

Sobre las noticias falsas, las consecuencias de su difusión y las herramientas para contrarrestarlas

About fake news, the consequences of its spread, and tools to counteract it and the tools to counteract them

Oscar R. Puccinelli*

FECHA DE RECEPCIÓN: FEBRERO DE 2023 | FECHA DE APROBACIÓN: JUNIO DE 2023

Para citar este artículo: Puccinelli, Oscar R. (2023). Sobre las noticias falsas, las consecuencias de su difusión y las herramientas para contrarrestarlas. *Criterios Cuadernos de Ciencias Jurídicas y Política Internacional*, 16(1), 1-22. <https://doi.org/10.21500/20115733.7784>

Resumen

El presente artículo examina el fenómeno de las noticias falsas (fake news), sus implicaciones para la democracia y los mecanismos disponibles para su regulación. La proliferación de información falsa ha sido potenciada por el desarrollo de la web 2.0, la inteligencia artificial generativa y la ingeniería social, factores que han facilitado su diseminación en el entorno digital. En este contexto, se analiza la tensión entre la protección de la libertad de expresión y la necesidad de mitigar los efectos nocivos de la desinformación en la sociedad.

* Doctor en Derecho Constitucional (Universidad de Buenos Aires) y doctor honoris causa en diversas universidades peruanas. Profesor de Derecho Constitucional y de Derecho Procesal Constitucional y Transnacional en las Facultades de Derecho de la Universidad Nacional de Rosario y de la Pontificia Universidad Católica Argentina de Rosario. Juez de la Sala II de la Cámara de Apelación en lo Civil y Comercial de Rosario. ORCID: <https://orcid.org/0000-0002-7978-7111>.

La tesis central del estudio plantea que las fake news representan un problema estructural en las sociedades contemporáneas debido a su capacidad para influir en la percepción pública, distorsionar la realidad y afectar la estabilidad democrática. A través de un análisis normativo y teórico, se argumenta que la desinformación no es un fenómeno meramente tecnológico, sino que responde a dinámicas políticas, económicas y sociales que refuerzan su impacto y la convierten en un instrumento de manipulación informativa.

En el apartado de discusión, se examina el conflicto entre la libertad de expresión y la regulación de la desinformación a partir de la teoría de la ponderación de derechos propuesta por Robert Alexy. Desde esta perspectiva, se plantea que cualquier intervención para combatir las fake news debe garantizar el respeto a los principios de adecuación, necesidad y proporcionalidad. Se analizan las distintas respuestas frente a la desinformación, destacando las estrategias implementadas por organismos internacionales, Estados y el sector privado. En particular, se identifican medidas como la autorregulación de plataformas digitales, la verificación de hechos (fact-checking), la moderación algorítmica de contenidos y la promoción de la alfabetización mediática.

Las conclusiones enfatizan la necesidad de adoptar un enfoque integral que combine estrategias regulatorias, tecnológicas y educativas para enfrentar la crisis de desinformación. Si bien la regulación estatal es fundamental, se advierte que cualquier medida debe ser compatible con los principios democráticos y la protección de los derechos fundamentales. Se resalta la importancia de la cooperación internacional para establecer estándares comunes en la lucha contra las fake news, así como el papel clave de la sociedad civil en la promoción de una ciudadanía digital informada y crítica.

Palabras clave

Noticias falsas, desinformación, libertad de expresión, verificación de hechos, regulación digital.

Abstract

This article examines the phenomenon of fake news, its implications for democracy and the mechanisms available for its regulation. The proliferation of false information has been boosted by the development of Web 2.0, generative artificial intelligence and social engineering, factors that have facilitated its dissemination in the digital environment. In this context, the tension between the protection of freedom of expression and the need to mitigate the harmful effects of disinformation in society is analyzed.

The central thesis of the study is that fake news represents a structural problem in contemporary societies due to its ability to influence public perception, distort reality and affect democratic stability. Through a normative and theoretical analysis, it is argued that disinformation is not merely a technological phenomenon, but responds to political, economic and social dynamics that reinforce its impact and turn it into an instrument of information manipulation.

The discussion section examines the conflict between freedom of expression and the regulation of disinformation based on the theory of the weighing of rights proposed by Robert Alexy. From this perspective, it is argued that any intervention to combat fake news must guarantee respect for the principles of appropriateness, necessity and proportionality. The different responses to disinformation are analyzed, highlighting the strategies implemented by international organizations, States and the private sector. In particular, measures such as self-regulation of digital platforms, fact-checking, algorithmic content moderation and the promotion of media literacy are identified.

The conclusions emphasize the need to adopt a comprehensive approach that combines regulatory, technological and educational strategies to address the disinformation crisis. While state regula-

tion is essential, it is cautioned that any measures must be compatible with democratic principles and the protection of fundamental rights. The importance of international cooperation to establish common standards in the fight against fake news is highlighted, as well as the key role of civil society in the promotion of an informed and critical digital citizenship.

Key words

Fake news, disinformation, freedom of expression, fact-checking, digital regulation.

1. Introducción: las *fake news* como un problema tecnológicamente reactualizado que afecta tanto a las personas como al sistema democrático

La difusión de información falsa, incompleta, errónea, distorsionada o intrusiva es un fenómeno que se remonta a la antigüedad, pero que se ha disparado inusitadamente en el marco de la web 2.0, de las redes sociales, de la ingeniería social que se realiza a través de ellas, del *big data* y de la inteligencia artificial generativa, que ya permite la creación directa de imágenes y sonidos cuya inautenticidad es prácticamente indetectable.

Dentro del amplio elenco de informaciones falsas o inexactas, se encuentran las denominadas *fake news*, que conceptualmente, y en su concepción más estricta, refieren a historias falsas que parecen ser noticias, difundidas en internet u otros medios a sabiendas de su falsedad, con la finalidad de influir sobre la población o sobre ciertos sectores de ella para lograr que se haga o se deje de hacer algo (p. ej., ir a votar por alguien o no ir a votar), generando así beneficios personales (p. ej., lograr la cancelación social de una persona determinada), económicos (desprestigiar a un producto) o políticos (p. *gr.*, ganar una elección, lo que significa un grave riesgo institu-

cional) para quienes realizan esa difusión o para quienes contratan esos servicios.

También, pueden incluirse, pero ya en una concepción más amplia de las *fake news*, a aquellas noticias que son las difundidas con ignorancia de su falsedad y las creadas en un contexto humorístico (donde se recurre a la parodia con fines de entretenimiento del receptor); pero nos centraremos en la primera tipología, que potenciada por la inteligencia artificial generativa conduce al fenómeno de las denominadas *deepfakes* (Oxford Advanced Learner's Dictionary, s.f.).

La reciente irrupción de las *fake news* a escala mundial se explica por factores tecnológicos, epistemológicos, económicos, afectivos y políticos que favorecen su difusión especialmente en las redes sociales a partir de mensajes dirigidos a personas influenciables, escogidas a través técnicas de ingeniería social –que incluyen las de perfilado–, a quienes se les ofrece información acorde con sus convicciones a través de las cascadas sociales (cuando se sigue lo dicho o hecho por personas influyentes) y de la polarización grupal (cuando grupos se unen para defender una versión más extremista que la que sostenían).

Frente a esta preocupante situación, tanto del sector público (los Estados y la comunidad internacional) como el privado (la sociedad civil, las asociaciones, los proveedores de servicios de la sociedad de la información), han adoptado una serie de medidas tendientes a detectar las informaciones falsas y operar sobre ellas, siempre a través de la denominada “ponderación de derechos” –con respeto a los subprincipios de adecuación, necesidad y proporcionalidad–, pues se enfrentan, por un lado, diversos derechos fundamentales de las personas afectadas por la difusión de esas noticias, y, por el otro, la libertad de expresión de sus emisores, que se amparan en una libertad preferida que protege incluso al discurso falso o erróneo, cuya detección además se complica dado que la noción de qué es “verdad” en un momento y lugar determinado depende de diversos factores

–*v. gr.*, el consenso social sobre algo, que puede variar con el tiempo, como ha ocurrido en el devenir histórico con los “terraplanistas”–.

2. El conflicto entre la libertad de expresión de quienes difunden las noticias falsas y los derechos de las personas afectadas por ellas

Como se anticipó, la libertad de expresión es amplia y esencialmente rehúye al control ideológico, de manera que, al entrar en conflicto con algún otro derecho o principio, la labor de quien debe interpretar, integrar o aplicar el derecho es más ardua que en los otros casos de conflicto. Cuando esto ocurre, tal como lo explica Alexy (2003), a fin de resolver el conflicto se presentan dos teorías básicas que se utilizan con diferentes matices: una estrecha y rigurosa, denominada “teoría de las reglas”, y una más amplia y comprensiva, la “teoría de los principios”, a las que podría rotulárselas mejor como modelos “de subsunción” y “de ponderación” o ponderativo.

Según la primera, las normas que garantizan los derechos fundamentales no se distinguen esencialmente de otras del sistema jurídico por más que las constitucionales tengan su lugar en el nivel más alto del sistema, pues no hay diferencia fundamental alguna de índole estructural: todas son normas jurídicas, y, como tales, son aplicables exactamente de la misma manera que todas las demás; su peculiaridad solamente consiste en que protegen al ciudadano frente al Estado en determinadas situaciones descritas en abstracto.

A tenor de la segunda, una teoría comprensiva u holística, las normas iusfundamentales no se agotan en esa función protectora frente al Estado, sino que se inserta en un marco más vasto, tal como lo ha venido desarrollando el Tribunal Constitucional Federal alemán desde 1958 cuando resolviera el caso Lüth, donde se enlazan tres ideas centrales: a) que la garantía constitucional de los derechos individuales no se limitan en su función de defensa frente al Estado,

sino que los derechos fundamentales encarnan también un orden de valores objetivos o principios, con lo cual estos no solo tienen el carácter de reglas, sino también el de principios; b) que los valores o principios iusfundamentales no valen únicamente para la relación entre el Estado y el ciudadano, sino mucho más allá de eso, “para todos los ámbitos del derecho”, de modo que existe un efecto de irradiación de los derechos fundamentales sobre todo el sistema jurídico, y, por ello, los derechos fundamentales contienen mandatos de optimización al ser normas de principio que ordenan la realización de algo en la más alta medida, relativamente a las posibilidades materiales y jurídicas, y c) de la estructura de los valores y los principios surge que unos como otros tienden a colisionar, de modo que una colisión de principios solo puede resolverse por la vía de la ponderación de los bienes en juego, ponderación que parte del principio comprensivo de proporcionalidad, compuesto, a su vez, por tres subprincipios que expresan la idea de optimización internalizada en los derechos: adecuación, necesidad y proporcionalidad en sentido estricto.

Así, los subprincipios de adecuación y necesidad imponen una optimización relativa a las posibilidades materiales, de modo que se excluye el empleo de medios que perjudican la realización de, aunque sea un principio, sin promover al menos un principio o meta a cuya realización sirven, y más de una opción de intervención, igualmente adecuada, debe optarse por el medio que interviene menos intensamente al derecho afectado. Finalmente, el subprincipio de proporcionalidad en sentido estricto refiere ya a las posibilidades jurídicas y encierra la máxima, según la cual cuanto más alto sea el grado de incumplimiento o perjuicio de un principio, tanto más tiene que ser la importancia de la realización del otro. Entonces:

La ley de la ponderación muestra que la ponderación se puede descomponer en tres pasos. En un primer paso se establece el grado de incumplimiento o menoscabo de un principio. A éste le sigue, en un segundo paso, la comprobación de la satisfacción del principio contrapuesto. En un tercer paso, finalmente, se determina si la

importancia de la satisfacción del principio contrapuesto justifica la afectación o incumplimiento del otro. (Alexy, 2003, p. 1510)

Esta teoría a la que nos adherimos –que obviamente no es la única, puesto que existen otros iusfilósofos relevantes que mantienen posiciones diferentes, entre ellos Dworkin, y Habermas– ha prestado y presta un gran servicio a la hora de resolver no solo los tradicionales conflictos entre derechos en el mundo físico, sino, especialmente, en el ámbito digital, donde los vertiginosos avances de las TIC han reeditado y potenciado aquellos antiguos conflictos, generado otros, ya propios de la interacción que se produce en sus ámbitos y de nuevos efectos que producen las tecnologías aplicadas (*v. gr.*, los buscadores de internet en tanto posicionan primeramente las noticias más buscadas, la inteligencia artificial como creadora de contenidos falsos que luego son diseminados por la red, etc.).

Estos nuevos efectos si bien comenzaron a producirse fundamentalmente luego de Arpanet¹, a partir del surgimiento de internet, en 1982, no se daban con tanta intensidad en el marco de la estática “web 1.0” de fin del siglo pasado. Esta web “1.0” tenía un carácter casi fijo al funcionar principalmente como un enorme repositorio de información donde los usuarios, en general, solo la “surfeaban” en la búsqueda de contenidos cargados por usuarios expertos, pero no podían contribuir con los propios, siendo así meros consumidores de información. El cambio se dio cuando vertiginosos avances tanto en hardware como en software la transformaron en una “web dinámica”, caracterizada especialmente por ser “participativa” o “colaborativa”, al permitir a los usuarios en general aportar información de todo tipo, especialmente a través de las redes sociales (la “web 2.0”), y, más en la actualidad, donde se discute si estamos en el marco de la “web semántica” (“web 3.0”) o ya ingresamos en la “web simbiótica” (“web 4.0”), donde máquinas y humanos se entremezclan y campea el *machine learning* y la inteligencia artificial.

1 ARPANET (Advanced Research Projects Agency Network) fue una red de computadoras que sentó las bases de internet. Fue creada por el Departamento de Defensa de Estados Unidos en 1966.

En lo que a este trabajo respecta y con independencia de en qué web estamos (tal vez hay todavía un poco de cada una), cabe destacar que se reeditaron y profundizaron tanto los más antiguos como los más recientes conflictos relacionados, por un lado, con el derecho a la información y más en particular con la libertad de expresión, y, por el otro, con los derechos individuales y colectivos que pueden verse afectados por la publicación de información en la red, conflictos que ya se habían exteriorizado en el marco de los medios de comunicación tradicionales, pero que aquí se han visto magnificados por las características particulares de la red, máxime cuando actualmente a partir de los notorios avances de la inteligencia artificial es muy difícil distinguir entre lo que es real y lo fabricado, que se torna prácticamente en indetectable.

Entre los principales puntos de conflicto se destacan los relacionados con la publicación de información falsa, incompleta, errónea, distorsionada o intrusiva, con los denominados “discursos de odio” y “peligrosos”, con el mantenimiento *sine die* de información verdadera que por su antigüedad carece de relevancia actual y perjudica innecesariamente a las personas referidas en ella –lo que da paso, no en todas las legislaciones, al controvertido “derecho al olvido”– y con ciertos contenidos que promueven, a veces justa y a veces injustamente, la denominada “cultura de la cancelación”. Nos referiremos, por cuestiones de espacio, solamente al caso de las noticias falsas.

3. Las respuestas de la comunidad internacional, de los Estados y del sector privado al fenómeno de las *fake news*

Las respuestas frente al auge de las *fake news* (resurgidas especialmente a partir de los escándalos Facebook-Cambridge Analítica en el marco del plebiscito por el Brexit y la elección de Donald Trump de 2016) parten necesariamente de la incorporación de herramientas tecnológicas que permitan detectarlas e identificar a sus autores. Será a partir de allí que se podrá sancionar administrativa, civil y hasta penalmente a sus autores y difusores y adoptar medidas

concretas respecto de los sitios utilizados para dicha difusión. Desde luego, también las medidas se deben dirigir a la concienciación de la población, intentando proteger la “identidad digital” de las personas y propender a la formación de una “ciudadanía digital” fuerte, que empodere a los miembros de la sociedad frente a quienes hacen uso de aquellas.

1.1. La comunidad internacional

Tanto en el plano internacional como en el regional europeo se han adoptado una cantidad prácticamente incontable de documentos.

Entre los emitidos de manera conjunta entre organismos internacionales y regionales se destacan los emitidos anualmente desde 2017 de manera conjunta por las relatorías especiales de la ONU y de la OEA para la Libertad de Expresión, los que, con mayor o menor especificidad, se refieren a las *fake news* y a la desinformación: en 2017, la *Declaración Conjunta sobre Libertad de Expresión y Fake News, Desinformación y Propaganda*, que incluye los “Estándares sobre Desinformación y Propaganda”; en 2018, la *Declaración conjunta sobre la independencia y diversidad de los medios en la era digital*; en 2019, el *Vigésimo Aniversario de la Declaración Conjunta: Desafíos a la Libertad de Expresión en la Próxima Década*; en 2020, la *Declaración conjunta sobre libertad de expresión y elecciones en la era digital*; en 2021, la *Declaración conjunta sobre políticos y funcionarios públicos y la libertad de expresión*, y en 2022 la *Declaración conjunta sobre libertad de expresión y justicia de género*.

Ya más específicamente en el ámbito europeo, se destacan la Comunicación de la Comisión Europea “Hacer frente a la desinformación en línea: Un enfoque europeo” y el Informe sobre la implementación de dicha comunicación, donde se trataron los desafíos que la desinformación en línea presenta para nuestras democracias,

y se describen diversos grupos de acciones que deberían llevar adelante las instituciones públicas y las entidades privadas, en respuesta a tales desafíos (European Commission, 2018b). En 2021, la Comisión presentó una guía para fortalecer el código de prácticas sobre desinformación y más recientemente se aprobó un vademécum sobre la organización y funcionamiento del proceso para dar forma y redactar el Código de Desinformación Reforzado (European Commission, 2018a).

1.2. El sector privado

Los medios de comunicación y las grandes plataformas al ser percibidos como “facilitadores” de las *fake news* en su mayoría comenzaron a autorregularse, crearon herramientas específicas en sus servicios, sitios web, blogs, extensiones para navegadores y aplicaciones, etiquetados de advertencia, eliminación de contenidos, suspensiones o eliminaciones de cuentas, etc.

En el caso de las plataformas de publicación de contenidos, el primer acercamiento de esas empresas frente a las recientes crisis electorales fue muy moderado, de modo que tuvieron un rol principalmente pasivo, pero la persistencia de los desórdenes informativos, ahondada ante la aparición del covid-19 y de la guerra entre Rusia y Ucrania provocó cambios importantes por las que pasaron a adoptar acciones proactivas, destinadas a luchar contra los efectos perniciosos de las *fake news*. En el escenario prepandemia, centraron su intervención en la autenticidad de las cuentas y en la visibilidad de las publicaciones, argumentando razonablemente que no podían convertirse en jueces del debate público y que arbitrar contenidos no solo era una tarea costosa y lenta, sino que los llevaba a confrontarse con los partidos políticos, los gobiernos y la sociedad civil en discusiones acerca de la veracidad de la información. La emergencia sanitaria provocada por el covid-19 llevó a las plataformas a enfocarse en identificar acciones no auténticas y juzgar el contenido a través

de procesos de moderación dependientes principalmente del análisis humano porque, si bien los algoritmos pueden detectar conductas sospechosas, determinar la falsedad depende de análisis complejos y contextuales sobre diversos contenidos que aún no desarrollan de manera efectiva.

En esta dirección, este tipo de actores adoptaron los siguientes tipos de medidas: a) acciones de sensibilización a través de políticas, incluso, compartidas con otros actores, campañas de educación, alfabetización digital y mediática, etc., todas dirigidas a construir un ecosistema positivo frente a la desinformación, empoderando a quienes decidan combatirla y aportando estrategias eficaces frente al fenómeno; b) cambios en el código de las plataformas, incluidos los de algoritmos, mediante recomendaciones; c) cambios de política y acciones de moderación tanto en el ámbito interno como en el externo para la eliminación de contenido reportado como ilegal o contrario a las pautas de la comunidad, y d) acciones de transparencia y relaciones públicas, que provean información sobre el funcionamiento de la plataforma y sobre cómo enfrentar el desafío de la desinformación (Álvarez Ugarte y Del Campo, 2021).

1.3. La sociedad civil

Por su parte, desde la sociedad civil, muchas organizaciones no gubernamentales ofrecen servicios gratuitos, principalmente en sitios de internet, donde comprueban la calidad de los contenidos que pueden tener un alto impacto social y advierten sobre su posible falsedad.

Finalmente, entre las iniciativas creadas sintónicamente en los medios de comunicación y la sociedad civil, se destaca la creación de herramientas tecnológicas dedicadas precisamente a combatir la desinformación, entre las cuales se destacan sitios web, extensiones para navegadores y aplicaciones.

4. Taxonomía de las respuestas recomendadas tanto para el sector público como para el privado

Entre los distintos tipos de respuestas que pueden adoptarse para prevenir y erradicar, en la mayor medida posible, este flagelo, merecen mención las “cuatro respuestas de desinformación de alto nivel”, clasificadas por la Comisión de Banda Ancha para el Desarrollo Sostenible (Unesco, 2020), en el reciente informe sobre “Libertad de expresión y tratamiento de la desinformación en Internet”, donde se menciona a las siguientes:

- a. respuestas de identificación de contenidos;
- b. respuestas legales y de políticas respecto de productores y distribuidores de información, que incluyen respuestas legislativas, prelegislativas y políticas;
- c. campañas nacionales e internacionales de “contradesinformación”, a partir de la construcción de contranarrativas;
- d. respuestas específicamente dirigidas a una problemática determinada, como las relacionadas con las elecciones, y
- e. respuestas dentro de los procesos de producción y distribución de los servicios de la sociedad de la información, concretamente:
 - e.1 las curatoriales (principalmente de políticas editoriales y de contenido que se reflejan en los denominados estándares de la comunidad);
 - e.2 las técnicas y algorítmicas implementadas por las plataformas de publicación de contenidos, los motores de búsqueda y otros terceros relacionados (por ejemplo, complementos de navegador), las que pueden incluir métodos experimentales de investigación con inteligencia artificial para detectar y limitar la difusión de desinformación, o proporcionar contexto o información adicional sobre artículos y publicaciones individuales);

- e.3 las de desmonetización y desincentivación, que eliminan las ganancias de los sitios infractores de las políticas de publicación, y
- e.4 las de apoyo a las víctimas de la desinformación, a través de respuestas éticas, normativas y educativas, incluyendo pautas, recomendaciones, resoluciones, alfabetización mediática y de datos y etiquetados de credibilidad de contenidos.

En esta dirección, este tipo de actores adoptaron los siguientes tipos de medidas:

- a. acciones de sensibilización: a través de políticas incluso compartidas con otros actores, campañas de educación, alfabetización digital y mediática, etc., todas dirigidas a construir un ecosistema positivo frente a la desinformación, empoderando a quienes decidan combatirla y aportando estrategias eficaces frente al fenómeno;
- b. cambios en el código de las plataformas, incluidos los de algoritmos, mediante recomendaciones;
- c. cambios de política y acciones de moderación tanto en el ámbito interno como en el externo para la eliminación de contenido reportado como ilegal o contrario a las pautas de la comunidad, y
- d. acciones de transparencia y relaciones públicas, que proveen información sobre el funcionamiento de la plataforma y sobre cómo enfrentar el desafío de la desinformación (Unesco, 2020).

5. Kit básico de herramientas para detectar *fake news*

Como primera medida, hay que tener claro que en la tipología de las informaciones falsas pueden encontrarse aquellas que encierran:

- a. contenido satírico o paródico (publicación realizada con potencial engañoso, aunque no se busque causar daño con la publicación);
- b. conexiones falsas (títulos, imágenes y citas infieles al contenido);
- c. contenido engañoso (información distorsionada para crear una realidad diferente);
- d. contextualización falsa (información genuina ubicada en un contexto fáctico o temporal diferente al real);
- e. contenido impostor (se suplanta la identidad de las fuentes genuinas),
- f. contenido manipulado (información o imágenes manipulados para engañar), y
- g. contenido fabricado (contenido totalmente falso, diseñado para engañar y causar daño) (AECOC, s/f).

Con esta distinción en mente, y a fin de poder discernir entre las noticias reales y las que no lo son, se aconseja poner atención a las siguientes características de la publicación dubitada:

- I. Mensaje y lingüística: a) factualidad: información verificada e imparcial, uso de apellidos para citar; b) evidencia: datos estadísticos, basados en investigaciones; c) calidad del mensaje: estilo periodístico, editado y corregido; d) léxico y sintáctico: uso frecuente de “hoy” o del tiempo pasado; e) interés de actualidad: conflicto, interés humano, protagonismo.
- II. Fuentes e intenciones: a) fuentes del contenido: fuentes contrastadas, citas o atribuciones, heterogeneidad de fuentes; b) pedigrí: sitio u organización conocido o no, reputación del autor; c) independencia: afiliación del periodista a la organización.
- III. Estructura: a) URL: tipo de formato escogido en la registración; b) sección “Acerca de nosotros” (*about us*): clari-

dad y verificabilidad de autores y editores, disponibilidad y formato de la sección “Contáctenos” (*contact us*), tipo y profesionalidad de los correos electrónicos y métodos de contactos de la organización profesional.

- IV. La red: a) metadatos: metadatos indicadores de autenticidad (Molina et al., 2021).

En la misma dirección, se sugiere poner especial atención a las siguientes claves para detectar *fake news*:

- a. Tener cuidado con los titulares. Las noticias falsas a menudo tienen titulares llamativos en mayúsculas con signos de exclamación y, a menudo, información impactante e inaudita.
- b. Examinar siempre la URL. Una dirección web falsa o una que copia una real puede indicar noticias falsas. Revisa bien los caracteres de la URL porque siempre son pequeños detalles.
- c. Investigar la fuente de la noticia, especialmente antes de compartirla o difundirla. Algunas redes sociales como Facebook o Google han habilitado el botón *Fact Checking* para que los usuarios puedan certificar la veracidad de la información.
- d. Prestar atención al formato. Muchos sitios de noticias falsas tienen errores ortográficos o diseños extraños.
- e. Mirar de cerca las fotos y hacer una búsqueda de imágenes. Las noticias falsas a menudo contienen imágenes o videos manipulados, incluso basados en fotos auténticas tomadas fuera de contexto.
- f. Consultar las fechas. Las noticias falsas pueden tener una línea de tiempo sin sentido o incluir fechas alteradas.
- g. Verificar los hechos y las fuentes del autor para confirmar que son exactos. Si no se menciona la identidad de supuestos expertos, es posible que la noticia sea falsa.
- h. Consultar otras noticias. Si ninguna otra fuente de noticias informa la misma historia, puede ser falsa.

- i. Considerar que la historia puede ser una broma, especialmente cuando la fuente de la noticia es conocida por sus parodias. Si los detalles y el tono sugieren que ha sido escrito con humor, no se trata de una *fake news*.
- j. Ser crítico. Algunas historias son falsas a propósito y con fines ocultos o maliciosos (Simple Lógica, 2017).

6. Balance conclusivo

Una intervención proactiva y eficiente para combatir la actual crisis desinfectómica, que a la vez sea respetuosa de los derechos en conflicto y, en particular, de la libertad de información, necesita, primeramente, determinar qué tipo de información debe ser intervenida sin afectar indebidamente la libertad de expresión y, posteriormente, establecer claramente las obligaciones y responsabilidades que cabe no solo al sector privado, sino, particularmente, al sector público, sobre el cual se hace especial hincapié en el rol de los funcionarios e incluso de los candidatos a cargos públicos electivos, sobre quienes se depositan obligaciones éticas y jurídicas, que son aún más intensas cuando la difusión de estas manifestaciones puede ocasionar graves perjuicios individuales o colectivos.

Haciendo referencia a las normas emergentes del sistema interamericano de derechos humanos y con base en lo expuesto anteriormente, Botero Marino (2017) indica que en el Sistema Interamericano de Derechos Humanos existen dos estándares importantes aplicables a cualquier esfuerzo estatal tendiente a prohibir o regular las noticias falsas: a) la simple falsedad objetiva de la expresión no puede ser objeto de prohibición o sanción estatal, y b) solo se justifica restringir la libertad de expresión para proteger los derechos de terceros o el orden público entendido desde una perspectiva democrática estrictamente definida.

Agrega que precisamente para atender sin restricciones indebidas a la protección de los derechos de terceros y del orden público

democrático, se debe realizar un test que involucre diversos factores relevantes para definir si se debe actuar sobre las expresiones falsas, y de ser así, establecer en qué medida puede intervenir sobre quienes las generan y distribuyen (Botero Marino, 2017).

Complementando esta perspectiva, Sunstein afirma que para establecer la viabilidad e intensidad de la intervención se debe considerar la intencionalidad del declarante (si miente a sabiendas, si es negligente, si duda y no le interesa corroborar la veracidad, si cree que lo que dice es verdad); la magnitud del daño que su expresión puede ocasionar (grave, moderada, leve, inexistente); la probabilidad de que el daño se materialice (cierta, probable, improbable, imposible), y el momento en que esto ocurriría (inmediato, futuro cercano, futuro no cercano, futuro lejano) (Sunstein, 2021).

Para medir la probabilidad de daño, también es muy relevante considerar la posición social del declarante, ya que afecta claramente su credibilidad (no es lo mismo un ministro de salud que se refiere a tratamientos contra el covid-19, a que lo proponga un particular que carece de estudios específicos). Y por supuesto, para medir las consecuencias personales sobre aquel, habrá que valorar si pesan sobre este obligaciones legales específicas de decir la verdad o de tener algún cuidado especial antes de emitir la declaración, ya sea por su función o por haberse sometido contractualmente a determinadas obligaciones (por ejemplo, la aceptación de los usos y condiciones de una red social que fijan pautas sobre cómo expresarse y las consecuencias que puede acarrear que el administrador de la red restrinja o elimine una cuenta, tal como sucedió recientemente con la sanción de Facebook al expresidente Donald Trump, cuya cuenta fue suspendida cuando estaba en el cargo y durante dos años, dado que las expresiones compartidas en la red incitaban al asalto al Capitolio).

A partir de la combinación de estos parámetros, es claro que tanto el sector público como el privado deben basar su actuación en una correcta ponderación de los intereses en conflicto y, así, partir de ella

para decidirse sobre la factibilidad y medida de la intervención. En tal sentido, resulta de utilidad el recurso de fórmulas matemáticas similares a la que proponen Susi (para la aplicación del derecho al olvido en línea o RTBF 2.0) (Susi, 2019), o aquellas propuestas para resolver tensiones entre el derecho al acceso a la información pública y el derecho a la protección de datos.

Enfrentar la crisis generada por la desinformación ha provocado tanto respuestas estatales -que buscan principalmente sancionar administrativa o penalmente a sus autores o difusores- como respuestas privadas, basadas en autorregulaciones desplegadas por las redes sociales (por ejemplo, diversas herramientas utilizadas por Facebook para facilitar la detección y denuncia; campañas de alfabetización y un seguimiento del trabajo realizado en *News Feed*), por otros proveedores de servicios de Internet (por ejemplo, el *Fast Check* de Google) y las organizaciones no gubernamentales que comprueban la calidad del contenido que puedan tener un alto impacto social y advertir sobre su posible falsedad (por ejemplo, a nivel global, la *Global Fact Check*).

Si bien una de las herramientas técnicas para combatir la difusión de *fake news* reside en el uso de algoritmos, Botero Marino (2017) advierte, en sintonía con lo que afirmaron los relatores especiales para la libertad de expresión en su Declaración Conjunta de 2017, que solo es compatible con la libertad de expresión si: a) se basa en criterios transparentes y objetivamente justificables, b) garantiza plenamente el derecho al debido proceso de las partes interesadas y c) incluye la participación de iniciativas ciudadanas dedicadas al *fact-checking* con base en códigos éticos transparentes.

También, recuerda que prohibir o regular las *fake news* a nivel estatal no solo no es el remedio menos restrictivo a la libertad de expresión para combatir este tipo particular de discurso, sino que es estructuralmente incompatible con el funcionamiento mismo de la democracia, ámbito donde el mejor remedio para la mentira es el debate libre. Y sobre el *fact-checking*, establece que, aunque es una

tendencia creciente, todavía son muy pocas las organizaciones que se dedican a ello, de modo que solo para algunas noticias está disponible en la red su correspondiente *fact-check*, que no siempre es fácil encontrar en la web, razón por la cual distintas organizaciones del campo han diseñado guías prácticas para que cualquier persona pueda verificar de forma independiente cualquier noticia (*v. gr.*, FactChecker.org, AfricaCheck, FullFact, Les Décodeurs; incluso PolitiFact.com ofrece una “lista negra” de portales web dedicados a difundir información deliberadamente falsa). La idea común detrás de estas guías es, en términos de la Suprema Corte de los Estados Unidos, que “cada persona es su propio guardián de la verdad” (Botero Marino, 2017, p. 77).

En síntesis de todo lo expuesto, es claro que si bien las *fake news* no son un fenómeno nuevo, al calor de los impensables cambios tecnológicos que se han producido en las últimas décadas (sobre todo en las dos últimas), se han convertido en un enorme problema para las sociedades, en general, y, en particular, para el funcionamiento de las democracias, constituyendo un enorme desafío para la comunidad internacional, los gobiernos, las empresas y la sociedad civil.

Su regulación se hizo imprescindible en los años recientes por diversas razones políticas y de salud pública global, pero aún es incompleta, y en algunos casos, confusa.

Lo que no se pone en duda, más allá de los conflictos que encuentra al tratarse de la libertad de expresión, es la necesidad de que la regulación legal existente hasta el momento (ya sea internacional, regional o nacional) y la autorregulación proveniente del sector privado profundicen, con lineamientos claros y homogéneos globales, buenas prácticas e iniciativas inteligentes, y en esto, la cooperación internacional se hace imprescindible.

La tarea no es fácil, pero tampoco imposible de llevar a cabo. Pongamos –o, en su caso, sigamos poniendo– manos a la obra.

Referencias

- Alexy, R. (2003). Derechos fundamentales, ponderación y racionalidad. En F. Fernández Segado (coord.), *The Spanish Constitution in the European Constitutional Context: La Constitución Española en el Contexto Constitucional Europeo* (pp. 1505-1514). Dykinson.
- Álvarez Ugarte, R. y Del Campo, A. (2021). *Noticias falsas en Internet: Acciones y reacciones de tres plataformas*. Universidad de Palermo, CELE.
- Unesco (2020). *Balancing act: Countering digital disinformation while respecting freedom of expression: Broadband Commission research report on 'Freedom of Expression and Addressing Disinformation on the Internet'*. <https://unesdoc.unesco.org/ark:/48223/pf0000379015>
- Oxford Advanced Learner's Dictionary (s.f.). Deepfake. En *OxfordLearnersDictionaries.com*. <https://www.oxfordlearnersdictionaries.com/definition/english/deepfake>
- European Commission (2018a). 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). <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52018DC0236>
- European Commission (2018b). Report from the commission to the European parliament, the European council, the council, the European economic and social committee and the committee of the regions on the Implementation of the Communication "Tackling Online Disinformation: A European Approach". <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52018DC0794>
- Simple Lógica. 2017. Grupo Nueva Pescanova, I Estudio sobre el Impacto de las Fake News en España. Recuperado el 20 de mayo de 2024, de <https://www.simplelogica.com/wp-content/uploads/2018/10/estudiopescanova.pdf>
- AECOC (s.f). *Facebook lucha contra las "Fake News"*. <https://www.aecoc.es/innovation-hub-noticias/facebook-lucha-contra-las-fake-news/>
- Botero Marino, C. (2017). "La regulación estatal de las llamadas "noticias falsas" desde la perspectiva del derecho a la libertad de expresión". En Trust for the Americas (eds.), *Libertad de expresión: A 30 años de la Opinión Consultiva sobre la colegiación obligatoria de periodistas* (pp. 65-84). Trust for the Americas. https://www.oas.org/es/cidh/expresion/docs/publicaciones/OC5_ESP.PDF

- Molina, M. D., Sundar, S. S., Le, T., & Lee, D. (2021). "Fake news" is not simply false information: A concept explication and taxonomy of online content. *American behavioral scientist*, 65(2), 180-212.
- Sunstein, C. R. (2021). *Liars: Falsehoods and free speech in an age of deception*. Oxford University Press.
- Susi, M. (2019). The internet balancing formula. *European Law Journal*, 25(2), 198-212. <https://doi.org/10.1111/eulj.12308>