



Ciencia Latina Revista Científica Multidisciplinar, Ciudad de México, México.
ISSN 2707-2207 / ISSN 2707-2215 (en línea), julio-agosto 2026,
Volumen 10, Número 4.

https://doi.org/10.37811/cl_rcm.v10i4

LIBERTAD DE EXPRESIÓN DIGITAL: DESAFÍOS JURÍDICOS, ÉTICOS Y TECNOLÓGICOS EN LA GOBERNANZA DE PLATAFORMAS E INTELIGENCIA ARTIFICIAL

**DIGITAL FREEDOM OF EXPRESSION: LEGAL, ETHICAL, AND
TECHNOLOGICAL CHALLENGES IN THE GOVERNANCE OF
PLATFORMS AND ARTIFICIAL INTELLIGENCE**

Elvis Elías Vallejos Bautista

Universidad para el Desarrollo Andino, Perú

Mara Lisseti Gonzales Fernández

Universidad Particular de Chiclayo, Perú

Diomédes Alexis Centurión Pérez

Universidad Particular de Chiclayo, Perú

Kemnedy Liliana Vela Frías

Universidad Particular de Chiclayo, Perú

Kyka Shaneet Santa Cruz Villegas

Universidad César Vallejo, Perú

Libertad de Expresión Digital: Desafíos Jurídicos, Éticos y Tecnológicos en la Gobernanza de Plataformas e Inteligencia Artificial

Elvis Elías Vallejos Bautista¹

valba260680@gmail.com

<https://orcid.org/0000-0002-7178-7336>

Universidad para el Desarrollo Andino
UDEA Perú

Diomédes Alexis Centurión Pérez

diomedescenper@gmail.com

<https://orcid.org/0009-0009-1290-2638>

Universidad Particular de Chiclayo
Perú

Kyka Shaneet Santa Cruz Villegas

shaneetsantacruz37@gmail.com

<https://orcid.org/0000-0003-0374-7639>

Universidad César Vallejo
Perú

Mara Lisseti Gonzales Fernández

maragf345@gmail.com

<https://orcid.org/0009-0007-6595-7566>

Universidad Particular de Chiclayo
Perú

Kennedy Liliana Vela Frías

kennedylilianavelafrias@gmail.com

<https://orcid.org/0009-0009-0018-8844>

Universidad Particular de Chiclayo
Perú

RESUMEN

La presente investigación analiza los desafíos jurídicos, tecnológicos y éticos que enfrenta la libertad de expresión en el entorno digital, considerando la influencia de las plataformas digitales, los algoritmos y la inteligencia artificial en la configuración del espacio público contemporáneo. Se desarrolló un estudio de enfoque cualitativo, de tipo exploratorio-descriptivo, con diseño no experimental y basado en una revisión documental de literatura científica reciente y documentos normativos internacionales relacionados con derechos humanos, gobernanza digital e inteligencia artificial. La selección y análisis de fuentes permitieron identificar las principales tensiones entre la protección de la libertad de expresión y los nuevos mecanismos de regulación tecnológica. Los resultados evidencian que la intermediación algorítmica, la moderación automatizada y la desinformación constituyen factores que pueden afectar la visibilidad, pluralidad y autonomía del debate público cuando operan bajo criterios poco transparentes. Asimismo, se determina que los modelos tradicionales de regulación resultan insuficientes frente al poder de las plataformas digitales, requiriéndose esquemas de gobernanza democrática sustentados en transparencia algorítmica, rendición de cuentas, supervisión humana, alfabetización digital y participación multisectorial. Se concluye que la protección efectiva de la libertad de expresión digital exige integrar garantías jurídicas, principios éticos y mecanismos tecnológicos responsables orientados a preservar el pluralismo democrático.

Palabras clave: libertad de expresión digital; plataformas digitales; inteligencia artificial; gobernanza digital; algoritmos; derechos humanos

¹ Autor principal

Correspondencia: valba260680@gmail.com

Digital Freedom of Expression: Legal, Ethical, and Technological Challenges in the Governance of Platforms and Artificial Intelligence

ABSTRACT

This research analyzes the legal, technological, and ethical challenges faced by freedom of expression in the digital environment, considering the influence of digital platforms, algorithms, and artificial intelligence on the configuration of the contemporary public sphere. A qualitative, exploratory-descriptive study was conducted using a non-experimental design based on a documentary review of recent scientific literature and international regulatory documents related to human rights, digital governance, and artificial intelligence. The selection and analysis of sources made it possible to identify the main tensions between the protection of freedom of expression and emerging forms of technological regulation. The findings show that algorithmic mediation, automated content moderation, and disinformation constitute factors that may affect the visibility, plurality, and autonomy of public debate when they operate under insufficient transparency standards. Likewise, the study determines that traditional regulatory approaches are inadequate to address the growing power of digital platforms, requiring democratic governance models based on algorithmic transparency, accountability, human oversight, digital literacy, and multi-stakeholder participation. It is concluded that the effective protection of digital freedom of expression requires the integration of legal safeguards, ethical principles, and responsible technological mechanisms aimed at preserving democratic pluralism.

Keywords: digital freedom of expression; digital platforms; artificial intelligence; digital governance; algorithms; human rights.

*Artículo recibido 20 junio 2026
Aceptado para publicación: 20 julio 2026*



INTRODUCCIÓN

Dentro de las sociedades democráticas, la libertad de expresión se consagra como uno de los derechos esenciales, pues garantiza la participación pública, el acceso a información plural y el control ciudadano. Sin embargo, la transformación digital ha modificado sus condiciones de ejercicio debido al creciente papel de plataformas digitales, algoritmos y sistemas automatizados en la selección, priorización y moderación de contenidos.

Aunque las tecnologías digitales han ampliado la participación y democratizado la producción informativa, también han generado nuevos riesgos asociados con la intermediación algorítmica, la desinformación, la moderación automatizada y la concentración del poder comunicativo en actores privados. Por ello, resulta necesario replantear la protección jurídica de la libertad de expresión frente a infraestructuras tecnológicas que repercuten en la visibilidad, alcance e impacto de los discursos.

La investigación parte del constitucionalismo democrático, el enfoque de derechos humanos, la gobernanza digital y la ética de la inteligencia artificial, reconociendo que la libertad de expresión posee una dimensión individual y colectiva, mientras que las tecnologías deben regirse por transparencia, explicabilidad, rendición de cuentas y supervisión humana. En este marco, las plataformas digitales adquieren responsabilidades por su capacidad de intervenir en la estructuración de la esfera pública.

Esta transformación es respaldada por estudios recientes. Seipp et al. (2023) sostienen que los algoritmos de recomendación tienen influencia en la visibilidad de los discursos; Gorwa y Veale (2024) advierten que la inteligencia artificial redistribuye decisiones entre desarrolladores, plataformas y usuarios, generando nuevos desafíos de responsabilidad; y Basanta y Azurmendi (2025) destacan que las plataformas poseen responsabilidades propias debido a su incidencia en la esfera pública. Asimismo, la investigación sobre desinformación resalta la necesidad de combinar regulación proporcional, alfabetización digital y participación ciudadana.

Pese al avance de los estudios sobre regulación tecnológica, persiste la necesidad de integrar las dimensiones jurídica, ética, tecnológica y democrática para comprender los desafíos de la libertad de expresión digital. En particular, se requiere superar enfoques limitados a eliminar contenidos y promover modelos de gobernanza que aseguren transparencia en los procesos que determinan la circulación de información.



En este escenario, caracterizado por el crecimiento sostenido de las plataformas digitales, la inteligencia artificial y normativas internacionales como las orientaciones de la UNESCO y el Digital Services Act de la Unión Europea, el objetivo de la investigación se fundamenta en el análisis de los desafíos jurídicos, tecnológicos y éticos que plantea la protección de la libertad de expresión digital, evaluando la influencia de las plataformas digitales y los sistemas de inteligencia artificial en la conformación de la esfera pública digital, con el objetivo de proponer principios orientados a una gobernanza democrática sustentada en la transparencia, la rendición de cuentas y la salvaguarda de los derechos humanos.

METODOLOGÍA

El enfoque de la investigación fue cualitativo, orientado a comprender e interpretar los desafíos jurídicos, tecnológicos y éticos que enfrenta la libertad de expresión en el ecosistema digital. Desde una perspectiva crítico-interpretativa, el estudio analizó el impacto de las plataformas digitales, los algoritmos y la inteligencia artificial en la configuración del espacio público contemporáneo, a través de la revisión de publicaciones científicas, instrumentos normativos y documentos institucionales.

También es una investigación documental de alcance descriptivo-interpretativo y diseño no experimental, transversal, centrada en el análisis de fuentes secundarias sin manipulación de variables ni intervención sobre una realidad empírica, que permitió identificar tendencias y convergencias presentes en la literatura reciente sobre gobernanza digital y protección de los derechos fundamentales. La unidad de análisis estuvo integrada por documentos científicos, jurídicos e institucionales relacionados con libertad de expresión digital, regulación de plataformas, gobernanza tecnológica, inteligencia artificial, algoritmos y desinformación. Se revisaron artículos publicados en revistas académicas indexadas, informes de organismos internacionales, instrumentos internacionales de derechos humanos y marcos regulatorios especializados, incluyendo documentos de la UNESCO y estudios de referencia sobre gobernanza algorítmica y ética digital.

La selección documental se realizó mediante muestreo intencional no probabilístico, aplicando criterios de pertinencia temática, actualidad y calidad académica. Se priorizaron publicaciones correspondientes al período 2022-2026, directamente vinculadas con el objeto de estudio, disponibles en español e inglés y provenientes de fuentes científicas o institucionales reconocidas.



Se excluyeron documentos duplicados, literatura divulgativa sin respaldo académico y publicaciones que no abordaban el fenómeno desde una perspectiva jurídica, tecnológica o ética.

La información se organizó mediante una matriz de análisis documental, estructurada en siete categorías temáticas: libertad de expresión, límites jurídicos digitales, plataformas tecnológicas, algoritmos, inteligencia artificial, desinformación y desafíos éticos y gobernanza digital, registrándose los principales aportes conceptuales, enfoques analíticos y contribuciones al objetivo de investigación, facilitando la identificación de tendencias, conexiones analíticas y aspectos aún pendientes de desarrollo en la producción académica existente. El procesamiento de la información se desarrolló mediante análisis hermenéutico y triangulación teórica, contrastando los aportes provenientes del derecho constitucional, la gobernanza digital, la ética de la inteligencia artificial y los estudios sobre plataformas digitales. Este procedimiento permitió interpretar de manera integrada las distintas perspectivas, identificar convergencias y divergencias entre los autores y construir un marco analítico orientado a explicar los desafíos contemporáneos de la libertad de expresión en el entorno digital.

En cuanto a las consideraciones éticas, la investigación respetó los principios de integridad científica, transparencia metodológica y reconocimiento de la autoría intelectual, a través del citado de fuentes según los estándares propuestos por las normas APA (7ª edición); además, no participaron seres humanos ni se recopilieron datos personales, debido a que se trata de una investigación documental.

Como principal limitación, se reconoce el carácter dinámico del ecosistema digital, especialmente en inteligencia artificial y regulación tecnológica, cuya rápida evolución puede modificar los escenarios analizados. Asimismo, el uso exclusivo de fuentes documentales limita la contrastación empírica de los hallazgos, por lo que futuras investigaciones podrían complementarlos mediante estudios de campo.

RESULTADOS

La revisión documental permitió identificar las principales tendencias, tensiones y vacíos de investigación sobre la libertad de expresión en entornos digitales, destacando los desafíos asociados a la regulación de plataformas, la gobernanza algorítmica, la inteligencia artificial y la protección de derechos fundamentales. Los resultados se organizan en categorías temáticas que integran los principales aportes de la literatura especializada y orientan el análisis de los hallazgos obtenidos.



1. Libertad de expresión como fundamento democrático

En los sistemas democráticos sustentados en el Estado constitucional de derecho, la libertad de expresión constituye un derecho esencial que permite la libre circulación de información, opiniones e ideas en el espacio público, sin interferencias arbitrarias. Su importancia trasciende la dimensión individual, al desempeñar una función esencial en la consolidación del pluralismo, la deliberación pública, la participación ciudadana y el control democrático del poder, elementos indispensables para la formación autónoma y plural de la opinión pública en una sociedad democrática.

La transformación digital ha ampliado el ejercicio de este derecho, pero también ha generado riesgos asociados al poder de las plataformas digitales y los sistemas algorítmicos, que condicionan la visibilidad y circulación de contenidos. Por ello, su protección requiere no solo limitar restricciones estatales ilegítimas, sino también garantizar un espacio digital plural, transparente y accesible.

En el ámbito internacional, tanto la Declaración Universal de Derechos Humanos (1948), como el Pacto Internacional de Derechos Civiles y Políticos (1966) y la Convención Americana sobre Derechos Humanos (1969) consolidaron la libertad de expresión como un derecho fundamental, estableciendo que sus limitaciones deben respetar los principios de legalidad, necesidad y proporcionalidad, además de prohibir la censura previa.

Esta evolución refleja el tránsito desde una protección frente a la intervención estatal hacia un enfoque que exige condiciones efectivas para el pluralismo informativo en entornos digitales. Por ello, la UNESCO (2023) considera que es necesario el fortalecimiento de la transparencia y la responsabilidad de las plataformas digitales. Asimismo, Knapp y Piszcz (2024), Gsenger (2025), Papaevangelou (2023) y Bayer (2022) coinciden en que la moderación de contenidos debe sustentarse en reglas claras, supervisión y garantías procedimentales.

Aunque las plataformas digitales han ampliado la participación ciudadana, también han facilitado la difusión de desinformación, generando tensiones entre la libertad de expresión y la protección del debate público. En efecto, Freiling et al. (2023) sostienen que la aceptación de medidas contra la desinformación depende del daño percibido sobre la democracia, mientras que Ó Fathaigh et al. (2025) y Atzori (2024) advierten que toda regulación debe guardar respeto por los principios de legalidad, proporcionalidad y necesidad.



En esa misma línea, los algoritmos influyen en la configuración de la opinión pública al determinar la visibilidad de la información y reproducir posibles sesgos (García y Calvo, 2022; Pedraza y Martínez, 2022), a lo cual se suman tecnologías de vigilancia y recopilación masiva de datos que pueden desalentar la participación ciudadana, afectando la privacidad y la autonomía individual (Palmiotto y Menéndez, 2023).

En síntesis, la literatura reciente coincide en que la libertad de expresión continúa siendo un pilar de la democracia, aunque su protección exige responder a los desafíos de la transformación digital, resultando indispensable fortalecer la transparencia algorítmica, la responsabilidad de las plataformas y modelos de gobernanza basados en supervisión y rendición de cuentas (Gorwa y Veale, 2024).

2. Límites jurídicos de la libertad de expresión en el entorno digital

La libertad de expresión representa un derecho esencial para la consolidación y funcionamiento de los sistemas democráticos; sin embargo, su ejercicio no es ilimitado, debido a que debe ser compatible con la protección de otros derechos fundamentales y valores constitucionalmente reconocidos, tales como la dignidad humana, la intimidad, la igualdad, el honor y la seguridad pública.

En el entorno digital, esta ponderación adquiere mayor complejidad debido al alcance global, la velocidad de difusión y la permanencia de los contenidos, así como por la creciente intervención de plataformas privadas en la organización del espacio comunicativo.

La transformación tecnológica ha desplazado el debate tradicional sobre la censura estatal hacia nuevas formas de regulación ejercidas por plataformas digitales mediante algoritmos y políticas de moderación. Barata (2022) sostiene que el análisis contemporáneo debe considerar el poder regulatorio de estos actores privados, cuya capacidad para determinar la visibilidad de los discursos exige estándares de transparencia, proporcionalidad y responsabilidad compatibles con los principios democráticos.

En un Estado constitucional, las limitaciones para ejercitar la libertad de expresión solo resulta legítima cuando cumple con los principios de legalidad, finalidad legítima, necesidad y proporcionalidad reconocidos por el Derecho Internacional de los Derechos Humanos, criterios que adquieren especial importancia en los entornos digitales, donde la amplia difusión de un mensaje no representa, por sí sola, un fundamento suficiente para limitar expresiones protegidas.

Entre los principales retos se encuentran los discursos de odio y la desinformación. Zacarias et al. (2022) señalan que el anonimato y la viralidad de las plataformas favorecen la expansión del discurso de odio, aunque cualquier restricción debe valorar el contexto, la intención y la magnitud del daño. Respecto de la desinformación, Freiling et al. (2023) destacan que la aceptación social de las medidas regulatorias depende de la percepción del perjuicio generado sobre bienes públicos, mientras que Ó Fathaigh et al. (2025) advierten que la aplicación del Digital Services Act debe respetar los estándares internacionales de derechos humanos para evitar mecanismos indirectos de censura.

Otro cambio relevante es la transferencia de decisiones sobre la libertad de expresión hacia plataformas privadas que aplican reglas comunitarias y sistemas automatizados de moderación. Vázquez (2022) sostiene que estas empresas han asumido funciones semejantes a reguladores privados, mientras que Martínez (2022) afirma que la moderación solo resulta compatible con la libertad de expresión cuando se desarrolla mediante procedimientos transparentes, previsibles y sujetos a revisión. En este sentido, la legitimidad de estas decisiones requiere garantizar un verdadero debido proceso digital.

El principio de proporcionalidad constituye el principal criterio para resolver las tensiones entre el derecho a expresarse y otros derechos esenciales. Su aplicación exige verificar que la medida restrictiva persiga un fin legítimo, sea adecuada para alcanzarlo, constituya la alternativa menos restrictiva y mantenga un equilibrio razonable entre beneficios y afectaciones.

Coral y Coral (2022) sostienen que reconocer límites razonables fortalece la democracia al impedir que la libertad de expresión sea utilizada para vulnerar otros derechos. Asimismo, Ordoñez (2022) destaca que cualquier regulación sobre plataformas digitales debe respetar las exigencias del Sistema Interamericano de Derechos Humanos, especialmente la prohibición de censura previa y la exigencia de responsabilidades ulteriores basadas en criterios objetivos.

En consecuencia, los límites jurídicos de la libertad de expresión digital deben orientarse hacia un equilibrio democrático que permita enfrentar daños reales sin afectar el pluralismo, la diversidad de opiniones ni el libre intercambio de información.

3. Plataformas digitales y nuevos desafíos regulatorios

La consolidación de las plataformas digitales ha transformado el ejercicio de la libertad de expresión al convertirlas en actores centrales del espacio público. Su capacidad para organizar, priorizar y moderar contenidos mediante algoritmos ha desplazado el enfoque tradicional, centrado en las restricciones estatales, hacia nuevas formas de regulación privada; en consecuencia, la protección de este derecho exige garantizar transparencia, pluralismo y responsabilidad en los mecanismos que regulan la circulación de la información.

Basanta y Azurmendi (2025) sostienen que las plataformas digitales han evolucionado desde una función meramente intermediaria a ocupar un rol activo en la gestión del ecosistema informativo, por su capacidad para incidir en los procesos de acceso a la información y construcción de la opinión pública; por ello, la regulación digital debe reconocer que el poder comunicativo de estas empresas genera obligaciones vinculadas con la protección de derechos fundamentales.

Plataformas digitales como actores de gobernanza

Las plataformas digitales han evolucionado de simples intermediarias a actores de gobernanza del espacio público digital, al administrar la circulación, moderación y visibilidad de los contenidos mediante decisiones, cada vez más, apoyadas en inteligencia artificial.

Papaevangelou (2023) sostiene que la legitimidad de estas funciones requiere mecanismos de participación ciudadana, mientras que Gorwa y Veale (2024) destacan que la IA redistribuye las responsabilidades entre desarrolladores, plataformas y usuarios. Asimismo, Çelik (2026) advierte que la regulación debe adaptarse a las particularidades institucionales de cada Estado para equilibrar la libertad de expresión, la innovación y la rendición de cuentas.

En este contexto, Vázquez (2022) sostiene que la denominada “censura privada” representa un desafío democrático, pues las plataformas pueden restringir contenidos sin control judicial directo, lo que exige una moderación basada en criterios transparentes y mecanismos efectivos de revisión.

Algoritmos, visibilidad informativa y configuración del debate público

En el ecosistema digital, la libertad de expresión no depende solo de publicar información, sino también de las condiciones tecnológicas que determinan su alcance y visibilidad.

García y Calvo (2022) señalan que los algoritmos modifican la deliberación democrática al influir en la selección de la información disponible, mientras que Seipp et al. (2023) sostienen que las plataformas ejercen un “poder de opinión” al condicionar indirectamente la agenda pública mediante la priorización de determinados discursos.

Esta influencia presenta efectos duales: puede ampliar la participación y facilitar el acceso a información diversa, pero también favorecer la desinformación, la polarización y los sesgos cuando los sistemas priorizan criterios de interacción o rentabilidad. Freiling et al. (2023) evidencian que la percepción ciudadana sobre los daños democráticos de la desinformación influye en el respaldo a medidas regulatorias, mientras que Mortenson et al. (2026) relacionan la percepción de una alta circulación de noticias falsas con mayores niveles de insatisfacción democrática.

Asimismo, la microsegmentación algorítmica genera riesgos al limitar la exposición a perspectivas diversas y favorecer entornos informativos cerrados. Pedraza y Martínez (2022) advierten que la opacidad de estos sistemas dificulta identificar sesgos y restringe el control democrático, mientras que Echeverría y Rodríguez (2023) resaltan que la alfabetización digital resulta esencial para desarrollar capacidades críticas frente al consumo de información.

Por tanto, el desafío regulatorio no se limita al control de contenidos ilícitos, sino a garantizar que la intermediación algorítmica opere bajo estándares de transparencia, explicabilidad, rendición de cuentas y supervisión democrática.

Desinformación, confianza democrática y alfabetización digital

Uno de los desafíos más relevantes del entorno digital es la desinformación ya que afecta la confianza en las entidades estatales y la calidad de los procesos deliberativos democráticos.

En efecto, Amin y Hong (2026) sostienen que la información falsa reduce la confianza ciudadana, mientras que Altay (2026) evidencia, a partir de un estudio en 46 países, que esta problemática es mayor en contextos de menor fortaleza institucional. En consecuencia, la respuesta no debe limitarse a la regulación, sino incorporar el fortalecimiento de las capacidades críticas de la ciudadanía.

En este sentido, Echeverría y Rodríguez (2023) destacan la alfabetización digital como complemento indispensable de la regulación, y Freiling et al. (2026) muestran que la percepción del daño potencial de la información refuerza la necesidad de desarrollar pensamiento crítico para evaluar su confiabilidad.



En conjunto, estos hallazgos evidencian que enfrentar la desinformación requiere combinar regulación, responsabilidad de las plataformas y alfabetización digital para fortalecer la confianza democrática.

Gobernanza democrática de las plataformas digitales

La complejidad del ecosistema digital evidencia que la regulación de plataformas requiere una gobernanza colaborativa entre el Estado, las empresas tecnológicas, la academia, la sociedad civil y los usuarios, más allá de la autorregulación empresarial o la intervención estatal directa.

Papaevangelou (2023) sostiene que la participación ciudadana fortalece la legitimidad de la moderación de contenidos, mientras que Çelik (2026) destaca que la regulación debe adaptarse a cada contexto institucional. En este marco, el Digital Services Act constituye un referente al promover transparencia, evaluación de riesgos y supervisión de grandes plataformas.

En síntesis, la gobernanza democrática de las plataformas digitales debe buscar el equilibrio entre la innovación tecnológica, la tutela de derechos fundamentales y responsabilidad empresarial mediante instrumentos que aseguren la transparencia, la supervisión y la rendición de cuentas.

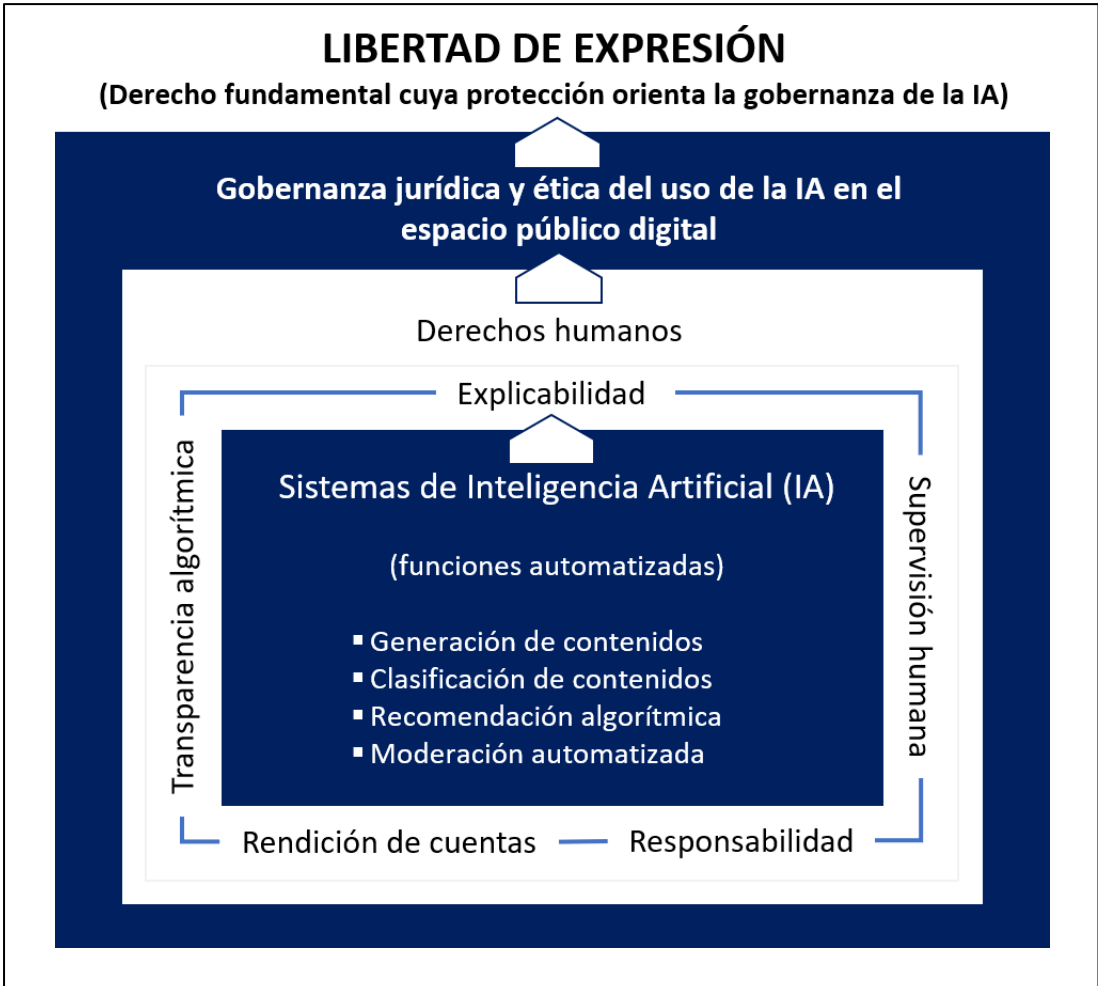
4. Inteligencia artificial y libertad de expresión

La incorporación de la inteligencia artificial (IA) ha intensificado los desafíos regulatorios del entorno digital al intervenir en la generación, clasificación, recomendación y moderación de contenidos. A diferencia de tecnologías anteriores, la IA no solo facilita la comunicación, sino que participa en la organización del discurso público mediante decisiones automatizadas que influyen en la visibilidad de las expresiones y en el acceso a la información.

Gorwa y Veale (2024) sostienen que la IA ha redistribuido la toma de decisiones entre desarrolladores, proveedores tecnológicos, plataformas y usuarios, generando dificultades para determinar responsabilidades cuando los sistemas automatizados afectan derechos fundamentales; por ello, debe comprenderse como una infraestructura de gobernanza y no solo como una herramienta tecnológica. Los sistemas inteligentes permiten gestionar grandes volúmenes de información y personalizar contenidos, pero también generan riesgos relacionados con sesgos algorítmicos, falta de explicabilidad y restricciones indebidas sobre discursos legítimos. Basanta y Azurmendi (2025) sostienen que las plataformas digitales han trascendido su condición de simples intermediarias para asumir responsabilidades en la configuración del espacio público y la protección del derecho a la información.

En la misma línea, Gsenger (2025) destaca que esta función exige fortalecer la transparencia de los sistemas de recomendación y moderación para generar confianza pública. Asimismo, Zheng, Zhou y Zeng (2023) plantean que la gobernanza de plataformas en la era de la inteligencia artificial debe sustentarse en principios de responsabilidad, transparencia y protección de los derechos humanos. Complementariamente, Bayer (2022) sostiene que las decisiones automatizadas que afectan la libertad de expresión deben garantizar procedimientos transparentes, motivación de las decisiones y mecanismos efectivos de revisión, como salvaguardas para la protección de los derechos fundamentales. La Figura 1 sintetiza el modelo de gobernanza jurídica y ética necesario para orientar el uso de la IA en la protección de la libertad de expresión digital.

Figura 1. Modelo conceptual de la gobernanza jurídica y ética del uso de la IA que protege la libertad de expresarse en el espacio público digital.



Nota: Elaboración propia.

La Figura 1 representa a la inteligencia artificial como un componente central del ecosistema comunicativo digital, destacando su intervención en la generación, clasificación, recomendación y

moderación automatizada de contenidos. Estos procesos influyen directamente en la organización del flujo informativo y en la configuración de la opinión pública.

Alrededor de este núcleo se articulan principios esenciales de gobernanza: transparencia algorítmica, explicabilidad, rendición de cuentas, supervisión humana, protección de derechos humanos y responsabilidad tecnológica. Estos elementos permiten orientar las decisiones automatizadas conforme a valores democráticos, reduciendo riesgos de arbitrariedad y fortaleciendo la legitimidad de las plataformas digitales.

Es decir, la legitimidad del uso de la IA depende de mecanismos permanentes de transparencia, supervisión y responsabilidad durante todo su ciclo de funcionamiento.

Por tanto, la protección de la libertad de expresión digital exige reconocer que los desafíos contemporáneos no se limitan al contenido de los discursos, sino también a las infraestructuras tecnológicas que determinan su circulación, visibilidad y alcance; en consecuencia, la gobernanza democrática de la IA requiere articular regulación jurídica, principios éticos, innovación tecnológica y participación ciudadana.

5. Tensiones éticas en la gobernanza de la libertad de expresión digital

La expansión del ecosistema digital evidencia que la protección de la libertad de expresión no constituye únicamente un desafío jurídico, sino también ético, debido a que algoritmos, plataformas digitales e inteligencia artificial intervienen en la selección, priorización y moderación de contenidos con efectos sobre los derechos fundamentales.

Entre los principales desafíos éticos están la transparencia algorítmica y la rendición de cuentas. Aunque los algoritmos facilitan la gestión de grandes volúmenes de información, la opacidad de sus criterios dificulta el control democrático. Bayer (2022) sostiene que las decisiones automatizadas deben ser comprensibles y revisables, mientras que Gsenger (2025) y Basanta y Azurmendi (2025) destacan la necesidad de garantizar información suficiente y mecanismos de supervisión sobre los sistemas de recomendación y moderación.

Otro reto central es la desinformación, cuya complejidad exige la corresponsabilidad del Estado, las plataformas, los medios, la academia y la ciudadanía. López-López et al. (2025) proponen estrategias integrales que incluyan alfabetización digital, fortalecimiento del periodismo y moderación

transparente. Asimismo, Freiling et al. (2023) señalan que la aceptación de estas medidas depende del daño percibido, mientras que Ó Fathaigh et al. (2025) advierten que toda restricción debe respetar los principios de necesidad, proporcionalidad y protección de los derechos humanos.

La legitimidad de la gobernanza digital también requiere participación ciudadana y cooperación entre instituciones públicas, plataformas, academia y sociedad civil. Papaevangelou (2023) sostiene que este enfoque fortalece la legitimidad de la moderación de contenidos, mientras que Trithara (2024) resalta su importancia para abordar los retos derivados de la inteligencia artificial y la desinformación.

Bajo este contexto, la protección de la libertad de expresión digital integra dimensiones jurídicas, éticas, tecnológicas y democráticas, que comprenden principios como legalidad, responsabilidad, explicabilidad, auditorías algorítmicas, inclusión y deliberación pública.

En consecuencia, una gobernanza ética debe sustentarse en la transparencia algorítmica, la supervisión humana, la rendición de cuentas, la protección del pluralismo y la mitigación responsable de la desinformación, combinando innovación tecnológica, responsabilidad ética y participación democrática para asegurar que se respeten los derechos fundamentales.

6. Tendencias regulatorias y perspectivas futuras

La transformación del ecosistema digital ha evidenciado la insuficiencia de los modelos regulatorios tradicionales frente a la influencia de las plataformas digitales y la inteligencia artificial en la circulación de información y el debate público. En respuesta, la regulación evoluciona hacia esquemas de gobernanza multinivel que articulan al Estado, empresas tecnológicas, organismos internacionales, academia y ciudadanía, buscando equilibrar innovación, libertad de expresión y protección de derechos fundamentales mediante transparencia, responsabilidad y legitimidad democrática.

De la autorregulación hacia la gobernanza democrática

Durante las primeras etapas de Internet predominó un modelo de amplia autonomía para las plataformas en la gestión de contenidos. Sin embargo, la expansión de la desinformación, los discursos de odio y la influencia algorítmica sobre la opinión pública evidenciaron sus limitaciones.

Basanta y Azurmendi (2025) sostienen que las plataformas deben asumir responsabilidades proporcionales a su impacto en la esfera pública, mientras que Papaevangelou (2023) destaca que la legitimidad de la moderación depende de procedimientos transparentes y participativos.



Por ello, la regulación contemporánea avanza hacia modelos de gobernanza democrática que integran obligaciones jurídicas, supervisión institucional y participación ciudadana, superando tanto la autorregulación absoluta como los mecanismos de control exclusivamente estatal.

Convergencia regulatoria internacional

La regulación digital presenta una tendencia hacia principios comunes vinculados con transparencia, rendición de cuentas, evaluación de riesgos y protección de derechos humanos. Instrumentos como el Digital Services Act de la Unión Europea y las recomendaciones de la UNESCO sobre gobernanza de plataformas reflejan esta convergencia regulatoria.

Ó Fathaigh et al. (2025) señalan que el Digital Services Act fortalece la responsabilidad de las grandes plataformas sin afectar injustificadamente la libertad de expresión, mientras que Gsenger (2025) destaca que la supervisión independiente y la transparencia algorítmica son esenciales para generar confianza pública.

No obstante, Çelik (2026) advierte que estos principios deben adaptarse a las particularidades constitucionales e institucionales de cada Estado, evitando modelos uniformes que desconozcan diferentes realidades jurídicas.

Inteligencia artificial y nuevos desafíos regulatorios

La inteligencia artificial constituye uno de los principales desafíos futuros de la regulación digital debido a su intervención en la generación, clasificación, recomendación y moderación de contenidos. Esta realidad exige garantías de transparencia, explicabilidad, supervisión humana y responsabilidad. Gorwa y Veale (2024) sostienen que la regulación de la IA debe considerar la complejidad de las decisiones automatizadas, donde participan desarrolladores, proveedores tecnológicos y plataformas digitales. Asimismo, Zheng et al. (2023) proponen una gobernanza basada en derechos humanos que permita compatibilizar innovación tecnológica y protección de libertades fundamentales.

Desde una perspectiva ética, Floridi (2023) plantea que la gobernanza de la inteligencia artificial debe sustentarse en principios éticos que garanticen la protección de la dignidad humana, los derechos fundamentales y la responsabilidad en el desarrollo y uso de las tecnologías digitales.

En consecuencia, los modelos regulatorios futuros deberán priorizar mecanismos preventivos basados en evaluación de riesgos, auditorías algorítmicas y supervisión continua.



Perspectivas futuras para la protección de la libertad de expresión

La protección futura de la libertad de expresión dependerá de la capacidad de los sistemas jurídicos y sociales para adaptarse a un entorno digital en constante transformación. Esto requiere fortalecer la cooperación internacional y consolidar modelos de gobernanza que articulen regulación jurídica, principios éticos y participación multisectorial.

López-López et al. (2025) sostienen que las medidas orientadas a contrarrestar la desinformación deben integrar regulación, alfabetización digital, fortalecimiento del periodismo y cooperación con plataformas tecnológicas. Asimismo, Freiling et al. (2023) evidencian que la legitimidad de las políticas regulatorias depende de su aceptación ciudadana basada en criterios de necesidad, proporcionalidad y respeto de derechos fundamentales.

En este sentido, las tendencias regulatorias muestran una transición desde modelos centrados en la eliminación de contenidos hacia enfoques integrales orientados a fortalecer la resiliencia informativa mediante transparencia, responsabilidad compartida y ciudadanía digital crítica.

Figura 2. Gobernanza colaborativa frente a los retos del ecosistema digital



Nota: Elaboración propia.

La Figura 2 representa la transición hacia un modelo de gobernanza colaborativa como respuesta a las dinámicas complejas del entorno digital. El modelo evidencia que no puede depender exclusivamente del Estado o de las plataformas, proteger la libertad de expresión sino de la articulación entre autoridades públicas, empresas tecnológicas, organismos internacionales, academia y ciudadanía.

En la parte central del modelo se ubican los principales desafíos del entorno digital: desinformación, inteligencia artificial, moderación automatizada, concentración del poder tecnológico y protección de derechos fundamentales. Estos factores justifican la necesidad de avanzar hacia mecanismos regulatorios preventivos, participativos y adaptativos.

Asimismo, la Figura 2 incorpora principios orientadores como transparencia algorítmica, rendición de cuentas, evaluación de riesgos, supervisión independiente, protección de derechos humanos y participación ciudadana, los cuales funcionan como criterios de legitimidad para las decisiones sobre circulación de información y uso de tecnologías inteligentes.

Finalmente, el modelo plantea tres resultados esperados: un ecosistema informativo más confiable, una protección efectiva de la libertad de expresarse y el fortalecimiento de la democracia digital, dado que la regulación digital requiere corresponsabilidad, transparencia y participación plural.

Por tanto, el futuro de la libertad de expresión digital dependerá de una gobernanza capaz de articular innovación tecnológica, regulación democrática y compromiso ético, garantizando que las transformaciones digitales permanezcan alineadas con la protección de los derechos esenciales.

DISCUSIÓN

Los resultados evidencian que la libertad de expresión atraviesa una transformación conceptual impulsada por la digitalización del espacio público. Mientras el paradigma clásico centraba su protección en la limitación del poder estatal, los principales desafíos actuales provienen también de las plataformas digitales, los algoritmos y la inteligencia artificial, que condicionan la circulación y visibilidad de la información; por tanto, la garantía de este derecho depende tanto de la facultad de expresar opiniones como de las condiciones tecnológicas que determinan su alcance.

Esta interpretación coincide con Seipp et al. (2023), quienes sostienen que los algoritmos de recomendación ejercen un “poder de opinión”, y con García y Calvo (2022), quienes destacan que la selección automatizada transforma las dinámicas de deliberación pública.



Asimismo, los resultados muestran que las plataformas digitales ya no son meras intermediarias, sino que asumen funciones con efectos regulatorios sobre el espacio público mediante la recomendación, priorización y moderación de contenidos, hallazgo que coincide con Basanta y Azurmendi (2025), quienes les atribuyen responsabilidades derivadas de su influencia en la esfera pública, y con Vázquez (2022), quien advierte que la “censura privada” plantea nuevos desafíos al Estado constitucional.

Ello justifica sustituir la autorregulación por mecanismos de transparencia, supervisión y responsabilidad compartida.

La investigación también evidencia que la desinformación y la inteligencia artificial constituyen riesgos complementarios para la libertad de expresión, al afectar tanto la confianza institucional como la producción, clasificación y difusión de la información. Esta interpretación coincide con Amin y Hong (2026), quienes destacan el impacto de la desinformación sobre la confianza pública, y con Gorwa y Veale (2024), quienes sostienen que la inteligencia artificial redistribuye la toma de decisiones entre desarrolladores, plataformas y usuarios. Asimismo, Zheng et al. (2023) y Bayer (2022) advierten que los sesgos algorítmicos pueden afectar el pluralismo informativo, por lo que proponen una gobernanza basada en transparencia, supervisión humana y responsabilidad.

Los resultados permiten afirmar que las acciones adoptadas para proteger la libertad de expresarse no deben limitarse a la eliminación de contenidos, ya que un enfoque exclusivamente restrictivo resulta insuficiente frente a la desinformación y la opacidad algorítmica. En esta línea, Ó Fathaigh et al. (2025) y Atzori (2024) sostienen que cualquier limitación debe encontrarse sustentada en criterios de legalidad, necesidad y proporcionalidad, mientras que Freiling et al. (2023) muestran que la legitimidad de estas políticas depende de su proporcionalidad y aceptación social. En consecuencia, resulta prioritario fortalecer la alfabetización digital, la transparencia algorítmica y la participación ciudadana.

Finalmente, un aporte de esta investigación consiste en integrar las dimensiones jurídica, ética y tecnológica en un marco de gobernanza digital. Los hallazgos muestran también que la protección de la libertad de expresarse depende de la interacción entre transparencia algorítmica, responsabilidad institucional, supervisión democrática y participación multisectorial, lo que permite comprender a las plataformas digitales como actores cuasirregulatorios sujetos a mecanismos permanentes de control.

En conjunto, la evidencia demuestra que la gobernanza colaborativa constituye el enfoque más adecuado para armonizar la innovación tecnológica con la protección de los derechos fundamentales y el fortalecimiento del pluralismo en el espacio público digital.

CONCLUSIONES

La investigación concluye que la libertad de expresión digital atraviesa una transformación estructural, debido a que su ejercicio ya no depende únicamente de la ausencia de restricciones estatales, sino también de las condiciones tecnológicas que determinan la circulación, visibilidad y alcance de los contenidos. Los resultados evidencian que las plataformas digitales y los sistemas algorítmicos influyen significativamente en la configuración del espacio público, por lo que su protección requiere integrar las garantías jurídicas tradicionales con una comprensión de los factores tecnológicos que intervienen en la comunicación democrática.

Se determina que los principales riesgos actuales no se limitan a la eliminación directa de contenidos, sino que incluyen la opacidad algorítmica, la priorización automatizada de información, la desinformación y la falta de mecanismos adecuados de revisión. En consecuencia, la protección de este derecho debe orientarse hacia modelos basados en transparencia, pluralismo y participación efectiva, superando enfoques centrados exclusivamente en evitar restricciones ilegítimas.

Asimismo, se concluye que la moderación de contenidos constituye un mecanismo necesario frente a la desinformación y otros daños potenciales; sin embargo, su legitimidad depende de criterios objetivos, procedimientos previsibles y garantías de revisión. Por ello, el reto actual consiste en asegurar que las decisiones restrictivas respeten los principios de legalidad, necesidad y proporcionalidad, más no eliminar la regulación digital.

Los hallazgos muestran que las plataformas digitales han asumido funciones con impacto directo en la construcción de la opinión y debate público mediante la organización, recomendación y moderación de contenidos. Esta capacidad genera responsabilidades vinculadas con la protección de derechos fundamentales, haciendo necesario superar modelos basados únicamente en la autorregulación y avanzar hacia esquemas de supervisión, transparencia y responsabilidad compartida.

Respecto a la inteligencia artificial, se concluye que su incorporación al ecosistema comunicativo redefine los desafíos de la libertad de expresión, debido a su intervención en la creación, clasificación

y distribución de información. Por ello, su aplicación democrática requiere garantías de explicabilidad, supervisión humana, trazabilidad y rendición de cuentas, especialmente cuando sus decisiones pueden afectar la diversidad informativa y la participación ciudadana.

Finalmente, la investigación establece que la respuesta frente a la desinformación debe superar la eliminación de contenidos como estrategia principal e incorporar alfabetización digital, responsabilidad de las plataformas y capacidades institucionales de supervisión. En este sentido, la protección sostenible de la libertad de expresión digital requiere una gobernanza colaborativa que articule dimensiones jurídicas, éticas, tecnológicas y democráticas, mediante la cooperación entre Estado, plataformas digitales, academia y ciudadanía bajo principios de transparencia, responsabilidad y respeto a los valores constitucionales.

REFERENCIAS BIBLIOGRÁFICAS

- Altay, S. (2026). Who is concerned about online false news and why? Evidence from 46 countries between 2018 and 2023. *Information, Communication & Society*, 29(7), 2141-2157. <https://doi.org/10.1080/1369118X.2025.2551615>
- Amin, A. & Hong, Y. (2026). Examining the impact of misinformation on governance efficacy in the health sector: The mediating role of public trust and the moderating effect of social media regulations. *Telematics and Informatics Reports*, 22(100326). <https://doi.org/10.1016/j.teler.2026.100326>
- Atzori, M. (2024). The Digital Services Act and the Freedom of Expression in the European Union: A Political Perspective. *Papers SSRN*, 1-23. <http://dx.doi.org/10.2139/ssrn.4887181>
- Barata, J. (2022). Libertad de expresión, regulación y moderación privada de contenidos. *Teoría & Derecho*, 32, 88-107. <https://doi.org/10.36151/TD.2022.039>
- Basanta, B. & Azurmendi, A. (2025). Las plataformas digitales y el derecho a la información: de meros alojadores a actores responsables. *Revista de Comunicación*, 24(1), 17-50. <https://doi.org/10.26441/RC24.1-2025-3641>
- Bayer, J. (2022). Procedural Rights as Safeguard for Human Rights in Platform Regulation. *Policy & Internet*, 14, 755-771. <https://doi.org/10.1002/poi3.298>



- Çelik, N. (2026). Who Controls the Digital Agora? Social Media Governance in the European Union, United States, and Turkey. *Global Society*, 40(2), 225-261. <https://doi.org/10.1080/13600826.2025.2571147>
- Coral, A. & Coral, A. (2022). Evolución histórica de la libertad de expresión en el Perú. *Llalliq*, 2(1), 55-69. <https://doi.org/10.32911/llalliq.2022.v2.n1.939>
- Echeverría, M., & Rodríguez Cano, C. (2023). ¿La alfabetización digital activa la incredulidad en noticias falsas? Eficacia de las actitudes y estrategias contra la desinformación en México. *Revista de Comunicación*, 22(2), 79-95. <https://doi.org/10.26441/RC22.2-2023-3246>
- Floridi, L. (2023). The Ethics of Artificial Intelligence: Principles, Challenges, and Opportunities *Oxford Academic*. <https://doi.org/10.1093/oso/9780198883098.001.0001>
- Freiling, I.; Stubenvoll, M. & Matthes, J. (2023). Support for misinformation regulation on social media: It is the perceived harm of misinformation that matters, not the perceived amount. *Policy & Internet*, 15, 731-749. <https://doi.org/10.1002/poi3.360>
- Freiling, I.; Yeo, S. & Xue, H. (2026). When AI and Humans Produce Partial Truths: Examining Acceptability of Perceived Error and Perceived Associated Harms. *Health Communication*, 1-12. <https://doi.org/10.1080/10410236.2025.2608202>
- García, D. & Calvo, P. (2022). Democracia algorítmica: ¿un nuevo cambio estructural de la opinión pública? *Isegoría*, (67), e17. <https://doi.org/10.3989/isegoria.2022.67.17>
- Gorwa, R. & Veale, M. (2024). Moderating model marketplaces: Platform governance puzzles for AI intermediaries. *Law, Innovation and Technology*, 16(2), 341-391. <https://doi.org/10.1080/17579961.2024.2388914>
- Gsenger, R. (2025). Platform governance under the Digital Services Act: a perspective on disinformation. *Information, Communication & Society*, 1-19. <https://doi.org/10.1080/1369118X.2025.2590561>
- Knapp, M. & Piszcz, A. (2024). Moving Towards More Transparent Online Platforms Under the Digital Services Act. *Springer*, 15, 105-123. https://doi.org/10.1007/978-3-031-69678-7_5



- Ordoñez, A. (2022). Plataformas de redes sociales: un enfoque para la legislación del sistema interamericano. *Revista Jurídica de la Universidad de San Andrés*, 14, 49-88. <https://revistasdigitales.udesa.edu.ar/index.php/revistajuridica/es/article/view/159>
- Ó Fathaigh, R.; Buijs, D. & Van Hoboken, J. (2025). The Regulation of Disinformation Under the Digital Services Act. *Media and Communication*, 13. <https://doi.org/10.17645/mac.9615>
- López-López, P.; Pereira-López, M.; Jaráiz-Gulías, E. & Lagares-Díez, N. (2025). Explanatory factors for the dissemination and control of fake news in the Latin American context. *Humanities & Social Sciences Communications*, 12(741). <https://doi.org/10.1057/s41599-025-05100-7>
- Martínez, P. (2022). Responsabilidad de los prestadores de servicios en línea por contenidos protegidos por la propiedad intelectual frente a libertad de expresión. *Revista CESCO de Derecho de Consumo*, 42, 168-181. https://doi.org/10.18239/RCDC_2022.42.3099
- Mortenson, C.; Nisbet, E. & Garrett, K. (2026). The Presumed Prevalence/Persuasiveness of Online Misinformation and Americans' Dissatisfaction With Democracy. *Social Media + Society*, 12(1). <https://doi.org/10.1177/20563051261420610>
- Palmiotto, F. & Menéndez, N. (2023). Facial recognition technology, democracy and human rights. *Computer Law & Security Review*, 50, 105857. <https://doi.org/10.1016/j.clsr.2023.105857>
- Papaevangelou, C. (2023). The role of citizens in platform governance: A case study on public consultations regarding online content regulation in the European Union. *Global Media and Communication*, 8(1). <https://doi.org/10.1177/20594364221150142>
- Pedraza, E. & Martínez, P. (2022). Algoritmos y democracia: impacto, peligro y propuestas regulatorias. *Pensamiento Jurídico*, (56), 91-113. <https://revistas.unal.edu.co/index.php/peju/article/view/106555>
- Seipp, T.; Helberger, N.; de Vreese, C. & Ausloos, J. (2023). Dealing with Opinion Power in the Platform World: Why We Really Have to Rethink Media Concentration Law. *Digital Journalism*, 11(8), 1542-1567. <https://doi.org/10.1080/21670811.2022.2161924>



- Trithara, D. (2024). Agents of platform governance: Analyzing U.S. civil society's role in contesting online content moderation. *Telecommunications Policy*, 48(4), 102735. <https://doi.org/10.1016/j.telpol.2024.102735>
- Vázquez, V. (2022). La censura “privada” de las grandes corporaciones digitales y el nuevo sistema de la libertad de expresión. *Teoría & Derecho. Revista de Pensamiento Jurídico*, 32, 108-129. <https://doi.org/10.36151/td.2022.040>
- Zacarias, F.; De la Corte, H. & Silva, M. (2022). O hate speech e os limites da liberdade de expressão nos meios digitais. *Revista Direito.UnB*, 6(3), 45-72. <https://periodicos.unb.br/index.php/revistadedireitounb/article/view/40647>
- Zheng, X.; Zhou, G. & Zeng, D. (2023). Platform governance in the era of AI and the digital economy. *Frontiers of Engineering Management*. 10, 177-182. <https://doi.org/10.1007/s42524-022-02411>

