

Estudios de Deusto

Revista de Derecho Público

Vol. 73/1 enero-junio 2025

DOI: <https://doi.org/10.18543/ed7312025>

ESTUDIOS

DE LA CENSURA A LA LEGITIMIDAD: LAS DIFERENTES ESTRATEGIAS DE COMBATE A LA DESINFORMACIÓN EN LÍNEA FRENTE A LA LIBERTAD DE EXPRESIÓN

*From censorship to legitimacy: competing strategies to
tackle online disinformation and their implications for
freedom of expression*

Marcelo Malheiros Cerqueira

Procurador da República, Ministério Público Federal. Brasil

Profesor y orientador pedagógico, Escola Superior

do Ministério Público da União (ESMPU). Brasil

<https://orcid.org/0000-0002-4724-6353>

<https://doi.org/10.18543/ed.3331>

Fecha de recepción: 01.05.2025

Fecha de aprobación: 10.06.2025

Fecha de publicación en línea: junio 2025

Derechos de autoría / Copyright

Estudios de Deusto. Revista de Derecho Público es una revista de acceso abierto, lo que significa que es de libre acceso en su integridad. Se permite su lectura, la búsqueda, descarga, distribución y reutilización legal en cualquier tipo de soporte sólo para fines no comerciales, sin la previa autorización del editor o el autor, siempre que la obra original sea debidamente citada y cualquier cambio en el original esté claramente indicado.

Estudios de Deusto. Revista de Derecho Público is an Open Access journal which means that it is free for full access, reading, search, download, distribution, and lawful reuse in any medium only for non-commercial purposes, without prior permission from the Publisher or the author; provided the original work is properly cited and any changes to the original are clearly indicated.

Estudios de Deusto. Revista de Derecho Público

© Universidad de Deusto • ISSN 0423-4847 • ISSN-e 2386-9062, Vol. 73/1, enero-junio 2025

<http://www.revista-estudios.deusto.es/>

DE LA CENSURA A LA LEGITIMIDAD: LAS DIFERENTES ESTRATEGIAS DE COMBATE A LA DESINFORMACIÓN EN LÍNEA FRENTE A LA LIBERTAD DE EXPRESIÓN

From censorship to legitimacy: competing strategies to tackle online disinformation and their implications for freedom of expression

Marcelo Malheiros Cerqueira

Procurador da República, Ministério Público Federal. Brasil

Profesor y orientador pedagógico, Escola Superior do Ministério Público da União (ESMPU). Brasil

<https://orcid.org/0000-0002-4724-6353>

<https://doi.org/10.18543/ed.3331>

Fecha de recepción: 01.05.2025

Fecha de aprobación: 10.06.2025

Fecha de publicación en línea: junio 2025

Resumen

En este artículo exploro los desafíos conceptuales, jurídicos y políticos que involucran la desinformación en el ambiente digital. A partir de la revisión de investigaciones empíricas y ejemplos concretos relacionados con el tema, discuto diferentes políticas y soluciones jurídicas de enfrentamiento a la diseminación de contenidos falsos o engañosos en el escenario europeo y brasileño. Al final, destaco los riesgos de censura e inseguridad jurídica en la imposición de límites a la libertad de expresión con base en la categorización de la desinformación o formulaciones análogas como “noticias falsas” (“fake news”) como ilícito.

Palabras clave

Libertad de expresión. Desinformación en línea. Estrategias de combate a la desinformación. Legitimidad. Censura.

Abstract

In this article, I explore the conceptual, legal, and political challenges surrounding disinformation in the digital environment. Based on a review of empirical research and concrete examples related to the topic, I discuss different policies and legal solutions to address the dissemination of false or misleading content in the European and Brazilian contexts. Finally, I highlight the risks of censorship and legal uncertainty in imposing limits on freedom of expression based on the categorization of disinformation or analogous formulations such as “fake news” as an unlawful act.

Keywords

Freedom of expression. Online disinformation. Strategies to tackle disinformation. Legitimacy. Censorship.

Sumario: I. INTRODUCCIÓN: LA CRISIS DE LA VERDAD EN LA ERA DIGITAL. II. DE LAS “FAKE NEWS” EN LAS ELECCIONES BRASILEÑAS A LAS INVESTIGACIONES SOBRE LA DESINFORMACIÓN EN LÍNEA. III. Las dificultades conceptuales que rodean a la desinformación. IV. LA DESINFORMACIÓN SOBRE LA DESINFORMACIÓN. V. EL COMBATE A LA DESINFORMACIÓN. 1. *Prebunking* y *debunking*. 2. La (controvertida) tendencia de categorización como contenido ilícito. 3. Separando el grano de la paja: qué es censurable y qué es legítimo. VI. CONCLUSIONES. VII. REFERENCIAS BIBLIOGRÁFICAS. JURISPRUDENCIA.

I. INTRODUCCIÓN

La distinción entre verdad y mentira ha sido una espina en las relaciones humanas. Los hechos, doctrinas e hipótesis, además de ser frecuentemente complejos, no siempre son unívocos. Sus múltiples interpretaciones dan lugar a discusiones que se oponen o se complementan mediante fragmentos de la verdad, lo que dificulta discernir lo que es cierto de lo que es erróneo.

Las evidencias empíricas son indicativas de la verdad, pero también pueden ser opuestas o disonantes entre sí, además de estar sujetas a eventual superación con el paso del tiempo. No existe, por lo tanto, monopolio de la verdad. Cualquier intento de arbitrarla coercitivamente en cuestiones bajo debate, ya sea por autoridades públicas o por entidades privadas, resbala en autoritarismo y censura.

Las tecnologías digitales del siglo XXI han permitido que la información –tanto verdadera cuanto falsa– circule de manera ágil, ininterrumpida y masiva, impulsada por sistemas de recomendación y diseños de plataformas que fomentan la interacción compulsiva. Ha habido una amplificación estructural, algorítmica y social en la difusión y el consumo de la información en el entorno digital, lo que ha añadido dificultades a la compleja tarea de discernir lo falso de lo verdadero.

En las redes sociales, los contenidos que se viralizan suelen ser breves o superficiales. Figuras influyentes, movidas por una obstinada preocupación narcisista, política o financiera por obtener “likes” y seguidores, recurren a tácticas retóricas que rechazan la complejidad, la relatividad y la dialéctica. Esto provoca la fragmentación del debate público, con información descontextualizada circulando sin un desarrollo discursivo que contemple todos los matices de una determinada cuestión polémica.

En este mismo escenario, personas malintencionadas utilizan la libertad de expresión como pretexto para difundir información falsa y engañosa. Sus propósitos pueden ir desde ofensas contra individuos hasta ataques dirigidos a instituciones públicas y procesos electorales. En un nivel superior, algunos

actores privados y gubernamentales organizan operaciones con cuentas y perfiles falsos para manipular el debate público en las plataformas digitales –lo que viene siendo tratado bajo las etiquetas de “operaciones de influencia” o “comportamiento coordinado no auténtico” (Gleicher et al. 2021, 3-4, 20-22; Franklin y Torrey 2024).

La proliferación de la desinformación en línea es el fenómeno que representa la crisis de la verdad en el entorno digital. Debido a sus posibles impactos negativos, existe un consenso razonable sobre la necesidad de establecer e implementar políticas para combatirla. Sin embargo, en muchas ocasiones, las estrategias desarrolladas con este propósito han aprovechado un cierto pánico moral ante las noticias falsas o fraudulentas para poner en práctica medidas censoras en asuntos de interés público.

En este artículo abordo los desafíos conceptuales, jurídicos y políticos relacionados con el enfrentamiento de la desinformación en el espacio digital. En el apartado 2, presento investigaciones y ejemplos concretos que demuestran el impacto de los contenidos falsos o engañosos en la vida online y offline, generando una percepción subjetiva negativa en la opinión pública, los medios de comunicación y el ámbito político. En el apartado 3, discuto la polisemia del concepto de desinformación y la dificultad en delimitarlo de manera clara, precisa y consistente; a la luz del debate sobre la confiabilidad de las urnas electrónicas en el contexto brasileño, problematizo los riesgos de censura y inseguridad jurídica en la imposición de límites a la libertad de expresión basados en la categorización de la desinformación como ilícito. En el apartado 4, destaco un conjunto de investigaciones empíricas que cuestionan la existencia o validez de evidencias sobre proposiciones como la de que las falsedades generan más atención o circulan más rápido que la verdad. En el apartado 5, presento diferentes estrategias adoptadas en el escenario europeo y brasileño para disminuir la circulación y la interacción con la desinformación en línea; al final, rechazo aquellas medidas desproporcionalmente restrictivas de la libre manifestación de opiniones e ideas en el debate público, en especial la categorización de la desinformación o formulaciones análogas como “noticias falsas” (*“fake news”*) como ilícito.

II. DE LAS “*FAKE NEWS*” EN LAS ELECCIONES BRASILEÑAS A LAS INVESTIGACIONES EMPÍRICAS SOBRE LOS PROBLEMAS DE LA DESINFORMACIÓN EN LA ERA DIGITAL: UN ESCENARIO PREOCUPANTE

En septiembre de 2021, camioneros bolsonaristas realizaron manifestaciones y bloqueos en territorio brasileño. Varios videos fueron grabados por los manifestantes, en ese momento, difundiendo de manera emocionada la

(falsa) noticia de que el presidente Bolsonaro habría decretado el estado de sitio¹.

Aproximadamente un año después, en las elecciones brasileñas de 2022, los partidarios del entonces presidente Jair Bolsonaro salieron a las calles para protestar contra el resultado de la segunda vuelta, en la que el candidato Luiz Inácio Lula da Silva resultó vencedor por un margen de votos muy ajustado. Durante las protestas y concentraciones, los manifestantes difundieron numerosas noticias fraudulentas sobre el proceso electoral y la posibilidad de intervención militar en el gobierno.

En Porto Alegre (RS), un manifestante leyó ante una multitud de bolsonaristas un texto que estaba siendo compartido en grupos de aplicaciones de mensajería. Según él, el ministro Alexandre de Moraes, del Supremo Tribunal Federal, habría tenido su arresto decretado². A pesar de la extrema fragilidad del relato, que era evidentemente falso, la reacción de los manifestantes fue unívoca e inmediata: gritaron como una hinchada en un estadio de fútbol celebrando un gol de su equipo, levantando las manos al cielo, saltando, abrazándose y aplaudiendo. Uno de los videos que se viralizó en las redes sociales mostraba a una mujer de rodillas, con las manos al cielo, gritando “¡Brasil es nuestro!” —mientras que otros manifestantes, justo detrás, agitaban una gran bandera de Brasil³.

Otra situación inusitada ocurrió frente a una dependencia del Ejército en São Paulo (SP). Uno de los organizadores transmitió a la multitud, por altavoz, un titular falso según el cual habría habido un error grave en el resultado de las elecciones, lo que podría favorecer el candidato Bolsonaro. Tal como en los casos anteriormente citados, los manifestantes presentes inmediatamente estallaron en celebración gritando, saltando, agitando banderas y tocando bocinas⁴. No hubo ninguna preocupación en averiguar cuál sería el supuesto error grave o cuál la credibilidad de la fuente de la noticia. Después,

¹ «Bolsonaristas comemoram falsa informação de fraude eleitoral». *Estado de Minas*, 1 de noviembre de 2022. Acceso el 10 de abril de 2025. https://www.em.com.br/app/noticia/politica/2022/11/01/interna_politica,1415626/bolsonaristas-comemoram-falsa-informacao-de-fraude-eleitoral-veja-video.shtml.

² El ministro Alexandre de Moraes, que en ese momento también presidía el Tribunal Superior Electoral, fue responsabilizado por los bolsonaristas por la derrota de su candidato en las elecciones por diversos motivos, entre ellos la forma en que actuó en la lucha contra la desinformación durante el período electoral.

³ «Era fake! Bolsonaro festeja prisão de Moraes». *Dol – Diário Online*, 1 de noviembre de 2022. Acceso el 30 de abril de 2025. <https://dol.com.br/noticias/politica/779179/era-fake-bolsonaristas-festejam-prisao-de-moraes?d=1>.

⁴ «Cenas bizarras em manifestações golpistas viralizam na web». *Dol – Diário Online*, 3 de noviembre de 2022. Acceso el 30 de abril de 2025. <https://dol.com.br/tuedoide/viral/779456/cenas-bizarras-em-manifestacoes-golpistas-viralizam-na-web?d=1>.

vehículos de prensa informaron que el mensaje leído en la camioneta con altavoces contenía un enlace, el cual direccionaba a la imagen de un humorista con la frase “es mentira, tonto”⁵.

Esos ejemplos caricaturescos, entre otros similares que involucran al espectro político opuesto⁶, ilustran cómo noticias fraudulentas pueden ser explotadas para influenciar a las personas y generar comportamiento de manada, incluso en casos en que los mensajes compartidos son notoriamente falsificados.

La desinformación ha sido señalada como una de las principales disfunciones del entorno digital. Son diversos los estudios y reportajes que la responsabilizan por el recrudecimiento de la polarización política, con la consecuente generación de riesgos para los regímenes democráticos y a la integridad del proceso electoral (Allcott y Gentzkow 2017; Suciu 2022). Además, la desinformación está relacionada con la proliferación de discursos de odio y al extremismo online (Efsas, 2021); al mal uso de remedios y otros agentes químicos durante la pandemia de la Covid-19, además de la reticencia en relación con la política de vacunación (Loomba et al. 2021; Enders et al. 2022; European Commission 2020, 4); a la volatilidad del mercado financiero y a pérdidas causadas a los inversores y sus negocios por depreciaciones en valores de acciones (Rapoza 2017).

La percepción subjetiva de la desinformación en la sociedad ratifica los estudios teóricos y empíricos sobre el asunto. En una investigación realizada por el Instituto Reuters en 2024, por ejemplo, se identificó una proporción del 59%, entre las personas entrevistadas en 47 países, que se preocupan por la desinformación en el ambiente digital. Las preocupaciones están menos motivadas por noticias completamente falsas que por opiniones o agendas con las que los entrevistados puedan discrepar. Los entrevistados relataron haber visualizado información falsa o engañosa en temas como política (36%), coronavirus (30%), economía y costo de vida (28%), conflicto entre Israel y Palestina (27%), guerra en Ucrania (24%), cambio climático (23%), inmigración (21%) y otras cuestiones de salud (18%). La proporción de usuarios que alegó dificultad para identificar noticias confiables varió entre plataformas y países, alcanzando el 27% para *TikTok* y el 24% para *X* (anteriormente

⁵ «Bolsonaristas comemoram informação falsa de que TSE identificou erro grave nas eleições». *O Povo*, 3 de noviembre de 2022. Acceso el 10 de abril de 2025. <https://www.opovo.com.br/eleicoes-2022/2022/11/03/bolsonaristas-comemoram-informacao-falsa-de-que-tse-identificou-erro-grave-nas-eleicoes.html>.

⁶ «Influenciador petista repete “gabinete do ódio”, gera 1 bilhão de views e fatura com desinformação». *Estadão*, 11 de junio de 2024. Acceso el 11 de junio de 2024. <https://www.estadao.com.br/politica/influenciador-petista-repete-gabinete-do-odio-gera-1-bilhao-de-views-e-fatura-com-desinformacao/>.

Twitter). A pesar de las preocupaciones externadas, cerca de la mitad de los encuestados que usan cada red afirmaron considerar fácil la distinción entre noticias confiables y falsas –percepción que, así como las demás registradas en la investigación, no necesariamente corresponde a la realidad (Newman et al. 2024, 17).

Este conjunto de ejemplos, evidencias empíricas y estudios de correlación y percepción subjetiva crea un escenario extremadamente favorable al avance de políticas represivas de la desinformación en el mundo online. Sin embargo, es necesario tener cautela tanto en el diagnóstico como en la dosis del remedio, empezando por la comprensión de los riesgos derivados de la vaguedad del término desinformación.

III. LAS DIFICULTADES CONCEPTUALES QUE RODEAN A LA DESINFORMACIÓN

El concepto de desinformación es esencialmente polisémico. Sus múltiples significados dependen de definiciones previas sobre dos elementos clave: uno de naturaleza objetiva, que abarca la delimitación de lo que debe entenderse como información falsa o engañosa, y otro de naturaleza subjetiva, que envuelve la exigencia o no de intencionalidad en la difusión de dicha información.

En inglés, tres términos buscan delimitar el fenómeno bajo diferentes acepciones técnicas: *misinformation*, que describe la divulgación no intencional de información falsa; *disinformation*, que indica la diseminación de dicha información con el intento de causar daños o engañar; *malinformation*, que corresponde a la información con origen en la verdad, pero propagada de forma deliberadamente exagerada o retirada de contexto, a fin de causar daño (Wardle y Derakhshan 2017, 20). Hay aún la expresión *fake news*, que normalmente es utilizada con menor rigor técnico para describir la fabricación y divulgación de noticias falsas con o sin intención, notablemente en el campo político –en el cual, debo añadir, comúnmente es capturada por políticos y sus apoyadores como una maniobra retórica, con la finalidad de retirar la credibilidad de materias periodísticas que les sean desfavorables.

Gran parte de las investigaciones empíricas aborda el problema de la desinformación desde una perspectiva más amplia (*misinformation*), de carácter predominantemente objetivo. En esta línea de razonamiento, la desinformación puede entenderse como “la información falsa o engañosa que es contradicha por evidencia empírica o que resulta inconsistente con el consenso común o de expertos” (Chuai et al. 2024, 1). Se trata de una opción metodológica que se justifica, por un lado, por la dificultad de evaluar la intención de cada usuario respecto al contenido falso o engañoso que se publica o comparte en las plataformas digitales. Por otro lado, esta opción impone

limitaciones claras a dichas investigaciones, en la medida en que dificulta discernir la magnitud del problema desde una perspectiva cognitiva y qué estrategias deben adoptarse para enfrentarlo.

Si bien este enfoque metodológico puede ser aceptable en el ámbito de la investigación científica, su aplicación resulta inaplicable cuando se trata de definir qué tipo de noticia u opinión informativa puede ver restringida o moderada su difusión por ser considerada desinformación. En un Estado de Derecho, los ciudadanos deben contar con seguridad jurídica respecto a lo que pueden expresar en el ejercicio de su derecho a la libertad de expresión, lo que excluye la posibilidad de prohibiciones amplias, vagas o inconsistentes.

El concepto de desinformación, en cualquiera de sus formas (*misinformation*, *disinformation* y *malinformation*), plantea dificultades como categoría limitadora de la libertad de expresión por dos motivos básicos.

Primeramente, como ya mencioné, evaluar la intención en la divulgación de una noticia o información falsa es una tarea extremadamente ardua. Las personas pueden diseminar falsedades sin considerarlas como tales, al mismo tiempo en que pueden ser deliberadamente ambiguas para borrar la línea distintiva entre un error legítimo y una mentira intolerable. Fuera de eso, incluso en casos en los cuales existen evidencias que apuntan de manera consistente a la falsedad de una información, es común que las personas la consideren verdadera. Se trata del fenómeno conocido como “sesgo de verdad” (Pantazi, Klein y Kissine 2020).

A ese conjunto de complejas peculiaridades que giran en torno al elemento subjetivo del concepto de desinformación se añade el segundo problema. Su elemento objetivo (falsedad o veracidad de una información) se encuentra frecuentemente en disputa epistémica, de tal modo que “no podemos presumir que las falsedades señaladas como desinformación sean, en realidad, falsas” (Eduardo y Goldstein 2024).

Explico mejor. Una información en disputa epistémica es un dato, conjunto de datos o conocimiento cuya veracidad o validez se debate debido a diferentes interpretaciones, a la falta de consenso entre especialistas o a la existencia de evidencias conflictivas al respecto –además de la inevitable influencia de aspectos políticos, sociales, culturales o ideológicos. Por ejemplo, cuestiones como los impactos reales de las noticias falsas en las elecciones, la eficacia de los sistemas de verificación de hechos y el origen de la Covid-19 pueden ser consideradas bajo disputa epistémica porque diferentes estudiosos, evidencias e interpretaciones compiten entre sí.

Los casos polémicos que involucran la restricción o moderación de contenido etiquetado como desinformación tienden a ocurrir precisamente en la zona gris de una información en disputa epistémica. Las declaraciones sobre hechos cuya constatación se encuentra fuera de toda duda razonable (ej.: “el sol es caliente”) o que coinciden con la opinión predominante en una época

(ej.: “las democracias son mejores que las autocracias”) no presentan ninguna dificultad. Pero, si un individuo afirma que “las urnas electrónicas son confiables” y otro afirma lo contrario, ambos están emitiendo opiniones sobre un hecho objeto de disputa epistémica. ¿De qué lado está la verdad?

El análisis de los diversos matices que involucran el argumento de una mayor confiabilidad del sistema de votación electrónica en detrimento de los sistemas manuales con papeletas excede los límites de este artículo. En el caso brasileño, por ejemplo, se adoptan múltiples capas de seguridad y transparencia que minimizan no solo los riesgos de fraudes electrónicos, accesos remotos no autorizados o ataques de hackers, sino también aquellos fraudes que eran frecuentes en las votaciones manuales, como la alteración de boletas y los recuentos incorrectos. Por estas razones, particularmente considero que el sistema electrónico es más robusto y seguro que cualquier otro que aún dependa del voto en papel, especialmente en el contexto brasileño, históricamente marcado por una alta incidencia de irregularidades en este tipo de modalidad.

Sin embargo, y aquí está el punto relevante, una mayor confiabilidad de las urnas electrónicas no es sinónimo de infalibilidad. Varios países aún prefieren sistemas de votación manual en lugar de urnas electrónicas por razones relacionadas con la seguridad, la transparencia, la confianza pública, los costos y el contexto político-cultural. En Alemania, por ejemplo, el Tribunal Constitucional decidió en 2009 que las urnas electrónicas, en la forma en que se utilizaban, no cumplían plenamente con el principio de transparencia electoral, según lo establecido en el artículo 38 de la Ley Fundamental. La decisión no prohibió totalmente la votación electrónica, pero estableció condiciones estrictas para su implementación –consistentes, en síntesis, en la posibilidad de una verificación confiable y comprensible, incluso sin conocimientos técnicos especializados, de las etapas de votación y del escrutinio de los resultados (Fitzpatrick y Jöst 2022).

Retomando el caso brasileño, es válido recordar la importancia de la dialéctica en situaciones bajo disputa epistémica. Incluso las opiniones populares que se revisten de una presunción de infalibilidad rara vez corresponden a toda la verdad. El choque entre dos posturas antagónicas permite que verdades parciales puedan ser perfeccionadas, ya sea por la incorporación de fragmentos que habían sido ignorados, ya sea por el desarrollo de argumentos más contundentes a su favor. Por todo ello es que Stuart Mill ya decía, hace más de dos siglos, que no existe un juez infalible de opiniones para decidir si una opinión es nociva o falsa; la opinión debe tener “total oportunidad para defenderse” (Mill 2017, 98).

Los cuestionamientos a las urnas electrónicas en el contexto de las elecciones brasileñas, especialmente durante el gobierno Bolsonaro (Ruediger 2022), llevaron a una mejora de la transparencia de los órganos electorales en

cuanto a las medidas de seguridad adoptadas desde la fabricación de esos equipos hasta el conteo de votos (TRE-GO, 2024). Al mismo tiempo, informaciones claramente falsas sobre su funcionamiento fueron sistemáticamente contradichas, tanto por el Tribunal Superior Electoral como por agencias de verificación de hechos, lo que es una estrategia válida y no censora de combate a la desinformación. Incluso con toda la movilización del bolsonarismo en torno al ideario de fraude en las urnas electrónicas, la propuesta de enmienda a la constitución para instituir el voto impreso auditable (PEC n.º 135/2019) fue rechazada en agosto de 2021.

Teniendo todo esto em mente, indago: ¿las críticas a la falta de seguridad de las urnas electrónicas deben siempre ser catalogadas como “*fake news*” o desinformación? ¿Cómo definir qué argumentos sobre el sistema de votación electrónica deben ser considerados intencionalmente falsos o equivocadamente imprecisos? ¿Están legitimadas las autoridades públicas para enfrentar la desinformación electoral –o formulaciones análogas como “noticias falsas” o divulgación de “hechos sabidamente inverídicos”– vía categorización como ilícito y eventual sanción civil o criminal? ¿Suprimir la posibilidad de escrutinio, aunque erróneo, en un tema de interés público no implica una restricción arbitraria de la libertad de expresión?

Estas cuestiones llevan a la discusión sobre los diferentes métodos que deben ser utilizados en el combate a la desinformación –y cómo ellos pueden eventualmente convertirse en mecanismos de censura. Antes de abordarlos más de cerca, no obstante, debo presentar algunos contrapuntos señalados por ciertas investigaciones empíricas frente a la percepción alarmista que se ha consolidado sobre los efectos negativos de la desinformación en el entorno digital.

IV. LA DESINFORMACIÓN SOBRE LA DESINFORMACIÓN

Al abordar lo que denomina “infocracia”, Byung-Chul Han afirma de manera inequívoca que “las *fake news*, noticias falsas, generan más atención que los hechos. Un solo tuit que contenga *fake news* o fragmentos de información descontextualizados es posiblemente más efectivo que un argumento fundamentado”. Así, concluye que la actuación de ejércitos de *trolls*, *bots* y cuentas falsas que difunden *fake news* y teorías conspirativas durante campañas electorales expone a los votantes a una influencia indebida y pone en peligro la democracia (Han 2022, 13, 42-43).

Este tipo de narrativa es respaldado, en cierta medida, por estudios que destacan la desinformación como una de las principales disfunciones del entorno digital. Como mencioné en el apartado 2 de este artículo, la difusión de contenidos falsos o engañosos se ha vinculado al agravamiento de la polarización política, la manipulación electoral, la proliferación de discursos de

odio y la reticencia a la vacunación durante la pandemia de la Covid-19, entre otros fenómenos.

La percepción negativa sobre la desinformación online se acentúa fuera del ámbito académico. La opinión pública, probablemente influenciada por el tono alarmista con el que el tema ha sido abordado en los medios de comunicación y en el sector político, ha manifestado preocupaciones relevantes al respecto. En una investigación realizada en Estados Unidos en 2019, por ejemplo, más norteamericanos dijeron que las noticias inventadas son un gran problema que aquellos que afirmaron lo mismo en relación con asuntos como crímenes violentos, cambio climático, racismo, inmigración ilegal, terrorismo y sexismo (Pew Research Center 2019). En Brasil, de manera similar, el 81% de los entrevistados en una investigación realizada en 2024 alegaron que la diseminación de noticias falsas puede impactar “mucho” los resultados electorales; además, el 78% evaluó como “muy importante” el control de noticias falsas en las redes sociales para garantizar una disputa justa en las elecciones (Instituto de Pesquisa Datasenado 2024, 16-19).

En sentido opuesto a esa visión pesimista, varias investigaciones cuestionan la existencia de evidencias empíricas que sustenten el “pánico moral” con la propagación de informaciones falsas o engañosas en el ambiente online. Ellas presentan constataciones en el sentido de que las noticias no confiables (falsas, de baja calidad, partidistas) representan una porción pequeña de las informaciones consumidas por las personas (Allen et al. 2020; Altay, Nielsen y Fletcher 2022); o que la mayor parte de las personas no comparte noticias no confiables (Grinberg et al. 2019; Guess, Nagler y Tucker 2019), considerándolas, en promedio, menos plausibles que las noticias verdaderas (Acerbi, Altay y Mercier 2022, 4); o también que las redes sociales no son las únicas culpables por la desinformación (Benkler, Faris y Roberts 2018, 37-38; Bennet y Livingston 2021, 3-40); o, finalmente, que la influencia de noticias falsas en comportamientos políticos es menor de lo que se imagina, excepto por el aumento de la creencia en afirmaciones falsas (Guess et al. 2020).

En el estudio que inspira el título de este apartado (“*Misinformation on misinformation: conceptual and methodological challenges*”), un trío de investigadores identifica dos conjuntos de equívocos comunes sobre la desinformación en el debate académico. El primero de ellos abarca las siguientes afirmaciones sobre la prevalencia y la circulación de la desinformación: i) se trata de un problema exclusivo de las redes sociales; ii) la internet está repleta de desinformación; iii) las falsedades circulan más rápido que la verdad. El segundo conjunto de equívocos comprende concepciones sobre el impacto y la recepción de la desinformación: iv) la gente cree todo lo que ve en internet; v) muchas personas están desinformadas; vi) la desinformación ejerce una

gran influencia en el comportamiento de las personas (Altay, Berriche y Acerbi 2023).

Varios de estos puntos se correlacionan con el uso de grandes volúmenes de datos digitales (*big data*) como metodología para evaluar la desinformación. Este enfoque tiende a sobreestimar la circulación, la prevalencia, el impacto y la recepción de la desinformación en el mundo online por diversos motivos. El hecho de que la internet permita la medición de datos sobre la difusión de noticias falsas o engañosas no implica que el fenómeno no existiera o haya dejado de existir en el universo offline. En la internet, los números siempre son superlativos, lo que hace necesario analizarlos de forma contextualizada dentro de todo el ecosistema de información, incluyendo los medios tradicionales como la televisión y la radio. También deben ser examinados en perspectiva, lo que significa que miles o millones de compartidos de una noticia falsa no necesariamente indican su predominancia o prevalencia en la dieta de consumo de los usuarios de las redes sociales (Altay, Berriche y Acerbi 2023, 2-4).

En la misma línea de razonamiento, grandes volúmenes de datos sobre noticias no confiables en internet no permiten presuponer que las personas creen en todo lo que ven, o que sean en su mayoría desinformadas, o incluso que la desinformación tenga una fuerte influencia en su comportamiento. Métricas como los “me gusta” y las comparticiones no proporcionan una comprensión precisa sobre cómo los individuos interpretan e influyen en los contenidos a los que acceden. Estos aspectos subjetivos están ligados a matices contextuales como la socialización, la ironía y la expresión de escepticismo o rabia –y no necesariamente implican un respaldo, total o parcial, de aquello que se indica con un “me gusta” o se comparte (Altay, Berriche y Acerbi 2023, 5-8).

Otro problema destacado por los investigadores involucra la definición de la desinformación, lo que puede influir tanto en la escala del problema como en las soluciones que se adopten. La afirmación de que “las falsedades se difunden más rápidamente que la verdad”, por ejemplo, es empíricamente controvertida⁷. Uno de los artículos frecuentemente citados en favor de esta hipótesis no es apto para demostrarla de forma tan generalizada, ya que analizó únicamente noticias contestadas que fueron clasificadas por verificadores de hechos como falsas o verdaderas (Vosoughi, Roy y Aral 2018; Sunstein 2021, 62-63) –lo que, evidentemente, deja fuera un gran número de noticias o publicaciones virales que ni siquiera son cuestionadas. Así, la confirmación o refutación de la tesis puede estar vinculada, respectivamente, a

⁷ Por ejemplo, una investigación de 2020 señaló que la información proveniente de fuentes confiables y cuestionables, en relación con el inicio de la pandemia del coronavirus, no presentó patrones de difusión significativamente diferentes (Cinelli et al. 2020).

concepciones más amplias o estrictas de lo que debe entenderse como información y desinformación (Altay, Berriche y Acerbi 2023, 4).

Finalmente, en cuanto a la alegación de que la desinformación ejerce una fuerte influencia en el comportamiento de las personas, los investigadores señalan que los datos correlacionales no indican necesariamente causalidad. Al contrario de lo que se puede presuponer, la exposición a contenido falaz no implica un cambio comportamental; la tendencia es que las personas consuman información con la que ya concuerdan, reforzando sus creencias pre-existentes (Altay, Berriche y Acerbi 2023, 7-8).

La importancia de estos contrapuntos a los efectos negativos de la desinformación en el ambiente digital no radica en negar la existencia del problema. Su complejidad no justifica un escepticismo técnico exacerbado, pero tampoco debe ser motivo para un abordaje excesivamente alarmista. Combatir la diseminación intencional de contenidos falsos o engañosos es necesario, pero no se puede, bajo tal pretexto, legitimar políticas públicas o decisiones judiciales de carácter censorio, que comprometan la libertad de expresión y la esfera pública discursiva.

Em una perspectiva holística, la desinformación debe entenderse como parte de problemas socioeconómicos complejos (Altay, Berriche y Acerbi 2023, 9), entre ellos la mayor propensión de las personas a no estar informadas o a ser ambiguas (Li y Wagner 2020, 656). Así, en lugar de ser vista como una explicación monocausal, la desinformación es más bien uno de los síntomas de estos problemas. Con esta salvedad en mente, trataré en el siguiente tópico las principales estrategias para combatirla.

V. EL COMBATE A LA DESINFORMACIÓN

1. *Prebunking y debunking*

Las políticas y soluciones jurídicas de combate a la desinformación pueden dividirse en tres enfoques principales: *prebunking* (“inmunización informativa”), *debunking* (“desenmascaramiento informativo”) y –lo que examinaré en el siguiente tópico– la categorización como contenido ilícito.

El *prebunking* es una estrategia preventiva que busca fortalecer la resiliencia de las personas contra la desinformación antes de que la encuentren. Esto puede hacerse por diferentes métodos de inmunización informativa. Uno de ellos consiste en exponer a los individuos a versiones controladas de falsedad, enseñándoles a reconocer técnicas comunes de manipulación y falacia lógica. En el mismo sentido, pueden realizarse campañas educativas sobre teorías de la conspiración y señales típicas de noticias falsas, como titulares alarmistas y la ausencia de fuentes confiables. El camino opuesto también es posible, preparando las personas con información confiable (teoría de

la inoculación) para que posean un filtro preventivo contra la desinformación subsiguiente.

El *debunking*, por su parte, involucra el conjunto de técnicas destinadas a desmentir o refutar la desinformación después de su diseminación. En lugar de centrarse en la prevención, el *debunking* actúa de manera reactiva, presentando explicaciones y evidencia confiable frente a información falsa o engañosa. Ejemplos de esta estrategia abarcan las verificaciones de hechos por agencias profesionales o por público lego (*crowdsourced*).

Ambas las intervenciones son demostradamente efectivas para disminuir la vulnerabilidad del público objetivo a la desinformación, reduciendo las creencias en alegaciones falsas y la probabilidad de compartirlas (Bruns et al. 2024). Las intervenciones preventivas contribuyen a que los individuos desarrollen resistencia a la desinformación en un contexto más amplio. Sin embargo, no suelen ser adecuadas para neutralizar noticias falsas o engañosas que puedan rápidamente volverse virales y causar preocupaciones por su potencial dañino (Chuai et al. 2024, 2). En estos casos, los mecanismos de *debunking* son imprescindibles –aunque con diversas limitaciones en cuanto a su alcance, rapidez y efectividad, ya sea en las verificaciones implementadas por profesionales (Walter et al. 2020; Graves 2016, Part III), ya sea en las realizadas por público lego, como las notas comunitarias (Drolsbach et al. 2024; Allen et al. 2021; Saed et al. 2022; Zhao y Naaman 2023; Chuai et al. 2024).

En el ámbito de la Unión Europea, el enfrentamiento a la desinformación online involucra principalmente la imposición de obligaciones y la formulación de incentivos a la adopción de mecanismos de *debunking* y *prebunking*. Se parte, para ello, de la premisa de que es necesario conciliar los desafíos planteados por la circulación de información falsa o engañosa a gran escala –que no necesariamente involucra contenidos ilegales– con la garantía de debates públicos libres, plurales e inclusivos en las sociedades democráticas. De acuerdo con la Comisión Europea, “el contenido legal, aunque se trate de contenido supuestamente nocivo, está generalmente protegido por la libertad de expresión y debe abordarse de manera diferente a los contenidos ilegales, en cuyo caso se puede justificar la eliminación del propio contenido” (European Commission 2018, 1). Volveré a este punto específico en el siguiente tópico.

En línea con estas directrices, el Reglamento de Servicios Digitales (*Digital Services Act – DSA*), adoptado oficialmente en 2022, prevé una serie de obligaciones para las plataformas y motores de búsqueda de “muy gran dimensión” relacionadas con el combate a la desinformación (en inglés, *Very Large Online Platforms and Very Large Online Search Engines – VLOP-SEs*). En síntesis, estos servicios deben realizar evaluaciones de riesgos sistémicos involucrando la diseminación o amplificación de contenidos

engañosos, incluso si no son ilegales (artículo 34 c/c considerando n.º 84). Deben, también, implementar medidas de mitigación de estos riesgos (artículo 35) y poseer un mecanismo de respuesta en caso de crisis (artículo 36). Considerando el papel que la publicidad dirigida puede desempeñar en la diseminación de la desinformación, el reglamento limita este tipo de publicidad (artículos 26.2 y 26.3) y exige que las plataformas y motores de búsqueda de “muy gran dimensión” mantengan un repositorio público de anuncios (considerando n.º 95). Finalmente, el reglamento fomenta a las plataformas a adherirse a un código de conducta voluntario sobre desinformación, a nivel de la Unión Europea (artículo 45 y considerando n.º 104).

En su versión de octubre de 2024, el Código de Conducta sobre Desinformación de la Unión Europea, basado en diversos documentos e iniciativas europeas, presenta una serie de medidas que deben adoptar sus signatarios, especialmente las plataformas digitales y los motores de búsqueda en internet (European Union 2025).

En síntesis, la adhesión al Código implica la asunción de compromisos como: el escrutinio de anuncios con el objetivo de desmonetizar y, en consecuencia, reducir la diseminación de contenidos desinformativos en formato de publicidad o junto con ella (compromisos 1 a 3); el establecimiento de mecanismos eficaces para garantizar la transparencia de la publicidad política, etiquetándola y distinguiéndola como contenido de pago (compromisos 4 a 13); la promoción de la integridad de los servicios, implementando o reforzando políticas para abordar acciones manipulativas de agentes maliciosos, como el uso de cuentas falsas, la compra de interacciones, la amplificación dirigida por *bots*, la coordinación de comportamientos inauténticos y la creación de *deepfakes* con inteligencia artificial generativa (compromisos 14 a 16); empoderamiento de los usuarios mediante el aumento de la educación mediática (*media literacy*), con el desarrollo e implementación de herramientas y campañas que fomenten el pensamiento crítico (compromiso 17).

Otros compromisos involucran: la adopción de un “diseño seguro” de la arquitectura de servicios, de modo que los sistemas de recomendación algorítmica estén “proyectados para mejorar la prominencia de información confiable y reducir la prominencia de desinformación” (compromiso 18, medida 18.1); el desarrollo y la aplicación de políticas proporcionales para limitar la difusión de información falsa o engañosa perjudicial, como prohibir, degradar o no recomendarla, de forma adaptada a la gravedad de los impactos y con respeto a la libertad de expresión (compromiso 18, medida 18.2) –asegurando, además, el acceso a un sistema de recursos transparente y ágil (compromiso 24); la garantía de transparencia y opciones disponibles para los usuarios con respecto a los sistemas de recomendación algorítmica (compromiso 19); la facilitación del acceso a herramientas para que los usuarios puedan evaluar la autenticidad y precisión factual de la información, incluyendo

verificaciones de hechos por organizaciones independientes y otras fuentes confiables (compromisos 20 a 23); la construcción e implementación de recursos o iniciativas en aplicaciones de mensajería que capaciten a los usuarios para pensar críticamente sobre la información que reciben y determinar si es precisa, sin debilitar la criptografía y con la debida consideración a la protección de la privacidad (compromiso 25); la divulgación y el acceso a los datos y estudios de los signatarios para la capacitación de la comunidad de investigación sobre desinformación (compromisos 26 a 29); el establecimiento de una estructura de cooperación transparente, estructurada, abierta, financieramente sostenible y no discriminatoria entre los signatarios y la comunidad de verificadores de hechos (compromisos 30 a 33).

Finalmente, el Código prevé mecanismos para supervisar su implementación y monitoreo, como la creación de un centro de transparencia y de un grupo de trabajo permanente, además de la presentación de informes periódicos por los firmantes (compromisos 34 a 43).

2. *La (controvertida) tendencia de categorización como contenido ilícito*

La gran polémica en torno al tema de la desinformación se refiere a su categorización como contenido ilícito, lo que puede acarrear diferentes consecuencias, desde la previsión de eliminación de contenido hasta la sanción penal. En este apartado, trazo un panorama de las disposiciones a este respecto en el ámbito de la Unión Europea y sus Estados miembros, así como de Brasil y de las plataformas digitales; en el apartado siguiente, examino críticamente estas restricciones.

En la Unión Europea, varios Estados miembros poseen leyes internas –no necesariamente creadas para el contexto digital– que prevén la desinformación o noticias falsas como contenido ilícito, incluso en el ámbito penal. Entre los ejemplos se encuentran la Ley sobre la Prestación de Información al Público de Lituania, que en su artículo 19 prohíbe explícitamente la difusión de desinformación; el Código Penal de Malta, cuyo artículo 82 criminaliza la divulgación maliciosa de noticias falsas “que puedan alarmar a la opinión pública, perturbar el orden público o la paz pública, o crear conmoción entre el público o entre determinadas clases de la población”; y la Ley sobre la Libertad de Prensa de Francia, que en su artículo 27 prohíbe la disseminación de noticias falsas cuando, “hechas de mala fe, hayan perturbado la paz pública o sean susceptibles de perturbarla”, sujetando a los responsables a una multa de hasta 45.000 euros (Ó Fathaigh, Helberger y Appelman 2021, p. 3, 7-11; Attenborough 2024).

En algunos países de la Unión Europea, la prohibición de la desinformación se dirige a su difusión en contextos específicos, como la pandemia de la Covid-19 o los periodos electorales. Francia, por ejemplo, aprobó en 2018 una ley

destinada al combate de la manipulación de la información durante las elecciones (en francés, *LOI n.º 2018-1202 du 22 décembre 2018, relative à la lutte contre la manipulation de l'information*)⁸. Su texto permite que alegaciones o imputaciones intencionales de hechos imprecisos o engañosos sean eliminadas por orden judicial, siempre que exista el potencial de daño a una elección y su diseminación ocurra de manera artificial o automatizada y masiva.

Estas y otras disposiciones nacionales son objeto de crítica por divergir “considerablemente en términos de alcance, destinatarios y sanciones jurídicas”, con “implicaciones potencialmente amplias para los derechos fundamentales” al considerar contenido falso, engañoso o desinformativo como ilegal (Ó Fathaigh, Helberger y Appelman 2021, p. 7-11). En este sentido, según la Comisión Europea, la desinformación debe ser tratada como contenido perjudicial debido a los riesgos que representa para los procesos políticos democráticos y otros valores (educación, ciencia, educación, finanzas), pero no debe confundirse con formas de contenido ilegales, como la pornografía infantil, la difamación, la incitación a la violencia y el discurso de odio (European Commission 2020a, 3-4; European Commission 2020b). Esta es también la posición adoptada por el Grupo de Expertos de Alto Nivel de la Unión Europea sobre Noticias Falsas y Desinformación en Línea –que excluye no solo la difusión en línea de contenido ilegal del concepto de desinformación, sino también las formas de distorsión deliberada de hechos no engañosas, como la sátira y la parodia (European Commission 2018a, 10-11).

Siguiendo estos parámetros, y buscando la uniformización de las reglas aplicables a los prestadores de servicios intermediarios, el Reglamento de Servicios Digitales europeo no considera la desinformación como ilegal por sí misma; en cambio, la trata como un riesgo a ser gestionado, dado su carácter perjudicial para los individuos y la sociedad en general. Como mencioné en el apartado anterior, el reglamento impone diversas obligaciones a las plataformas digitales para lidiar con contenido desinformativo (cf. considerandos n.ºs 2, 9, 69, 83, 84, 88, 95, 104, 106 y 108, además de los artículos 34 a 36), pero rechaza la influencia del derecho de los Estados miembros en la categorización de la desinformación como contenido ilícito (art. 3º, h) –excepto cuando involucre simultáneamente otras formas de discurso prohibidas por el derecho de la Unión Europea, como difamación, discurso de odio e incitación a la violencia.

Es curioso notar, sin embargo, que el Código de Conducta sobre Desinformación de la Unión Europea establece en su medida 18.2 que sus signatarios deben desarrollar y adoptar “políticas públicamente documentadas y proporcionales para limitar la difusión de información falsa o engañosa perjudicial” –entre ellas, dependiendo del servicio, “prohibir, degradar o no

⁸ Acceso el 28 de febrero de 2025. «<https://www.legifrance.gouv.fr/jorf/id/JORF-TEXT000037847559>».

recomendar dicha información”, según la gravedad de sus impactos, además de tomar “medidas contra páginas web o actores que violen estas políticas de forma persistente” (European Union 2025, 26). Aunque la adopción del Código no sea obligatoria y él contenga un aviso legal en el sentido de que no debe considerarse representativo de la posición oficial de la Comisión Europea, su firma es estimulada por el Reglamento de Servicios Digitales (artículo 45 y considerando n.º 104).

En Brasil, existe una fuerte tendencia represiva de la desinformación en el ámbito electoral. Se considera delito la divulgación, en la propaganda electoral o durante el período de campaña electoral, de hechos que se saben inveraces en relación con partidos o candidatos y que puedan ejercer influencia ante el electorado (artículo 323 del Código Electoral – Ley n.º 4.737/65). Además, tres resoluciones del Tribunal Superior Electoral (TSE) abordan la cuestión de la desinformación y las exigencias impuestas por el órgano a las plataformas digitales: Resolución n.º 23.610/2019 (propaganda electoral), modificada por la Resolución n.º 23.732/2024; Resolución n.º 23.714/2022 (enfrentamiento a la desinformación que afecte la integridad del proceso electoral); y Resolución n.º 23.735/2024 (ilícitos electorales).

A modo de ejemplo de las diversas disposiciones incluidas en estos actos normativos, el artículo 27, § 1º, de la Resolución TSE n.º 23.610/2019 prohíbe “divulgar hechos sabidamente inveraces” en la propaganda electoral, lo que puede dar lugar a la aplicación de una multa (art. 57-D, § 2º, de la Ley n.º 9.504/97). Ya la Resolución TSE n.º 23.714/2022 prohíbe “la divulgación o compartir hechos sabidamente inveraces o gravemente descontextualizados que afecten la integridad del proceso electoral, incluidos los procesos de votación, escrutinio y totalización de votos” –autorizando al TSE a determinar a las plataformas digitales que eliminan el contenido desinformativo, bajo pena de multa, así como a suspender temporalmente perfiles, cuentas o canales mantenidos en medios sociales, en caso de producción sistemática de desinformación.

Las resoluciones del TSE, pese a su carácter general y abstracto, son justificadas por la corte como parte de su poder de policía sobre la propaganda electoral, de acuerdo con lo establecido en el art. 23, IX, del Código Electoral y en los arts. 57-J y 105 de la Ley n.º 9.504/97. Al analizar la constitucionalidad de la Resolución TSE n.º 23.714/2022, el ministro Edson Fachin, del Supremo Tribunal Federal, resaltó que dicho acto normativo formaba parte de un “arco de experimentación regulatoria en lo que respecta al enfrentamiento del complejo fenómeno de la desinformación y sus impactos electorales”. Los potenciales perjuicios de la desinformación para la integridad del proceso electoral constituyeron la justificación central de la decisión⁹:

⁹ Supremo Tribunal Federal (Brasil), *Acción Directa de Inconstitucionalidad n.º 7261*, rel. min. Edson Fachin, 6 de marzo de 2024.

“[...] es bien sabido que, en el caso de ‘fake news’ difundidas a través de internet, existe un vacío y un desfase entre el conocimiento del hecho y la eliminación de su contenido (*notice and take down*). De este modo, los perfiles falsos pueden amplificar el alcance de la desinformación, en un claro abuso de poder. Mientras el tiempo de reacción es corto, el daño potencial a la integridad del proceso electoral es inmensurable.

[...] la normalidad de las elecciones está en cuestión cuando la *libertad* se convierte en ausencia de libertad, ya que se encuentra desconectada de la realidad, la verdad y los hechos. Este ejercicio abusivo pone en riesgo a la propia sociedad libre y al Estado de Derecho democrático.

[...] La difusión de noticias falsas, en el corto período del proceso electoral, puede tener la fuerza de ocupar todo el espacio público, restringiendo la libre circulación de ideas.”

El análisis de diversos precedentes del TSE revela su actuación, mediante la imposición de multas, las órdenes de eliminación de contenido y las prohibiciones de nuevas publicaciones con el mismo tenor, frente a la desinformación en línea dirigida: a) a la integridad del proceso electoral, como manifestaciones falsas o engañosas sobre fraude electoral, adulteración de votos, violación o manipulación de las urnas electrónicas y la imposibilidad de ser auditadas; b) a integrantes de la Justicia Electoral, como alegaciones sobre su actuación sospechosa, parcial o abusiva, así como ataques injuriosos, calumniosos o difamatorios; c) a candidatos y partidos políticos, cuando sea capaz de afectar la legitimidad del proceso electoral (por ejemplo, la difusión de un video que inducía a los votantes a creer que un candidato apoyaba el aborto, a pesar de que la prensa ya había informado que la afirmación era falsa) (Lucena 2024, 28, 87-90).

Finalmente, destaco que el tema de la desinformación también es abordado por las plataformas digitales en la esfera privada. Independientemente de la existencia de una ley nacional que regule el asunto¹⁰, varias de estas plataformas prohíben contenido desinformativo con mayor o menor amplitud en sus términos de uso, lo que les permite la moderación mediante la eliminación de publicaciones o la suspensión de usuarios. Según la política adoptada por Meta, por ejemplo, la desinformación se elimina “cuando existe la posibilidad de que contribuya directamente al riesgo de lesiones corporales inminentes”, así como en los casos en que “pueda contribuir directamente a la interferencia del funcionamiento de procesos políticos” (Meta 2025).

¹⁰ El Código de Conducta sobre Desinformación de la Unión Europea, como se ha visto, enumera entre sus compromisos el que se adopten políticas por las plataformas para prohibir, degradar o no recomendar información falsa o engañosa perjudicial.

3. *Separando el grano de la paja: qué es censurable y qué es legítimo*

Las restricciones sobre la desinformación, de orden público o privado, criminales o no, ¿pueden ser consideradas constitucionalmente compatibles con la libertad de expresión? ¿Existen contextos, como el periodo electoral, o situaciones claras de abuso de derecho, como la producción sistemática y masiva de contenido desinformativo, que puedan legitimar limitaciones al derecho de manifestarse libremente? ¿La solución de la Unión Europea, consistente en imponer obligaciones a las plataformas digitales para lidiar con contenido desinformativo sin considerarlo ilícito, puede ser considerada apta para los objetivos deseados y proporcionalmente menos intrusiva a la libertad de expresión?

El primer desafío que se opone a la categorización de la desinformación como contenido ilícito consiste en la dificultad de establecer legalmente, de manera clara y precisa, lo que debe entenderse como falso o engañoso. Como traté en el apartado 3 de este artículo, los hechos y la información están frecuentemente sujetos a disputa epistémica o a revisión con el paso del tiempo, lo que dificulta su clasificación como falsos o verdaderos. Además, incluso si un contenido puede ser considerado desinformativo más allá de toda duda razonable, la verificación de la intención en su divulgación jamás puede presumirse, especialmente para fines penales.

Estas dificultades, que ya son significativamente impeditivas de la prohibición legal de la desinformación, se suman a otro desafío: si noticias e informaciones falsas o engañosas no están protegidas, ¿quién debe controlarlas, arbitrando lo que debe entenderse como verdad?

Las respuestas fáciles oscilan entre la atribución de tales poderes al gobierno, al Poder Judicial, a entidades supranacionales o a las grandes corporaciones tecnológicas propietarias de las plataformas digitales. Sucede que estas opciones son incompatibles con el paradigma del Estado de Derecho, que es un principio fundamental de gobernanza de las democracias liberales.

El reconocimiento de la libertad de expresión como derecho fundamental en los regímenes democráticos fue el resultado de un largo proceso histórico de oposición a la censura y a la intolerancia que eran practicadas por autoridades públicas y religiosas. La legitimidad de los poderes censorios de esas autoridades derivaba de la reivindicación tiránica que hacían, para sí, del monopolio de la verdad. En el Estado de Derecho, diversamente, el carácter relativo de la verdad es uno de los factores de limitación del poder, ya sea en la esfera pública, ya sea en la esfera de las plataformas digitales —que, aunque privadas, funcionan como foros híbridos, de naturaleza público-privada¹¹.

¹¹ Plataformas como *X* y *Facebook*, por su diseño abierto a la interacción y expansión permanente, integrando en sí diferentes foros (públicos y privados), tienen una naturaleza híbrida. Aunque les corresponde moderar contenido con base en sus propias políticas in-

Como consecuencia tanto de la eficacia vertical como horizontal de los derechos fundamentales¹², los individuos poseen la libertad de manifestarse en busca de la verdad en temas de interés público y contrariamente a cualquier intento de su captura por órganos públicos o privados.

Estos dos argumentos (vaguedad de la desinformación e imposibilidad de control o captura de la verdad) se suman a un tercero, de igual o mayor importancia: la libertad de expresión, como derecho o libertad pública fundamental, ocupa una posición preferente cuando entra en conflicto con otros valores. En el derecho estadounidense, donde se concibió tal doctrina, su formulación no implica una prohibición absoluta de restricciones a la libertad de expresión (Linera y Teruel Lozano 2017, 21-22); significa, más bien, el condicionamiento de eventuales limitaciones a la demostración de un interés estatal imperioso (*compelling interest*), que sea capaz de prevalecer en la situación conflictiva. Más específicamente, la presencia de un interés estatal imperioso (*compelling interest*) es necesaria para que las restricciones legales basadas en el contenido (*content-based*) puedan superar la prueba de razonabilidad en lo que se denomina estándar de escrutinio estricto. Para las restricciones que no afectan el contenido de las manifestaciones, tales como las de tiempo, lugar y modo (*time, place and manner*), el nivel de escrutinio es intermedio (Terra, 2016, 168-181).

La doctrina de la posición preferente, aunque originada en Estados Unidos, influye fuertemente en la jurisprudencia de las cortes internacionales de derechos humanos y de varios tribunales constitucionales. El Tribunal Europeo de Derechos Humanos reiteradamente afirma la libertad de expresión como “uno de los pilares fundamentales de una sociedad democrática”¹³, razón por la cual es menor el margen para restricciones que involucran manifestaciones sobre temas de interés público o político¹⁴. En ese sentido, en el famoso caso *Handyside c. Reino Unido* (1976), el tribunal estableció el

ternas (“términos de uso”), su peculiar naturaleza y la eficacia de los derechos fundamentales en la relación entre particulares imponen límites a esas políticas, con el fin de proteger el pluralismo democrático y las libertades de expresión e información (Vázquez Alonso 2021, 141-165; Vázquez Alonso 2022, 115-118).

¹² Aunque no es aceptada en el derecho estadounidense, la doctrina de la eficacia de los derechos fundamentales en la relación entre particulares se ha consolidado en Europa y en varios países, entre los cuales Brasil (Barroso, 2011, 394-395).

¹³ *Caso Fressoz y Roire c. Francia*, demanda n.º 29183/95, sentencia de 21 de enero de 1999, § 45. En el mismo sentido, *Caso Observer y Guardian c. Reino Unido*, demanda n.º 13585/88, sentencia de 26 de noviembre de 1991, § 59.

¹⁴ *Caso Sunday Times c. Reino Unido*, demanda n.º 6538/74, sentencia de 29 de abril de 1979, § 65; *Caso of Incal c. Turquía*, n.º 41/1997/825/1031, sentencia de 9 de junio de 1998, § 46; *Caso Lingens c. Austria*, demanda n.º 9815/82, sentencia de 8 de julio de 1986, §§ 41-43.

entendimiento de que la libertad de expresión se aplica no solo a información e ideas que son bien recibidas o inofensivas, sino también a aquellas que “ofenden, chocan o perturban”¹⁵.

Aplicando estas directrices al combate a la desinformación en línea, la Comisión Europea subraya que los mecanismos diseñados con este propósito “deben respetar rigurosamente la libertad de expresión e incluir salvaguardias que impidan su uso abusivo, como la censura de discursos críticos, satíricos, disidentes o impactantes” (European Commission 2018b, 8). Esta misma orientación se reitera en el Código de Conducta sobre Desinformación de la Unión Europea, al enfatizar que las acciones adoptadas por las plataformas digitales con respecto a los contenidos señalados como perjudiciales, falsos o engañosos deben respetar la libertad de expresión, el derecho a la privacidad y la protección de los datos personales (European Union 2025, 24).

Considerando la vaguedad conceptual de la desinformación, la imposibilidad de controlar o capturar la verdad en un Estado de derecho y la posición preferente de la libertad de expresión en los regímenes democráticos, entiendo que la categorización de contenido desinformativo como ilícito no es constitucionalmente válida, salvo en los siguientes casos: a) cuando el contenido involucre otras formas de discurso prohibido (incitación, difamación, amenazas, etc.); b) cuando sea el resultado de estrategias organizadas con el objetivo de manipular artificialmente el debate público, como el uso de perfiles falsos o automatizados en operaciones de influencia.

La primera excepción no impone una restricción sobre la desinformación propiamente dicha, sino que simplemente reconoce la extensión de contenidos ya prohibidos a aquellas publicaciones o noticias que también sean señaladas como falsas o engañosas.

La segunda excepción, por su parte, tampoco apunta directamente a una restricción sobre el contenido desinformativo, sino más bien al modo (perfiles falsos o automatizados) y al lugar (entorno en línea) por el cual se produce. En las operaciones de influencia y en los comportamientos coordinados inauténticos, como observé, no hay debate público entre usuarios reales. Tampoco se utiliza la libertad de expresión con el propósito de búsqueda de la verdad o en nombre de la autodeterminación de la persona humana. Los objetivos son la manipulación del mercado de ideas para avanzar agendas bien definidas, como las de atacar instituciones e individuos. Por lo tanto, las restricciones a este tipo de conducta no se enfocan en el contenido propiamente dicho, sino en un comportamiento ilícito, con el fin de preservar la integridad del debate público¹⁶. Esto permite el uso de un criterio objetivo y neutral para la eliminación de

¹⁵ *Caso Handyside c. Reino Unido*, demanda n.º 5493/72, sentencia de 7 de diciembre de 1976, §§ 49.

¹⁶ En el mismo sentido, Vázquez Alonso (2022, 119).

publicaciones y la suspensión o el bloqueo de perfiles, consistente en la comprobación del uso coordinado de cuentas falsas.

Salvadas estas excepciones —es decir, en casos en los que la verdad de una afirmación esté en debate entre individuos reales en el mercado de ideas, tanto online como offline, sin involucrar otras formas de expresiones prohibidas—, el contradiscurso constituye, en regla, una solución preferible a la prohibición o restricción del contenido considerado falso o engañoso.

Las estrategias de *prebunking* y *debunking*, siempre que no impliquen la retirada de contenido o la suspensión de perfiles únicamente por la difusión de desinformación, se consideran modalidades de contradiscurso constitucionales. En esos términos, incluso, no existe obstáculo para que el Estado, dentro de un marco regulatorio del entorno digital como el propuesto por el Reglamento de Servicios Digitales europeo, imponga obligaciones a las plataformas para que adopten estas estrategias y presenten informes periódicos al respecto; lo que no es posible es la censura estatal indirecta, mediante la amenaza de sanciones a las plataformas que no actúen activamente en la supresión de contenido desinformativo.

Una situación más delicada se presenta cuando la ley exige limitar los mecanismos de recomendación algorítmica y el impulso financiero de publicaciones señaladas como falsas o engañosas por sistemas de verificación de hechos. Aunque una disposición de este tipo implique cierto grado de restricción de contenido, en la medida en que afecta su circulación, debe considerarse constitucionalmente válida. La acción combinada de algoritmos y los mecanismos de impulso hace que determinadas ideas y opiniones se difundan artificialmente, sin una correlación necesaria con su sofisticación o calidad. En este contexto, la ausencia de condiciones propias del libre mercado de ideas compromete la efectividad del contradiscurso como medida típicamente menos intrusiva para contener los efectos negativos de la desinformación.

La degradación del posicionamiento de contenido (*downranking*) ciertamente afecta la libertad de promover el alcance de una publicación, pero, como restricción a la libertad de expresión, constituye una medida menos gravosa que la eliminación. En un mercado donde las ideas no compiten en condiciones equitativas, sino a través de recursos monetarios y algorítmicos aplicados de forma desigual, resulta ingenuo sostener que los contenidos desinformativos promovidos de manera dolosa puedan ser combatidos únicamente mediante el contradiscurso. Así, cuando un contenido es señalado como falso o engañoso por sistemas de verificación de hechos, la reducción de posicionamiento puede considerarse proporcional si se observan ciertos criterios: claridad y previsibilidad de las normas de la plataforma (transparencia); notificación al usuario sobre la limitación del alcance de la publicación, con una justificación adecuada que permita evitar una discrecionalidad excesiva; posibilidad de apelación y revisión del acto; y consideración del

contexto en el que se publicó el contenido, a fin de diferenciar entre un error fáctico, una opinión o una desinformación intencional.

VI. CONCLUSIONES

A lo largo de este artículo, he explorado las complejidades inherentes al tema de la lucha contra la desinformación online en relación con la libertad de expresión. La tenue línea entre la legítima búsqueda de la verdad en el debate público y la imposición de medidas censorias exige cautela, tanto en el diagnóstico como en la implementación de políticas para enfrentar la difusión de contenidos falsos y engañosos.

Aunque investigaciones, noticias y una serie de acontecimientos potencian una percepción subjetiva preponderantemente negativa con respecto a las falsedades en el entorno digital, las evidencias empíricas sobre su prevalencia, circulación, impacto y recepción son conflictivas. Además, problemas socioeconómicos complejos, como la baja educación mediática de parte de la población, son más abarcadores que el de la desinformación propiamente dicha.

En diversas legislaciones nacionales, el asunto se ha tratado con cierto alarmismo, lo que ha llevado a la categorización de la desinformación en línea como contenido ilícito, sujeto a sanción en diversas esferas, incluso la penal. Esta solución es desproporcionadamente intrusiva para la libertad de expresión, generando el riesgo de enfriamiento del debate público libre, plural e inclusivo, algo fundamental para las sociedades democráticas. Tres motivos concurren para tal conclusión: la vaguedad conceptual de la desinformación; la imposibilidad de control o captura de la verdad por órganos públicos o privados; la posición preferencial de la libertad de expresión, especialmente en asuntos de interés público o político.

Por otro lado, son admisibles dos excepciones: contenidos desinformativos que involucren otras formas de discursos prohibidos (incitación, difamación, amenazas, etc.) y expedientes organizados que tengan como objetivo la manipulación artificial del debate público en el entorno digital (uso de perfiles falsos o automatizados en operaciones de influencia, compra de interacciones, amplificación dirigida por bots, etc.). En ambos casos, las restricciones no se dirigen directamente al contenido por su carácter desinformativo, sino, respectivamente, a otras categorías ya prohibidas o a acciones manipulativas de agentes maliciosos contra la integridad del mercado de ideas.

Fuera de estos casos, la lucha contra la desinformación online debe llevarse a cabo principalmente mediante estrategias de *prebunking* y *debunking*, ya que constituyen modalidades de contradiscurso constitucionales que no limitan el derecho a la libre expresión. Las previsiones del Reglamento de Servicios Digitales de la Unión Europea se ajustan a estos parámetros, lo que

lo califica, en este aspecto específico, como un modelo de marco regulatorio de los servicios prestados por intermediarios de internet que debe ser replicado fuera del contexto europeo.

VII. REFERENCIAS BIBLIOGRÁFICAS

- Acerbi, A.; Altay, S.; Mercier, H. (2022). «Research note: Fighting misinformation or fighting for information?», *Harvard Kennedy School Misinformation Review*, v. 3 (1): 1-15.
- Allcott, Hunt y Gentzkow, Matthew (2017). «Social media and fake news in the 2016 election», *Journal of Economic Perspectives*, v. 31 (2): 211-236.
- Allen, Jennifer et al. (2020). «Evaluating the fake news problem at the scale of the information ecosystem». *Science Advances*, v. 6 (14), eaay3539: 1-6.
- Allen, Jennifer et al. (2021). «Scaling up fact-checking using the wisdom of crowds». *Science Advances*, v. 7: 1-10.
- Altay, S.; Berriche, M.; Acerbi, A. (2023). «Misinformation on misinformation: conceptual and methodological challenges». *Social Media + Society*, v. 9: 1-13.
- Altay, S.; Nielsen, R. K.; Fletcher, R. (2022). «Quantifying the “infodemic”: people turned to trustworthy news outlets during the 2020 coronavirus pandemic». *Journal of Quantitative Description: Digital Media*, v. 2: 1-29.
- Attenborough, Frederick (2024). Cyprus proposes five-year prison sentences for spreading “fake news”. *The Free Speech Union*. Acceso el 12 de abril de 2025. <https://freespeechunion.org/cyprus-proposes-five-year-prison-sentences-for-spreading-fake-news/>.
- Barroso, Luís Roberto (2011). *Curso de Direito Constitucional contemporâneo: os conceitos fundamentais e a construção do novo modelo*, 3. ed. São Paulo: Saraiva.
- Benkler, Y.; Faris, R.; Roberts, H. (2018). *Network propaganda: manipulation, disinformation, and radicalization in American politics*. New York: Oxford University Press.
- Bennet, W. L. y Livingston, S. (2021). «A brief history of the disinformation age: information wars and the decline of institutional authority». En *The disinformation age: politics, technology, and disruptive communication in the United States*, editado por Bennet, W. L. y Livingston, S., 3-40. Cambridge: Cambridge University Press.
- Bruns, Hendrik et al. (2024). «Investigating the role of source and source trust in prebunks and debunks of misinformation in online experiments across four EU countries». *Scientific Reports*, v. 14 (20723): 1-15.
- Chuai, Yuwei et al. (2024). «Did the roll-out of community notes reduce engagement with misinformation on X/Twitter?». *Proc. ACM Hum.-Comput. Interact.*, v. 8, CSCW 2, article 428: 1-52.
- Cinelli, Matteo et al. (2020). «The COVID-19 social media infodemic». *Scientific Reports*, v. 10(1): 1-10.
- Drolsbach, C. P. et al. (2024). «Community notes increase trust in fact-checking on social media». *PNAS Nexus*, v. 3 (7): 1-14.

- Eduardo, Angel y Goldstein, Adam (2024). «Misinformation: a flawed concept». *Quillette*. Acceso el 12 de enero de 2025. <https://quillette.com/2024/10/28/misinformation-a-flawed-concept/>.
- Enders, Adam M. et al. (2022). «On the relationship between conspiracy theory beliefs, misinformation, and vaccine hesitancy». *Plos One*, v. 17 (10).
- European Commission (2018a). *A multi-dimensional approach to disinformation: report of the independent High Level Group on Fake News and Online Disinformation*. Luxembourg: Publications Office of the European Union. Edición en PDF.
- European Commission (2018b). *Tackling online disinformation: a European approach*. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. Acceso el 12 de marzo de 2025. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:52018DC0236>.
- European Commission (2020a). *Communication on tackling COVID-19 disinformation – getting the facts right*. Acceso el 20 de marzo de 2025. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020JC0008>.
- European Commission (2020b). *Digital Services Act package: deepening the internal market and clarifying responsibilities for digital services*. Acceso el 20 de enero de 2025. [https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=pi_com:Ares\(2020\)2877686](https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=pi_com:Ares(2020)2877686).
- Efsas – European Foundation For South Asian Studies (2021). *The role of fake news in fueling hate speech and extremism online; promoting adequate measures for tackling the phenomenon*. Amsterdam: EFSAS. Edición en PDF.
- European Union (2025). *Code of Conduct on Disinformation [as amended in October 2024]*. Luxembourg, Publications Office of the European Union. Edición en PDF.
- Fitzpatrick, Jasmin y Jöst, Paula (2022). «“The high mass of democracy” – why Germany remains aloof to the idea of electronic voting». *Frontiers in Political Science*, v. 4, 876476: 1-14.
- Franklin, Margarita y Torrey, Mike (2024). *Adversarial threat report: third quarter*. Meta. Edición en PDF.
- Gleicher, Nathaniel et al. (2021). *Threat report: the state of influence operations 2017-2020*. Facebook. Edición en PDF.
- Graves, L (2016). *Deciding what's true: the rise of political fact-checking in American journalism*. New York: Columbia University Press. Edición en PDF.
- Grinberg, N. et al. (2019). «Fake news on Twitter during the 2016 U.S. presidential election». *Science*, v. 363 (6425): 374-378.
- Guess, A. et al. (2020). «“Fake news” may have limited effects beyond increasing beliefs in false claims». *Harvard Kennedy School Misinformation Review*, v. 1 (1): 1-12.
- Guess, A.; Nagler, J.; Tucker, J. (2019). «Less than you think: prevalence and predictors of fake news dissemination on Facebook». *Science Advances*, v. 5 (1).
- Han, Byung-Chul (2022). *Infocracia: digitalização e a crise da democracia*. Petrópolis: Vozes.
- Instituto de Pesquisa Datasenado (2024). *Panorama político 2024: notícias falsas e democracia*. Brasília: Senado Federal. Edición en PDF.

- Li, J. y Wagner, M. W. (2020). «The value of not knowing: partisan cue-taking and belief updating of the uninformed, the ambiguous, and the misinformed». *Journal of Communication*, v. 70 (5): 646-669.
- Línera, Miguel Ángel Presno; Teruel Lozano, Germán M. (2017). *La libertad de expresión en América y Europa: teoría y práctica*. Curitiba: Juruá.
- Loomba, Sahil et al. (2021). «Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA». *Nature Human Behaviour*, v. 5 (3): 337-348.
- Lucena, Marina et al. (2024). *Cartilha orientativa TSE e desinformação: conceitos relevantes e sua compreensão no Brasil*. Vol. 1. Rio de Janeiro: FGV Direito Rio.
- Meta (2025). *Desinformação: 7 de jan. de 2025*. Acceso el 28 de febrero de 2025. <https://transparency.meta.com/pt-br/policies/community-standards/misinformation/>.
- Mill, John Stuart (2017). *Sobre a Liberdade*. Tradução de Paulo Geiger. São Paulo: Penguin Classics Companhia das Letras.
- Newman, Nic et al. (2024). *Reuters Institute Digital News Report 2024*. Oxford: Reuters Institute for the Study of Journalism. Edición en PDF.
- Ó Fathaigh, R; Helberger, N; Appelman, N. (2021). «The perils of legally defining disinformation». *Internet Policy Review*, v. 10(4): 1-26.
- Pantazi, Myrto; Klein, Olivier; Kissine, Mikhail (2020). «Is Justice blind or myopic? An examination of the effects of meta-cognitive myopia and truth bias on mock jurors and judges». *Judgement and Decision Making*, v. 15 (2): 214-229.
- Pew Research Center (2019). *Many Americans say made-up news is a critical problem that needs to be fixed*. Edición en PDF.
- Rapoza, Kenneth (2017). Can 'fake news' impact the stock market? *Forbes*. Acceso el 10 de marzo de 2025. <https://www.forbes.com/sites/kenrapoza/2017/02/26/can-fake-news-impact-the-stock-market/>.
- Ruediger, Marco Aurélio (ed. 2022). *Desinformação on-line e contestação das eleições: quinze meses de postagens sobre fraude nas urnas eletrônicas e voto impresso auditável no Facebook*. Policy paper. Rio de Janeiro: FGV DAPP. Edición en PDF.
- Saeed, Mohammed et al. (2022). «Crowdsourced fact-checking at Twitter: how does the crowd compare with experts?». *arXiv:2208.09214v1 [cs.IR]*.
- Suciu, Peter (2022). Bots and misinformation on social media could undermine the midterm elections. *Forbes*. Acceso el 15 de febrero de 2025. <https://www.forbes.com/sites/petersuciu/2022/11/02/bots-and-misinformation-on-social-media-could-undermine-the-midterm-elections/>.
- Sunstein, Cass R. (2021). *Liars: falsehoods and free speech in an age of deception*. New York: Oxford University Press.
- Terra, Felipe Mendonça (2016). «Razão ou sensibilidade? Decidindo casos sobre liberdade de expressão: lições do cenário norte-americano». *Universitas JUS*, v. 27, n. 1: 163-187.
- TRE-GO (2024). *Entenda por que não é possível fraudar a urna eletrônica*. Tribunal Regional Eleitoral de Goiás. Acceso el 4 de abril de 2025. <https://www.tre-go.jus.br/comunicacao/noticias/2024/Fevereiro/entenda-por-que-nao-e-possivel-fraudar-a-urna-eletronica>.

- Vázquez Alonso, Víctor J. (2021). «¿Puede @realDonaldTrump bloquearte en Twitter? Foros públicos digitales y libertad de expresión». En *Estudios de casos líderes europeos y nacionales*, v. XIV. Editado por Víctor J. Vázquez Alonso, 141-165. Ciudad de México: Tirant lo Blanch.
- Vázquez Alonso, Víctor J (2022). «La censura «privada» de las grandes corporaciones digitales y el nuevo sistema de la libertad de expresión». *Teoría & Derecho: Revista de Pensamiento Jurídico*, v. 32: 108-129.
- Vosoughi, S.; Roy, D.; Aral, S. (2018). «The spread of true and false news online». *Science*, v. 359 (6380): 1146-1151.
- Walter, N. et al. (2020). «Fact-checking: a meta-analysis of what works and for whom». *Political Communication*, v. 37: 350-375.
- Wardle, Claire; Derakhshan, Hossein (2017). *Information disorder: toward an interdisciplinary framework for research and policy making*. Strasbourg: Council of Europe. Edición en PDF.
- Zhao, Andy; Naaman, Mor (2023). «Unraveling the potential of crowds: insights from a comparative study on the variety, velocity, veracity, and viability of crowdsourced and professional fact-checking services». *Journal of Online Trust and Safety*: 1-24.

JURISPRUDENCIA

- Tribunal Europeo de Derechos Humanos. *Caso Handyside c. Reino Unido*, demanda n.º 5493/72, sentencia de 7 de diciembre de 1976. ECLI:CE:ECHR:1976:1207JUD000549372.
- Tribunal Europeo de Derechos Humanos. *Caso The Sunday Times c. Reino Unido*, demanda n.º 6538/74, sentencia de 29 de abril de 1979. ECLI:CE:ECHR:1979:0426JUD000653874.
- Tribunal Europeo de Derechos Humanos. *Caso Lingens c. Austria*, demanda n.º 9815/82, sentencia de 8 de julio de 1986. ECLI:CE:ECHR:1986:0708JUD000981582.
- Tribunal Europeo de Derechos Humanos. *Caso Observer y Guardian c. Reino Unido*, demanda n.º 13585/88, sentencia de 26 de noviembre de 1991. ECLI:CE:EC HR:1991:1126JUD001358588.
- Tribunal Europeo de Derechos Humanos. *Caso Incal c. Turquía*, n.º 41/1997/825/1031, sentencia de 9 de junio de 1998. ECLI:CE:ECHR:1998:0609JUD002267893.
- Tribunal Europeo de Derechos Humanos. *Caso Fressoz y Roire c. Francia*, demanda n.º 29183/95, sentencia de 21 de enero de 1999. ECLI:CE:ECHR:1999:0121JUD002918395.
- Supremo Tribunal Federal (Brasil). *Acción Directa de Inconstitucionalidad n.º 7261*, rel. min. Edson Fachin, 6 de marzo de 2024.