

Democracia en riesgo: IA, desinformación y discursos de odio en los procesos políticos latinoamericanos

Cómo citar este artículo [Chicago]: Astudillo Muñoz, Jorge y María Lorena Rossel Castagneto. "Democracia en riesgo: IA, desinformación y discursos de odio en los procesos políticos latinoamericanos". *Novum Jus* 20, no. 2 (2026): 315-340. <https://doi.org/10.14718/NovumJus.2026.20.2.11>

Jorge Astudillo Muñoz /
María Lorena Rossel Castagneto



Código: 2264368504 • Autor: istockphoto.com

Democracia en riesgo: IA, desinformación y discursos de odio en los procesos políticos latinoamericanos

Jorge Astudillo Muñoz^{*}
María Lorena Rossel Castagneto^{**}

Recibido: 11 de marzo de 2026 | **Evaluated:** 21 de abril de 2026 | **Aceptado:** 22 de abril de 2026

Resumen

El presente estudio examina la conexión compleja que existe entre la inteligencia artificial, la desinformación, y los discursos de odio; factores cruciales que han proliferado en recientes campañas electorales en Latinoamérica. En efecto, en procesos políticos como los de Brasil, Colombia, Ecuador y Chile, la propagación de noticias falsas y narrativas agresivas se han convertido en un recurso político habitual. Adicionalmente, en varias de esas circunstancias se aprecian técnicas muy específicas, como el uso de *bots*, el *astroturfing* y, los *deepfakes*, las cuales distorsionan el debate público y socavan la equidad en la competencia electoral. La investigación combina revisión documental y análisis comparado de experiencias nacionales para evaluar cómo la manipulación digital intensifica las dinámicas de polarización y deteriora la confianza en las instituciones. A partir de estos hallazgos se derivan propuestas para orientar la regulación y la educación; todo ello con el fin de facultar a los Estados para abordar los desafíos de la era digital sin que esto vulnere los derechos fundamentales ni la protección de la democracia.

Palabras clave: democracia; desinformación; inteligencia artificial; polarización; regulación; discursos de odio.

^{*} Dr. Jorge Astudillo Muñoz, Facultad de Derecho, Universidad Andrés Bello (Chile); correo: jastudillo@unab.cl; ORCID: <https://orcid.org/0000-0001-6562-1930>; SCOPUS ID: 57223086456

^{**} Dra. María Lorena Rossel Castagneto, Facultad de Derecho, Universidad de Las Américas (Chile); correo: mrossel@udla.cl; ORCID: <https://orcid.org/0000-0003-4085-3000>; SCOPUS ID: 57813195800.

Democracy at Risk: Artificial Intelligence, Disinformation, and Hate Speech in Latin American Political Processes

Jorge Astudillo Muñoz
María Lorena Rossel Castagneto

Received: March 11, 2026 | **Evaluated:** April 21, 2026 | **Accepted:** April 22, 2026

Abstract

This study examines the complex relationship between artificial intelligence, disinformation, and hate speech, all of which have become crucial factors in recent electoral campaigns across Latin America. Indeed, in political processes such as those in Brazil, Colombia, Ecuador, and Chile, the dissemination of fake news and aggressive narratives has become a common political tool. Moreover, several of these cases have involved highly specific techniques, including the use of bots, astroturfing, and deepfakes, all of which distort public debate and undermine fairness in electoral competition. The research combines documentary review and a comparative analysis of national experiences to assess how digital manipulation intensifies polarization and erodes public trust in institutions. Based on these findings, it proposes measures to guide regulation and education, with the aim of enabling States to address the challenges of the digital age without compromising fundamental rights or the protection of democracy.

Keywords: democracy; disinformation; artificial intelligence; polarization; regulation; hate speech.

Introducción

La ahora palpable facultad de los *softwares* para asimilar datos, procesarlos y aprender a determinar constituye un hito crucial para entender las sociedades actuales. La inteligencia artificial (en adelante IA), concretamente la generativa, se ha afianzado como un instrumento con un potencial transformador que puede impactar la rutina de los individuos y reconfigurar dinámicas sociales, culturales, y políticas. Pese a que sus ventajas en áreas como la medicina, la educación o la industria resultan indiscutibles, conlleva también riesgos notorios, dentro de los cuales resalta la producción y propagación de desinformación y discursos de odio.

En América Latina, región que se ha visto influenciada por procesos de alta polarización política y giros en los proyectos de gobierno del siglo XXI, el uso de IA, la desinformación y los discursos de odio plantean una amenaza directa a la calidad de la democracia. Las campañas electorales han sido un espacio particularmente fértil para la utilización de estas tecnologías, tanto mediante la difusión de información falsa o de *fake news* como a través de la creación de discursos de odio contra colectivos vulnerables o adversarios políticos. Este fenómeno no solo afecta el derecho a acceder a información veraz, también incide en el ejercicio libre del voto y en la estabilidad de los sistemas políticos.

De este modo, esta investigación tiene como objetivo responder a la pregunta: ¿cómo la desinformación y los discursos de odio amplificados por inteligencia artificial han promovido giros políticos en América Latina, especialmente en contextos de alta polarización electoral?

Se postula que, al potenciar los discursos de odio y promover la desinformación durante las campañas políticas, la inteligencia artificial exacerba la polarización política y afectiva en América Latina, y con ello debilita la deliberación democrática e influye en la conformación de cambios políticos que favorecen a los líderes populistas y a los proyectos autoritarios en el siglo XXI.

En este estudio, el propósito primordial radica en analizar la repercusión de la inteligencia artificial en la amplificación de la desinformación y los discursos de odio en América Latina, y su influencia en los giros políticos del siglo XXI, particularmente en contextos de campañas electorales marcadas por la polarización.

Para alcanzar el objetivo general hemos considerado necesario examinar el vínculo conceptual y práctico entre inteligencia artificial, desinformación y discursos de odio en el entorno digital contemporáneo. Después, llevaremos a cabo un análisis comparativo de distintas campañas políticas latinoamericanas en las que los discursos de odio y la desinformación hayan tenido un rol decisivo (se analizan en particular los casos de Brasil [2018 y 2022], Chile [2022 y 2025], Colombia [2022] y Ecuador [2021 y 2025]). Finalmente, se ofrecen una serie de lineamientos normativos y políticos orientados a mitigar el uso de la IA para fines de manipulación electoral y protección de los derechos fundamentales.

Desde el punto de vista metodológico, esta investigación opta por un método cualitativo y documental, basándose en el análisis pormenorizado de bibliografía académica, informes de organismos internacionales y estudios previos para estudiar el impacto de la IA en la propagación de desinformación y mensajes de odio. Así las cosas, se examinaron documentos académicos, leyes vigentes y estudios de caso para entender el fenómeno en su totalidad.

Inteligencia artificial, desinformación y discursos de odio en el entorno digital contemporáneo

La desinformación, especialmente aquella realizada con la intención deliberada de tener efectos perjudiciales, encuentra en la IA —y muy particularmente en la IA generativa— una poderosa herramienta tanto para producir contenido más sofisticado como para contribuir a su difusión y propagación de manera más rápida y con más alcance.

De esta manera, herramientas de última generación como los algoritmos de aprendizaje automático y la IA generativa posibilitan la producción de contenido dotado de un gran realismo que se traduce en textos, imágenes o videos (conocidos como *deepfakes*) utilizados para difundir mensajes falsos. Todo este contenido falso o distorsionado, aunque dotado de un gran realismo, resulta cada vez más complejo de distinguir de la información genuina. Como consecuencia de lo anterior, la opinión pública recibe información errónea y termina siendo confundida, y además, se va creando una sensación de desconfianza en las fuentes informativas tradicionales. A continuación, se analizará la repercusión de estos fenómenos en la comunicación política y en las elecciones.

Sobre la desinformación

Desde un enfoque amplio y restringido es posible abordar el concepto de desinformación según el elemento de intencionalidad que subyace a este fenómeno.

Así, recientes investigaciones consideran que la intencionalidad no siempre debe estar presente como elemento esencial para definir la desinformación.¹ En consecuencia, para Bachmann et al. la utilización del término desinformación resulta adecuada para dar cuenta de la complejidad que implica la difusión de información incorrecta, carente de verificación o de carácter ambiguo, aun cuando no exista una intención deliberada de engañar.²

Desde una perspectiva restringida, la intencionalidad puede ser entendida como la difusión intencional de información inexacta, alterada, engañosa o descontextualizada, dirigida a la consecución de objetivos específicos, sobre todo en ámbitos políticos y económicos, mediante la afectación de la confianza social y la construcción de percepciones distorsionadas, con la intención de confundir o manipular a las personas para alcanzar diversos objetivos^{3,4}.

Con todo, y a pesar de la tipología que acabamos de indicar, algunos autores aluden a la expresión ‘desórdenes informativos’ para referirse al fenómeno de producir y divulgar información manipulada, distorsionada o descontextualizada e incluso innecesaria, que produce efectos negativos en la percepción ciudadana y en su consustancial proceso de toma de decisiones.⁵ De esta forma, tomando en cuenta la magnitud del daño se distingue entre misinformación, desinformación y malinformación.

¹ Marian Slipczuk, *Una guía para principiantes para establecer una organización de fact-checking en América Latina y el Caribe* (Unesco, 2022), 8.

² Ingrid Bachmann et al., *El fenómeno de la desinformación y el desafío de corregirla: Una revisión bibliográfica*, 2021, 2, <https://bit.ly/4e9KnCn>.

³ Jorge Astudillo, “Desinformación: aproximación conceptual, riesgos y remedios”, *Revista Derecho PUCP* 94 (2024): 64.

⁴ Esta noción es aceptada por parte de la doctrina especializada y organizaciones estatales e internacionales como ocurre en el Plan contra la Desinformación de la Comisión Europea y el Alto Representante de la Unión para Asuntos Exteriores y Política de Seguridad (2018), el Procedimiento de actuación contra la desinformación del Consejo de Seguridad Nacional Español (2020), la Relatoría Especial para la Libertad de Expresión de la Comisión Interamericana de Derechos Humanos (2019), etc. Véase: (European Commission: Directorate-General for Communications Networks 2018).

⁵ Cherylyn Ireton y Julie Posetti, *Journalism, fake news & disinformation: handbook for journalism education and training* (Unesco, 2018), 45-46.

La “misinformación” es un término que alude a la difusión de información falsa, pero que se da sin ánimo de provocar daño. Ireton y Posetti⁶ ofrecen el siguiente ejemplo: durante abril de 2017 un ataque terrorista en los Campos Elíseos de París despertó el interés de gran parte de la comunidad europea y mundial. En distintas plataformas digitales circularon diversos rumores, entre ellos la supuesta información de que un segundo funcionario policial había sido asesinado. Las personas que comparten este tipo de contenido rara vez lo hacen con la finalidad de causar daño, más bien se dejan llevar por el momento, tratando de ser útiles, pero sin verificar de manera adecuada la información que comparten.

La “malinformación” alude a la circunstancia de divulgar información verdadera utilizada con el objetivo de provocar afectación o daño. Por lo general, la malinformación se concreta por medio de la difusión de una información en principio privada hacia el ámbito público. Según Ireton y Posetti⁷, un ejemplo llamativo de malinformación acaeció con la filtración de correos de Emmanuel Macron horas antes de la segunda vuelta electoral del 7 de mayo de 2017. Si bien estos correos se consideraron auténticos, la divulgación de mensajes privados se hizo con la intención de perjudicar sus opciones electorales.

La “desinformación” constituye la divulgación de información errónea o claramente falsa que se crea con la exclusiva y deliberada intención de engañar y dañar. Ireton y Posetti⁸ ejemplifican este fenómeno con la divulgación de una copia sofisticada del periódico belga *Le Soir* que incluía un artículo falso que informaba que el candidato Emmanuel Macron recibía financiamiento de Arabia Saudita.

La desinformación no solo afecta el derecho de los ciudadanos a recibir y a acceder información veraz, también puede afectar de manera importante la estabilidad política y social de una sociedad, especialmente porque ella es un factor que puede desembocar en polarización política y afectiva que termine colapsando pilares fundamentales de un sistema social y político⁹¹⁰. En efecto, la desinformación puede ser utilizada para hacer uso de un lenguaje agresivo y polarizante, que conduzca a una profunda división entre las personas y agudice los antagonismos

⁶ Ireton y Posetti, *Journalism, fake news*, 47.

⁷ Ireton y Posetti, *Journalism, fake news*, 47.

⁸ Ireton y Posetti, *Journalism, fake news*, 47.

⁹ Pratibha N. Vasist et al., “The Polarizing Impact of Political Disinformation and Hate Speech: A Cross-Country Configurational Narrative”, *Information Systems Frontiers* 26, n.º 2 (2024): 670.

¹⁰ Matteo Cinelli et al., “Online Hate: Behavioural Dynamics and Relationship with Misinformation”, preprint, arXiv, 2021, 3, <https://doi.org/10.48550/ARXIV.2105.14005>.

y diferencias y dificulte el diálogo y los consensos sociales y políticos. Diversos estudios han demostrado que estas dinámicas se alimentan de entornos digitales fragmentados —como las cámaras de eco algorítmicas— donde discursos de odio y desinformación se refuerzan mutuamente^{11, 12}, mientras que el discurso agresivo se difunde con mayor alcance dentro de redes emocionalmente polarizadas.¹³

A la tradicional polarización política hoy se agrega una polarización afectiva que se caracteriza principalmente por su carácter socioemocional, su potencialidad de ser un vector generador de desconfianzas, sesgos, hostilidades y discriminación según la identidad política, partidista o no.¹⁴ Una de las técnicas propias de la polarización afectiva es el uso de estrategias que persiguen instalar una especie de deshumanización de aquellos a los que se quiere atacar. Este fenómeno es de común ocurrencia en redes sociales en sociedades con conflictos derivados de los flujos migratorios. Esto es particularmente grave toda vez que ciertos grupos, por ejemplo, migrantes, al ser constantemente desacreditados, estigmatizados y en definitiva deshumanizados, serán víctimas de discursos del odio proferidos en su contra, los que serán más efectivos en la medida que existan supuestas razones que los justifiquen.¹⁵

La desinformación se ha consolidado como una herramienta central en las campañas políticas, utilizada tanto por partidos como por sujetos externos, con la finalidad de influir en la opinión pública y condicionar los procesos democráticos. Según Ruiz-Mora et al., los partidos políticos recurren a “teorías conspiratorias y publicaciones sensacionalistas” como parte de sus estrategias de comunicación, lo que les permite instalar en la agenda mediática marcos discursivos alternativos y favorables a sus intereses¹⁶. Estas prácticas generan narrativas antagonistas que, al difundirse en medios y redes, buscan deslegitimar propuestas institucionales y consolidar proyectos partidistas.

¹¹ Cinelli et al., “Online Hate”.

¹² Berta Chulvi et al., “Fake news and hate speech: Language in common”, *Information Processing & Management* 59, n.º 6 (2022): 103093, <https://doi.org/10.48550/arXiv.2212.02352>.

¹³ Libby Jenke, “Affective polarization and misinformation belief”, *Political Behavior* 46, n.º 2 (2024): 830, <https://doi.org/10.1007/s11109-022-09851-w>.

¹⁴ Juan Manuel Rojo-Martínez y Isabel Crespo-Martínez, “Lo político como algo personal: Una revisión teórica sobre la polarización afectiva”, *Revista de Ciencia Política* 43, n.º 1 (2023): 42.

¹⁵ María Lorena Rossel Castagneto, “Regular el odio digital: Entre la libertad de expresión y la protección de colectivos vulnerables en Chile”, *Universitas XXI* 43 (2025): 95-122, <https://doi.org/10.17163/uni.n43.2025.04>.

¹⁶ Isabel Ruiz-Mora et al., “Estrategias de desinformación de VOX: Cobertura mediática de su discurso parlamentario sobre la Agenda 2030”, *Palabra Clave* 28, n.º s1 (2025): 10, <https://doi.org/10.5294/pacla.2025.28.s1.12>.

Un mecanismo específico empleado es el *astroturfing*, entendido como la creación artificial de apoyo popular a través de la coordinación de cuentas y mensajes que aparentan ser orgánicos. En campañas políticas, estas estrategias han permitido amplificar la polarización y condicionar la percepción ciudadana sobre temas relevantes o sobre la real adhesión hacia uno o varios candidatos. Arce-García et al. sostienen que el *astroturfing* se caracteriza por impulsar temas en redes sociales a través de la propagación masiva y la amplificación de contenidos desinformativos, constituyendo una técnica avanzada de manipulación en contextos electorales¹⁷.

Del mismo modo, Ballesteros-Aguayo et al. señalan que en el debate sobre la Ley Trans en España se desplegaron campañas de *astroturfing* en TikTok con el propósito de “crear una falsa impresión de voluntad popular” y legitimar discursos populistas contrarios a la norma y que constituían un retroceso respecto de ese colectivo.¹⁸ Esta situación demuestra cómo la desinformación no se limita a falsear hechos, sino que se articula como una estrategia política para debilitar los consensos democráticos e instalar un falso apoyo a ciertas posiciones.

Finalmente, Torres et al. advierten que la desinformación puede “erosionar la confianza pública en la credibilidad de los procesos”, y crear un escenario en el cual los resultados electorales son percibidos con sospecha, lo cual debilitaría la cohesión social, pero agrega que en la tarea de regular estas prácticas se debe evitar promulgar leyes que limiten la libertad de expresión.¹⁹

Asimismo, la desinformación, especialmente aquella que se difunde en contextos de posverdad y polarización, afecta de manera importante los derechos de determinados grupos en situación de vulnerabilidad, como los migrantes o los miembros de ciertos colectivos como las personas LGTBIQ+, y puede ser utilizada para promover y potenciar discursos de odio que pueden desembocar en discriminaciones e incluso agresiones físicas en contra de estas personas.²⁰

¹⁷ Sergio Arce-García et al., “Tipos de campaña Astroturfing de contenidos desinformativos y polarizados en tiempos de pandemia en España”, *ICONO14. Revista Científica de Comunicación y Tecnologías Emergentes* 21, n.º 1 (2023): 2, <https://doi.org/10.7195/ri14.v21i1.1890>.

¹⁸ Laura Ballesteros-Aguayo et al., “Desinformación y fact-checking en las elecciones en Francia de 2022: Guerra de Ucrania y polarización”, *Estudos em Comunicação* 38 (2024): 268, <https://doi.org/10.5209/hics.92238>.

¹⁹ Gabriel Torres Espinoza, “El nuevo frente de las fake news en las democracias: formas de contrarrestar los efectos de la desinformación”, en *Impactos de la desinformación en un mundo cambiante*, ed. Gabriel Torres Espinoza y Omar Guillermo García Santiago (Innovación Educativa, 2025), 251.

²⁰ Rossel Castagneto, “Regular el odio digital: Entre la libertad de expresión y la protección de colectivos vulnerables en Chile”, 104.

Analizada la complejidad que encierra el fenómeno de la desinformación y sus efectos sobre la democracia y el ejercicio de derechos humanos, analizaremos a continuación los discursos de odio. En efecto, muchas veces la desinformación se convierte en un puente para la propagación de expresiones agresivas, hostiles y discriminatorias dirigidas en contra de determinados grupos. En las páginas que siguen se ofrece una aproximación conceptual al discurso del odio y una reflexión respecto de los desafíos que encierra su regulación.

Discursos de odio

La doctrina y los instrumentos internacionales reconocen al menos tres tendencias para definir el discurso del odio: i) el criterio de la violencia o agresión inminente, ii) el criterio del fomento, promoción o instigación de animadversión hacia ciertos grupos, y iii) el criterio de la sola manifestación de un mensaje agresivo o humillante contra un grupo o personas que pertenezcan a este.

El discurso del odio bajo el criterio de la violencia inminente

Este criterio considera discurso del odio solo a las manifestaciones que inciten de forma directa y concreta a la violencia, caso en el cual el Gobierno puede restringir la libertad de expresión. Surge de la jurisprudencia de la Corte Suprema de EE. UU., en *Brandenburg v. Ohio* (1969)²¹, que estableció que la expresión solo es sancionable si busca incitar a una acción ilícita inminente, y es probable que lo logre. Este estándar protege expresiones ofensivas mientras estas no impliquen un riesgo real de violencia, consolidándose en fallos como *Collin vs. Smith* (1978)²² y *Snyder vs. Phelps* (2011).²³ Dworkin defendió esta postura al considerar autoritario sancionar más allá de este límite.²⁴

El discurso del odio bajo el criterio de la instigación o fomento

Aquí el discurso del odio incluye manifestaciones que promuevan o justifiquen la exclusión, la estigmatización o la intolerancia, aunque no llamen directamente a la violencia. La Comisión Europea contra el Racismo y la Intolerancia (2015) y el Consejo de Europa (1997) lo han definido en esta clave, vinculándolo a expresiones

²¹ Corte Suprema de Estados Unidos, *Brandenburg v. Ohio*, 395 U.S. 444, 1969.

²² *Collin v. Smith* (Supreme Court of the United States 16 de octubre de 1978).

²³ *Snyder v. Phelps*, 562 U.S. 443 (Supreme Court of the United States, 2 de marzo de 2011).

²⁴ Ronald Dworkin, *A matter of principle* (Harvard University Press, 1985), 364.

que difunden estereotipos negativos o legitiman la discriminación. El Tribunal Europeo de Derechos Humanos también ha adoptado este enfoque.

El discurso del odio bajo el criterio de la sola expresión de mensajes humillantes

Una tercera visión propone sancionar la sola manifestación pública de mensajes racistas o humillantes, aun sin instigación al odio o a la violencia. La Recomendación N.º 7 de la Comisión Europea contra el Racismo y la Intolerancia (2002) y la Convención internacional sobre la eliminación de todas las formas de discriminación racial (1965) exigen sancionar expresiones que difundan ideas de superioridad racial o étnica. Basta la manifestación pública de estas ideologías para configurar una responsabilidad penal.

En suma, el análisis de las distintas concepciones del discurso del odio muestra la necesidad de adoptar criterios normativos equilibrados que protejan a los colectivos vulnerables sin menoscabar de forma desproporcionada la libertad de expresión. Ahora bien, estas categorías jurídicas tradicionales enfrentan hoy un nuevo desafío: la irrupción de la IA en procesos de producción y difusión de contenidos. La IA, especialmente en su vertiente generativa, ha introducido un escenario inédito en el que los discursos de odio y la desinformación se ven potenciados y adquieren una capacidad de propagación y sofisticación sin precedentes, lo que exige repensar los marcos regulatorios y las estrategias de protección de la democracia y los derechos humanos, como se analizará a continuación.

Inteligencia artificial

Resulta complejo ofrecer una definición de la IA, en primer lugar, porque se trata de un fenómeno que con toda seguridad irá variando al ritmo de los avances tecnológicos. En segundo lugar, porque reúne tecnologías muy diversas, que van desde el aprendizaje automático y las redes neuronales hasta modelos de deducción, técnicas estadísticas y enfoques psicológicos.²⁵

Para los efectos de esta investigación diremos, en términos muy generales, que la IA está asociada a ciertos sistemas que de alguna manera tienden a igualar la

²⁵ Gabriele Vestri, *Diccionario de términos para comprender la transformación digital* (Aranzadi, 2023), 197-98.

capacidad de las personas naturales para nutrirse de información, procesarla y tomar decisiones^{26, 27, 28}.

Estudios recientes destacan que la IA, particularmente la generativa, se ha convertido en un factor clave para amplificar, automatizar y sofisticar la desinformación en campañas políticas, lo cual aumenta los riesgos para la democracia. García-Orosa²⁹ observa que la irrupción de la IA, junto con el *big data* y los algoritmos, marca una “cuarta ola de la democracia digital”, donde fenómenos como las *fake news*, los *deep fakes* y el *astroturfing* suponen una validación de la mentira como estrategia política y potencian la polarización social.³⁰

En la misma línea, Ballesteros-Aguayo et al.³¹, al estudiar el caso de la Ley Trans en España, evidencian que el uso de *astroturfing* en TikTok se apoya en la coordinación de cuentas humanas y automatizadas que simulan apoyo ciudadano, lo cual intensifica la polarización ideológica y refuerza liderazgos populistas. Los autores subrayan que esta práctica no se limita a la propaganda clásica, sino que utiliza herramientas tecnológicas difíciles de rastrear, al tiempo que crean una falsa impresión de legitimidad popular.³²

Estrechamente vinculados al *astroturfing*, como consecuencia de la IA y, en especial de la IA generativa, están los llamados bots. De manera general, los bots pueden definirse como sistemas automatizados diseñados para ejecutar tareas específicas en el entorno digital, tales como la distribución masiva de imágenes, videos o textos hacia públicos determinados o indeterminados; asimismo, pueden emplearse para identificar y analizar patrones de consumo. Ahora bien, desde una perspectiva más específica, el bot social corresponde a un tipo de programa automatizado que actúa en plataformas digitales y se emplea para respaldar posturas, difundir contenidos, posicionar mensajes, impulsar estrategias comunicacionales y crear tendencias, entre

²⁶ Vestri, *Diccionario de términos para comprender la transformación digital*, 197.

²⁷ Laura Rebollo Delgado, *Inteligencia artificial y derechos fundamentales* (Dykinson, 2023), 19.

²⁸ José Manuel Girón Sierra, *Introducción a la inteligencia artificial: La tecnología que nos cambiará para siempre* (Sekotia, 2023), 27.

²⁹ Berta García-Orosa, “Disinformation, social media, bots, and astroturfing: the fourth wave of digital democracy”, *Profesional de la información* 30, n.º 6 (2021): 2, <https://doi.org/10.3145/epi.2021.nov.03>.

³⁰ García-Orosa, “Disinformation”, 2.

³¹ Ballesteros-Aguayo et al., “Desinformación y fact-checking en las elecciones en Francia de 2022: Guerra de Ucrania y polarización”, 268.

³² Ballesteros-Aguayo et al., “Desinformación y fact-checking en las elecciones en Francia de 2022: Guerra de Ucrania y polarización”, 268.

otros fines.³³ En definitiva ambos fenómenos —los bots y el *astroturfing*— tienen en común que validan la mentira como estrategia política y refuerzan dinámicas de polarización y manipulación en campañas electorales.³⁴

Otra herramienta recurrente asociada a la Inteligencia Artificial Generativa (IAG) es la *deepfake*. La tecnología de *deepfakes* se caracteriza por el empleo de técnicas de aprendizaje automático para crear contenido multimedia asociado a personas reales mostrándose en situaciones que nunca sucedieron. Estas representaciones están cargadas de una fuerte dosis de realismo que cada vez son más difíciles de detectar. La creación de las *deepfakes* está asociada al uso de redes neuronales avanzadas, como Generative Adversarial Networks (GAN), lo que permite una alta sofisticación técnica y potencialidad para crear representaciones realistas.³⁵

Por su parte, Cervantes et al.³⁶ ponen de relieve que la IA permite la viralización de contenidos fabricados o descontextualizados en redes sociales, lo que contribuye a reforzar tácticas de manipulación política que buscan moldear la opinión pública, afectando tanto la deliberación democrática como el acceso a la información veraz. En este sentido, remarcan la necesidad de regular y promover la alfabetización digital como mecanismos para contrarrestar estos riesgos, como se analizará en el apartado 4 del presente artículo³⁷.

Del mismo modo, Fontenla et al.³⁸ señalan que los avances en IA han complejizado las campañas de desinformación, y han permitido con ello un grado de segmentación y viralización lo que contribuye a erosionar la confianza en las instituciones, así como a debilitar el discurso público y aumentar la división social. Así, agrega que la desinformación potenciada por IA constituye una amenaza directa a la estabilidad democrática.³⁹

³³ Emilio Ferrara et al., “The rise of social bots”, *Communications of the ACM* 59, n.º 7 (2016): 96.

³⁴ García-Orosa, “Disinformation”, 2.

³⁵ Danielle Keats Citron y Robert Chesney, “Deep fakes: A looming challenge for privacy, democracy, and national security”, *California Law Review* 107 (2019): 1753, <https://doi.org/10.15779/Z38RV0D15J>.

³⁶ Rubén Cervantes Hernández et al., “Estrategias contra la desinformación en redes socio digitales”, *Revista Cuestiones Políticas* 9, n.º 3 (2025): 2123-30, https://doi.org/10.37811/cl_rcm.v9i3.17837.

³⁷ Cervantes Hernández et al., “Estrategias contra la desinformación”.

³⁸ Julia Fontenla-Pedreira et al., “Regulación de la desinformación: enfoques multidisciplinares y desafíos legislativos globales en 2024”, *Revista Mediterránea de Comunicación* 16, n.º 2 (2025): 2, <https://doi.org/10.14198/MEDCOM.29218>.

³⁹ Fontenla-Pedreira et al., “Regulación de la desinformación”, 2.

En este sentido, las relaciones entre inteligencia artificial, la desinformación y los discursos de odio socavan la confianza pública en las instituciones y, a la vez, incitan la polarización política y social, con lo cual debilitan la deliberación pública con información, y a veces, pueden contribuir a legitimar prácticas discriminatorias que dañan gravemente a colectivos vulnerables. Estos fenómenos, combinados no solo expanden los discursos de odio en el ámbito digital, también moldean la percepción pública, con el fin de impactar la calidad del voto y, por tanto, la validez de las elecciones. Así, es esencial estudiar cómo diferentes países de las Américas han afrontado los efectos de los discursos de odio y la desinformación, ambos impulsados y difundidos por la inteligencia artificial, un tema de análisis en el siguiente capítulo.

Procesos de desinformación y discursos de odio en campañas políticas latinoamericanas

Brasil

Las elecciones presidenciales de Brasil en 2018 y 2022 evidencian cómo la desinformación, las *fake news* y los discursos de odio se consolidaron como estrategias electorales con efectos corrosivos en la democracia. En 2018, WhatsApp fue la principal plataforma: más del 60 % de los contenidos en grupos familiares eran falsos o solo parcialmente verdaderos, que favorecían a Jair Bolsonaro frente a Fernando Haddad⁴⁰, ⁴¹. Estos mensajes, basados en la defensa de la familia tradicional, el rechazo a la “ideología de género” y el inexistente “kit gay”, reforzaron narrativas religiosas y militares que legitimaban la desinformación.⁴²

En 2022, líderes evangelistas difundieron en redes mensajes políticos en clave religiosa, el 70 % de ellos con desinformación,⁴³ consolidando cámaras de eco que intensificaron la polarización y la estigmatización de la izquierda, el feminismo y las minorías sexuales. Frente a ello, el Proyecto de Ley 2630/2020 y acciones del

⁴⁰ Juliana Colussi, “Desinformación, política y religión: Un análisis de los contenidos compartidos por WhatsApp durante la campaña presidencial de Brasil en 2018”, *Revista Ibérica de Sistemas e Tecnologias de Informação (RISTI)*, n.º E35 (2020): 478.

⁴¹ João Canavilhas y Juliana Colussi, “WhatsApp como plataforma de desinformación: Estudio de caso de las elecciones presidenciales brasileñas”, *Visual Review* 9 (2022): 3, <https://doi.org/10.37467/revvisual.v9.3519>.

⁴² Colussi, “Desinformación, política y religión: Un análisis de los contenidos compartidos por WhatsApp durante la campaña presidencial de Brasil en 2018”, 480.

⁴³ Breno Moreira, “Desinformación en las declaraciones políticas de los líderes evangelistas en las elecciones de 2022 en Brasil”, *Miguel Hernández Communication Journal* 14, n.º 1 (2023): 109-31, <https://doi.org/10.21134/mhjourn.v14i.1813>.

TSE, como la Red Nacional de Combate a la Desinformación, buscaron mitigar estos riesgos^{44,45}.

Ecuador

En Ecuador, las elecciones recientes han estado marcadas por desinformación, las *fake news* y los discursos de odio, con lo cual se afectan la opinión pública y la confianza en la democracia. Guadarrama et al. señalan que, en la lógica de la posverdad, la comunicación política apela más a las emociones que a los hechos, de manera que genera polarización y “opinorofobia”, es decir, miedo a expresar opiniones en contextos de linchamiento mediático.⁴⁶ Del mismo modo, Tomalá muestra que desde 2017 las redes sociales se consolidaron como espacios centrales de campaña, donde las *fake news* moldearon percepciones y fragmentaron el debate político.⁴⁷

En 2025, Viñán et al. pusieron en evidencia que las campañas digitales y las “cámaras de eco” reforzaron el aislamiento ideológico, y deterioraron la deliberación plural y la estabilidad democrática.⁴⁸ Ante ello, el Estado ha impulsado alfabetización digital⁴⁹ y apoyado iniciativas de verificación como Ecuador Chequea.⁵⁰ Estas respuestas buscan mitigar riesgos que amenazan la calidad democrática.

⁴⁴ Denysson Axel Ribeiro Mota et al., “Régimen de información y propagación de la desinformación: Los desafíos del Marco Civil de la Internet y la Ley de Fake News en Brasil”, *Actas de las 7ma Jornadas de Intercambio y Reflexión acerca de la Investigación en Bibliotecología* (La Plata), 2024, 4-5.

⁴⁵ Antonia Isabel Nogales-Bocio et al., “Estrategias locales de verificación y fact-checking en Brasil: La red nacional de combate a la desinformación, COAR Noticias, Colectivo Bereia, Verifica RS y Agencia Tatu”, *Doxa Comunicación* 41 (2025): 400, <https://doi.org/10.31921/doxacom.n41a2906>.

⁴⁶ Jorge J. M. Guadarrama Torres et al., “Desinformación en la era digital: abordaje de la fórmula de la posverdad en la comunicación política”, *Scripta Mundi* 4, n.º 1 (2025): 12-13, <https://doi.org/10.53591/scmu.v4i1.2179>.

⁴⁷ Francisco Tomalá Medina, “Campañas electorales, redes sociales y fake news”, *Justicia Electoral y Democracia* 5, n.º 6 (2022): 55-59.

⁴⁸ Lourdes Viñán Carrasco et al., “Construcción de la Opinión Pública en un Contexto de Desinformación y Polarización Previo a las Elecciones Generales 2025 de Ecuador”, *Polo del Conocimiento* 10, n.º 1 (2025): 2072-74.

⁴⁹ Marco López-Paredes y Andrea Carrillo-Andrade, “Ultramediaciones en contextos de analfabetismo digital en Ecuador: aproximación a la desinformación y malinformación”, *Austral Comunicación* 14, n.º 3 (2025): e01430, <https://doi.org/10.26422/aucom.2025.1403.lop>.

⁵⁰ Agustín Alexander Sánchez Moscoso et al., “Desinformación y migración venezolana. El caso Ecuador”, *Cuaderno del Centro de Estudios en Diseño y Comunicación* 161 (2023 de 2022): 107-23.

Colombia

Las elecciones presidenciales de 2018 y 2022 en Colombia se desarrollaron en un escenario de “guerra digital” dominado por bodegas políticas que coordinan narrativas polarizantes, desinformación y ataques afectivos para controlar la agenda y debilitar la deliberación pública.⁵¹ Estas estructuras operan como dispositivos tecnopolíticos que refuerzan burbujas ideológicas y reducen el debate a una lógica binaria de confrontación.

Entre 2019 y 2022, la convergencia de bulos y narrativas distorsionadas derivó en ataques de odio con efectos reputacionales y episodios de violencia, instrumentalizando *fake news* para activar rechazo social y cancelación digital.⁵² Ello profundizó la polarización y minó la confianza en autoridades, lo cual supuso una ingobernabilidad simbólica.

En 2023, un panel Delphi destacó el uso de IA para crear contenidos y segmentar mensajes, lo que intensificó la propaganda computacional y la microsegmentación emocional en campañas.⁵³ Así, bodegas y desinformación potenciadas por IA erosionan la libertad de expresión y la integridad electoral.

Chile

En el caso de Chile, diversos estudios permiten identificar cómo la desinformación, los discursos de odio y las *fake news* incidieron en los procesos electorales recientes.

Por un lado, investigaciones sobre la elección presidencial de 2021 muestran que plataformas como WhatsApp fueron utilizadas para difundir desinformación y tácticas de “campana sucia”, en particular desde sectores de derecha contra la candidatura de Gabriel Boric. Se documentaron estrategias de sabotaje digital, desinformación y ataques coordinados que buscaron obstaculizar la deliberación democrática mediante

⁵¹ Jairo Eduardo Soto-Molina y Margarita Rosa De La Hoz Pertuz, “Bodegas Políticas y Guerra Digital en Colombia: Una Lectura Comparativa de las Estrategias Comunicativas de Derecha e Izquierda”, *Ciencia Latina Revista Científica Multidisciplinar* 9, n.º 4 (2025): 3441-63, https://doi.org/10.37811/cl_rcm.v9i4.19000.

⁵² Erynn Norza Céspedes, “PaZCiencia: Ataques de odio y PoLHicidio moral”, *Revista Logos Ciencia & Tecnología* 16, n.º 1 (2024): 6-7, <https://doi.org/10.22335/rlct.v16i1.1909>.

⁵³ Marcela López-Ponce et al., “Usos y riesgos de la inteligencia artificial en las campañas electorales 2023: Encuesta Delphi a expertos estratégicos de Colombia”, *ICONO14. Revista Científica de Comunicación y Tecnologías Emergentes* 22, n.º 2 (2024): 2-3, <https://doi.org/10.7195/ri14.v22i2.2078>.

rumores y mensajes polarizantes.⁵⁴ Este fenómeno se consolidó como parte de un repertorio de manipulación política caracterizado por la instrumentalización del discurso de odio y el recurso a prácticas antidemocráticas que aprovecharon vacíos regulatorios en torno al uso de redes de mensajería instantánea.

Asimismo, durante el proceso constituyente de 2020-2022, el uso de Twitter/X se convirtió en un espacio central para la propagación de los discursos de odio y desinformación dirigidos contra la Convención Constitucional y sus integrantes, particularmente vinculados a pueblos originarios o mujeres.⁵⁵

Del mismo modo, un reportaje de un canal de televisión en Chile (CHV) difundido en septiembre de 2025 puso en evidencia el despliegue de estrategias de desinformación digital en la campaña presidencial en Chile, a través del uso de bots y *astroturfing*, consistente en la creación y coordinación de cuentas automatizadas o falsas que viralizaban mensajes que daban la apariencia de respaldo ciudadano genuino. Si bien estas prácticas están siendo objeto de análisis por parte de la autoridad, se cree que el objetivo es afectar la reputación de las candidatas a la elección presidencial —las afectadas eran mujeres— influir en la agenda pública y condicionar la percepción electoral mediante la construcción de consensos ficticios.

Esta situación da cuenta de la vulnerabilidad institucional frente a estas estrategias que combinan desinformación y odio, lo que ha motivado la presentación de denuncias por eventuales delitos informáticos, pero también ponen de relieve la necesidad de tener marcos regulatorios claros y equilibrados que permitan prevenir, reprimir y sancionar estas prácticas, con el fin de proteger la democracia y garantizar la libertad de expresión. En definitiva, el caso chileno evidencia cómo el uso instrumental de bots y *astroturfing* en campañas presidenciales no es un fenómeno aislado, sino parte de una estrategia transnacional que combina tecnología, desinformación y odio político para socavar los cimientos mismos de la democracia.

⁵⁴ Marcelo Santos et al., “Jiji, Jaja, Sabotaje y Campaña Sucia: WhatsApp en las Elecciones Presidenciales Chilenas de 2021”, *Chasqui. Revista Latinoamericana de Comunicación* 1, n.º 156 (2024): 229-30, <https://doi.org/10.16921/chasqui.v1i156.5054>.

⁵⁵ Amalia Sepúlveda Koch, “Nueva Constitución para Chile: Twitter, Trending Topics y Plebiscito de Salida de 2022” (Tesis de Maestría, Universidad de Chile, 2025), 12-14.

Estrategias para mitigar el impacto de la IA como factor de desinformación y de propagación de discursos del odio

La irrupción de tecnologías basadas en IA, en particular la generativa, ha ampliado de manera significativa la capacidad de producir contenidos falsos y, con ello, difundir discursos de odio o desinformar. Este escenario obliga a diseñar respuestas articuladas que permitan conciliar los avances tecnológicos y, a su vez, proteger la democracia y los derechos humanos. Las estrategias más relevantes pueden agruparse en tres ejes: normativas, tecnológicas y educativas.

Estrategias normativas

Las estrategias normativas constituyen marcos jurídicos que delimitan responsabilidades y orientan la acción tanto estatal como de los actores privados frente a la desinformación y el odio digital.

Responsabilidad de los intermediarios

La responsabilidad de los intermediarios constituye un aspecto clave en la regulación de la desinformación, pues resulta más eficaz accionar contra proveedores o gestores de contenido que contra los autores materiales, muchas veces anónimos o ubicados en otros países.⁵⁶ En esta materia existen dos enfoques principales: el modelo europeo, plasmado en la Ley de Servicios Digitales (2022)⁵⁷, que clasifica a los intermediarios, impone deberes de diligencia y auditorías y concibe las plataformas como espacios cuasi públicos; y el modelo estadounidense, basado en el artículo 230 de la *Telecommunications Act* (1996),⁵⁸ que exime a los proveedores de responsabilidad por el contenido de terceros y favorece la autorregulación. A estos marcos se añaden las normas privadas de las propias plataformas, como las de Meta, YouTube o X, las cuales, aunque influyen globalmente, plantean riesgos de censura, opacidad y ausencia de garantías procesales.

⁵⁶ Jack M. Balkin, "Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation", *Davis Law Review* 51, n.º 3 (2018): 50.

⁵⁷ Reglamento (UE) 2022/2065 relativo a un mercado único de servicios digitales (Digital Services Act) (2022).

⁵⁸ Telecommunications Act of 1996 (Pub. L. No. 104-104 1996).

Regulación de algoritmos y transparencia algorítmica

Los algoritmos de personalización permiten amplificar los discursos de odio y la desinformación, por lo que se requiere establecer obligaciones legales, que concreten el principio de la transparencia algorítmica mediante auditorías independientes. Según Pauner, la importancia de la transparencia algorítmica tiene una dimensión externa e interna. La dimensión externa permite a cualquiera verificar la actuación de los programadores y verificar que se cumplan las obligaciones de transparencia, información y derecho de acceso. En su dimensión interna, la transparencia algorítmica permite comprobar el funcionamiento correcto del algoritmo.⁵⁹ La Corte Constitucional de Colombia⁶⁰ (2025) sostuvo que la transparencia algorítmica es una garantía esencial para justicia, equidad y dignidad.

Sistema de sanciones

Como una manera de combatir y mitigar el impacto de la IA y su influencia en la elaboración, difusión, y expansión de desinformación y discursos de odio, la aplicación de sanciones se vuelve crucial, aunque no excluyente. Así, por ejemplo, desde el derecho administrativo se puede imponer multas u órdenes de suspensión de servicios, entre otras medidas. Por otro lado, desde la perspectiva del derecho civil, se puede impetrar un sistema normativo que permita obtener una reparación integral del daño sufrido por quienes sean afectados directamente por la desinformación y los discursos del odio, tanto por quienes los producen como por aquellos que hayan facilitado el uso de tecnologías por omitir las normas que imponen deberes de diligencia. El derecho penal, en cambio, debe quedar reservado para los casos de mayor gravedad y siempre respetando los principios de tipicidad, legalidad, y proporcionalidad, cuidando que no ocurran afectaciones inadecuadas a la libertad de expresión. En materia penal, urge modernizar tipos penales o establecer nuevas figuras que sancionen, por ejemplo, la desinformación intencional y maliciosa con propósitos electorales o la propagación automatizada de discursos de odio. Del mismo modo, se puede extender la responsabilidad penal a personas jurídicas cuando estas se benefician económicamente del empleo inapropiado de la IA, entre otros escenarios.

⁵⁹ Cristina Pauner Chulvi, “Transparencia algorítmica en los medios de comunicación y las plataformas digitales”, *Revista Española de la Transparencia*, n.º 17 (2023): 123, <https://doi.org/10.51915/ret.308>.

⁶⁰ Sentencia T-067 de 2025 (Corte Constitucional de Colombia 26 de febrero de 2025).

Estrategias tecnológicas

Las estrategias tecnológicas que se diseñan para neutralizar o mitigar los efectos negativos del uso abusivo de la IA mediante filtros y sistemas de detección automatizada, etiquetado de contenidos generados por IA, auditorías externas y *fact-checking* automatizado. También incluyen la detección de bots maliciosos, la cual, aunque no elimina el problema, proporciona herramientas adicionales para hacer frente e intentar neutralizar campañas de desinformación y discursos del odio.

Estrategias educativas y de alfabetización digital

Por último, se debe avanzar hacia la profundización de la alfabetización digital mediante campañas ciudadanas que permitan fortalecer el pensamiento crítico en la educación y actualización de las mallas curriculares de periodistas y comunicadores sociales. La educación es esencial para promover resiliencia social y reforzar la cultura democrática frente a la manipulación informativa.

Conclusiones

Las nuevas tecnologías, especialmente de la IA y sus variantes generativas en los últimos años, plantean nuevos y significativos desafíos que, de no ser abordados de manera oportuna y profunda, pueden afectar de manera irreversible los cimientos de nuestras democracias y sistemas de derechos fundamentales. El análisis de los casos de Brasil, Ecuador, Colombia y Chile revela que estas prácticas, al ser utilizadas estratégicamente en contextos electorales, no solo inciden en la configuración de giros políticos recientes, también debilitan pilares esenciales como la confianza institucional, la deliberación pública y la igualdad en la competencia electoral.

A lo largo del presente trabajo hemos intentado exponer cómo la IA, en especial la generativa, ha elevado exponencialmente la capacidad no solo de producir contenidos falsos, sino también de difundirlos de manera masiva y difícil de detectar. Este fenómeno ha consolidado estrategias más o menos coordinadas, como el *astroturfing* y el uso de bots que permiten simular apoyo ciudadano y condicionar la agenda pública y ha impactado campañas políticas en prácticamente todo el continente. En paralelo, los discursos de odio se han intensificado, dirigidos con frecuencia contra colectivos vulnerables, reforzando la polarización política y afectiva.

Ante estos nuevos desafíos, la investigación ha destacado la necesidad de respuestas regulatorias sólidas y proporcionales según las cuales la responsabilidad de los intermediarios emerja como un componente clave. El modelo europeo, con la Ley de Servicios Digitales, ofrece un marco que impone obligaciones diferenciadas, sistemas de notificación, apelación y auditorías algorítmicas. Por su parte, el modelo estadounidense, basado en el artículo 230 de la *Telecommunications Act*, privilegia la autorregulación, lo que muestra la importancia de que los países latinoamericanos evalúen críticamente qué aproximación adoptar, asegurando mecanismos que no se traduzcan en censura arbitraria ni en impunidad digital.

Dentro de este esquema, resulta imperioso avanzar en un marco conceptual común en el cual se identifiquen prácticas como la desinformación, discriminación y violencia digital y en el que se establezca la transparencia algorítmica como principio rector, ya que esta constituye una garantía esencial para que los sistemas automatizados sean compatibles con la justicia y la dignidad humana. De este modo, las plataformas deberían estar obligadas a someterse a auditorías independientes, lo que permitiría identificar sesgos, errores o prácticas discriminatorias que amplifican la desinformación y el odio.

Asimismo, resulta indispensable articular un sistema de sanciones graduales y proporcionales que contemple medidas administrativas y civiles —como multas o reparación de daños— aunque limitando la intervención penal a los supuestos de mayor gravedad, tales como la propagación automatizada de discursos de odio o la desinformación maliciosa con fines electorales. Del mismo modo debería evaluarse la extensión de responsabilidad penal a personas jurídicas que se beneficien de la manipulación informativa.

A las medidas regulatorias indicadas deben implementarse estrategias tecnológicas tales como *fact-checking* automatizado, la detección de bots, el etiquetado de contenidos generados por IA, y, que a su vez, estas se complementen con políticas educativas de alfabetización digital que fortalezcan la resiliencia ciudadana frente a la manipulación.

En suma, el estudio concluye que la regulación no puede limitarse a reaccionar ante los riesgos de la IA, sino que debe integrar prevención, reparación y represión de manera equilibrada, con el fin de evitar la sobreregulación excesiva e irracional. Solo mediante la combinación de regulación normativa, control tecnológico y

educación crítica será posible enfrentar los desafíos que plantea la era digital y resguardar la integridad de los procesos políticos en el siglo XXI.

Referencias

- Arce-García, Sergio, Elías Said-Hung, y Daria Mottareale-Calvanese. "Tipos de campaña Astroturfing de contenidos desinformativos y polarizados en tiempos de pandemia en España". *ICONO14. Revista Científica de Comunicación y Tecnologías Emergentes* 21, n.º 1 (2023). <https://doi.org/10.7195/ri14.v21i1.1890>.
- Astudillo, Jorge. "Desinformación: aproximación conceptual, riesgos y remedios". *Revista Derecho UCP* 94 (2024): 55-97.
- Bachmann, Ingrid, Sabela Valenzuela, Lionel Brossi, et al. *El fenómeno de la desinformación y el desafío de corregirla: Una revisión bibliográfica*. 2021. <https://bit.ly/4e9KnCn>.
- Balkin, Jack M. "Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation". *Davis Law Review* 51, n.º 3 (2018).
- Ballesteros-Aguayo, Laura, Rosa Magallón-Rosa, y María Lamuedra-Graván. "Desinformación y fact-checking en las elecciones en Francia de 2022: Guerra de Ucrania y polarización". *Estudos em Comunicação* 38 (2024). <https://doi.org/10.5209/hics.92238>.
- Canavilhas, João, y Juliana Colussi. "WhatsApp como plataforma de desinformación: Estudio de caso de las elecciones presidenciales brasileñas". *Visual Review* 9 (2022): 2-10. <https://doi.org/10.37467/revvisual.v9.3519>.
- Cervantes Hernández, Rubén, Perla Ivonne Gallegos Flores, Jorge Alberto Esponda Pérez, Paola Margarita Chaparro-Medina, y Adso Eduardo Gutiérrez Espinoza. "Estrategias contra la desinformación en redes socio digitales". *Revista Cuestiones Políticas* 9, n.º 3 (2025): 2123-30. https://doi.org/10.37811/cl_rcm.v9i3.17837.
- Chulvi, Berta, Alejandro Toselli, y Paolo Rosso. "Fake news and hate speech: Language in common". *Information Processing & Management* 59, n.º 6 (2022): 103093. <https://doi.org/10.48550/arXiv.2212.02352>.
- Cinelli, Matteo, Andraž Pelicon, Igor Mozetič, Walter Quattrociocchi, Petra Kralj Novak, y Fabiana Zollo. "Online Hate: Behavioural Dynamics and Relationship with Misinformation". Preprint, arXiv, 2021. <https://doi.org/10.48550/ARXIV.2105.14005>.
- Citron, Danielle Keats, y Robert Chesney. "Deep fakes: A looming challenge for privacy, democracy, and national security". *California Law Review* 107 (2019): 1753. <https://doi.org/10.15779/Z38RV0D15J>.
- Collin v. Smith (Supreme Court of the United States 16 de octubre de 1978).
- Colussi, Juliana. "Desinformación, política y religión: Un análisis de los contenidos compartidos por WhatsApp durante la campaña presidencial de Brasil en 2018". *Revista Ibérica de Sistemas e Tecnologias de Informação (RISTI)*, n.º E35 (2020): 477-89.

- Corte Suprema de Estados Unidos. *Brandenburg v. Ohio*, 395 U.S. 444. 1969.
- Dworkin, Ronald. *A matter of principle*. Harvard University Press, 1985.
- European Commission: Directorate-General for Communications Networks, Content and Technology. 2018. *A multi-dimensional approach to disinformation – Report of the independent High level Group on fake news and online disinformation*. Publications Office. <https://doi.org/10.2759/739290>.
- Ferrara, Emilio, Onur Varol, Clayton B. Davis, Filippo Menczer, y Alessandro Flammini. “The rise of social bots”. *Communications of the ACM* 59, n.º 7 (2016): 96-104.
- Fontenla-Pedreira, Julia, Juan-Manuel Corbacho-Valencia, y Alba Silva-Rodríguez. “Regulación de la desinformación: enfoques multidisciplinares y desafíos legislativos globales en 2024”. *Revista Mediterránea de Comunicación* 16, n.º 2 (2025): e29218. <https://doi.org/10.14198/MEDCOM.29218>.
- García-Orosa, Berta. “Disinformation, social media, bots, and astroturfing: the fourth wave of digital democracy”. *Profesional de la información* 30, n.º 6 (2021): e300603. <https://doi.org/10.3145/epi.2021.nov.03>.
- Girón Sierra, José Manuel. *Introducción a la inteligencia artificial: La tecnología que nos cambiará para siempre*. Sekotia, 2023.
- Guadarrama Torres, Jorge J. M., Laura Cristina Álvarez Garzón, y María Ángeles Molina Picazo. “Desinformación en la era digital: abordaje de la fórmula de la posverdad en la comunicación política”. *Scripta Mundi* 4, n.º 1 (2025): 10-30. <https://doi.org/10.53591/scmu.v4i1.2179>.
- Ireton, Cherilyn, y Julie Posetti. *Journalism, fake news & disinformation: handbook for journalism education and training*. Unesco, 2018.
- Jenke, Libby. “Affective polarization and misinformation belief”. *Political Behavior* 46, n.º 2 (2024): 825-84. <https://doi.org/10.1007/s11109-022-09851-w>.
- López-Paredes, Marco, y Andrea Carrillo-Andrade. “Ultramediaciones en contextos de analfabetismo digital en Ecuador: aproximación a la desinformación y malinformación”. *Austral Comunicación* 14, n.º 3 (2025): e01430. <https://doi.org/10.26422/aucom.2025.1403.lap>.
- López-Ponce, Marcela, Daniel Barredo-Ibáñez, y Hada Sánchez Gonzales. “Usos y riesgos de la inteligencia artificial en las campañas electorales 2023: Encuesta Delphi a expertos estratégicos de Colombia”. *ICONO14. Revista Científica de Comunicación y Tecnologías Emergentes* 22, n.º 2 (2024): e2078. <https://doi.org/10.7195/ri14.v22i2.2078>.
- Moreira, Breno. “Desinformación en las declaraciones políticas de los líderes evangelistas en las elecciones de 2022 en Brasil”. *Miguel Hernández Communication Journal* 14, n.º 1 (2023): 109-31. <https://doi.org/10.21134/mhjourn.v14i.1813>.
- Nogales-Bocio, Antonia Isabel, Julieti Sussi de Oliveira, y Carmen Marta-Lazo. “Estrategias locales de verificación y fact-checking en Brasil: La red nacional de combate a la

- desinformación, COAR Noticias, Colectivo Bereia, Verifica RS y Agencia Tatu”. *Doxa Comunicación* 41 (2025): 395-415. <https://doi.org/10.31921/doxacom.n41a2906>.
- Norza Céspedes, Eryvn. “PaZCiencia: Ataques de odio y PoLHicidio moral”. *Revista Logos Ciencia & Tecnología* 16, n.º 1 (2024): 5-14. <https://doi.org/10.22335/rclt.v16i1.1909>.
- Pauner Chulvi, Cristina. “Transparencia algorítmica en los medios de comunicación y las plataformas digitales”. *Revista Española de la Transparencia*, n.º 17 (2023): 107-36. <https://doi.org/10.51915/ret.308>.
- Rebollo Delgado, Laura. *Inteligencia artificial y derechos fundamentales*. Dykinson, 2023.
- Reglamento (UE) 2022/2065 relativo a un mercado único de servicios digitales (Digital Services Act) (2022).
- Ribeiro Mota, Denysson Axel, Gracy Kelli Martins, y Denise Braga Sampaio. “Régimen de información y propagación de la desinformación: Los desafíos del Marco Civil de la Internet y la Ley de Fake News en Brasil”. *Actas de las 7ma Jornadas de Intercambio y Reflexión acerca de la Investigación en Bibliotecología* (La Plata), 2024, 1-13.
- Rojó-Martínez, Juan Manuel, y Isabel Crespo-Martínez. “Lo político como algo personal: Una revisión teórica sobre la polarización afectiva”. *Revista de Ciencia Política* 43, n.º 1 (2023): 25-48.
- Rossel Castagneto, María Lorena. “Regular el odio digital: Entre la libertad de expresión y la protección de colectivos vulnerables en Chile”. *Universitas XXI* 43 (2025): 95-122. <https://doi.org/10.17163/uni.n43.2025.04>.
- Ruiz-Mora, Isabel, Antonio Enrique López-Carrión, y David Guerrero-Navarro. “Estrategias de desinformación de VOX: Cobertura mediática de su discurso parlamentario sobre la Agenda 2030”. *Palabra Clave* 28, n.º s1 (2025). <https://doi.org/10.5294/pacla.2025.28.s1.12>.
- Sánchez Moscoso, Agustín Alexander, Sabrina Candela, y Ángel Torres Toukoumidis. “Desinformación y migración venezolana. El caso Ecuador”. *Cuaderno del Centro de Estudios en Diseño y Comunicación* 161 (2023 de 2022): 107-23.
- Santos, Marcelo, Jorge Valdebenito Allendes, y Jorge Luis Ortiz Fuentes. “Jiji, Jaja, Sabotaje y Campaña Sucia: WhatsApp en las Elecciones Presidenciales Chilenas de 2021”. *Chasqui. Revista Latinoamericana de Comunicación* 1, n.º 156 (2024): 227-48. <https://doi.org/10.16921/chasqui.v1i156.5054>.
- Sentencia T-067 de 2025 (Corte Constitucional de Colombia 26 de febrero de 2025).
- Sepúlveda Koch, Amalia. “Nueva Constitución para Chile: Twitter, Trending Topics y Plebiscito de Salida de 2022”. Tesis de Maestría, Universidad de Chile, 2025.
- Slipczuk, Marian. *Una guía para principiantes para establecer una organización de fact-checking en América Latina y el Caribe*. Unesco, 2022.
- Soto-Molina, Jairo Eduardo, y Margarita Rosa De La Hoz Pertuz. “Bodegas Políticas y Guerra Digital en Colombia: Una Lectura Comparativa de las Estrategias Comunicativas de

- Derecha e Izquierda”. *Ciencia Latina Revista Científica Multidisciplinar* 9, n.º 4 (2025): 3441-63. https://doi.org/10.37811/cl_rcm.v9i4.19000.
- Telecommunications Act of 1996 (Pub. L. No. 104-104 1996).
- Tomalá Medina, Francisco. “Campañas electorales, redes sociales y fake news”. *Justicia Electoral y Democracia* 5, n.º 6 (2022): 55-59.
- Torres Espinoza, Gabriel. “El nuevo frente de las fake news en las democracias: formas de contrarrestar los efectos de la desinformación”. En *Impactos de la desinformación en un mundo cambiante*, editado por Gabriel Torres Espinoza y Omar Guillermo García Santiago. Innovación Educativa, 2025.
- Vasist, Pratibha N., Dipjyoti Chatterjee, y Sreya Krishnan. “The Polarizing Impact of Political Disinformation and Hate Speech: A Cross-Country Configurational Narrative”. *Information Systems Frontiers* 26, n.º 2 (2024): 670.
- Vestri, Gabriele. *Diccionario de términos para comprender la transformación digital*. Aranzadi, 2023.
- Viñán Carrasco, Lourdes, Mariam Murillo Naranjo, Paula Méndez Naranjo, y William X. Castro Ortiz. “Construcción de la Opinión Pública en un Contexto de Desinformación y Polarización Previo a las Elecciones Generales 2025 de Ecuador”. *Polo del Conocimiento* 10, n.º 1 (2025): 2072-74.