Roozenbeek, J., Berriche, M., Uenal, F., McClanahan, W. P. & Linden, S. van der. Towards psychological herd immunity: Cross-cultural evidence for two prebunking interventions against COVID-19 misinformation. *Big Data & Society* **8**, 20539517211013868 (2021) www.doi.org/10.1177/20539517211013868.
446. Schmid, P. & Betsch, C. Effective strategies for rebutting science denialism in public discussions. *Nature Human Behaviour* **3**, 931–939 (2019) www.doi.org/10.1038/s41562-019-0632-4.
447. Maldita.es. Falacias lógicas que pueden hacer que te la cuelen: aprende a identificarlas. *Maldita.es — Periodismo para que no te la cuelen* https://maldita.es/malditateexplica/20211213/falacias-tipos-discusion-argumentos/ [19/10/2023].
448. Cook, J. Deconstructing climate science denial. *Research Handbook on Communicating Climate Change* 62–78 (Edward Elgar Publishing, 2020). ISBN: 978-1-78990-040-8.
449. McPhedran, R., Ratajczak, M., Mawby, M., King, E., Yang, Y. & Gold, N. Psychological inoculation protects against the social media infodemic. *Scientific Reports* **13**, 5780 (2023) www.doi.org/10.1038/s41598-023-32962-1.
450. Cook, J. *et al.* The cranky uncle game—combining humor and gamification to build student resilience against climate misinformation. *Environmental Education Research* **29**, 607–623 (2023) www.doi.org/10.1080/13504622.2022.2085671.
451. Roozenbeek, J. & van der Linden, S. Fake news game confers psychological resistance against online misinformation. *Palgrave Communications* **5**, 1–10 (2019) www.doi.org/10.1057/s41599-019-0279-9.
452. Butler, L. H. *et al.* The (Mis)Information Game: A social media simulator. *Behavior Research Methods* (2023) www.doi.org/10.3758/s13428-023-02153-x.
453. Roozenbeek, J., Traberg, C. S. & van der Linden, S. Technique-based inoculation against real-world misinformation. *Royal Society Open Science* **9**, 211719 (2022) www.doi.org/10.1098/rsos.211719.
454. Lewsey, F. How to 'inoculate' millions against misinformation on social media. *University of Cambridge* https://www.cam.ac.uk/stories/inoculateexperiment [24/09/2023].
455. Swire-Thompson, B., Cook, J., Butler, L. H., Sanderson, J. A., Lewandowsky, S. & Ecker, U. K. H. Correction format has a limited role when debunking misinformation. *Cognitive Research: Principles and Implications* **6**, (2021) www.doi.org/10.1186/s41235-021-00346-6.
456. Lewandowsky, S. *et al. The Debunking Handbook 2020*. (DOI:10.17910/b7.1182, 2020).
457. Paynter, J. *et al.* Evaluation of a template for countering misinformation—Real-world Autism treatment myth debunking. *PLOS ONE* **14**, e0210746 (2019) www.doi.org/10.1371/journal.pone.0210746.
458. Casero-Ripollés, A., Tuñón, J. & Bouza-García, L. The European approach to online disinformation: geopolitical and regulatory dissonance. *Humanities and Social Sciences Communications* **10**, 1–10 (2023) www.doi.org/10.1057/s41599-023-02179-8.
459. StratCom | NATO Strategic Communications Centre of Excellence Riga, Latvia. https://stratcomcoe.org/ [07/03/2022].
460. The European Centre of Excellence for Countering Hybrid Threats. Hybrid CoE. https://www.hybridcoe.fi/ [27/09/2023].
461. Naciones Unidas. A/HRC/47/25: La desinformación y la libertad de opinión y de expresión Informe de la Relatora Especial sobre la promoción y protección del derecho a la libertad de opinión y de expresión, Irene Khan. *OHCHR* https://www.ohchr.org/es/documents/reports/disinformation-and-freedom-opinion-and-expression-report-special-rapporteur [06/07/2023].
462. Diario Oficial de la Unión Europea. *Carta de los Derechos Fundamentales de la Unión Europea*. (2016).
463. Constitución Española. *Título I. De los derechos y deberes fundamentales*. https://app.congreso.es/consti/constitucion/indice/titulos/articulos.jsp?ini=20&tipo=2 (1978).
464. Ministerio de la Presidencia, Relaciones con las Cortes y Memoria Democrática. *Orden PCM/1030/2020, de 30 de octubre, por la que se publica el Procedimiento de actuación contra la desinformación aprobado por el Consejo de Seguridad Nacional*. vol. BOE-A-2020-13663 96673–96680 (2020).
465. Transparency Centre. Reports Archive. https://disinfocode.eu/reports-archive/ [27/09/2023].
466. European Commison. Code of Practice on Disinformation: new reports available in the Transparency Centre | Shaping Europe's digital future. https://digital-strategy.ec.europa.eu/en/news/code-practice-disinformation-new-reports-available-transparency-centre [27/09/2023].
467. Corredoira y Alfonso, L. Como dijera cicerón, no todo lo molesto es delito: (pese al art. 510 del código penal). *OTROSÍ.: Revista del Colegio de Abogados de Madrid* 28–31 (2021).
468. Tribunal de Cuentas Europeo. *El impacto de la desinformación en la UE: una cuestión abordada, pero no atajada*. https://www.eca.europa.eu/es/publications?did=58682 (2021).
469. *Assessment of the Code of Practice on Disinformation – Achievements and areas for further improvement | Shaping Europe's digital future*. https://digital-strategy.ec.europa.eu/en/library/assessment-code-practice-disinformation-achievements-and-areas-further-improvement (2020).
470. Plasilova, I., Hill, J., Carlberg, M., Goubet, M. W. & Procee, R. Study for the "Assessment of the implementation of the Code of Practice on Disinformation". *Comisión Europea* (2020) www.doi.org/10.2759/188091.
471. European Commison. Code of Practice on Disinformation | Shaping Europe's digital future. https://digital-strategy.ec.europa.eu/en/news/code-practice-disinformation [05/06/2023].
472. *Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on a Single Market For Digital Services (Digital Services Act) and amending Directive 2000/31/EC*. (2020).
473. European Union External Action. Factsheet: Rapid Alert System | EEAS. https://www.eeas.europa.eu/node/59644_en [07/07/2023].
474. *COMUNICACIÓN CONJUNTA AL PARLAMENTO EUROPEO, AL CONSEJO EUROPEO, AL CONSEJO, AL COMITÉ ECONÓMICO Y SOCIAL EUROPEO Y AL COMITÉ DE LAS REGIONES La lucha contra la desinformación acerca de la COVID-19: contrastando los datos*. (2020).
475. *COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS European Commission Guidance on Strengthening the Code of Practice on Disinformation*. (2021).
476. *Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act) (Text with EEA relevance). OJ L* vol. 277 (2022).
477. European Parliament. EU Digital Markets Act and Digital Services Act explained. https://www.europarl.europa.eu/news/en/headlines/society/20211209STO19124/eu-digital-markets-act-and-digital-services-act-explained [25/09/2023].
478. European Commission. State of the Union 2018: European Commission proposes measures for securing free and fair European elections. https://ec.europa.eu/commission/presscorner/detail/en/IP_18_5681 [26/09/2023].
479. Parlamento Europeo. Textos aprobados - Injerencias extranjeras en todos los procesos democráticos de la Unión Europea. https://www.europarl.europa.eu/doceo/document/TA-9-2022-0064_ES.html [26/09/2023].
480. Parlamento Europeo. Comisión Especial sobre Injerencias Extranjeras en Todos los Procesos Democráticos de la Unión Europea, en particular la Desinformación, y sobre el Refuerzo de la Integridad, la Transparencia y la Rendición de Cuentas en el Parlamento Europeo. ING2. https://www.europarl.europa.eu/committees/es/ing2/home/highlights [26/09/2023].
481. République Française. *LOI organique n° 2018-1201 du 22 décembre 2018 relative à la lutte contre la manipulation de l'information (1)*.
482. *Alemania (2017). Network Enforcement Act (Netzdurchsetzunggesetz, NetzDG) de 1 de septiembre (Federal Law Gazette I, p. 3352), https:// germanlawarchive.iuscomp.org/?p=1245*.
483. Departamento de Seguridad Nacional. Gobierno de España. *Estrategia Nacional de Ciberseguridad 2019*. https://www.dsn.gob.es/es/documento/estrategia-nacional-ciberseguridad-2019 (2019).
484. Departamento de Seguridad Nacional. *Informe Anual de Seguridad Nacional 2021*. https://www.dsn.gob.es/es/documento/informe-anual-seguridad-nacional-2021 (2022).
485. Presidencia del Gobierno. *Informe Anual de Seguridad Nacional 2020*. https://www.dsn.gob.es/es/documento/informe-anual-seguridad-nacional-2020 (2021).
486. Saltz, E., Barari, S., Leibowicz, C. & Wardle, C. Misinformation interventions are common, divisive, and poorly understood. *Harvard Kennedy School Misinformation Review* (2021) www.doi.org/10.37016/mr-2020-81.
487. La Moncloa. Interior activa la Red de Coordinación para la Seguridad en Procesos Electorales para el 23J. https://www.lamoncloa.gob.es/serviciosdeprensa/notasprensa/interior/paginas/2023/040723-red-coordinacion-seguridad-elecciones.aspx [26/09/2023].
488. Government of Canada, P. S. and P. C. *Political communications in the digital age. Discussion paper 1: the regulation of political communications under the Canada Elections Act*. https://publications.gc.ca/site/eng/9.886736/publication.html (2020).
489. Elections Canada. Registry Requirements for Political Ads on Online Platforms. https://www.elections.ca/content.aspx?section=pol&dir=regifaq&document=index&lang=e [26/09/2023].
490. Senate of the United States. S.1989 - Honest Ads Act. 115th Congress (2017-2018). https://www.congress.gov/bill/115th-congress/senate-bill/1989/text [26/09/2023].
491. Electoral Commission in New Zealand. About election advertising. *Elections* https://elections.nz/guidance-and-rules/advertising-and-campaigning/about-election-advertising/ [26/09/2023].

492. Furnémont, J. F. & Deirdre, K. *Regulation of Political advertising - A comparative study with reflections on the situation in South-East Europe -*. Council of Europe https://rm.coe.int/study-on-political-advertising-eng-final/1680a0c6e0 (2020).
493. Assemblée Nationale. *Proposition de loi n°419 - 15e législature visant à lutter contre les contenus haineux sur internet*. (2020).
494. Government of Ireland. Government publishes first Report of the Interdepartmental Group on security of Ireland's Electoral Process and Disinformation. https://www.gov.ie/ga/preasraitis/37e936-government-publishes-first-report-of-the-interdepartmental-group-on-/ [26/09/2023].
495. Gov.UK. Transparency in digital campaigning: technical consultation on digital imprints. https://www.gov.uk/government/consultations/transparency-in-digital-campaigning-technical-consultation-on-digital-imprints [26/09/2023].