

**PROTECCIÓN DEL DERECHO A LA LIBERTAD DE EXPRESIÓN ANTE
DESINFORMACIÓN POLÍTICA POR IA EN COLOMBIA**

Luis Dayan Silva Gallego

Arnol Stiven Silva Gallego

Juan José Trujillo Méndez



UNIVERSIDAD
La Gran Colombia

Vigilada MINEDUCACIÓN

Programa académico, Facultad de Derecho

Universidad La Gran Colombia

Bogotá D.C.

2026

**Protección del derecho a la libertad de expresión ante
desinformación política por
IA en Colombia**

Luis Dayan Silva Gallego

Arnol Stiven Silva Gallego

Juan José Trujillo Méndez

Trabajo de Grado presentado como requisito para optar al título de Abogado



UNIVERSIDAD
La Gran Colombia

Vigilada MINEDUCACIÓN

Programa académico, Facultad de Derecho

Universidad La Gran Colombia

Bogotá D.C.

2026

Dedicatoria

Dedicamos este trabajo a nuestras familias por su apoyo constante, comprensión y motivación durante todo el proceso académico, asimismo, a quienes contribuyeron con sus enseñanzas y acompañamiento para hacer posible el desarrollo de esta investigación.

Agradecimientos

En primer lugar, expresamos agradecimientos a Dios, a nuestras familias y a los docentes de la Universidad por el acompañamiento brindado durante este proceso formativo. De manera especial, agradecemos al director del trabajo de grado por su orientación académica, observaciones y apoyo permanente en el desarrollo de la investigación.

Igualmente agradecemos a la Universidad La Gran Colombia y a la facultad por los espacios de formación jurídica Investigativa y crítica, que permitieron fortalecer el análisis sobre la libertad de expresión, inteligencia artificial y desinformación política dentro del contexto normativo y constitucional colombiano. A la vez, extendemos nuestra gratitud a las instituciones, doctrina y jurisprudencia consultadas, las cuales constituyeron una base fundamental para la construcción y desarrollo del presente trabajo investigativo.

Tabla de contenido

Resumen	11
Abstract	12
Introducción	13
Objetivos	14
OBJETIVO GENERAL	14
OBJETIVOS ESPECÍFICOS	14
CAPÍTULO I: PROBLEMÁTICA	15
PROBLEMA DE INVESTIGACIÓN	15
HIPÓTESIS	18
CAPÍTULO II: MARCO TEÓRICO	19
MARCO REFERENCIAL	19
MARCO CONCEPTUAL	27
<i>Inteligencia artificial (IA)</i>	27
<i>Desinformación política</i>	28
<i>Libertad de expresión</i>	30
<i>Democracia deliberativa</i>	31
<i>Regulación jurídica digital</i>	32
MARCO LEGAL	33
<i>Marco internacional</i>	33
<i>Marco institucional colombiano</i>	34
<i>Instituciones democráticas y necesidad de articulación regulatoria</i>	37

IA, DESINFORMACIÓN Y LIBERTAD DE EXPRESIÓN	6
CAPÍTULO III: ASPECTOS METODOLÓGICOS	39
ENFOQUE DE LA INVESTIGACIÓN	39
TIPO DE INVESTIGACIÓN	40
MÉTODO DE INVESTIGACIÓN	41
DISEÑO DE LA INVESTIGACIÓN	42
TÉCNICAS DE RECOLECCIÓN DE INFORMACIÓN	43
CRITERIOS DE SELECCIÓN DE LA INFORMACIÓN	46
TÉCNICAS DE ANÁLISIS DE LA INFORMACIÓN	47
CAPÍTULO IV: ANÁLISIS Y DISCUSIÓN DE RESULTADOS	50
ANÁLISIS DOCTRINAL SOBRE LIBERTAD EXPRESIÓN, DESINFORMACIÓN POLÍTICA E INTELIGENCIA ARTIFICIAL	50
<i>Inteligencia artificial generativa, deepfakes y automatización del discurso político</i>	52
<i>Libertad de expresión y regulación de la desinformación política</i>	56
<i>Vacíos doctrinales</i>	57
ANÁLISIS JURISPRUDENCIAL Y NORMATIVO	59
<i>Protección constitucional y jurisprudencial de la libertad de expresión</i>	60
<i>Aplicabilidad de las categorías jurisprudenciales frente a la inteligencia artificial</i>	64
<i>Vacíos regulatorios y desafíos jurídicos.</i>	67
CRITERIOS JURÍDICOS PARA LA PROTECCIÓN	74
<i>Protección reforzada del discurso político en entornos digitales</i>	74
<i>La desinformación automatizada como riesgo para el orden democrático</i>	77
<i>Necesidad de regulación proporcional y responsabilidad ulterior</i>	78

IA, DESINFORMACIÓN Y LIBERTAD DE EXPRESIÓN	7
<i>Propuesta de criterios jurídico-normativos aplicables al contexto colombiano</i>	82
DISCUSIÓN JURÍDICA DE RESULTADOS	87
CONCLUSIONES Y RECOMENDACIONES	91
LISTA DE REFERENCIA O BIBLIOGRAFÍA	92

Lista de Tablas

Tabla 1. <i>Clasificación temática de la literatura analizada</i>	50
Tabla 2. <i>Tensiones jurídicas identificadas en la literatura especializada</i>	53
Tabla 3. <i>Líneas jurisprudenciales sobre libertad de expresión digital en Colombia</i>	62
Tabla 4. <i>Principales vacíos regulatorios y desafíos jurídicos</i>	71

Declaración de autoría y uso de inteligencia artificial

Los autores declaramos que el presente trabajo de grado es una obra original fruto de nuestra investigación y análisis intelectual exclusivo, asumiendo plena y total responsabilidad sobre su contenido, interpretaciones, resultados, conclusiones y rigor ético.

Asimismo, en cumplimiento de las buenas prácticas de integridad académica y transparencia científica, manifestamos que se emplearon modelos de inteligencia artificial generativa (Gemini de Google y Claude de Anthropic) de manera estrictamente complementaria. Su uso se limitó a tareas de asistencia lingüística, revisión de estilo, corrección sintáctica y ortográfica, sin delegar en estas herramientas la concepción metodológica, el análisis crítico, la redacción sustancial de argumentos ni la toma de decisiones investigativas.

La versión final ha sido revisada, contrastada y avalada íntegramente por los autores.

Glosario

Inteligencia artificial (IA): Son sistemas creados para realizar tareas que tradicionalmente requerían la intervención de las personas, como aprender a partir de los datos, analizar información y generar contenidos de manera cada vez más autónoma.

Libertad de expresión: Es el derecho esencial que permite a cualquier persona compartir y recibir ideas u opiniones libremente, protegiendo el debate público frente a la censura.

Microtargeting político: Es una técnica de precisión digital. Utiliza algoritmos y datos personales para "radiografiar" a los ciudadanos, permitiendo que las campañas les envíen mensajes diseñados específicamente para tocar sus fibras emocionales o intereses particulares.

Propaganda computacional: Se trata del uso de automatización y herramientas digitales para inclinar la balanza de la opinión pública de forma artificial, manipulando el debate político mediante algoritmos.

Transparencia algorítmica: Es el derecho a comprender cómo funcionan las máquinas y por qué toman determinadas decisiones. Busca que los criterios que utilizan para mostrar información sean claros, transparentes y fáciles de entender para todas las personas.

Resumen

La investigación analiza el problema jurídico que enfrenta Colombia ante la desinformación política que puede generarse mediante el uso de la inteligencia artificial. El estudio se centra especialmente en herramientas como los deepfakes, los bots y el microtargeting, debido a que su uso puede afectar la manera en que circula la información política y, con ello, influir en el debate público y en el funcionamiento de la democracia en los espacios digitales. Para analizar esta situación, se revisó el marco normativo vigente en Colombia y diferentes decisiones y antecedentes de la jurisprudencia nacional e internacional relacionados con la libertad de expresión y el discurso político en internet. A partir de esta revisión, se encontró que la inteligencia artificial facilita la creación y difusión de información que puede ser manipulada o presentada de manera engañosa. La investigación permitió identificar que en Colombia todavía existen vacíos jurídicos frente al uso de la IA; entre los principales retos se encuentran la falta de transparencia en el funcionamiento de los algoritmos y la dificultad para establecer responsabilidades por la creación y difusión de contenidos generados artificialmente. Estos desafíos requieren una respuesta jurídica que permita proteger el debate público y, al mismo tiempo, garantizar los derechos fundamentales en los entornos digitales.

Palabras clave: libertad de expresión, inteligencia artificial, desinformación política, democracia digital.

Abstract

This research analyzes the legal challenge Colombia faces considering political disinformation that can be generated using artificial intelligence. The study focuses particularly on tools such as deepfakes, bots, and microtargeting, as their use can affect the way political information circulates and, in turn, influence public debate and the functioning of democracy in digital spaces. To analyze this situation, the researchers reviewed Colombia's current regulatory framework and various national and international legal precedents and case law related to freedom of expression and political discourse on the internet. Based on this review, they found that artificial intelligence facilitates the creation and dissemination of information that can be manipulated or presented in a misleading manner. The research revealed that legal gaps regarding the use of AI still exist in Colombia; among the main challenges are the lack of transparency in how algorithms operate and the difficulty in establishing accountability for the creation and dissemination of artificially generated content. These challenges require a legal response that protects public debate while also guaranteeing fundamental rights in digital environments.

Keywords: *freedom of expression, artificial intelligence, political disinformation, digital democracy.*

Introducción

El crecimiento de las tecnologías digitales y el uso de la inteligencia artificial han cambiado la forma en que se produce y circula la información política en Colombia. En ese sentido, el debate público y los procesos electorales tienen cada vez mayor presencia en espacios digitales donde los contenidos pueden difundirse de manera rápida y alcanzar un gran número de personas. En este escenario, herramientas como los *deepfakes*, los bots, la propaganda computacional y el microtargeting generan nuevos riesgos para la libertad de expresión, el acceso a información confiable y la democracia. Y por esta razón, es necesario revisar la forma en que el derecho responde a estas nuevas formas de creación y difusión de contenidos.

En este sentido, la investigación analiza si el marco constitucional, legal y jurisprudencial colombiano resulta suficiente para proteger la libertad de expresión frente a la desinformación política generada o difundida mediante inteligencia artificial. Para ello se considera la constitución política de 1991, los pronunciamientos de la Corte Constitucional y los principales estándares internacionales sobre la libertad de expresión y circulación de la información. Es así como el análisis parte de la atención entre la necesidad de adoptar medidas frente a la manipulación informativa y el deber de evitar restricciones injustificadas al discurso político, especialmente cuando se trata de asuntos de interés público.

Asimismo, se examinan los principales vacíos jurídicos relacionados con la creación y difusión de estos contenidos y con la responsabilidad de quienes intervienen en dichos procesos. Finalmente, se plantean algunos criterios jurídicos que pueden contribuir a encontrar un equilibrio entre la protección de la libertad de expresión, el derecho de la ciudadanía a recibir información y la protección del orden democrático en los entornos digitales.

En últimas, el objetivo consiste en comprender cómo estos cambios tecnológicos plantean nuevos problemas de aplicación para categorías constitucionales como la prohibición de la censura previa, la responsabilidad ulterior y la proporcionalidad de las restricciones a la libertad de expresión.

Objetivos

Objetivo General

Analizar la suficiencia del marco constitucional, legal y jurisprudencial colombiano para proteger la libertad de expresión frente a la desinformación política generada o amplificada mediante inteligencia artificial, con el fin de establecer criterios jurídicos que permitan armonizar este derecho con la protección del orden democrático.

Objetivos Específicos

(1) Organizar y revisar las investigaciones actuales, tanto de Colombia como del exterior, sobre libertad de expresión y desinformación política generada o amplificada mediante inteligencia artificial, para identificar las principales categorías jurídicas y problemáticas que plantea este fenómeno.

(2) Evaluar las leyes colombianas, las sentencias de la Corte Constitucional y las normas del sistema interamericano relacionadas con la libertad de expresión, la responsabilidad ulterior, la prohibición de la censura previa y las restricciones al discurso político, con el propósito de determinar hasta qué punto estas categorías resultan aplicables a los escenarios de desinformación política generada o amplificada mediante inteligencia artificial.

(3) Identificar los criterios jurídicos derivados del análisis normativo, jurisprudencial y doctrinal, especialmente en materia de necesidad, proporcionalidad, atribución de responsabilidad, transparencia y garantías frente a la intervención estatal, para formular una propuesta que permita armonizar la protección de la libertad de expresión con la salvaguarda del orden democrático.

CAPÍTULO I: PROBLEMÁTICA

Problema de investigación

El espacio público contemporáneo se caracteriza por una creciente mediación tecnológica en la circulación de información política, y en este contexto, el desarrollo y la adopción de sistemas de inteligencia artificial han ampliado las posibilidades de producción, modificación y difusión de contenidos, lo que ha incidido directamente en las dinámicas de formación de la opinión pública.

Cuando estas tecnologías se utilizan para propagar información falsa o engañosa de incidencia política, se generan escenarios capaces de afectar el proceso de formación de la opinión pública y, con ello, la deliberación democrática. En ese sentido, el reto jurídico consiste tanto en reconocer y delimitar el contenido desinformativo, como en definir hasta dónde puede llegar la actuación del Estado sin vulnerar un derecho fundamental que, en los sistemas constitucionales, cuenta con una protección especialmente reforzada.

En consecuencia, toda respuesta de carácter normativo o jurisprudencial debe sustentarse en los principios de necesidad, proporcionalidad y razonabilidad; además dentro del ordenamiento jurídico colombiano, la libertad de expresión ha sido objeto de un amplio desarrollo constitucional y jurisprudencial, orientado a la protección de un debate público abierto, plural y democrático. El desafío actual consiste en determinar hasta qué punto las categorías jurídicas desarrolladas por este marco constitucional y jurisprudencial pueden ser aplicadas a escenarios en los que intervienen sistemas de inteligencia artificial capaces de producir, modificar, segmentar y difundir contenidos políticos a gran escala.

Por ello, es necesaria una revisión crítica del marco normativo existente y de los estándares establecidos tanto por la jurisprudencia constitucional como por el sistema interamericano de derechos humanos, no solo para identificar posibles vacíos, sino también para determinar la aplicabilidad de las categorías jurídicas existentes frente a las nuevas dinámicas tecnológicas. Lo anterior, con el propósito de

evitar, por un lado, los escenarios de inacción estatal, y por otro, la adopción de restricciones desproporcionadas a la libertad de expresión.

Uno de los problemas centrales que plantea este escenario corresponde a la atribución de responsabilidad jurídica por los contenidos políticos generados o modificados mediante inteligencia artificial. La intervención de diferentes actores como usuarios, desarrolladores, plataformas y otros sujetos que participan en la producción o difusión del contenido puede dificultar la identificación de quién debe responder jurídicamente por los efectos de una determinada pieza de desinformación. Y esta situación exige examinar si las categorías tradicionales de responsabilidad ulterior y los criterios jurisprudenciales existentes permiten establecer dicha atribución sin desconocer las garantías propias de la libertad de expresión

Bajo esa perspectiva, surge la siguiente pregunta:

¿Es suficiente el marco constitucional, legal y jurisprudencial colombiano para enfrentar la desinformación política generada o amplificada mediante inteligencia artificial sin imponer restricciones desproporcionadas al derecho a la libertad de expresión?

La presente investigación surge de la necesidad de realizar un análisis jurídico actualizado sobre la forma en que los principios constitucionales que protegen la libertad de expresión pueden aplicarse en un contexto donde la inteligencia artificial tiene cada vez mayor presencia en la manera en que se produce, circula y consume la información. La importancia del estudio está relacionada con los cambios que estas tecnologías han generado en la producción y difusión de contenidos, especialmente aquellos que pueden influir en el debate público y afectar la confianza de los ciudadanos en las instituciones democráticas.

En el ámbito académico, la investigación aporta el estudio del Derecho constitucional al revisar temas como la responsabilidad posterior, la previsión de la censura previa y la protección de los derechos fundamentales frente al uso a la inteligencia artificial, ya que estas cuestiones adquieren especial importancia cuando la nueva tecnología permiten crear, modificar y fundir contenidos políticos con gran

rapidez. Por ello, el análisis busca establecer si las herramientas jurídicas actuales son suficientes para enfrentar la desinformación política y, en caso contrario, identificar los vacíos que todavía existen en el ordenamiento colombiano.

Desde el punto de vista social y jurídico, el estudio permite revisar hasta dónde puede intervenir el Estado frente a la desinformación política sin afectar de forma injustificada la libertad de expresión. Además, se considera si es necesario mantener un equilibrio entre la protección frente a la manipulación informativa y el respeto por el debate público, evitando tanto la ausencia de regulación como medidas excesivas que puedan restringir derechos fundamentales.

La investigación también pretende establecer criterios jurídicos que permitan encontrar un equilibrio entre la protección de la democracia y el respeto de las libertades reconocidas en la Constitución; este planteamiento cobra importancia si se tienen en cuenta que una parte significativa de la opinión pública y de la participación ciudadana tiene lugar actualmente en espacios digitales, en los cuales la información se difunde con gran rapidez y, al mismo tiempo, surgen nuevas formas de influencia asociadas al uso de la tecnología. Igualmente, el análisis considera cómo puede responder el sistema jurídico frente a situaciones que afecten la transparencia electoral y la diversidad de opiniones, sin desconocer los principios de un Estado democrático.

De igual manera, el estudio pretende aportar elementos que permitan determinar cuándo una intervención estatal frente a contenidos políticos creados o difundidos con inteligencia artificial puede ser legítima, necesaria y proporcional. Como tal, el análisis puede resultar de utilidad para juristas y legisladores y jueces que deben interpretar estos nuevos fenómenos desde el marco constitucional, pero también para la ciudadanía, en la medida en que permite comprender mejor los riesgos que estas tecnologías pueden generar sobre los derechos fundamentales y el debate público.

Por último, la investigación es viable porque se desarrolla bajo un enfoque cualitativo y documental, con un alcance descriptivo y exploratorio, ya que para ello se revisan fuentes normativas

jurisprudenciales y doctrinales relacionadas con el problema planteado, lo que permite desarrollar el estudio de acuerdo con los recursos y el tiempo disponible.

Para su desarrollo se realiza la revisión y análisis de fuentes normativas, jurisprudenciales y doctrinales seleccionados de acuerdo con su relación y aporte al problema planteado y este tipo de metodología permite abordar el objeto de estudio de manera organizada y rigurosa, teniendo en cuenta los recursos y el tiempo disponible para el desarrollo de la investigación jurídica.

Hipótesis

Se plantea que el marco jurídico colombiano ofrece herramientas para proteger la libertad de expresión frente a la desinformación política, pero presenta vacíos interpretativos ante los contenidos generados o amplificados mediante inteligencia artificial, especialmente en la atribución de responsabilidades y en la definición de límites legítimos al discurso político.

CAPÍTULO II: MARCO TEÓRICO

Marco Referencial

Para abordar el problema de investigación se definió su enfoque epistemológico y disciplinario, ya que el fenómeno estudiado presenta diferentes dimensiones que no pueden ser analizadas únicamente del ámbito jurídico. En ese sentido, la desinformación política generada mediante sistemas de inteligencia artificial se configura como un fenómeno complejo que articula dimensiones tecnológicas, comunicativas y jurídicas, por lo que resulta necesario adoptar un enfoque interdisciplinario que permita analizar sus efectos sobre el orden democrático, y sobre el ejercicio de la libertad de expresión en entornos digitales (Solar, 2020).

Desde una perspectiva epistemológica, la investigación se desarrolla desde un enfoque jurídico constitucional de carácter analítico, ya que esto se debe a que busca examinar de manera en que el ordenamiento jurídico colombiano protege los derechos fundamentales frente a los cambios que ha traído consigo el uso de tecnologías digitales. En particular, se presta atención a aquellos cambios que tienen relación con la libertad de expresión, la formación de la opinión pública y el funcionamiento del sistema democrático. En este contexto, se parte la idea de que el derecho debe responder a las nuevas formas de comunicación que surgen en los aspectos digitales, especialmente cuando éstas pueden afectar el ejercicio de derechos fundamentales (Ibáñez, et al. 2024).

De igual forma, el estudio entiende el Derecho como un sistema normativo abierto cuya interpretación puede cambiar de acuerdo con las transformaciones sociales y con las nuevas formas de comunicación e interacción que surgen en la sociedad. Esta concepción evolutiva del ordenamiento jurídico se nutre de la perspectiva holmesiana, la cual entiende el Derecho como un instrumento dinámico, empírico y generacional puesto al servicio incondicional de las demandas cambiantes de la sociedad (Vargas & Álzate, 2017). Bajo esta aproximación pragmática, la vida de la norma jurídica no ha sido la lógica formalista, sino la experiencia histórica y fáctica que surge de las tensiones colectivas reales. Trasladada esta premisa teórica a la coyuntura contemporánea, se hace patente que las instituciones

constitucionales no pueden permanecer ancladas a realidades comunicativas analógicas, sino que deben adaptarse con elasticidad para salvaguardar la deliberación democrática frente a la disrupción de la inteligencia artificial.

Asimismo, se analizan estudios en el campo de la comunicación política que evidencian que la circulación de información y la configuración de opinión pública en entornos digitales no obedecen únicamente a decisiones individuales, sino que pueden estar influenciadas por procesos tecnológicos y sociales integrados que condicionan la recepción y la interpretación de los mensajes políticos (Guerrero, 2024).

Aunado a lo anterior, existen investigaciones sobre la interacción entre inteligencia artificial y comunicación política que señalan que los modelos generativos de la inteligencia artificial facilitan tanto la creación de contenidos automatizados como la ampliación algorítmica de mensajes, lo que afecta la dinámica de la agenda pública y plantea desafíos para las normas tradicionales sobre discurso y verdad en la esfera pública. Por ejemplo, Anstead et al. (2025) evidencian que la integración de contenidos generados por la inteligencia artificial en la vida cotidiana ha permitido a los actores con diversas motivaciones a diseminar información a gran escala, lo que intensifica la complejidad de la regulación y la gobernanza de la información política en plataformas digitales.

De igual manera, los estudios sobre el papel de la inteligencia artificial en el contexto de la libertad de expresión subrayan que la implementación de tecnologías automatizadas en el periodismo y en los sistemas de moderación de contenidos genera nuevos interrogantes sobre la protección de derechos fundamentales; en ese caso, Sierra et al. (2025) sostiene que la creciente presencia de la inteligencia artificial en los procesos editoriales y en la verificación de información plantea la necesidad de diseñar marcos institucionales que garanticen el uso responsable de estas tecnologías sin llegar a restringir indebidamente los derechos de los sujetos comunicantes, en particular la libertad de expresión y de prensa.

Por lo tanto, la incorporación de estudios provenientes de las ciencias de la comunicación y sobre desinformación no desplaza el análisis propio del derecho constitucional, sino que de una u otra manera contribuye a contextualizar empíricamente el fenómeno estudiado. De ese modo, se permite comprender mejor cómo las transformaciones tecnológicas alteran las prácticas comunicativas y a su vez, cómo esas transformaciones interactúan con las normas jurídicas existentes.

El análisis de la libertad de expresión en el contexto contemporáneo exige partir del reconocimiento de que el espacio público se encuentra crecientemente mediado por tecnologías digitales (Caballero, 2020). En particular, el desarrollo reciente de sistemas de inteligencia artificial ha transformado las dinámicas de producción, circulación y consumo de información política, al permitir la generación automatizada de contenidos con altos niveles de verosimilitud, y esta transformación tecnológica incide directamente en la formación de la opinión pública y plantea nuevos retos para la protección jurídica de la libertad de expresión en entornos digitales (Bailon, 2024).

Dentro del sistema jurídico colombiano la libertad de expresión se encuentra reconocida en el artículo 20 de la Constitución Política de 1991, donde establece y garantiza la facultad de manifestar y divulgar pensamientos y opiniones por parte de la ciudadanía, así como de comunicar y recibir información que sea veraz e imparcial, además de que incorpora de manera explícita la prohibición de cualquier forma de censura; de igual manera, la jurisprudencia de la corte constitucional reitera que este derecho goza de una protección reforzada cuando se ejerce en contextos de debate político o en asuntos de interés público, por su atención a su estrecha relación con el pluralismo y con los fundamentos del sistema democrático.

Sin embargo, dicha protección no tiene un carácter ilimitado, pues la doctrina y la jurisprudencia han admitido la procedencia de restricciones legítimas cuando entran en tensión otros derechos fundamentales o bienes constitucionales relevantes, y en esta línea, la teoría de los derechos fundamentales propuesta por Alexy (2002), la cual es ampliamente acogida en el constitucionalismo contemporáneo, concibe estos derechos como principios susceptibles de ponderación, lo que permite

resolver conflictos normativos mediante criterios de proporcionalidad, necesidad y razonabilidad; y esta metodología resulta especialmente pertinente para examinar las tensiones entre la libertad de expresión y la salvaguarda del orden democrático frente a fenómenos de desinformación.

En el ámbito de la desinformación política, los efectos diferenciados de la difusión de información falsa en entornos digitales. Investigaciones como la de Vosoughi et al. (2018) han demostrado que las noticias falsas tienden a propagarse más rápido y alcanzar a un mayor número de personas que la información verificada, especialmente en contextos políticos, lo que puede incidir negativamente en la deliberación pública y en la confianza institucional; y estos hallazgos contribuyen a que la desinformación sea comprendida tanto como un problema informativo, como un desafío jurídico y democrático.

A ese panorama se suma la influencia de la inteligencia artificial en la generación y la difusión de diferentes mensajes políticos, ya que el empleo de sistemas son capaces de crear textos, imágenes y videos que facilitan la producción masiva de contenidos automatizados y que pueden llegar a confundirse como material genuino; en ese caso, Aral (2021) advierte que estas herramientas llegan a identificar los incentivos económicos y políticos para la rápida propagación de información engañosa, al mismo tiempo que hace más difícil la identificación de los responsables y la verificación de las fuentes, lo que impone desafíos regulatorios adicionales a marcos jurídicos diseñados para realidades tecnológicas diferentes.

En el plano internacional, distintos organismos han alertado sobre los riesgos que la desinformación apoyada en inteligencia artificial supone para los derechos humanos y el normal desarrollo de los procesos democráticos. En particular, la UNESCO (2021) indica que, aunque la inteligencia artificial puede contribuir al fortalecimiento de la libertad de expresión, también puede ser empleada para incidir de forma indebida en el debate público.

En el sistema interamericano de derechos humanos, la libertad de expresión es interpretada como piedra angular de la democracia, además, la Corte Interamericana de Derechos Humanos ha establecido que cualquier restricción a este derecho debe cumplir estrictamente con los principios de legalidad,

finalidad legítima, necesidad y proporcionalidad (Corte IDH, 1985). De igual manera, la relatoría especial para la libertad de expresión señala que las actuaciones estatales orientadas a contrarrestar la desinformación no pueden derivar en prácticas de censura previa ni en instrumentos que impliquen un control ilegítimo del discurso político (OEA, 2019).

En el contexto colombiano, si bien existe un desarrollo jurisprudencial consistente en materia de libertad de expresión, es elaborado, en su mayor parte, antes de la masificación de las tecnologías de IA generativa, y esta circunstancia plantea interrogantes acerca de la idoneidad de los criterios actualmente vigentes para abordar fenómenos como la desinformación producida de manera automatizada, sin menoscabar las garantías constitucionales (Castro & Martínez, 2025).

Es así que, en Colombia, la libertad de expresión se reconoce como un derecho fundamental de especial importancia dentro del orden constitucional, conforme a lo dispuesto en la constitución política de 1991; su desarrollo jurisprudencial es ampliamente consolidado por la corte constitucional, lo cual ha sostenido de manera reiterada que el discurso político goza de una protección reforzada por su relación directa con el principio democrático y la participación ciudadana; y estas decisiones establecen criterios como la prohibición de la censura previa, la prevalencia del control posterior y la necesidad de aplicar un test estricto de proporcionalidad cuando se limita a este derecho.

Aun así, dicho desarrollo jurisprudencial se configuró en un contexto tecnológico caracterizado por la producción y difusión de información atribuible de forma directa a sujetos humanos identificables; por ello, la evolución reciente del entorno digital el cual está marcada por el uso de sistemas de inteligencia artificial generativa que son capaces de producir contenidos informativos de manera automatizada, introduce dinámicas que no han sido abordadas de forma expresa por la jurisprudencia constitucional colombiana.

Por otro lado, la libertad de expresión, además de proteger las opiniones que son aceptadas socialmente, también protege aquellas que pueden resultar polémicas, críticas o incluso incómodas. Por ello, su importancia entre una democracia está precisamente en garantizar la diversidad de opiniones y

evitar que desde el poder se determine una única forma de pensar o una sola versión de la verdad (Loretti & Lozano, 2019). Hoy en día, este derecho no es solo una garantía para el individuo, sino la base que sostiene el debate público, permitiendo que los ciudadanos participen en política y vigilen a quienes mandan (Ferrajoli, 2024).

Desde el punto de vista de la doctrina, autores como Pulido (2014) y Yepes et al. (2015) explican que para proteger este derecho es necesario entender el contexto social y tecnológico en el que vivimos. Por eso, estudiar la desinformación creada con inteligencia artificial nos obliga a revisar críticamente las reglas actuales. El reto es saber si lo que tenemos hoy alcanza para enfrentar estos nuevos riesgos sin terminar afectando la esencia de la libertad de expresión o el desarrollo del debate político en Colombia.

Continuando con el desarrollo previo, es necesario analizar críticamente un aspecto que se presenta como un problema estructural en la era de la información automatizada, y en ese caso sería la identificación de sujetos responsables por contenidos políticos generados o amplificados por sistemas de inteligencia artificial y las implicaciones de esta automatización en la protección jurídica de la libertad de expresión (Bardelli, 2024). En ese sentido, el derecho ha supuesto la existencia de un emisor humano como sujeto portador de derechos y obligaciones relacionados con el discurso político; sin embargo, la proliferación de modelos generativos de la inteligencia artificial plantea la pregunta de qué entendemos por discurso político cuando las máquinas producen mensajes con alta verosimilitud comunicativa, y cómo puede asignarse responsabilidad jurídica de manera compatible con los principios constitucionales que rigen la libertad de expresión (Almache, 2025).

Hay que tener en cuenta que una parte importante de la literatura contemporánea sobre derecho y tecnología aborda este problema desde la reflexión sobre la categoría de discurso generado por inteligencia artificial, y su lugar dentro del marco jurídico tradicional. En ese sentido, Kaminski & Jones (2024) sostienen que el fenómeno debe entenderse tanto desde la ruptura tecnológica como desde el enfoque que es un objeto sujeto a construcción jurídica y social, en el que los lenguajes legales y las instituciones deben definir cómo se incorporan estas tecnologías en los sistemas normativos existentes.

Desde esta perspectiva, la ley participa activamente en la configuración del significado jurídico de las prácticas comunicativas mediadas por la inteligencia artificial.

Cuando la inteligencia artificial se mete en la creación y difusión de mensajes políticos, nos obliga a repensar las categorías legales de siempre sobre la libertad de expresión, sobre todo en lo que tiene que ver con quién es el autor, qué intención hay detrás del mensaje y quién responde por lo que se publica. En este sentido, Bassini (2025) advierte que el hecho de que no haya un humano redactando directamente no le quita peso constitucional a lo que la IA genera. Al final, lo que importa es cómo ese contenido afecta el derecho de la gente a estar informada y a participar activamente en el debate público. Así, la discusión jurídica se traslada del emisor hacia el impacto del mensaje en los receptores y hacia el papel de las plataformas, los desarrolladores y los intermediarios tecnológicos en la circulación de la información.

Azgin y Kiralp (2024) señalan que la inteligencia artificial ha ampliado el uso de herramientas como la propaganda digital y el microtargeting, lo que permite dirigir mensajes a grupos específicos y facilita, en algunos casos, la difusión de contenidos falsos o manipulados que pueden afectar la autonomía de los votantes y los procesos democráticos. Frente a esta problemática, diversos autores sugieren la necesidad de contar con marcos normativos bien delimitados que, por una parte, protejan la libertad de expresión y, por otra, establezcan responsabilidades para los sujetos que intervienen en la creación y difusión de estos contenidos.

Junto a esta dificultad se añade la obligación de decidir a quién corresponde atribuir la responsabilidad cuando un contenido generado por inteligencia artificial genera un daño. Por ejemplo, Cheong et al. (2024) destacan que la complejidad técnica de estos sistemas dificulta saber, en última instancia, el grado de control, la intencionalidad y las relaciones existentes entre la mera generación de los contenidos y el daño que se causa. La presente situación evidencia la necesidad de llevar a cabo una reevaluación de los modelos tradicionales de responsabilidad ante las nuevas formas de interacción entre las personas y los sistemas de inteligencia artificial. Desde el punto de vista constitucional, la aplicación

de la inteligencia artificial al discurso político pone de relieve importantes desafíos relacionados con principios como el de legalidad y con la distribución de responsabilidades. La ausencia de criterios jurídicos claros y vinculantes para atribuir responsabilidad por los contenidos generados o difundidos mediante sistemas automatizados puede dar lugar a decisiones arbitrarias y, en consecuencia, afectar la libertad de expresión (Cifuentes, 2025).

Por ello, resulta necesario considerar tanto el funcionamiento de los sistemas de inteligencia artificial como el tipo de responsabilidad que corresponde a los sujetos que producen o distribuyen dichos contenidos, especialmente cuando estos pueden afectar derechos fundamentales. En este sentido, Almache (2025) plantea la necesidad de revisar las categorías jurídicas tradicionales frente a las nuevas formas de comunicación digital.

La inteligencia artificial engloba aquellos sistemas capaces de procesar información, aprender a partir de los datos y modificar sus respuestas en función de la consecución de determinados objetivos (Kaplan, 2020). Cuando se integra con el big data y las plataformas digitales, la inteligencia artificial transforma las dinámicas de comunicación en el entorno de las redes sociales, al permitir nuevas formas de segmentación y potencial manipulación de los contenidos. Herramientas como los bots y la inteligencia artificial generativa facilitan la producción y distribución de contenidos con diferentes grados de intervención humana y, al mismo tiempo, incrementan la presencia de estos sistemas en los procesos comunicativos, particularmente en la circulación de contenidos políticos.

En este contexto, la desinformación política consiste en la difusión deliberada de información falsa o manipulada con el propósito de influir en las percepciones y decisiones de los ciudadanos, lo que a diferencia de la información incorrecta por la existencia de una intención expresa (Vaccari & Chadwick, 2020; Dan et al., 2021). Por lo tanto, esta problemática adquiere una dimensión mayor en los entornos digitales debido al uso de herramientas tecnológicas que facilitan la creación, difusión y amplificación de este tipo de contenidos. En relación con ello, García (2021) señala que los bots, el astroturfing, los

deepfakes y otros mecanismos algorítmicos han generado nuevas formas de intervención y manipulación que pueden incidir en el debate democrático.

Aunado a lo anterior, el entorno digital ha transformado las formas de influencia política y ha generado nuevos mecanismos capaces de afectar la circulación de información y el desarrollo del debate democrático. En ese contexto, la desinformación política no se limita a la difusión de noticias falsas, sino que puede utilizar condiciones sociales y tecnológicas para aumentar la polarización, debilitar la confianza en las instituciones e influir en las decisiones de los ciudadanos, incluso en escenarios electorales (McKay & Tenove, 2021).

Ante estas formas de intervención y manipulación en los entornos digitales, el análisis jurídico debe partir del reconocimiento de la libertad de expresión como un derecho fundamental y un componente esencial de la democracia; sin embargo, su ejercicio no es absoluto, pues puede estar sujeto a límites cuando existen razones legítimas para establecerlos y estos cumplen con los principios de necesidad y proporcionalidad. Por ejemplo, Carson & Wright (2022) destacan que uno de los principales desafíos consiste en adoptar medidas frente a la desinformación sin generar restricciones excesivas que terminen afectando el debate público.

En la misma línea, McKay y Tenove (2021) plantean que la respuesta jurídica no debería centrarse únicamente en retirar contenidos, sino también en las condiciones que favorecen la polarización y el deterioro del debate público. De forma similar, Hunter (2023) señala que los problemas para la democracia no provienen solamente de las plataformas digitales, sino también de las prácticas de manipulación que pueden desarrollarse mediante ellas. Esto plantea para Colombia la necesidad de fortalecer los marcos jurídicos centrados en la transparencia, la responsabilidad compartida y las garantías sólidas para la deliberación política.

Marco conceptual

Inteligencia artificial (IA)

La inteligencia artificial (IA) puede entenderse como un campo interdisciplinario dedicado al desarrollo de sistemas capaces de realizar tareas que tradicionalmente se han relacionado con las capacidades cognitivas humanas, entre ellas el aprendizaje, el razonamiento, la toma de decisiones y la resolución de problemas (Vásquez et al., 2024). Además, esta concepción tiene uno de sus principales antecedentes en McCarthy (2007), quien plantea la inteligencia artificial como la ciencia y la ingeniería orientadas a construir máquinas inteligentes y programas capaces de desarrollar actividades, que, habitualmente, requieren formas de razonamiento humano.

Desde esta perspectiva, la IA comprende las tecnologías que son capaces de procesar grandes cantidades de datos, identificar patrones y producir respuestas a partir de la información disponible. Por lo tanto, el desarrollo creciente de modelos algorítmicos cada vez más sofisticados ha ampliado aún más la posibilidad, permitiendo tanto analizar la información como generar contenidos y reconocer determinados comportamientos con un nivel de complejidad que transforma la manera en que estas tecnologías se relacionan con la sociedad.

Sanz (2025) plantea que la inteligencia artificial ya tiene presencia en ámbitos que van más allá de la tecnología y su incorporación en la vida cotidiana también está generando cambios en las relaciones sociales, en la actividad política y en la manera en que el derecho de responder a los nuevos escenarios derivados de su uso.

En el contexto digital contemporáneo, la inteligencia artificial ha adquirido un papel cada vez más relevante en espacios como las redes sociales, los motores de búsqueda y los sistemas generativos, interviniendo de manera en que circula la información pública (Lassi, 2025). Desde una perspectiva funcional, Brandao (2025) la concibe como el desarrollo de agentes racionales orientados al cumplimiento de objetivos específicos, lo que permite examinar críticamente su aplicación en escenarios políticos y mediáticos. Sin embargo, estos sistemas no operan de manera completamente neutral, pues dependen de

datos, objetivos programados y modelos de optimización que pueden incidir en procesos de segmentación, recomendación y amplificación de contenidos (Sera, 2024). Asimismo, la expansión de la IA plantea nuevas formas de mediación informativa capaces de influir en el debate político, las condiciones de acceso a la información y la protección de derechos fundamentales en las sociedades democráticas.

Desinformación política

La desinformación política puede definirse como la difusión deliberada de información falsa o manipulada con el propósito de influir en percepciones públicas, alterar procesos de deliberación ciudadana o favorecer intereses políticos específicos (Muñoz, 2024). La modificación de la información no constituye una problemática exclusiva de la era digital. De acuerdo con lo señalado por Arendt (1972), la modificación sistemática de los hechos podría desdibujar el límite entre lo que puede entenderse como verdad y aquello que el propio sujeto construye a partir de determinados intereses ideológicos. Así las cosas, la desinformación no debe entenderse como la mera transmisión involuntaria de información errónea o falsa, sino como una práctica intencionada orientada a modificar o influir en las creencias y decisiones de las personas.

Este problema es particularmente importante en el ámbito político, en la medida en que la difusión descontrolada de información manipulada puede afectar los procesos electorales, generar desconfianza hacia las instituciones y dificultar el acceso de la ciudadanía a información relacionada con asuntos de interés general (Ros et al., 2025). Con el incremento de las tecnologías digitales también ha aumentado la velocidad y el alcance con lo que circula la información manipulada. Así, Wardle y Derakhshan (2017) consideran que uno de los atributos de este tipo de información es, precisamente, su capacidad para generar determinados efectos o daños en los ámbitos político y social.

Por lo tanto, poder comprender la desinformación política es fundamental ya que dentro de un entorno digital que es cada vez más mediado por algoritmos, donde herramientas como los bots, la microsegmentación y los contenidos generados mediante inteligencia artificial han ampliado

considerablemente la capacidad de producir y difundir narrativas engañosas a gran escala. En este caso, la desinformación política puede concebirse como una forma organizada de manipulación comunicativa que aprovecha tanto las posibilidades tecnológicas como determinadas vulnerabilidades cognitivas para influir en la percepción ciudadana y orientar el debate público (Kaplan, 2020). Su alcance, debe de analizarse y requiere considerarse desde la interacción entre la tecnología, el poder y el derecho, especialmente por sus posibles efectos sobre la información de la opinión pública y el funcionamiento de las sociedades democráticas.

Libertad de expresión

La libertad de expresión constituye uno de los fundamentos esenciales de las sociedades democráticas, pues garantiza la posibilidad de expresar, buscar y recibir ideas e informaciones sin interferencias arbitrarias. En este sentido, López (2025) señala que la circulación libre de opiniones resulta indispensable para la construcción del conocimiento colectivo, incluso cuando se trata de ideas equivocadas, ya que su confrontación permite fortalecer la búsqueda de la verdad.

Por ejemplo, dentro de los Estados constitucionales, este derecho protege tanto las expresiones socialmente aceptadas como aquellas críticas, incómodas o controversiales, en la medida en que el pluralismo existe la coexistencia de diferentes perspectivas y evita la imposición de una visión (Loretto & Lozano, 2019). Así, la libertad de expresión adquiere una doble dimensión, una como derecho individual y la otra como garantía necesaria para el debate público, la participación política y el control ciudadano del poder (Ferrajoli, 2024). Aun así, su reconocimiento jurídico no implica que se trate de un derecho limitado; como sostiene Alexy (2010) cuando entra en conflicto con otros principios constitucionales, como la dignidad humana, la protección de la democracia o la prevención de determinados daños sociales resulta necesario acudir a un ejercicio de ponderación que permita establecer en cada caso qué de hecho, principio debe prevalecer.

Por lo tanto, esa atención adquiere especial relevancia en el entorno digital, ya que está caracterizado por la velocidad y el amplio alcance de los contenidos, donde la protección del discurso

libre debe coexistir con la necesidad de enfrentar fenómenos como la desinformación política. Por ello, la libertad de expresión debe comprenderse como un derecho fundamental de especial relevancia constitucional, cuya protección demanda conciliar la máxima garantía del debate político con aquellos límites jurídicos que bajo criterios de legalidad, necesidad y proporcional, resulten legítimos para proteger otros bienes y derechos fundamentales.

Democracia deliberativa

La democracia deliberativa se entiende como una concepción de la democracia que va más allá de la celebración periódica de elecciones, al reconocer que la legitimidad de las decisiones colectivas también dependen de la existencia de espacios públicos en los que puedan confrontarse ideas, argumentos, y distintas fuentes de información.

Habermas (2015) desarrolló este enfoque al sostener que la legitimidad democrática depende de procesos comunicativos inclusivos donde la ciudadanía pueda participar en condiciones de libertad e igualdad discursiva. Bajo esta mirada, el espacio público es mucho más que un lugar físico o virtual; es el escenario donde se construye, de forma racional, la voluntad política de una sociedad. Por eso, para que la democracia funcione, necesita una esfera comunicativa que garantice que las decisiones de todos nazcan de un debate bien informado y no de la manipulación, las presiones o el cambio sistemático de los datos (Ros et al., 2025).

Sin embargo, en nuestra realidad digital, este ideal se choca con retos enormes. Hoy vemos cómo fenómenos como la polarización (alimentada por algoritmos), la manipulación informativa y la desinformación que corre sola están rompiendo esa capacidad de debatir con libertad (Palma et al., 2025).

Ya lo advertía Dryzek (2002): la salud de una democracia depende directamente de qué tan limpios sean los canales de comunicación donde se da el debate público.

Cualquier alteración significativa en las formas de circulación de la información puede afectar la capacidad de los ciudadanos para formarse un criterio propio y participar de manera autónoma en los asuntos públicos. En ese contexto, cuando la inteligencia artificial es utilizada para producir y difundir de

manera estratégica contenidos falsos engañosos, sus efectos pasan de la simple propagación de información incorrecta, pues pueden comprometer las condiciones necesarias para un debate público libre e informado.

De acuerdo con Limia (2025), este tipo de prácticas representa un riesgo para los fundamentos de la democracia deliberativa, en la medida en que puede distorsionar la información de la opinión pública y afectar la calidad de las decisiones colectivas. Por tal razón, la democracia deliberativa constituye un referente fundamental para analizar los efectos que la desinformación política generada mediante inteligencia artificial puede producir sobre las dinámicas comunicativas que sustentan la legitimidad democrática.

Regulación jurídica digital

La regulación jurídica digital comprende los principios y normas destinados a responder a los efectos que las tecnologías generan sobre los derechos, las libertades y el funcionamiento democrático. En este campo, Lessig (2009) advierte que la conducta en el entorno digital no depende únicamente de la ley, sino también de la arquitectura tecnológica, el mercado y las normas sociales, por lo que resulta necesario considerar la manera en que el diseño de las plataformas condiciona las actuaciones individuales y colectivas.

Las redes sociales y los sistemas de inteligencia artificial deben analizarse como espacios en los que confluyen las decisiones tecnológicas y jurídicas que pueden incidir en el ejercicio de los derechos, por lo que su funcionamiento no puede considerarse neutral (Almache, 2025; García, 2025). En este escenario, la regulación digital no debería limitarse a establecer sanciones, sino de incorporar medidas de prevención, transparencia y responsabilidad que involucren a las plataformas, intermediarios y desarrolladores tecnológicos, especialmente cuando sus estructuras puedan afectar los derechos humanos o el funcionamiento democrático (Balkin, 2018).

Lo anterior cobra especial importancia frente al uso de la inteligencia artificial para producir y difundir contenidos, ya que se requieren reglas que permitan identificar responsabilidades y enfrentar la

desinformación sin imponer restricciones injustificadas a la libertad de expresión y al pluralismo democrático.

Esta exigencia reguladora encuentra un anclaje axiológico en el principio de seguridad jurídica, cuya vigencia dentro del Estado de Derecho contemporáneo exige salvaguardar la certidumbre procesal y material de los ciudadanos frente al impacto disruptivo de las tecnologías emergentes (Acosta et al., 2026). En efecto, la resignificación de esta garantía procesal y constitucional demuestra que la formulación de límites a la desinformación no puede sustentarse en fórmulas vagas o discrecionales que desdibujan la predictibilidad jurídica. Por el contrario, la legitimidad democrática de cualquier control normativo sobre los entornos algorítmicos depende de reglas procesales claras, proporcionales y respetuosas de las garantías sustanciales de los sujetos comunicantes frente a la intervención punitiva o administrativa estatal.

Marco legal

La presente investigación se sustenta en el marco constitucional, legal, jurisprudencial y de política pública colombiana sobre la libertad de expresión, entorno digital e inteligencia artificial, así como en los estándares internacionales de derechos humanos aplicables al discurso político.

Marco internacional

El análisis de la libertad de expresión frente a la desinformación política producida mediante inteligencia artificial en Colombia debe ubicarse dentro del marco institucional conformado por autoridades estatales, órganos judiciales y administrativos y organismos internacionales encargados de proteger los derechos fundamentales. En el ámbito internacional, el Sistema Interamericano de Derechos Humanos constituye un referente para el país a partir de la ratificación de la Convención Americana sobre Derechos Humanos mediante la Ley 16 de 1972.

Su artículo 13 protege la libertad de pensamiento y expresión, prohíbe la censura previa y establece que las restricciones posteriores deben responder a criterios de legalidad, necesidad y proporcionalidad (Convención Americana sobre Derechos Humanos, 1969). Estos principios han sido desarrollados por la Corte Interamericana de Derechos Humanos, que en la Opinión Consultiva OC-5/85 reconoció la libertad de expresión como una condición fundamental para la vida democrática, tanto por el derecho de cada persona a expresar sus ideas como por la dimensión colectiva relacionada con el acceso a la información (Corte IDH, 1985).

Desde esta perspectiva, la desinformación política puede afectar las condiciones necesarias para la formación de una opinión pública libre e informada. Por su parte, la Comisión Interamericana de Derechos Humanos ha advertido que las respuestas estatales frente a la desinformación, particularmente durante los procesos electorales, no deben convertirse en instrumentos de censura o control del discurso público (OEA, 2019). En la misma dirección, la UNESCO (2021) resalta la necesidad de una gobernanza tecnológica basada en la transparencia, la supervisión ética y la protección de los derechos humanos ante los riesgos que plantea la automatización de la comunicación.

La obligatoriedad de estos compromisos convencionales debe articularse a través de las fuentes y principios rectores del Derecho Internacional público, cuyo corpus iuris vincula al Estado colombiano en la protección efectiva de las garantías fundamentales consagradas en los tratados suscritos (Vargas, 2025c). En virtud del principio *pacta sunt servanda* y de la integración de las cláusulas de derechos humanos al bloque de constitucionalidad, las obligaciones internacionales sobre libertad de expresión constituyen barreras infranqueables frente a cualquier tentativa de regulación restrictiva interna. De este modo, la soberanía regulatoria nacional frente a los riesgos tecnológicos debe armonizarse estrictamente con la doctrina y la jurisprudencia de los tribunales internacionales de derechos humanos.

Marco institucional colombiano

En Colombia, la protección de la libertad de expresión tiene como principal referente a la Corte Constitucional y a las garantías previstas en la Constitución Política de 1991, particularmente en los

artículos 20, 73 y 74. La jurisprudencia ha establecido una protección especial para este derecho y ha señalado que cualquier restricción debe estar debidamente justificada. En la Sentencia C-150 de 2015, la Corte estableció reglas relacionadas con su posición preferente y con la aplicación de un juicio estricto de proporcionalidad frente a las limitaciones que se pretendan imponer (Corte Constitucional de Colombia, 2007).

Esta protección adquiere especial importancia cuando se trata del discurso político, debido a su relación con la participación ciudadana y el pluralismo, aspecto desarrollado posteriormente en la Sentencia C-150 de 2015 (Corte Constitucional de Colombia, 2015). Junto con la actuación judicial, el Ministerio de Tecnologías de la Información y las Comunicaciones (MinTIC) tiene incidencia en la regulación y orientación del entorno digital colombiano, particularmente en los asuntos relacionados con el desarrollo y uso de las tecnologías de la información y las comunicaciones.

Creado mediante la Ley 1341 de 2009 y fortalecido por la Ley 1978 de 2019, este marco institucional establece competencias estatales relacionadas con la conectividad, el desarrollo de la sociedad de la información y la transformación digital (Ley 1341, 2009; Ley 1978, 2019). Su actuación resulta pertinente para comprender el contexto tecnológico en el que actualmente circulan contenidos políticos y se desarrollan nuevas formas de comunicación mediadas por inteligencia artificial. Su incorporación al marco institucional responde a que la infraestructura digital influye directamente en la circulación y amplificación de información.

Dicha infraestructura y el despliegue tecnológico estatal y privado no pueden desvincularse de la dinámica de los contratos de transferencia de tecnología, los cuales constituyen instrumentos estratégicos para la adopción e innovación de herramientas operativas en la economía digital (Martínez et al., 2018). La comprensión contractual y funcional de estos esquemas resulta determinante al evaluar cómo las plataformas, los centros de investigación y las corporaciones adquieren o licencian sistemas de procesamiento masivo de datos. En el terreno de la deliberación pública, la absorción estratégica de tecnología sin cláusulas estrictas de gobernanza, transparencia e idoneidad técnica expone al ecosistema

nacional a la importación acrítica de modelos algorítmicos que potencian la desinformación sin salvaguardas contextuales.

De manera paralela, el Departamento Nacional de Planeación (DNP) adquiere especial relevancia dentro del marco institucional colombiano a través de una evolución progresiva de política pública que trasciende la sola expedición del Documento CONPES 4144 de 2025. En primer lugar, este instrumento surge como resultado de un proceso escalonado iniciado con el CONPES 3975 de 2019 (Política Nacional para la Transformación Digital e Inteligencia Artificial), primera política pública nacional que reconoció formalmente a la inteligencia artificial como tecnología estratégica para la Cuarta Revolución Industrial, orientada a generar condiciones habilitantes en infraestructura, talento, innovación y gobernanza (DNP, 2019).

Hay que destacar que el CONPES inicial, aunque finalizó en el año 2022, dejó las bases institucionales para incorporar la inteligencia artificial en la agenda pública colombiana y evidenció la necesidad de actualizar la política estatal frente a los riesgos emergentes como la automatización intensiva, la opacidad algorítmica y la manipulación informática; un proceso que fue fortalecido por la Misión de Expertos en Inteligencia Artificial (2021-2022), coordinada por el DNP, que formuló los lineamientos prospectivos sobre talento, futuro del trabajo, sostenibilidad y desarrollo ecológico (DNP, 2023).

Posteriormente, la estrategia nacional digital 2023 - 2026 (Ley 2294 de 2023) y el Plan Nacional de Desarrollo (2022-2026) consolidaron el reconocimiento de la inteligencia artificial como eje estratégico para la modernización institucional, la productividad y la construcción de una sociedad del conocimiento, mientras que la Hoja de Ruta para el desarrollo y aplicación de la inteligencia artificial en Colombia (Congreso de Colombia, 2023). Presentada por el Ministerio de Ciencia, Tecnología e Innovación (2024), funcionó como antecedente técnico y político inmediato el CONPES 4144.

La Hoja de Ruta definió de manera directa y trazó cinco áreas estratégicas para consolidar una política integral: desde la ética y la gobernanza hasta la educación, la innovación, la infraestructura de datos y la ciberseguridad. Asimismo, propuso mecanismos de articulación entre el Estado, la academia y el sector empresarial bajo principios de transparencia. En este escenario, el **CONPES 4144 de 2025** se consolida como un instrumento orientado a fortalecer la política nacional en materia de inteligencia artificial. Su finalidad es orientar una política pública que permita al Estado aprovechar las posibilidades de la inteligencia artificial, sin dejar de atender los riesgos que puede generar su implementación.

Si bien el CONPES no tiene el carácter de una ley, sí constituye un referente importante para definir las actuaciones estatales en esta materia. Su desarrollo ha permitido incorporar asuntos como los sesgos algorítmicos, la evaluación ética de los sistemas automatizados, la protección de los derechos fundamentales y los riesgos asociados con la desinformación. En este sentido, la evolución del CONPES 3975 (2019), la Misión de Expertos (2023), la Hoja de Ruta (2024) y el CONPES (2025) muestra cómo la discusión sobre inteligencia artificial ha ido adquiriendo mayor alcance en Colombia. Se trata de establecer condiciones para su desarrollo y uso de manera compatible con los derechos humanos y las garantías propias de un sistema democrático.

De la misma manera, la Superintendencia de Industria y Comercio cumple un papel institucional clave como autoridad nacional encargada de proteger los datos personales bajo la Ley 1581 de 2012, especialmente porque el uso automatizado de información personal puede aumentar las prácticas de segmentación política, perfilamiento y microtargeting, relacionadas con las nuevas formas de manipulación informativa (Ley 1581, 2012)

Instituciones democráticas y necesidad de articulación regulatoria

Por la naturaleza política del objeto de estudio, también resulta necesario incorporar el análisis de las instituciones responsables de salvaguardar la integridad democrática. En ese caso, las entidades como la Registraduría Nacional del Estado civil y el Consejo Nacional Electoral, que históricamente están vinculadas a la organización y vigilancia de procesos electorales; amplían su relevancia ante contextos en

los que nuevas tecnologías pueden influir en campañas políticas, circulación de desinformación o modificación de percepciones ciudadanas.

En este panorama, el Congreso de la República es un actor institucional clave, ya que tiene la facultad de legislar sobre los vacíos que dejan las plataformas digitales, la comunicación automatizada y la responsabilidad de lo que se publica en la red. Como bien señalan Kaminski y Jones (2024), la entrada de nuevas tecnologías al espacio público obliga a los legisladores a repensar las categorías legales de siempre, con el fin de adaptarse a los cambios profundos que está viviendo el discurso político hoy en día.

Por lo tanto, la inclusión del marco institucional en esta investigación se sustenta en que la protección de la libertad de expresión frente a la desinformación política generada por inteligencia artificial depende tanto del contenido normativo como de las capacidades institucionales para aplicar, interpretar y adaptar dichas normas (Solar, 2020). Desde esta perspectiva del análisis institucional, permite reconocer cómo actúan y se relacionan las autoridades judiciales, administrativas y regulatorias, junto con los organismos internacionales, frente a los riesgos que plantea el uso de nuevas tecnologías.

De ese modo, el enfoque refuerza el sustento jurídico de la presente investigación, pues evidencia que los desafíos de la inteligencia artificial no pueden abordarse desde una sola institución. Como tal, se requiere la participación coordinada de distintos niveles de gobierno y organismos internacionales.

CAPÍTULO III: ASPECTOS METODOLÓGICOS

Este capítulo presenta la metodología empleada en la investigación “Protección del derecho a la libertad de expresión ante desinformación política por IA en Colombia”. Donde se precisa el tipo y enfoque, el estudio, las técnicas de recolección de información, los criterios, criterios para seleccionar las fuentes y el procedimiento de análisis. Por lo tanto, estos elementos buscan mantener la coherencia entre el problema jurídico, la hipótesis y los objetivos de la investigación, y asegurar un análisis sistemático del fenómeno estudiado (Atencio et al., 2021).

Asimismo, se establece la relación entre las categorías jurídicas, las normas, la jurisprudencia y las doctrinas revisadas, junto con los criterios empleados para avalar y el marco jurídico colombiano ofrece una respuesta suficiente frente a la desinformación política general amplificada mediante la inteligencia artificial.

Enfoque de la investigación

La investigación estudia el fenómeno en cuestión a través de un enfoque cualitativo porque busca comprender e interpretar las normas, los cuerpos de sentencias y los aportes doctrinales que se relacionan con la libertad de expresión y la generación de desinformación política a partir de la inteligencia artificial.

En este sentido, la perspectiva que se estudia realiza un examen sobre las disposiciones del ordenamiento jurídico y de las posibles tensiones entre los derechos fundamentales y los principios constitucionales (Borjas, 2020). Desde el examen de lo anterior se llevan a cabo reflexiones sobre aspectos tales como la prohibición de la censura previa o la responsabilidad posterior, la proporcionalidad o la ponderación, con el objetivo de determinar cómo puede reaccionar el ordenamiento jurídico ante la producción y circulación automatizadas de contenidos políticos (Valladolid & Chávez, 2020).

También se considera la atribución de responsabilidad, especialmente en la relación con los sujetos que participan en la generación, modificación, difusión o ampliación de contenidos mediante la inteligencia artificial y con las circunstancias que podrían derivar en consecuencias jurídicas.

La complejidad de dicho abordaje metodológico se fundamenta en la perspectiva tridimensional del derecho, la cual exige articular el análisis formal de las normas con la dimensión fáctica y la valoración axiológica inherente a las transformaciones sociales contemporáneas (Vargas, 2025b). Dentro de la investigación jurídica, esta tríada epistemológica permite superar el dogmatismo formalista para comprender cómo la desinformación algorítmica incide en la realidad política y en la eficacia material de los mandatos constitucionales. Asumir este enfoque resulta decisivo para examinar la libertad de expresión digital no como una abstracción normativa estática, sino como un valor democrático vivo que demanda un método interpretativo dialéctico ante el avance tecnológico.

Tipo de investigación

El estudio se enmarca en una investigación de tipo jurídica doctrinal, con alcance descriptivo y exploratorio (Sampieri, 2023); donde el carácter descriptivo se evidencia en la sistematización del marco normativo colombiano, la jurisprudencia constitucional y los estándares internacionales aplicables a la libertad de expresión, particularmente aquellos relacionados con el discurso político y el funcionamiento del sistema democrático en entornos digitales (Hernández & Mendoza, 2020).

El carácter judicial-doctrinal se refleja en la revisión de normas constitucionales y legales, decisiones judiciales y fuentes doctrinales, con el fin de identificar los criterios que regulan la libertad de expresión y sus límites. A su vez, el alcance descriptivo permite identificar el contenido y alcance de dichos criterios, mientras que el alcance exploratorio permite examinar los problemas de aplicación que surgen cuando estas categorías se proyectan sobre contenidos políticos generados o amplificadas mediante inteligencia artificial.

Por su parte, el alcance exploratorio responde a que la relación entre inteligencia artificial, desinformación política y libertad de expresión constituye un campo de estudio reciente dentro del derecho constitucional colombiano, en el que se presentan vacíos interpretativos y desafíos de aplicación frente a la automatización del discurso político (Galarza, 2020).

Aunado a la anterior, la investigación busca identificar los principales debates jurídicos y doctrinales, así como los criterios que permitan interpretar el ordenamiento colombiano frente a los nuevos desafíos tecnológicos. Además, incorpora un análisis comparativo de estándares y enfoques presentes en instrumentos internacionales y fuentes doctrinales que serán contrastadas en el marco jurídico nacional, lo que permite reconocer coincidencias, diferencias y criterios que pueden resultar útiles para abordar el problema de investigación.

Método de investigación

El método principal es jurídico analítico, utilizado para examinar normas, sentencias y fuentes doctrinales e identificar los principios aplicables al problema de investigación (Maldonado et al., 2021). A esto se suma el método sistemático que permite comprender las normas dentro del conjunto del ordenamiento jurídico y establecer su relación con los principios constitucionales que protegen la libertad de expresión.

También se aplica el método hermenéutico jurídico para interpretar las normas y decisiones judiciales, considerando su contexto, finalidad, evolución frente a los cambios tecnológicos; es así como su uso resulta importante ante los nuevos escenarios generados por la inteligencia artificial, especialmente en materia de desinformación política (Vásquez, 2024). De esta forma, se busca establecer el alcance de las reglas y principios existentes y determinar si requieren ser aplicados, reinterpretados o complementados frente a estas nuevas situaciones.

Finalmente, la comparación entre el marco jurídico colombiano y los estándares internacionales seleccionados permite establecer los principales puntos de coincidencia y diferencia entre ambos referentes, y este contraste contribuye a identificar los elementos que resultan importantes para comprender la respuesta jurídica frente a la libertad de expresión y la desinformación política en entornos digitales.

Diseño de la investigación

La investigación se desarrolla bajo un diseño documental, centrado en la revisión y análisis de fuentes normativas, jurisprudenciales, doctrinales e institucionales relacionadas con el problema de estudio. Como tal, el corpus documental está integrado por fuentes primarias y secundarias seleccionadas de acuerdo con criterios de pertinencia jurídica, relación directa con el problema de investigación, autoridad de la fuente y actualidad, sin que la fecha de publicación constituya por sí sola un criterio de inclusión o exclusión.

Es así como el diseño documental permite examinar de manera sistemática:

- **Normas constitucionales y legales:** Las bases que rigen actualmente el ordenamiento jurídico en Colombia, con especial atención a las disposiciones relacionadas con la libertad de expresión, el discurso político, la responsabilidad y la protección del orden democrático.
- **Jurisprudencia de la Corte Constitucional:** Específicamente las sentencias y fallos relacionados con libertad de expresión, discurso político y restricciones a este derecho, responsabilidad ulterior, censura previa, proporcionalidad y protección del debate público.
- **Estándares del Sistema Interamericano:** Los tratados y lineamientos de derechos humanos que vinculan al país, especialmente aquellos relacionados con la libertad de expresión, las restricciones legítimas al discurso y la circulación de información en contextos democráticos.
- **Literatura académica especializada:** Investigaciones sobre inteligencia artificial, desinformación y regulación del discurso en entornos digitales, así como estudios jurídicos que permitan identificar los principales debates sobre responsabilidad y regulación tecnológica.

- **Estudios interdisciplinarios:** Tal contribución se la puede considerar como proveniente de áreas como el derecho, la comunicación política y la ética tecnológica, puesto que nos han permitido identificar, de manera directa, cómo se produce o amplifica la desinformación política por medio de la inteligencia artificial y cuáles son los efectos que esto tiene en el debate democrático.

De este modo, el diseño de la investigación es adecuado a la hora de detectar los fundamentos legales ya existentes y de analizar si son suficientes para responder a los retos que trae aparejada la inteligencia artificial en la elaboración y difusión de los mensajes políticos.

La selección del corpus no se limita, por tanto, a reunir documentos relacionados con inteligencia artificial, ya que responde a la necesidad de reconstruir las categorías jurídicas que permiten analizar el problema. Asimismo, se examinan las fuentes que aportan elementos para determinar: (i) el alcance de la libertad de expresión en el discurso político; (ii) la prohibición de censura previa; (iii) las condiciones de la responsabilidad ulterior; (iv) la aplicación de los principios de necesidad y proporcionalidad frente a restricciones estatales; (v) la protección del orden democrático; y (vi) la atribución de responsabilidad por contenidos generados, modificados o amplificados mediante inteligencia artificial.

Técnicas de recolección de información

La técnica principal de recolección de información es la revisión documental, la cual permite identificar, seleccionar y analizar fuentes relevantes para el desarrollo de la investigación. Dentro de esta técnica se emplean las siguientes herramientas:

a) Fichas de análisis documental (Fichas RAE)

Se utilizan para sistematizar información proveniente de artículos científicos, libros, documentos institucionales y estudios doctrinales relacionados con libertad de expresión, inteligencia artificial y desinformación política, las cuales permiten identificar objetivos, enfoques teóricos, metodologías empleadas y principales hallazgos de cada estudio.

Además de estos elementos, las fichas permiten registrar la definición o tratamiento que cada fuente realiza de las categorías relevantes para la investigación, así como sus principales argumentos sobre la relación entre libertad de expresión, desinformación política, inteligencia artificial y protección del orden democrático.

b) Matriz de sistematización de información

Permite organizar la información recopilada a partir de criterios como:

- Autor.
- Año de publicación.
- País de estudio.
- Problema jurídico abordado.
- Enfoque metodológico.
- Principales resultados.
- Aportes al estudio de la libertad de expresión frente a la IA.

De ese modo, la matriz facilita la trazabilidad de la información y permite identificar tendencias doctrinales y vacíos investigativos. La principal herramienta de recolección organización de la información será una matriz de sistematización temática y jurídica, lo cual su función es ordenar las fuentes seleccionadas y facilitar el análisis de los estudios y documentos que abordan la relación entre inteligencia artificial, desinformación, política, libertad de expresión y democracia.

Para ello, la matriz registró datos como autor, año y contexto, además de los conceptos de desinformación, las tecnologías empleadas y los riesgos identificados para derechos fundamentales y el orden democrático, y para comparar los principales aportes jurídicos de cada fuente. La matriz incluye categorías centrales como la libertad de expresión, censura, previa responsabilidad, ulterior, atribución de responsabilidad, necesidad, proporcionalidad y protección del ordenamiento democrático.

Por otro lado, en las fuentes jurisprudenciales se registra la sentencia de los hechos relevantes, el problema jurídico, los derechos involucrados, la regla o su regla aplicable y los criterios utilizados para

proteger o restringir la libertad de expresión. Asimismo, se tuvo en cuenta la responsabilidad atribuida a los sujetos involucrados y aquellos elementos que puedan resultar pertinentes para analizar la generación o difusión automatizada de contenidos.

En definitiva, esta herramienta permitirá identificar patrones, reconocer vacíos normativos y comprender las tensiones constitucionales asociadas al fenómeno estudiado. Asimismo, permitirá comparar los criterios encontrados en las distintas fuentes y establecer cuáles de ellos pueden ser trasladados al análisis de la desinformación política generada o amplificada mediante inteligencia artificial y cuáles presentan dificultades de aplicación. De este modo, se garantiza que el análisis mantenga criterios de rigurosidad, coherencia y trazabilidad desde la revisión documental inicial hasta el cumplimiento de los objetivos de la investigación.

c) Análisis jurisprudencial

Para este caso, se examinan sentencias relevantes de la Corte Constitucional colombiana relacionadas con libertad de expresión, discurso político y restricciones legítimas a derechos fundamentales, con el fin de identificar criterios aplicables al fenómeno de la desinformación generada por inteligencia artificial.

El análisis jurisprudencial se basa tanto en la identificación de las sentencias que mencionen la libertad de expresión, y buscar a la vez, reconstruir las reglas y subreglas jurídicas relevantes para el problema de investigación. Para ello, en cada decisión seleccionada se identifican: el problema jurídico, los derechos y principios involucrados, la regla o subregla jurisprudencial establecida, los criterios utilizados por la Corte para determinar la legitimidad de una restricción a la libertad de expresión, el tratamiento de la censura previa y la responsabilidad ulterior, y la aplicación de los principios de necesidad y proporcionalidad.

La selección de las decisiones jurisprudenciales se justifica por su relevancia material para estas categorías y no únicamente por su fecha de expedición. De esta manera, se incorporan decisiones anteriores a 2018 cuando constituyan precedentes fundamentales para comprender la doctrina

constitucional sobre libertad de expresión, y decisiones posteriores cuando aporten elementos relacionados con los nuevos desafíos tecnológicos o con la aplicación contemporánea de dichas categorías.

Criterios de selección de la información

Para garantizar la calidad y pertinencia de la información analizada, se establecen los siguientes criterios de selección:

- Publicaciones académicas en revistas indexadas o editoriales reconocidas.
- Estudios publicados preferiblemente entre los años 2018 y 2026, con el fin de incorporar desarrollos recientes sobre inteligencia artificial. Este rango temporal se utilizará como criterio de actualización para la literatura especializada, pero no como límite absoluto, debido a que determinadas obras o decisiones anteriores pueden resultar indispensables para reconstruir las categorías jurídicas y los precedentes relevantes para la investigación.
- Documentos institucionales emitidos por organismos internacionales como la OEA, UNESCO y el sistema interamericano de derechos humanos, especialmente aquellos que desarrollen estándares sobre libertad de expresión, desinformación, democracia y regulación de contenidos digitales.
- Normativa vigente en el ordenamiento jurídico colombiano.
- Jurisprudencia constitucional relevante para la interpretación de la libertad de expresión, seleccionada por su aporte al desarrollo de las categorías jurídicas objeto de análisis.
- Literatura doctrinal relacionada con derecho constitucional, derecho digital, comunicación política e inteligencia artificial.

Adicionalmente, se aplicarán criterios de inclusión relacionados con la pertinencia temática, autoridad de la fuente, relación directa con el problema jurídico y capacidad de aportar elementos para el análisis de las categorías definidas.

Técnicas de análisis de la información

La información recopilada fue examinada mediante el análisis de contenido jurídico, con el propósito de reconocer los conceptos, categorías y principios relacionados con la libertad de expresión y la desinformación política generada o amplificada mediante inteligencia artificial. Para ello, se realizó una revisión sistemática de las fuentes seleccionadas y se organizaron los hallazgos de acuerdo con las categorías definidas previamente. Como tal, estas funcionaron como unidades de análisis para determinar la presencia, alcance y tratamiento de los aspectos estudiados dentro del corpus documental. El procedimiento se desarrolló en las siguientes etapas.

- **Rastreo de categorías jurídicas:** Se identifican y organizan las categorías centrales de la investigación, entre ellas la libertad de expresión, la censura previa a la responsabilidad ulterior, la atribución de responsabilidad y la proporcionalidad y la protección del orden democrático. También se consideran aspectos como la transparencia, la trazabilidad, el papel de las plataformas digitales y la automatización de contenidos.
- **Operacionalización de las categorías:** Cada categoría se examina mediante indicadores jurídicos específicos. En el caso de la libertad de expresión, se identifica su alcance y el nivel de protección otorgado al discurso político; respecto de la censura previa, se determina si la medida analizada supone una prohibición anticipada de la circulación del contenido; frente a la responsabilidad ulterior, se examinan las condiciones jurídicas para atribuir consecuencias posteriores a la difusión; respecto de la proporcionalidad, se identifican la finalidad legítima, la idoneidad, la necesidad y la proporcionalidad en sentido estricto de la medida restrictiva; y, en materia de atribución de responsabilidad, se identifica el sujeto al que se atribuye la conducta, su grado de intervención y el fundamento jurídico de dicha atribución.
- **Jerarquización por relevancia:** Se clasifica toda la información obtenida según su relación con la pregunta de investigación y su aporte al cumplimiento de los objetivos

específicos, priorizando las fuentes que permitan establecer reglas, principios, criterios jurisprudenciales o argumentos directamente relacionados con las categorías de análisis.

- **Lectura integral de normas y fallos:** Se examinan las leyes y la jurisprudencia de manera conjunta, tomando como referencia los principios constitucionales y los tratados internacionales sobre los derechos humanos. En el caso de las decisiones judiciales, la lectura se orienta a identificar tanto las consideraciones generales de la sentencia, como también las reglas o subreglas que puedan ser aplicables al problema investigado, prestando especial atención a los criterios utilizados para proteger o restringir la libertad de expresión, al tratamiento de la censura previa y la responsabilidad ulterior, a la aplicación del principio de proporcionalidad y a los elementos relevantes para determinar la atribución de responsabilidad.
- **Debate de posturas doctrinales:** Se contrastan las diferentes visiones académicas que existen sobre cómo debería regularse la inteligencia artificial en el marco del discurso político, identificando coincidencias, diferencias y argumentos relevantes para la protección de la libertad de expresión y del orden democrático.
- **Análisis comparativo:** Se confrontan los criterios del ordenamiento colombiano en los estándares del sistema Interamericano de Derechos Humanos y las fuentes doctrinales institucionales seleccionadas, con el fin de establecer coincidencias, diferencias y posibles criterios de interpretación frente al problema estudiado.
- **Identificación de vacíos y dificultades de aplicación:** Se revisan posibles vacíos normativos, ambigüedades y dificultades relacionadas con la atribución de responsabilidades por contenidos generados o modificados mediante inteligencia artificial, el papel de las plataformas digitales, los límites estatales a la libertad de expresión de acuerdo con los principios de necesidad y proporcionalidad.

De ese modo, los resultados de estas etapas se organizan en una matriz que relaciona los hallazgos con las categorías jurídicas, los objetivos específicos y la pregunta de investigación. Con ello se establece el nivel de suficiencia del marco constitucional, legal y jurisprudencial colombiano y se identifican los elementos que sirven de base para formular los criterios jurídicos propuestos.

Además, la articulación entre las categorías de análisis, las fuentes seleccionadas, el procedimiento de interpretación y los objetivos específicos se estableció de manera transversal durante el proceso de revisión documental. Las categorías relacionadas con la libertad de expresión, la censura previa, la responsabilidad ulterior, la atribución de responsabilidad, la necesidad y proporcionalidad, la transparencia, la trazabilidad y la protección del orden democrático orientaron la selección y lectura de las normas, decisiones jurisprudenciales, estándares internacionales, fuentes doctrinales e institucionales.

A partir de estas categorías, las fuentes fueron examinadas mediante el análisis de contenido jurídico, la lectura sistemática y hermenéutica, el contraste de posturas y el análisis comparativo, con el propósito de identificar los criterios jurídicos, coincidencias, diferencias, vacíos y dificultades de aplicación vinculados con cada objetivo específico. De esta manera, el proceso de interpretación permitió relacionar los hallazgos documentales con los objetivos planteados y establecer los elementos jurídicos necesarios para responder la pregunta de investigación.

CAPÍTULO IV: ANÁLISIS Y DISCUSIÓN DE RESULTADOS

Análisis doctrinal sobre libertad expresión, desinformación política e inteligencia artificial

El análisis de la literatura nacional e internacional permitió identificar las principales discusiones doctrinales relacionadas con la protección de la libertad de expresión frente a los fenómenos de desinformación política asociados al uso de tecnologías digitales e inteligencia artificial. La revisión realizada muestra que el debate académico ha pasado de centrarse en las *fake news* y la posverdad en redes sociales a abordar fenómenos más complejos como la inteligencia artificial generativa, los *deepfakes*, la automatización de contenidos y la manipulación algorítmica del debate público.

Como tal, el problema jurídico, además de establecer una información si es verdadera o falsa, también implica analizar cómo se producen, modifican y difunden los contenidos, quién interviene en estos procesos y qué efectos pueden tener sobre la libertad de expresión y el funcionamiento democrático.

La cuestión adquiere relevancia constitucional cuando la producción y circulación automatizada de contenidos puede afectar las condiciones de formación de la opinión pública y, simultáneamente, generar respuestas estatales susceptibles de restringir la libertad de expresión. De ese modo, la literatura revisada evidencia una tensión entre la protección de la libertad de expresión, especialmente en el ámbito del discurso político, y la necesidad de preservar las condiciones que sostienen el funcionamiento democrático.

La revisión permitió identificar cuatro líneas temáticas principales que orientan el análisis: en primer lugar, la transformación del debate democrático en entornos digitales; en segundo lugar, la inteligencia artificial como mecanismo de generación, automatización y amplificación de contenidos políticos; en tercer lugar, las tensiones jurídicas derivadas de la regulación de contenidos y su relación con la libertad de expresión; y, finalmente, las dificultades normativas relacionadas con la atribución de responsabilidad, la transparencia y la protección del orden democrático frente a la utilización política de sistemas de inteligencia artificial.

Tabla 1.*Clasificación temática de la literatura analizada*

Categoría temática	Autores	Hallazgo principal
Fake news y democracia digital	Estupiñán et al. (2020); Escobar (2023); Ortiz y Agudelo (2023); Manfredi et al. (2021)	Las redes sociales transformaron la comunicación política y facilitaron la circulación masiva de desinformación electoral
IA generativa y deepfakes	Battista (2024, 2025); Vélez (2026); Pawelec y Labuz (2026); Beacken y Woolley (2026)	La IA amplifica riesgos de manipulación política mediante contenidos sintéticos y automatización discursiva
Libertad de expresión y regulación	Arango (2021); Rossel (2025); Jiménez (2025); Filho et al. (2022)	La regulación de contenidos digitales debe evitar restricciones desproporcionadas y censura previa
Gobernanza digital y transparencia algorítmica	UNESCO (2021); Carta Iberoamericana (2023); CONPES 3975 de 2019	Se requieren mecanismos de supervisión, transparencia y responsabilidad en sistemas de IA

Nota. Elaboración propia a partir de la literatura sistematizada.

Uno de los hallazgos identificados en la literatura es que las redes sociales no constituyen únicamente nuevos canales de comunicación, sino que han modificado las condiciones de producción, circulación y recepción del discurso político.

La expansión de Internet amplió las posibilidades de participación ciudadana y la diversidad de actores que intervienen en el debate público, pero también facilitó una circulación más rápida de la información, así como fenómenos de polarización y exposición selectiva a determinados contenidos. En Colombia, Estupiñán et al. (2020) evidenciaron, a partir del análisis de las elecciones regionales de 2019, que las fake news pueden aprovechar emociones, prejuicios y creencias previamente arraigadas para

alcanzar una mayor difusión. Este aspecto resulta relevante desde el ámbito jurídico, pues muestra que los efectos de la desinformación no dependen únicamente de que un contenido sea falso, sino también de las condiciones sociales y tecnológicas que favorecen su propagación.

En la misma línea, Manfredi et al. (2021) encontraron que las campañas políticas digitales en Colombia se desarrollan en entornos vulnerables a la desinformación, entre otras razones por los bajos costos de difusión y las dificultades para establecer mecanismos eficaces de control y verificación. Estas condiciones pueden afectar la formación libre de la opinión pública, especialmente cuando la información política circula de manera rápida y resulta difícil comprobar su veracidad.

Ortiz y Agudelo (2023) señalan que las redes sociales adquirieron un papel central en la comunicación política durante las elecciones presidenciales de 2022, al facilitar la interacción entre candidatos y ciudadanos. Sin embargo, estos espacios también presentan riesgos relacionados con la polarización, la difusión de información falsa y la radicalización del discurso. Escobar (2023) añade que los sistemas de recomendación y los mecanismos que favorecen la viralización pueden determinar qué contenidos alcanzan mayor visibilidad, de modo que el problema no se limita al contenido que circula, sino también a las condiciones tecnológicas que favorecen su difusión.

Inteligencia artificial generativa, deepfakes y automatización del discurso político

La segunda gran línea doctrinal identificada hace referencia al campo de estudio dedicado al impacto que la inteligencia artificial tiene en la creación y difusión de la desinformación política.

Esta línea doctrinal se sitúa frente a la anterior, centrada en las fake news de la política tradicional, ya que, en contraposición con la noción ampliamente utilizada de fake news tradicionales, los estudios recientes advierten sobre un notable incremento de las posibilidades de manipulación mediante inteligencia artificial generativa. Estas herramientas son capaces de generar textos, imágenes, audios y vídeos sintéticos con un alto grado de realismo, lo que plantea nuevos desafíos para la identificación y control de los contenidos desinformativos en el ámbito político.

Battista (2024) sostiene que la inteligencia artificial generativa introduce nuevas posibilidades de automatización de la comunicación política y de producción de contenidos sintéticos. El elemento relevante para esta investigación consiste en la capacidad técnica de generar contenido y en la posibilidad de producirlo a gran escala y adaptarlo a diferentes públicos, aumentando potencialmente su capacidad de persuasión.

En un trabajo posterior, Battista (2025) advierte que la expansión de la IA generativa aumenta los riesgos de desinformación y también transforma la relación entre verdad, opinión pública y legitimidad democrática. Este planteamiento resulta relevante porque desplaza la discusión desde la falsedad individual de un mensaje hacia la capacidad estructural de los sistemas tecnológicos para alterar las condiciones bajo las cuales los ciudadanos reciben y valoran información política.

En la misma línea, Vélez (2026), al estudiar el contexto electoral ecuatoriano, identifica que la inteligencia artificial puede contribuir a ampliar la participación y el acceso a la información, pero también facilitar procesos de desinformación y manipulación electoral. Este doble efecto confirma que la tecnología no puede ser jurídicamente valorada de manera abstracta como positiva o negativa; la respuesta debe atender al uso concreto del sistema, al grado de afectación de los derechos y a la finalidad perseguida.

Asimismo, Pawelec y Labuz (2026) identifican que los deepfakes y los llamados “*softfakes*” generan riesgos éticos y jurídicos importantes, especialmente porque permiten crear representaciones audiovisuales falsas capaces de modificar la percepción ciudadana sobre actores políticos. Bajo esta lógica, el fenómeno no sólo fragmenta el debate público, sino que termina por normalizar la manipulación informativa como algo cotidiano. En una línea similar, Beacken y Woolley (2026) relacionan la IA generativa con una expansión de la propaganda computacional, debido a la posibilidad de automatizar la producción y distribución coordinada de contenidos.

La gravedad de este fenómeno encuentra un referente normativo crítico en el estudio de iniciativas como el Take it Down Act en el derecho estadounidense, estructurado para sancionar la

difusión no consentida de representaciones sintéticas y deepfakes que lesionan la intimidad y la dignidad (Vargas, 2025a). Aunque dicha legislación se concentra preponderantemente en la violencia digital y la explotación de la imagen personal, su lógica regulatoria arroja lecciones esenciales sobre la dificultad de imputar dolo, identificar la cadena de custodia digital y neutralizar la viralización. Proyectada esta experiencia hacia el ámbito electoral, se constata que la producción deliberada de falsificaciones políticas demanda tipificaciones y remedios sancionatorios que no vacíen el debate público, pero que castiguen la usurpación fraudulenta de la identidad democrática.

Las revisiones de la literatura considerada permiten establecer tres dimensiones que incrementan la relevancia jurídica de la información generada a través de la inteligencia artificial: la escalabilidad, asociada a la capacidad de producir muy grandes volúmenes de contenido; la personalización, vinculada al empleo de datos y herramientas de segmentación; y las dificultades para establecer si hay responsables de la participación del usuario, del desarrollador, de las plataformas y de los sistemas automatizados. Todas estas características, en su conjunto, establecen un cambio respecto a los tradicionales contenidos desinformativos, ya que se amplían tanto las características del alcance de los contenidos como las dificultades para conocer su procedencia y establecer quién debe hacerse responsable por el resultado de estos.

Mientras un contenido falso convencional puede ser atribuido con relativa facilidad a una persona o grupo determinado, los sistemas de inteligencia artificial pueden introducir diferentes niveles de intervención humana y tecnológica en la producción, modificación, distribución y amplificación del mensaje y esta circunstancia convierte la atribución de responsabilidad en una categoría jurídica central para la investigación.

Por consiguiente, la doctrina revisada no permite afirmar que la inteligencia artificial haya eliminado las categorías jurídicas existentes, pero sí evidencia que su aplicación requiere ser examinada frente a nuevas formas de producción y circulación del discurso político. En este sentido, el principal

desafío consiste en determinar hasta qué punto las categorías tradicionales de libertad de expresión, responsabilidad ulterior, proporcionalidad y protección de derechos de terceros pueden responder adecuadamente a contenidos producidos o amplificados mediante sistemas automatizados.

La exigencia de salvaguardar el discurso público frente a determinaciones discrecionales remite al valor de la argumentación jurídica como eje articulador del control judicial y garantía insoslayable frente a la arbitrariedad decisonal (Vargas, 2013). En el ámbito de la deliberación constitucional, la sólida justificación dogmática y racional de las providencias judiciales opera como un límite infranqueable para evitar que la ponderación derive en criterios meramente subjetivos de conveniencia. Al abordar la desinformación automatizada, los tribunales colombianos deben robustecer su carga argumentativa de motivación, garantizando que toda calificación restrictiva sobre una emisión digital esté sustentada en juicios rigurosos de necesidad, razonabilidad y respeto irrestricto al principio democrático.

Tabla 2.

Tensiones jurídicas identificadas en la literatura especializada

Derecho o principio involucrado	Riesgo derivado de la IA	Posible respuesta jurídica identificada en la doctrina
Libertad de expresión	Restricciones excesivas o censura digital	Regulación proporcional y prohibición de censura previa
Derecho a la información	Deepfakes y desinformación automatizada	Transparencia y verificación de contenidos
Democracia participativa	Manipulación algorítmica y polarización	Supervisión institucional y control electoral
Pluralismo informativo	Segmentación automatizada	Transparencia algorítmica

Honra y buen nombre	Difusión masiva de contenido sintético falso	Responsabilidad judicial	ulterior y protección
---------------------	---	-----------------------------	-----------------------

Nota. Elaboración propia con base en la literatura revisada.

Libertad de expresión y regulación de la desinformación política

Otro de los hallazgos relevantes corresponde a la tensión existente entre la necesidad de enfrentar la desinformación política y la obligación estatal de proteger la libertad de expresión. La literatura revisada permite identificar una coincidencia fundamental: la existencia de riesgos asociados a la inteligencia artificial no autoriza, por sí misma, la adopción de mecanismos generales de censura ni restricciones desproporcionadas sobre el discurso político.

Arango (2021) plantea que las *fake news* y la posverdad no son solo mentiras, sino que golpean derechos tan profundos como la verdad, la memoria histórica y la propia democracia. Sin embargo, su planteamiento también resulta relevante para establecer un límite a la intervención estatal: las medidas destinadas a combatir la desinformación deben preservar el núcleo de la libertad de expresión y evitar que conceptos jurídicamente indeterminados sean utilizados para restringir opiniones o críticas legítimas.

Filho et al. (2022) sostienen que el uso de redes automatizadas y *deepfakes* distorsiona el pluralismo y altera lo que se conoce como el libre mercado de ideas. Desde esta perspectiva, la respuesta jurídica puede orientarse hacia la identificación de comportamientos coordinados o mecanismos de amplificación artificial, en lugar de establecer controles basados exclusivamente en la valoración del contenido ideológico de los mensajes.

Por su parte, Rossel (2025) destaca la necesidad de que la regulación de las plataformas digitales sea clara, proporcional y compatible con los derechos fundamentales. Asimismo, Jiménez (2025) identifica posibles afectaciones a la honra, la imagen, los datos personales y la participación política derivadas de los deepfakes, pero advierte igualmente sobre la necesidad de preservar las garantías propias de la libertad de expresión. Como tal, estos planteamientos permiten identificar que el criterio de

proporcionalidad adquiere una función transversal; y de ese modo, la regulación puede evaluarse tanto por su finalidad de combatir la desinformación, como también por la intensidad de la restricción que produce sobre la libertad de expresión y por la existencia de alternativas menos restrictivas.

Por lo tanto, la doctrina revisada permite establecer que una intervención jurídica legítima debería diferenciar entre: (i) la expresión de opiniones, incluso aquellas que puedan resultar polémicas; (ii) la difusión de información susceptible de afectar derechos de terceros; (iii) la producción deliberada de contenidos sintéticos destinados a generar engaño; y (iv) las estrategias coordinadas de manipulación automatizada; y esta diferenciación resulta necesaria para evitar que la categoría general de “desinformación” termine comprendiendo expresiones constitucionalmente protegidas.

Vacíos doctrinales

La revisión doctrinal permitió identificar importantes avances en materia de ética, gobernanza y regulación de inteligencia artificial; sin embargo, también evidenció que existe una distancia entre los principios generales de protección de derechos y la definición de mecanismos jurídicos concretos aplicables a la desinformación política generada o amplificada mediante IA. En distintos países y organismos internacionales se han planteado principios orientados a la transparencia, la gestión de riesgos, la supervisión humana y la protección de los derechos fundamentales frente al uso de la inteligencia artificial.

Sin embargo, estos avances aún dejan abierta una cuestión de especial importancia para Colombia: determinar quién debe responder jurídicamente cuando un contenido político generado mediante IA causa una afectación y bajo qué criterios puede atribuirse dicha responsabilidad. Esta preocupación también está presente en países como Estados Unidos, Italia, España, Brasil, Chile y Ecuador, así como en varios Estados de la Unión Europea, donde se han reconocido los desafíos que plantea el desarrollo de estas tecnologías para la protección de los derechos fundamentales y el funcionamiento democrático.

En Estados Unidos, por ejemplo, autores como Beacken y Woolley (2026), Downes (2025) y Kelani y Lepadatu (2026) han documentado cómo la IA generativa y los *deepfakes* ya están moldeando las campañas electorales, alimentando la polarización y sofisticando la propaganda computacional. Por otro lado, en Italia, Battista (2024/2025) advierte que estos contenidos sintéticos tienen un poder real para quebrar la confianza en las instituciones. Una preocupación que comparte Jiménez (2025) en España, quien señala que los *deepfakes* son una amenaza directa no solo para la información, sino para derechos tan personales como la honra y la propia imagen de los actores políticos.

En América Latina el panorama es igual de complejo. En Brasil, Filho et al. (2022) subrayan cómo las redes automatizadas distorsionan el pluralismo y ensucian el "mercado de ideas" con mentiras que parecen verdades. Mientras tanto, en Chile, Rossel (2025) destaca que la IA aplicada a la política está empujando discursos de odio y contenidos discriminatorios, lo que nos obliga a sentarnos a repensar dónde termina la libertad de expresión y dónde empieza la protección de nuestros derechos fundamentales. Por su parte, en Ecuador, Vélez (2026) analiza el impacto de la IA en las elecciones presidenciales de 2025 y advierte que la automatización de contenidos políticos incrementa los riesgos de manipulación electoral y desinformación masiva.

En el caso colombiano, los estudios revisados coinciden en que el ordenamiento jurídico posee mecanismos generales de protección de la libertad de expresión, la honra, el habeas data y la participación democrática; sin embargo, no cuenta con herramientas suficientemente precisas para enfrentar fenómenos contemporáneos asociados a la inteligencia artificial generativa. Investigaciones clave en el país, como las de Arango (2021), Álvarez (2021) y Estupiñán et al. (2020), dejan claro que las *fake news*, la posverdad y la manipulación digital ya no son amenazas teóricas, sino riesgos reales para la democracia colombiana.

Las fuentes revisadas demuestran que el avance de la inteligencia artificial supera la capacidad de la regulación colombiana para atender los nuevos riesgos que plantea su utilización. Entre estos riesgos se encuentran la transparencia de los sistemas algorítmicos, la identificación de contenidos generados artificialmente y el control de campañas automatizadas. Ante esta situación, organismos como la

UNESCO (2021), la OEA (2017), la ONU (2025) y la OSCE (2025) han propuesto y recomendado orientaciones relacionadas con la necesidad de garantizar el control humano, la trazabilidad y la protección de los derechos humanos, haciendo hincapié en que las medidas adoptadas para combatir la desinformación no deben convertirse en mecanismos de censura.

De forma similar, la Carta Iberoamericana de Derechos en Entornos Digitales (2023), así como los proyectos del Consejo de Europa y de la Unión Europea, desarrollan estos criterios de transparencia, gestión de riesgos y protección de los derechos fundamentales. Desde la perspectiva del análisis doctrinal, una de las vías señaladas consiste en atribuir la responsabilidad que pueden tener los usuarios en interacción con los desarrolladores, las plataformas y los sistemas automatizados. Para ello, se debe considerar el grado de intervención, el conocimiento del riesgo, el control sobre el contenido y su previsibilidad.

De este modo, la trazabilidad y la transparencia pueden ayudar a establecer cómo se genera y cómo circula un contenido, aunque no son suficientes por sí solas para justificar cualquier intervención estatal. En definitiva, la revisión pone en evidencia que, para articular adecuadamente el binomio libertad de expresión-responsabilidad ulterior, se requieren criterios que permitan integrar aspectos como la trazabilidad, la transparencia, la identificación de contenidos sintéticos y la atribución de responsabilidades en el contexto del uso de la inteligencia artificial.

El debate en torno a los mecanismos técnicos de intervención inmediata guarda correspondencia con las discusiones contemporáneas sobre el bloqueo dinámico en telecomunicaciones y propiedad intelectual, donde las medidas cautelares en tiempo real buscan contener infracciones continuadas en la red (Vargas, 2026). Sin embargo, el traslado de estas órdenes cautelares de interrupción al terreno de la circulación de información y del debate político impone una frontera infranqueable por el riesgo inminente de consumir actos de censura previa. La doctrina advierte que, si bien la tutela efectiva exige dinamismo tecnológico, en materia de opinión pública y derechos humanos la respuesta estatal debe

prescindir de bloqueos automáticos, priorizando la celeridad probatoria judicial y la atribución de responsabilidades ulteriores.

Análisis jurisprudencial y normativo

En el marco del segundo objetivo específico, se realizó el estudio del esquema jurídico colombiano, la jurisprudencia constitucional y los estándares internacionales relativos a la libertad de expresión y la desinformación política, considerando aspectos como la prohibición de la censura previa, la responsabilidad ulterior por el ejercicio de la expresión, la proporcionalidad, la protección del orden democrático y la atribución de responsabilidades.

Este análisis también se desarrolló a partir de la observación de fenómenos recientes relacionados con la inteligencia artificial generativa y su uso en la producción de deepfakes, contenidos automatizados, propaganda computacional y segmentación política mediante algoritmos.

El objetivo del análisis no fue cuestionar las decisiones judiciales, sino identificar en qué medida las categorías jurídicas existentes resultan útiles frente a estas nuevas situaciones y en qué aspectos sería necesaria una interpretación adicional, un mayor desarrollo jurisprudencial o una revisión de las normas.

Protección constitucional y jurisprudencial de la libertad de expresión

La revisión del bloque de constitucionalidad y de la jurisprudencia colombiana muestra que la libertad de expresión tiene una protección especial dentro del sistema democrático, principalmente cuando se relaciona con asuntos de interés público y con el debate político. Esta salvaguarda también llega a los espacios de Internet y de las redes sociales, espacios estos que hoy en día juegan un papel fundamental como espacios de generación de opiniones, como espacios de la participación ciudadana y como espacios de la deliberación.

Por poner un ejemplo, el artículo 20 de la Constitución Política de 1991, al reconocer de forma expresa la libertad de expresión, la libertad de la difusión y de la recepción de informaciones,

pensamientos y opiniones a través de cualquier tipo de medio, sirve de base fundamental para dicha garantía, puesto que establecido bajo el principio de que la información tiene que ser veraz e imparcial. De la misma manera, esta disposición prohíbe la censura y protege la libertad de los medios de comunicación.

En otro orden de ideas, para el caso de la presente investigación, estas garantías adquieren particular importancia, ya que se rigen como un límite frente a las actuaciones estatales y, paralelamente, como un parámetro para valorar las medidas que pudieran adoptarse frente a la desinformación política. En este sentido, el reto jurídico consiste en sopesar adecuadamente la tensión entre la protección de la libertad de expresión y la necesidad de desplegar respuestas frente a contenidos que podrían afectar otros bienes constitucionales.

Al mismo tiempo, esta problemática cobra más importancia cuando se trata de expresiones referidas a cuestiones políticas, ya que las opiniones, valoraciones y posturas relacionadas con los temas del interés público son parte del pluralismo y de la participación democrática, de modo que cualquier restricción realizada en este tipo de discursos debe tener un fundamento jurídico bastante sólido, y conseguir demostrar que, persigue un objetivo legítimo, además tendrá que satisfacer los criterios de necesidad (o idoneidad) y proporcionalidad.

Por otra parte, los fallos revisados muestran cómo la Corte ha ido adaptando estas reglas al contexto de las tecnologías digitales. Un caso clave es, sin duda, la Sentencia T-155 de 2019, en la que se reconoce que las redes sociales pueden constituirse, de alguna manera, en espacios públicos que permiten el desarrollo del debate social legítimo. En la precitada ocasión, en la que se analizó una publicación en Facebook, la Corte señaló que los temas de interés general cuentan con un estándar de protección reforzado, independientemente de que la información se divulgue a través de plataformas digitales. La trascendencia de esta sentencia radica en que reconoce que Internet transforma las dinámicas de la comunicación, ya que permite la difusión masiva de información, lo que favorece el pluralismo, pero, a su vez, puede generar afectaciones más intensas a derechos como el buen nombre. En este sentido, la Corte

no optó por la censura, sino por una solución basada en la ponderación de los derechos e intereses constitucionales en conflicto.

Su importancia para la presente investigación no consiste en que haya resuelto directamente un problema de inteligencia artificial, sino en que permite identificar una regla de interpretación aplicable al entorno digital: la naturaleza de la plataforma no elimina la protección constitucional de la expresión y los conflictos deben analizarse atendiendo a los derechos involucrados y a las circunstancias concretas del caso.

En esa misma línea, la **Sentencia SU-420 de 2019** profundiza el análisis de la responsabilidad derivada del ejercicio de la libertad de expresión en internet. Para la investigación, su relevancia se encuentra especialmente relacionada con la categoría de responsabilidad ulterior, puesto que permite diferenciar la protección de la expresión frente a intervenciones preventivas de aquellas consecuencias jurídicas que pueden derivarse de un ejercicio abusivo del derecho cuando concurren los presupuestos constitucionales y legales correspondientes.

La **Sentencia T-361 de 2019** también resulta pertinente para examinar los efectos que puede producir la circulación de contenidos en redes sociales. Su incorporación al análisis permite considerar que el alcance y las características propias de los medios digitales pueden incrementar la capacidad de difusión de un mensaje y, con ello, la eventual afectación de derechos de terceros.

Por su parte, la **Sentencia T-179 de 2019** reafirma la especial protección que merece el discurso relacionado con asuntos de interés público y con la actuación de funcionarios públicos. Como tal, este criterio resulta especialmente relevante frente a la desinformación política, porque impide equiparar automáticamente cualquier contenido desfavorable, crítico o polémico con una expresión jurídicamente ilegítima.

Finalmente, fallos como la **Sentencia T-356 de 2021** y la **T-061 de 2022** permiten complementar esta línea jurisprudencial en relación con el ejercicio de la libertad de expresión y el debate democrático en entornos digitales. El análisis conjunto de estas providencias permite identificar una estructura

jurisprudencial compuesta por cuatro elementos principales: (i) protección reforzada de la libertad de expresión en asuntos de interés público; (ii) rechazo de la censura previa; (iii) posibilidad de establecer responsabilidades ulteriores bajo condiciones constitucionales; y (iv) necesidad de ponderar la libertad de expresión con otros derechos y principios constitucionales en los casos concretos.

Por lo tanto, esta estructura conserva relevancia frente a los contenidos generados mediante IA. La novedad se encuentra en la complejidad que adquiere su aplicación cuando el contenido ha sido producido, modificado o amplificado mediante sistemas automatizados.

Tabla 3.

Líneas jurisprudenciales sobre libertad de expresión digital en Colombia

Sentencia	Tema principal	Regla jurisprudencial	Relación con entornos digitales	Proyección frente a IA
T-155 de 2019	Libertad de expresión en Facebook	Protección reforzada del discurso político	Reconoce redes sociales como escenarios democráticos	Permite analizar impacto de contenidos automatizados
SU-420 de 2019	Libertad de expresión internet	Restricciones deben ser excepcionales y proporcionales	Desarrolla responsabilidad digital y neutralidad de red	Sirve como base frente a difusión algorítmica
T-361 de 2019	Redes sociales y derechos fundamentales	Existen responsabilidades ulteriores por abuso del derecho	Reconoce alcance masivo del discurso digital	Relacionable con viralización automatizada

T-179 de 2019	Discurso político digital	Protección reforzada del debate público	Amplía protección del discurso político en internet	Aplicable a contenidos políticos generados por IA
T-356 de 2021	Libertad de expresión en medios digitales	Necesidad de ponderación constitucional	Reitera protección del entorno digital	Conecta jurisprudencia clásica con IA generativa
T-061 de 2022	Debate democrático digital	Mayor tolerancia frente a críticas contra funcionarios públicos	Fortalece deliberación democrática en redes	Relacionable con campañas automatizadas y desinformación

Nota. Elaboración propia a partir de jurisprudencia de la Corte Constitucional de Colombia

Aplicabilidad de las categorías jurisprudenciales frente a la inteligencia artificial

Aunque las decisiones estudiadas ofrecen herramientas importantes para interpretar conflictos derivados del debate digital, todavía persisten vacíos regulatorios y conceptuales frente a fenómenos asociados a inteligencia artificial generativa, *deepfakes*, automatización política y manipulación algorítmica de la información. De esa manera, el principal desafío jurídico contemporáneo consiste en garantizar que el ecosistema digital continúe siendo compatible con los principios democráticos, la transparencia informativa y la formación libre de la opinión pública.

Uno de los principales hallazgos derivados del análisis jurisprudencial se basa en que las decisiones de la Corte Constitucional fueron construidas para resolver conflictos asociados al ejercicio de la libertad de expresión por parte de sujetos humanos dentro de plataformas digitales; sin embargo, los problemas jurídicos derivados de la inteligencia artificial generativa modifican sustancialmente el contexto sobre el cual se edificaron dichas reglas. Mientras las sentencias T-155 de 2019, SU-420 de 2019 y T-361 de 2019 parten de la existencia de un emisor identificable y de contenidos atribuibles a una

persona determinada, los sistemas actuales de inteligencia artificial permiten producir mensajes, imágenes, audios y videos sintéticos cuya autoría, trazabilidad y nivel de automatización resultan considerablemente más complejos de determinar.

En consecuencia, las categorías tradicionales sobre responsabilidad por el ejercicio abusivo de la libertad de expresión requieren ser reinterpretadas frente a escenarios donde intervienen desarrolladores, plataformas digitales, usuarios y sistemas algorítmicos. De igual forma, la investigación permitió advertir que la protección reforzada del discurso político, reiterada por la Corte Constitucional en las sentencias T-179 de 2019 y T-061 de 2022, continúa siendo plenamente válida dentro de los entornos digitales.

La aparición de herramientas de inteligencia artificial generativa introduce un desafío que no fue considerado de una manera directa a las decisiones judiciales examinadas. Si bien la jurisprudencia constitucional reconoce la necesidad de preservar la circulación amplia de ideas como elemento fundamental del sistema democrático, el desarrollo estratégico también facilita la producción de contenidos manipulados, que pueden presentar una apariencia de veracidad.

Además, la difusión de este tipo de materiales puede alterar artificialmente la percepción de determinados asuntos y con ello afectar la manera en que la ciudadanía construye su opinión. Por esta razón, la discusión jurídica no puede limitarse únicamente a proteger la facultad de expresar ideas, sino que debe considerar también las condiciones que permiten que la liberación pública se produzca de forma libre y forma ingenua.

En ese contexto, la prohibición de la censura previa adquiere particular relevancia. Por ejemplo, la Corte Constitucional establece que cualquier limitación al discurso político debe aplicarse de manera excepcional y encontrarse sustentada en razones suficientes que permitan demostrar su necesidad y proporcionalidad. Y este criterio guarda correspondencia con los parámetros desarrollados por el Sistema Interamericano de Derechos Humanos y con los planteamientos identificados en la doctrina examinada. Como tal, aunque la utilización de la inteligencia artificial para producir y difundir información engañosa

representa un riesgo para el debate público, dichas circunstancias no resultan suficientes, por sí solas, para establecer sistemas de bloqueo, restricción o supresión anticipada de contenidos.

La incorporación de la inteligencia artificial en los procesos de producción y difusión de información no modifica por sí misma la protección constitucional que ampara la libertad de expresión. Sin embargo, sí generar nuevos escenarios que requiere respuestas jurídicas distintas de aquellas basadas en la censura. En ese sentido, adquieren importancia mecanismos relacionados con la transferencia de los sistemas, la posibilidad de identificar el origen y recorrer los contenidos y la determinación de responsabilidades posteriores cuando se produzcan afectaciones a derechos de terceros.

Como tal, esta situación resulta importante, compleja frente a fenómenos deepfakes, la propaganda computacional y la microsegmentación política automatizada. Debido a que estas herramientas permiten generar mensajes de manera masiva, adaptarlos a públicos específicos y distribuirlos con gran rapidez, tales características hacen necesario considerar las formas tradicionales de establecer responsabilidades y de garantizar un espacio adecuado para la deliberación democrática.

Aunado a lo anterior, el análisis efectuado permite advertir que los criterios desarrollados por la Corte Constitucional continúan siendo útiles para borrar controversias relacionadas con la libertad de expresión, la honra y el buen nombre, aunque la expansión de la inteligencia artificial generativa incorpora desafíos que no estaban presentes cuando se establecieron buena parte de las reglas jurisprudenciales vigentes. Asimismo, la posibilidad de elaborar contenidos que parecen auténticos, el desconocimiento sobre el funcionamiento de determinado sistema algorítmico y la facilidad con la que éstos pueden distribuir información a grandes cantidades de personas, modifican las condiciones en las que se desarrolla la comunicación pública.

Ahora, si estos cambios tecnológicos no implica que los principios constitucionales hayan dejado de ser explicables, por el contrario, garantías como la libertad de expresión, el pluralismo al principio de proporcionalidad y la prohibición de la censura previa mantiene su importancia, pero deben ser interpretadas considerando las características y los riesgos propios de las nuevas tecnologías.

El contraste entre los referentes internacionales y el desarrollo jurisprudencial colombiano permite identificar una coincidencia importante frente a los desafíos que plantea la inteligencia artificial en el ámbito de la comunicación política. En ese sentido, tanto los documentos de la UNESCO (2021) como los planteamientos de Battista (2024/2025), Rossel (2025) y Jiménez (2025) destacan la necesidad de incorporar mecanismos de transparencia algorítmica, supervisión humana y responsabilidad por los efectos derivados del uso de estas tecnologías.

De ese modo, estos elementos pueden relacionarse con el criterio constitucional, según el cual las restricciones a la libertad de expresión, particularmente cuando se involucra discursos políticos o asuntos de interés público, deben encontrarse debidamente justificadas y superar los criterios de necesidad y proporcionalidad. Es así como el análisis permite interpretar que la respuesta frente a la desinformación generada mediante inteligencia artificial no debería orientarse principalmente hacia la eliminación anticipada de contenidos, sino hacia la creación de condiciones que permitan a los ciudadanos reconocer cuándo se encuentran ante material sintético, conocer las dinámicas que intervienen en su circulación y establecer quién debe responder cuando su utilización produce afectaciones jurídicamente relevantes.

Esta comparación también permite advertir que el principal desafío no se encuentra en reemplazar el marco constitucional exigente, sino en determinar cómo puede aplicarse frente a unas condiciones tecnológicas diferentes. En ese caso, los resultados sugieren que principios como la libertad de expresión, el pluralismo y la proporcionalidad continúan ofreciendo criterios suficientes para evitar respuestas excesivas, pero requieren ser acompañados de instrumentos que atiendan las particularidades y características de la inteligencia artificial generativa.

En particular, la identificación de contenidos artificialmente producidos, la transferencia sobre los sistemas utilizados y la definición de responsabilidades entre quienes generan, distribuyen o amplifican dichos contenidos aparecen como elementos necesarios para fortalecer la protección del debate público. Desde esta perspectiva de la evolución del marco jurídico, debería buscar un equilibrio entre dos intereses que pudieran entrar en tensión: preservar la circulación advertida de ideas y al mismo tiempo, evitar que

la manipulación automatizada de información deteriore las condiciones en las que la ciudadanía forma sus opiniones.

Así, el análisis realizado permitió sostener que enfrentar la desinformación mediante inteligencia artificial no exige debilitar la libertad de expresión, sino desarrollar mecanismos que permitan proteger la calidad y autenticidad del debate democrático, sin convertir la prevención de sus riesgos en forma indirecta de censura.

Vacíos regulatorios y desafíos jurídicos.

La revisión del marco jurídico colombiano y los estándares internacionales permite establecer que, aunque existen disposiciones que pueden ser utilizadas para abordar situaciones relacionadas con la información política mediante inteligencia artificial, ninguna de ellas fue concebida específicamente para responder a las características de este fenómeno.

En ese sentido, los artículos 15, 20 y 40 de la constitución política proporcionan una base importante para proteger la libertad de expresión y de información, la privacidad y los datos personales, así como el derecho de los ciudadanos a participar en la conformación, ejercicio y control del poder político. Sin embargo, al analizar estas garantías frente a las nuevas dinámicas de producción y circulación de información, se observa que su aplicación puede resultar insuficiente cuando se trata de establecer reglas concretas para contenidos generados artificialmente o para mecanismos automatizados capaces de intervenir en la formación de la opinión pública.

Otra situación similar se presenta con la Ley 1581 del 2012, cuyo régimen de protección de datos adquiere importancia cuando la inteligencia artificial utiliza información personal para identificar perfiles, establecer preferencias y dirigir contenidos políticos hacia determinados grupos de usuarios. Aun así, el alcance de esta regulación fue definido en un contexto tecnológico distinto al actual, por lo que no contempla de manera detallada fenómenos como la generación automatizada de mensajes, la microsegmentación política o la amplificación algorítmica.

Por tanto, el vacío no necesariamente consiste en la ausencia absoluta de normas, sino en la dificultad de aplicar categorías jurídicas tradicionales a prácticas que han adquirido una escala, velocidad y capacidad de personalización que no existían cuando fueron formuladas dichas disposiciones.

Aunado a ello, esta situación adquiere relevancia si se considera la capacidad de los sistemas de inteligencia artificial para procesar grandes cantidades de datos y utilizarlos en la evaluación de estrategias de comunicación altamente dirigidas. Por ejemplo, la experiencia de Cambridge Analytica constituye un referente para comprender cómo la utilización de información personal y técnicas de segmentación puede emplearse para orientar políticos hacia públicos específicos y generar posibles efectos sobre sus decisiones. A partir de ese tipo de experiencias, el problema jurídico es al estar relacionado exclusivamente con la facilidad de un contenido y pasa a involucrar también la manera en que esté producido, dirigido, distribuido y amplificado mediante sistemas tecnológicos.

Desde esta perspectiva, uno de los principales desafíos regulatorios se encuentra en establecer mecanismos que permitan conocer el origen y las características de los contenidos generales mediante inteligencia artificial, identificar las dinámicas que intervienen en su difusión y determinar quién debe asumir responsabilidad cuando su utilización ocasiona una afectación jurídica relevante. Además, la transparencia y la trazabilidad adquieren así una función preventiva, mientras que la responsabilidad ulterior permite responder frente a los daños que efectivamente se produzcan, sin convertir la regulación de la desinformación en una herramienta de censura previa.

De esa manera, el análisis sugiere que el reto del ordenamiento colombiano consiste en adaptar sus mecanismos de protección a un entorno en el que la manipulación informativa puede desarrollarse de manera automatizada y a gran escala, ya que esta adaptación debe procurar un equilibrio entre la protección de la participación política libre e informada y la preservación de la libertad de expresión como condición esencial del sistema democrático.

En cuanto al entorno tecnológico, la Ley 1341 de 2009 y la Ley 1978 de 2019 establecen disposiciones generales relacionadas con las tecnologías de la información y las comunicaciones. Sin

embargo, su objeto y estructura no proporcionan por sí solos un régimen específico de obligaciones para enfrentar la producción y amplificación automatizada de desinformación política mediante sistemas de IA.

En esa misma línea, la Ley 1273 de 2009 incorporó al ordenamiento colombiano herramientas penales relacionadas con determinadas formas de criminalidad informática. Aun así, el análisis evidencia que la existencia de delitos informáticos no equivale a la existencia de una regulación integral sobre desinformación política generada mediante IA. Las conductas de manipulación política pueden presentar características que no necesariamente encajan en las categorías tradicionales de criminalidad informática.

En el contexto colombiano también se han desarrollado instrumentos dirigidos a promover un uso responsable de la inteligencia artificial. Entre ellos se encuentran el CONPES 3975 de 2019 y la Guía Ética para la Inteligencia Artificial del Ministerio de Tecnologías de la Información y las Comunicaciones (2021), documentos que incorporan principios relacionados con la transparencia, la explicabilidad y la intervención humana en el funcionamiento de los sistemas de IA. Estos referentes representan un avance importante para orientar la gobernanza tecnológica en el país; sin embargo, su alcance principalmente orientador permite identificar una de las dificultades señaladas por esta investigación: existen principios generales para promover un uso responsable de estas tecnologías, pero no necesariamente mecanismos jurídicos específicos para enfrentar situaciones en las que la inteligencia artificial es utilizada para producir contenidos falsos, construir perfiles políticos o distribuir información de manera automatizada. En consecuencia, el marco existente ofrece criterios relevantes, aunque todavía presenta dificultades para responder de manera directa a las particularidades de la desinformación generada mediante IA.

El examen de los referentes internacionales permite advertir que la problemática no se resuelve únicamente mediante el establecimiento de nuevas medidas de protección. El artículo 13 de la Convención Americana sobre Derechos Humanos y la Declaración Conjunta de 2017 sobre Libertad de Expresión y las Fake News hacen hincapié en la necesidad de que tales medidas no limiten injustificadamente la libertad de expresión, el pluralismo ni el debate político. Este punto cobra especial

interés en relación con la inteligencia artificial, ya que una regulación excesivamente alegada podría demostrar limitaciones desproporcionadas en la justificación de una prevención de las eventuales limitaciones. Desde aquí, la revisión de los mecanismos mencionados permite postular que cualquier respuesta jurídica debe contrastar la protección de aquellos supuestos de prácticas que fueron efectivamente lesivas, con la imposición general de controles a la circulación de la información.

Del mismo modo, la Recomendación sobre Ética de la Inteligencia Artificial (2021) proclamada por la UNESCO y la Carta Iberoamericana de Derechos en Entornos Digitales (2023) destacan la importancia de conceptos como la transparencia, la supervisión humana y la responsabilidad en el desarrollo y uso de sistemas que contengan algoritmos, que son de referentes que pueden acoplarse al marco normativo colombiano, pues el debate ya no se ciñe al plano del contenido sino que permite centrarse en otros aspectos que también remiten al funcionamiento de los sistemas en virtud de los cuales se produce y se difunde contenido. En esta línea, el reto regulador se centra en pensar en mecanismos que de alguna forma permitieran rastrear y mitigar los riesgos que nos traen los usos políticos de la inteligencia artificial, pero sin someter la libertad de expresión a algún tipo de restricción desproporcionada.

De este modo, la inspección de la normatividad a nivel nacional e internacional puede mostrar que existen fundamentos suficientemente consistentes para dar respuesta jurídica a los riesgos que la inteligencia artificial puede conllevar, aunque todavía quedan algunos aspectos que hay que trabajar en mayor profundidad. La principal de las dificultades está en traducir estos principios en procedimientos que permitan dar respuestas a fenómenos concretos como el contenido sintético, el perfilado político o la amplificación automatizada. Por ello, más que establecer una regulación basada exclusivamente en prohibiciones, los resultados sugieren la necesidad de fortalecer las obligaciones de quienes desarrollan, utilizan o distribuyen estas tecnologías, de manera que sea posible conocer cómo intervienen los sistemas de IA en la comunicación política y establecer responsabilidades cuando su utilización afecte derechos o condiciones esenciales del debate democrático.

A continuación, se presenta la síntesis de los principales vacíos regulatorios y desafíos jurídicos frente a la desinformación política generada por inteligencia artificial que se hallaron en la investigación:

Tabla 4.

Principales vacíos regulatorios y desafíos jurídicos

Norma / Estándar	Aspecto regulado	Vacío identificado frente a IA	Riesgo jurídico-democrático asociado
Constitución Política Colombia – Art. 20	Libertad de de expresión y derecho a la información	No regula contenidos sintéticos, deepfakes ni automatización del discurso político	Manipulación del debate público y dificultad para distinguir información auténtica
Constitución Política Colombia – Art. 15	Habeas data y protección de datos personales	No contempla expresamente perfilamiento algorítmico ni microsegmentación política automatizada	Uso masivo de datos para influir electoralmente sobre ciudadanos
Constitución Política Colombia – Art. 40	Participación de política y democracia participativa	No prevé mecanismos frente a manipulación electoral mediante IA	Distorsión de la deliberación democrática y afectación del voto informado
Ley 1341 de 2009	Funcionamiento del ecosistema TIC	Regula infraestructura tecnológica, pero no propaganda automatizada ni algoritmos políticos	Circulación masiva de contenidos manipulados sin control específico

Ley 1581 de 2012	Protección de datos personales	No regula expresamente uso político de datos mediante IA generativa	Microtargeting y segmentación ideológica masiva
Ley 1273 de 2009	Delitos informáticos	No tipifica desinformación automatizada ni manipulación política algorítmica	Difusión coordinada de contenidos falsos sin respuesta jurídica específica
Ley 1978 de 2019	Modernización TIC y plataformas digitales	No establece obligaciones concretas sobre transparencia algorítmica	Amplificación opaca de contenidos políticos y desinformativos
CONPES 3975 de 2019	Política pública de IA	Enfoque ético-programático sin regulación específica sobre desinformación política	Desarrollo tecnológico sin controles jurídicos suficientes
Guía Ética para IA – MinTIC (2021)	Principios éticos de IA	Carece de fuerza vinculante y mecanismos sancionatorios	Uso irresponsable de sistemas algorítmicos en comunicación política
Convención			
Americana sobre Derechos Humanos – Art. 13	Libertad de pensamiento y expresión	No desarrolla específicamente riesgos de IA generativa	Posibles restricciones estatales desproporcionadas al discurso digital
Declaración ONU (2025), OEA (2017) y OSCE (2025)	Libertad de expresión y fake news	Advierte riesgos de regulación excesiva, pero no define mecanismos concretos frente a IA	Censura indirecta bajo pretexto de combatir desinformación

UNESCO	–	No establece obligaciones jurídicas obligatorias para los Estados	Persistencia de riesgos sin supervisión efectiva
Recomendación sobre Ética de la IA (2021)	Gobernanza ética y derechos humanos		
Carta Iberoamericana de Derechos Entornos Digitales (2023)	Democracia digital y transparencia algorítmica	Carece de implementación normativa obligatoria en los Estados	Débil protección frente a manipulación digital transnacional

Nota. Elaboración propia a partir del análisis normativo, jurisprudencial y doctrinal desarrollado en la investigación.

El análisis de los vacíos permite advertir que el principal problema jurídico colombiano no corresponde a la inexistencia de protección constitucional de la libertad de expresión. El problema se encuentra en la ausencia de mecanismos suficientemente específicos que permitan aplicar dicha protección y, simultáneamente, responder frente a determinadas prácticas de manipulación automatizada.

Criterios jurídicos para la protección

Protección reforzada del discurso político en entornos digitales

La revisión normativa y jurisprudencial permite reconocer que la libertad de expresión constituye una garantía esencial para el funcionamiento democrático, con especial importancia cuando se encuentra vinculada al debate político y a los asuntos de interés público. La libertad de expresión e información reconocida en el artículo 20 de la Constitución Política, junto con el imperativo de no llevar a cabo censura previa, quedan garantizados. Esta garantía comprende tanto una dimensión personal, en la medida en que cada persona tiene la capacidad de expresar sus opiniones, como una dimensión social, relacionada con el derecho de la comunidad a conocer, contrastar y valorar diferentes posiciones. Su contenido

también se extiende a los espacios digitales, donde las redes sociales y las plataformas digitales se han convertido en espacios relevantes para el desarrollo de la participación y la deliberación pública.

Desde esta óptica, las iniciativas dirigidas a combatir la desinformación política deben evitar que la protección del debate democrático termine justificando la prohibición de opiniones, críticas o posiciones relacionadas con asuntos de relevancia social. Las Sentencias T-155 de 2019 y T-061 de 2022 permiten evidenciar la importancia que han adquirido los espacios digitales para el ejercicio de la libertad de expresión, así como el mayor nivel de escrutinio al que están sujetos quienes desempeñan cargos públicos. De este modo, se extrae un primer criterio jurídico, según el cual la lucha contra la desinformación no da lugar, por sí sola, a la imposición de restricciones generales al discurso político. Cualquier restricción que se pretenda establecer debe contar con la correspondiente base legal, perseguir una finalidad legítima y superar los juicios de necesidad y proporcionalidad.

El desarrollo de la inteligencia artificial generativa introduce factores que incrementan la complejidad en la aplicación de tales criterios. La posibilidad de generar contenidos en masa, de manera rápida y personalizada, así como la capacidad de producir información falsa con apariencia auténtica, amplían el potencial de afectar la formación de la opinión pública. Battista (2024) señala que estas nuevas tecnologías habilitan modalidades de manipulación política mediante medios sintéticos capaces de representar hechos, declaraciones e imágenes de manera muy verosímil. Por otra parte, Filho et al. (2022) ponen en duda si las redes automatizadas y los deepfakes pueden producir alteraciones en la circulación de ideas, así como el hecho de que la distinción entre información veraz y manipulada sea cada vez más problemática, con posibles efectos sobre el pluralismo democrático. En este contexto, el problema para el derecho consiste en dar respuesta a estos riesgos sin desconocer la protección reforzada que corresponde al discurso político.

El uso de la inteligencia artificial permite abordar ante todo los grupos orientados a grupos de contenido mediante el tratamiento de datos personales, preferencias ideológicas y, por tanto, patrones de

conducta, lo que puede facilitar tácticas de segmentación que permiten profundizar en la polarización y en sofisticadas manifestaciones de influencia política (Baltezarević et al., 2025).

La incidencia de los intermediarios y de las funciones automatizadas sobre el flujo informativo ya había sido advertida a partir del análisis de herramientas como el autocompletar en motores de búsqueda, cuya arquitectura algorítmica condiciona el acceso a la información y tensiona la libertad de expresión digital (Guzmán et al., 2014). Este antecedente evidencia que las sugerencias y clasificaciones generadas por código pueden pre configurar el imaginario del usuario y lesionar derechos de terceros cuando difunden términos difamatorios o imprecisos. Trasladado al ámbito de la desinformación política actual, se constata que los algoritmos de predicción y ordenación profundizan las asimetrías comunicativas, exigiendo criterios de transparencia que impidan la manipulación silenciosa del debate democrático.

Ahora bien, estas amenazas no justifican que, en términos generales, se excluyan de la protección constitucional los contenidos generados o bien la difusión de estos mediante estas tecnologías. El análisis jurídico debe ser capaz de distinguir entre el legítimo ejercicio de la libertad de expresión y aquellas conductas encaminadas de forma dolosa a manipular, suplantar, engañar o a afectar derechos fundamentales.

Por tanto, para establecer una eventual restricción no basta con sostener que una expresión sea falsa, infructuosa o que no genere la reacción deseada, puesto que también deben considerarse aspectos como su finalidad, el recurso a mecanismos de engaño o suplantación, el daño concreto que produce, el grado de automatización y coordinación de la acción y su eventual incidencia electoral. A partir de estos criterios, es posible adoptar respuestas proporcionales frente a prácticas que puedan afectar una participación política libre e informada, sin incurrir en una restricción excesivamente amplia de lo que puede denominarse discurso público.

Los criterios identificados encuentran soporte en la doctrina del Sistema Interamericano de Derechos Humanos, pues se encuentran sustentados en el artículo 13 de la Convención Americana sobre Derechos Humanos, que reconoce la libertad de pensamiento y de expresión y establece que las limitaciones a este derecho deben cumplir determinadas condiciones. En particular, las restricciones deben estar previstas legalmente, perseguir una finalidad legítima y superar los criterios de necesidad y proporcionalidad, de manera que las medidas adoptadas no impliquen una afectación injustificada de la libertad de expresión.

Esta protección también se refleja en la Declaración Conjunta sobre Libertad de Expresión y Fake News de 2017, emitida por la ONU (2025), la OEA (2017) y la OSCE (2025), donde se advierte que las prohibiciones generales sobre las denominadas “noticias falsas” pueden afectar expresiones legítimas cuando se sustentan en conceptos amplios o imprecisos. A estos referentes se suman la Recomendación de la UNESCO sobre la Ética de la Inteligencia Artificial (2021) y la Carta Iberoamericana de Derechos en Entornos Digitales (2023), que destacan la importancia de la transparencia, la supervisión humana y la protección de los derechos fundamentales. En conjunto, estos instrumentos muestran una orientación hacia respuestas regulatorias que permitan enfrentar los riesgos tecnológicos sin convertir la regulación de los contenidos políticos en una forma de censura.

Desde esta perspectiva, la protección reforzada del discurso político debe preservarse también en los espacios digitales, pero ello parte, a su vez, de la búsqueda de herramientas jurídicas que tengan en cuenta las nuevas formas de manipulación que permite la inteligencia artificial. El reto para el ordenamiento colombiano no consiste simplemente en establecer límites a la desinformación, sino en definir criterios que permitan determinar cuáles son las conductas efectivamente lesivas y cuáles son las responsabilidades derivadas de acciones como los deepfakes, la suplantación digital, la segmentación política abusiva o la automatización y amplificación de contenidos. Para ello, la legalidad, la necesidad, la proporcionalidad, la transparencia y la responsabilidad pueden servir como criterios orientadores de una

regulación que busque garantizar, a la vez, la libertad de expresión, los derechos humanos y las condiciones necesarias para un debate político libre e informado.

La desinformación automatizada como riesgo para el orden democrático

Uno de los principales hallazgos de la investigación permite considerar la desinformación política no como la mera propagación de información errónea por parte de usuarios o medios digitales, sino como un fenómeno que incorpora procesos de automatización capaces de incidir en la formación de la opinión pública y en la deliberación democrática. Esto se debe a que la inteligencia artificial permite producir contenidos sintéticos, como los deepfakes , con niveles muy altos de realismo, lo que dificulta la distinción entre hechos y simulaciones. Jiménez (2025) ya advierte sobre esta situación en relación con la pérdida de confianza pública y las dificultades para identificar información cierta. Battista (2024), por su parte, también la vincula con una “crisis de verificabilidad” que afecta la capacidad de distinguir entre hechos y simulaciones. Así, la relevancia jurídica de esta forma de desinformación política ya no se limita a la posible vulneración de derechos individuales, como la honra o la imagen, sino que también puede comprometer las condiciones necesarias para un debate público informado y basado en información veraz.

Como tal, este fenómeno se intensifica mediante bots, redes automatizadas y estrategias de microtargeting, que permiten amplificar determinados discursos y dirigir contenidos políticos hacia grupos específicos a partir de datos, preferencias y patrones de comportamiento. Filho y colaboradores (2022) afirman que las redes automatizadas pueden engañar la percepción con la que los usuarios cuentan respecto a las condiciones de apoyo social de ciertos conceptos, mientras que Rossel (2025) menciona las consecuencias de la fragmentación algorítmica del discurso que afectan la existencia de un debate público común. De forma similar, Vélez (2026) advierte de una conexión entre el uso de inteligencia artificial y la comunicación política, indicativa de la polarización y manipulación emocional. Los problemas por tanto no están solo al nivel de la información que se divulga, sino también en las condiciones de visibilidad, la

segmentación y la ampliación que determinan qué información le llega a cada ciudadano y en qué cantidad.

A la vez, la investigación permite sostener que la automatización de la desinformación supone un riesgo estructural para el orden democrático pues afecta a tres condiciones fundamentales de la deliberación pública: la veracidad de la información, la publicidad del debate político y la libertad de la formación de la opinión ciudadana. La infraestructura algorítmica de las plataformas permite la circulación de información muy emocional o extrema como pone de relieve Battista (2025), y a su vez la segmentación dificulta el contraste público y la posibilidad de que diferentes ciudadanos puedan acceder al mismo escenario informativo. En consecuencia, el reto jurídico no es únicamente la falta de contenidos falsos, sino la reacción frente a ecosistemas tecnológicos que habilitan la intervención sobre la percepción y participación política mediante automatización, segmentación o amplificación algorítmica, garantizando al mismo tiempo las garantías propias de una sociedad democrática.

Necesidad de regulación proporcional y responsabilidad ulterior

Los resultados de la investigación permiten identificar que la existencia de desinformación política producida o amplificada mediante inteligencia artificial representa un riesgo para el funcionamiento del orden democrático, por cuanto puede generar efectos sobre la autenticidad de la información, la transparencia del debate público y el proceso libre de formación de la opinión pública. La existencia de fenómenos como los deepfakes, las redes automatizadas, la propaganda programática y el microtargeting modifica las dinámicas de circulación de la información política, ya que permite producir contenidos altamente realistas, amplificarlos y dirigirlos hacia grupos específicos. En este sentido, Battista (2024) advierte sobre la existencia de una “crisis de verificabilidad” vinculada con la capacidad de la inteligencia artificial para producir contenidos sintéticos que resultan difíciles de distinguir de los auténticos, mientras que Filho et al. (2022) señalan que las redes automatizadas pueden alterar artificialmente la percepción sobre el respaldo social de determinadas ideas. Estos resultados indican que

el problema trasciende la existencia de contenidos aislados completamente falsos y se relaciona con el grado en que la tecnología puede intervenir en las dinámicas de deliberación democrática.

Ante esta realidad, el segundo criterio de orden jurídico identificado corresponde a la necesidad de salvaguardar la diferencia entre **censura previa y responsabilidad ulterior**. La prohibición de la censura previa impide que el Estado pueda establecer procedimientos de autorización o control previo sobre los contenidos políticos únicamente por el hecho de que puedan ser falsos, controvertidos o potencialmente manipuladores. Esta garantía resulta relevante precisamente porque, como advierten la ONU (2025), la OEA (2017) y la OSCE (2025), las prohibiciones generales sobre las denominadas “fake news” podrían dar lugar a restricciones incompatibles con la libertad de expresión.

Sin embargo, la plena protección constitucional que se hace extensiva al discurso político no conlleva la inexistencia de conflictos de responsabilidad. Siempre que se produzca alguna conducta que genere una determinada e importante afectación colisionando con la regla del discurso político, el ordenamiento jurídico puede desviar el contenido de la libertad de expresión habilitando mecanismos de posterior responsabilidad, siempre que tenga por fundamento una norma legal y que la medida de intervención se ajuste a las garantías constitucionales y, a la vez, a los parámetros de necesidad y proporcionalidad. En otras palabras, la respuesta del ordenamiento jurídico deberá orientarse hacia conductas lesionantes concretas y no al control preventivo general del discurso.

El tercer criterio que se designa es el **principio de proporcionalidad**, que cobra especial importancia en la medida en que las medidas reguladoras pueden restringir la libertad de expresión. La intervención por parte del Estado para hacer frente a los riesgos derivados de la inteligencia artificial debe perseguir una finalidad legítima, resulta adecuada para alcanzar la misma, no debe ser más gravosa; es decir, no debe existir una medida menos restrictiva y la relación que se establece entre la protección obtenida y la afectación producida sobre el derecho fundamental.

Este análisis resulta particularmente relevante frente a medidas como la eliminación de contenidos, las restricciones de alcance, las obligaciones impuestas a las plataformas o la identificación de

materiales generados artificialmente. No todas estas intervenciones tienen la misma intensidad. Por ejemplo, exigir la identificación de un contenido generado mediante inteligencia artificial constituye una medida sustancialmente diferente de impedir anticipadamente su publicación o disponer su eliminación. Por ello, las medidas de transparencia, advertencia e identificación pueden constituir alternativas menos restrictivas frente a los mecanismos de bloqueo o eliminación generalizada.

Un cuarto criterio se refiere a la **atribución de responsabilidad**, que adquiere especial complejidad debido a la participación de diversos actores en la producción y circulación de contenidos generados mediante inteligencia artificial. Una conducta puede ser llevada a cabo por usuarios, desarrolladores, proveedores de tecnología, plataformas e intermediarios; sin embargo, esta pluralidad de actores no permite establecer automáticamente una responsabilidad conjunta por la publicación de los contenidos generados. Por el contrario, corresponde evaluar la intervención de cada sujeto concreto en la conducta, su finalidad, el grado de intervención humana, el conocimiento o la previsibilidad del riesgo y la capacidad efectiva para ejercer control sobre la producción, difusión o amplificación del contenido. También debe considerarse la existencia de un daño concreto jurídicamente relevante y de una norma que sustenta la atribución de responsabilidad.

De esta manera, se busca evitar tanto la atribución automática de responsabilidad a sujetos que no tenían un control efectivo sobre el contenido como la posibilidad de utilizar la complejidad de estos sistemas como argumento para obstaculizar la identificación de los actores que participaron en la conducta causante del daño.

Esta problemática de atribución subjetiva guarda estrecha sintonía con las discusiones sobre el rol de los proveedores de servicios de internet (ISP) en el entorno digital, donde su grado de involucramiento y deberes de colaboración han sido objeto de constante delimitación jurídica (Vargas, 2016). En el ecosistema de la comunicación automatizada, no es jurídicamente viable asimilar mecánicamente al mero transmisor de datos con el creador del contenido lesivo, pues ello desincentivaría la neutralidad de la red y fomentaría la censura privada. Por ende, la delimitación de responsabilidades ulteriores frente a campañas

desinformativas debe distinguir con claridad entre la infraestructura de intermediación y la intervención activa en la producción algorítmica de la falsedad política.

La **transparencia y la trazabilidad** son, a su vez, criterios complementarios para robustecer la respuesta jurídica frente a estos riesgos de forma adecuada. La transparencia algorítmica no es ir hacia la exigencia de una divulgación sin límites en lo que se refiere a códigos fuente o información protegida, sino la de garantizar que pueda existir una información suficiente sobre aquellos aspectos de los sistemas automatizados que puedan tener una incidencia notable en la circulación del discurso político. En particular, resulta relevante poder conocer cuándo entran en juego los mecanismos automatizados de recomendación, de segmentación o de amplificación de contenidos y cuándo un material audiovisual ha sido generado o alterado de forma relevante por inteligencia artificial. Por su parte, la trazabilidad puede ayudar a reconstruir, dentro de los límites técnica y jurídicamente razonables de determinados elementos sobre la generación, la modificación y la difusión de un contenido, ayudando a la identificación de responsabilidades, sin transformar, sin embargo, esos mecanismos en medios de vigilancia general sobre la expresión política.

Estos criterios deben analizarse conjuntamente con la **protección reforzada del discurso político** reconocida por el orden constitucional colombiano y por los estándares interamericanos. El artículo 20 de la Constitución Política protege la libertad de expresión y prohíbe la censura previa, mientras que la jurisprudencia constitucional ha reconocido la especial importancia del discurso relacionado con asuntos de interés público. En el ámbito interamericano, el artículo 13 de la Convención Americana sobre Derechos Humanos establece que las restricciones a la libertad de expresión deben responder a responsabilidades ulteriores previstas por la ley y cumplir condiciones de necesidad y proporcionalidad. En consecuencia, la lucha contra la desinformación no puede utilizarse como fundamento para establecer controles generales sobre las opiniones o críticas políticas. La regulación debe concentrarse en prácticas

que, por su finalidad, medios empleados o efectos concretos, puedan afectar derechos fundamentales o las condiciones esenciales de participación democrática.

Finalmente, la **protección del orden democrático** debe constituir la finalidad sustantiva que oriente la respuesta jurídica, pero sin convertirse en una categoría abierta que permita restringir cualquier contenido considerado inconveniente por las autoridades. Su aplicación debe vincularse con afectaciones concretas a elementos esenciales de la democracia, como el pluralismo, la participación política, la formación libre de la opinión pública y la integridad de los procesos electorales. En este sentido, los aportes de la UNESCO (2021) y de la Carta Iberoamericana de Derechos en Entornos Digitales (2023) permiten reforzar la importancia de la transparencia, la supervisión humana y la protección de los derechos fundamentales frente al desarrollo de sistemas algorítmicos. Por tanto, el desafío jurídico colombiano no consiste únicamente en crear restricciones frente a la desinformación generada mediante IA, sino en construir mecanismos de regulación que articulen responsabilidad ulterior, proporcionalidad, atribución clara de responsabilidades, transparencia y trazabilidad, garantizando al mismo tiempo la protección reforzada del discurso político y las condiciones necesarias para un debate democrático libre, plural e informado.

Propuesta de criterios jurídico-normativos aplicables al contexto colombiano

A partir del análisis doctrinal, jurisprudencial y normativo que ha sido elaborado durante la investigación, fue posible identificar la necesidad de construir unos criterios jurídicos para poder responder ante los riesgos que surgen de la desinformación política generada con inteligencia artificial sin que ello comprometa la protección reforzada que brinda la libertad de expresión y el debate democrático. En este sentido, la investigación ha podido evidenciar que el ordenamiento jurídico colombiano cuenta con bases constitucionales sólidas para la protección de los derechos fundamentales en los entornos digitales, pero todavía no cuenta con mecanismos específicos que permitan hacer frente a fenómenos contemporáneos como los deepfakes, la propaganda computacional, el microtargeting político y la manipulación algorítmica de los contenidos electorales.

De ahí que, en el caso de una futura regulación, ésta deba basarse en un principio de especial importancia, aunque pueda parecer evidente: la desinformación automatizada no puede combatirse mediante mecanismos de censura previa ni a través de controles extensivos sobre el contenido político que circula por Internet. En este sentido, la transparencia algorítmica no debe entenderse como una respuesta meramente técnica, sino como un criterio jurídico orientado a garantizar que las personas puedan conocer aquellos elementos del funcionamiento de los sistemas automatizados que tengan una incidencia relevante en la producción, circulación y amplificación del discurso político.

La defensa de la libertad de expresión, tanto en el marco constitucional como en el sistema interamericano, establece un estándar exigente para cualquier intervención estatal: esta debe contar con fundamento legal, responder a una finalidad legítima, ser necesaria y resultar proporcional. Este estándar adquiere especial importancia cuando se trata del discurso político, debido a su relación con la participación ciudadana, el pluralismo y la formación de la opinión pública. Por ello, cualquier regulación dirigida a enfrentar los riesgos asociados con la inteligencia artificial debe procurar un equilibrio entre la protección frente a prácticas efectivamente lesivas y la preservación de las condiciones necesarias para un debate político libre e informado.

Es por ello por lo que el primer gran aporte que se espera de esta investigación es la propuesta de un criterio central: **la transparencia algorítmica**. Se requieren reglas claras dirigidas a las plataformas digitales y a los sistemas que, de manera automatizada, determinan qué contenidos políticos aparecen ante los usuarios. Se trata de “abrir la caja negra” de los algoritmos para comprender, dentro de límites técnica y jurídicamente razonables, por qué determinados contenidos adquieren mayor visibilidad, sin que ello implique controlar lo que las personas expresan. Solo mediante mayores niveles de transparencia es posible evaluar si estas tecnologías contribuyen al funcionamiento democrático o, por el contrario, pueden favorecer formas de manipulación mediante procedimientos que permanecen ocultos para los usuarios.

En este sentido, como hemos dicho, una gran parte de la circulación de la información queda determinada por algoritmos cada vez más sofisticados, que pueden hacer prevalecer, segmentar o amplificar unas determinadas narrativas en el entorno digital. En esa línea, no obstante, los propios usuarios suelen desconocer los criterios explicativos que funcionan como motores de tales sistemas a la hora de establecer la visibilidad de los contenidos políticos. Por ello, se nos hace necesario avanzar hacia instancias normativas que nos sitúen en unos mínimos necesarios de información sobre cuestiones como la publicidad política digital, los contenidos automatizados o las propias especificaciones de funcionamiento de los sistemas algorítmicos para los procesos electorales.

Por otro lado, se ha identificado la necesidad de reforzar los mecanismos de detección de contenidos sintéticos generados mediante inteligencia artificial. En efecto, la investigación llega a la conclusión de que los deepfakes son uno de los principales riesgos contemporáneos para la autenticidad de la información y para la confianza pública dentro del debate democrático. De ahí que más que prohibiciones generales sobre los contenidos digitales, los futuros desarrollos normativos podrían en todo caso construir senderos de deberes de trazabilidad y etiquetado de los contenidos generados de forma sintéticamente, especialmente cuando estén relacionados con campañas políticas, propaganda electoral o representaciones audiovisuales de figuras públicas.

De la misma manera, la investigación permite establecer una conclusión clara: cualquier futura regulación debe sujetarse a un modelo de **responsabilidad ulterior**. El Estado no debe convertirse en una autoridad encargada de determinar previamente qué contenidos pueden publicarse y cuáles deben ser excluidos, pues ello podría conducir a formas de censura o control preventivo. Su intervención debe producirse cuando exista una lesión concreta y jurídicamente relevante. Ya sea frente a situaciones de fraude electoral, afectaciones a la honra o manipulación deliberada de la información, los mecanismos de responsabilidad deben operar con posterioridad a los hechos, sin generar restricciones desproporcionadas sobre la libertad de prensa y la libertad de expresión.

En otro orden de cosas, es fundamental que nuestras instituciones electorales se pongan al día. No podemos estar vigilando las elecciones del siglo XXI con instrumentos de finales del siglo XX. La utilización de bots y la microsegmentación algorítmica fuerzan a las autoridades que vigilan las elecciones a armarse de una suerte de "inteligencia técnica" que les permita percibir patrones de manipulación que ensucien el voto. Eso sí, esta vigilancia no puede ser una carta blanca. La intervención del Estado será justificada siempre que tenga una finalidad democrática transparente y no se adentre en el campo del pluralismo político. La amenaza real es que, amparándose en la defensa, acabemos por proporcionar al Estado un control excesivo sobre lo que decimos en Internet, lo que contradiría los estándares internacionales de derechos humanos.

En definitiva, Colombia necesita dar el salto hacia un modelo regulatorio integral. No basta con hablar de tecnología o economía; la inteligencia artificial en la política debe gestionarse bajo principios de transparencia, trazabilidad y, sobre todo, supervisión humana. Al final del día, la IA no es solo un tema de *software*, es un tema de supervivencia democrática y de protección a nuestros derechos fundamentales.

En términos normativo-constitucionales, la implementación del principio de transparencia algorítmica en Colombia podría estructurarse bajo un modelo de obligaciones de información y no de autorización previa de contenidos. Donde las plataformas digitales permiten la difusión masiva de publicidad política o contenidos electorales deberían informar a las autoridades competentes y a los usuarios, como mínimo, los criterios generales utilizados para la priorización de contenidos, la existencia de mecanismos automatizados de recomendación y la utilización de sistemas de segmentación política basados en datos personales. Lo anterior no implicaría la intervención estatal sobre el contenido difundido, sino el establecimiento de deberes de transparencia compatibles con el artículo 20 de la Constitución Política

De manera complementaria, el etiquetado de contenidos sintéticos generados mediante inteligencia artificial podría operar bajo un esquema de trazabilidad tecnológica similar al adoptado en

experiencias regulatorias internacionales. De este modo, siempre que se produzca una variación relevante de un contenido audiovisual mediante la intervención de la inteligencia artificial y que se emplee en el ámbito político o electoral, las plataformas deberían anotar la advertencia visible sobre la modificación de los materiales generados o alterados artificialmente y esto no implica inhibir o eliminar el contenido, sino garantizar al ciudadano, en este ámbito del derecho, el acceso al origen y naturaleza de la información que maneja, de este modo se contribuye a garantizar el derecho a la formación libre e informada de la opinión pública.

El examen crítico de la implementación de mecanismos técnicos de control en entornos digitales halla un antecedente significativo en el estudio de la gestión digital de derechos (DRM), donde la adopción de barreras automatizadas generó amplias tensiones entre la salvaguarda de prerrogativas y el acceso abierto a bienes informativos (Ciusa & Vargas, 2013). La experiencia comparada enseña que las soluciones tecnológicas basadas en filtros de bloqueo automático tienden a reproducir dinámicas asimétricas si no cuentan con contrapesos jurídicos transparentes que impidan restricciones colaterales desmedidas. Por consiguiente, el diseño de herramientas de trazabilidad frente a la desinformación política debe evitar la rigidez ciega de las soluciones algorítmicas propietarias, privilegiando en su lugar estándares de supervisión humana orientados a no vulnerar prerrogativas fundamentales del usuario digital.

En igual sentido, la investigación permite sostener que la supervisión institucional de estas obligaciones recae en autoridades preexistentes, pero también impide que surjan organismos con potestades amplias de control del discurso. Concretamente, el Consejo Nacional Electoral podría asumir competencias respecto de la publicidad política automatizada por períodos electorales y la Superintendencia de Industria y Comercio conservaría competencias relacionadas con el tratamiento de datos de carácter personal y el ejercicio de mecanismos de microsegmentación política. Al mismo tiempo, los jueces constitucionales deberían atender las controversias relacionadas con las posibles afectaciones a la libertad de expresión y a otros derechos fundamentales.

A la vez, resulta indispensable precisar que los criterios propuestos son compatibles con la prohibición constitucional de censura previa debido a que no condicionan la publicación de contenidos a una autorización estatal ni facultan a las autoridades para determinar la veracidad de las opiniones políticas. Por el contrario, se trata de mecanismos de responsabilidad ulterior y transparencia informativa orientados a preservar la autenticidad del debate democrático.

Por lo tanto, la intervención estatal únicamente procedería de manera posterior y excepcional, cuando se acrediten afectaciones concretas a derechos fundamentales, fraude electoral o conductas expresamente previstas por la ley. Por ello, el principal desafío jurídico contemporáneo consiste en construir mecanismos normativos capaces de garantizar simultáneamente libertad de expresión, autenticidad informativa y protección democrática dentro de entornos digitales cada vez más automatizados.

Discusión jurídica de resultados

Los resultados que se han conseguido mediante el tipo de análisis que hemos llevado a cabo, ahora bien, indican que la desinformación política asociada a la utilización de la inteligencia artificial es un problema que no puede quedar circunscrito a considerarse un problema solamente tecnológico o comunicacional, sino también jurídico y democrático, ya que altera las condiciones de producción relacionadas con la realización del debate público, puesto que los entornos digitales son utilizados a través de sistemas algorítmicos que pueden favorecer algunos contenidos, segmentar las audiencias y aumentar la visibilidad de los mensajes políticos, lo cual, por consiguiente, está generando toda una serie de retos para el derecho constitucional, en el sentido en que el sacrificio de la libertad de expresión debe tener en cuenta que no únicamente se trata de la posibilidad de emitir opiniones, sino también de las condiciones con que las y los ciudadanos reciben información y participan en la deliberación democrática.

Los aportes que han sido extraídos como consecuencia de la revisión doctrinal y normativa permiten advertir cómo la inteligencia artificial generativa está generando situaciones que no son acatables a las categorías jurídicas tradicionales. El uso de deepfakes, bots, microtargeting hace que la

distinción entre información, contenido falseado y propaganda política se vea realmente muy dificultada; su uso permite, además, la producción y difusión de mensajes en términos masivos, en discursos dirigidos a públicos determinados y con una capacidad de influencia que supera las formas tradicionales de comunicación política. Por ello, el riesgo no es sólo la existencia de información falsa sino una utilización de la tecnología capaz de alterar las formas de concebir las cosas, la posibilidad de hacer variar la forma en que se construye la opinión pública.

En este sentido, los hallazgos indican que la existencia de estos riesgos no justifica el establecimiento de controles estatales ilimitados. La jurisprudencia constitucional colombiana y los estándares interamericanos han reconocido una protección reforzada de la libertad de expresión frente a contenidos políticos y opiniones relacionadas con asuntos de interés público. Por lo tanto, cualquier forma de actuación destinada a combatir la desinformación debe respetar los principios de legalidad, necesidad y proporcionalidad. De igual forma, la regulación no debe recurrir a conceptos excesivamente amplios que puedan restringir de manera injustificada opiniones, críticas o posturas políticas legítimas.

Un segundo aspecto importante es que las amenazas a la libertad de expresión ya no provienen exclusivamente de la acción estatal. Las plataformas digitales y sus sistemas automatizados también tienen una incidencia importante en la determinación de qué información será vista, compartida y consumida por los usuarios. Los algoritmos pueden determinar qué contenidos adquieren mayor visibilidad frente a otros, lo que plantea nuevos problemas jurídicos relacionados con la transparencia algorítmica, la trazabilidad de los contenidos y la responsabilidad de las plataformas frente a las dinámicas de manipulación que puedan surgir en el ámbito político.

En Colombia, los resultados reflejan una situación particular. El orden constitucional ofrece una protección extensa a la libertad de expresión y existen criterios jurisprudenciales que tienen cabida para los contenidos de internet. Pero, se presentan dificultades para la respuesta a fenómenos concretos relacionados con la inteligencia artificial como deepfakes en elecciones, campañas automatizadas y el uso de bots que amplifiquen discursos concretos. Las situaciones que se dan con las tecnologías muestran que

algunas categorías jurídicas tradicionales son insuficientes para hacer frente a tecnologías que producen y distribuyen de manera masiva y automática contenidos.

Frente a estos vacíos normativos, las tendencias modernas de armonización procesal y justicia concertada en el Derecho Comunitario Europeo ilustran la trascendencia de fijar garantías mínimas inderogables para dotar de legitimidad a los procedimientos sancionatorios y de resolución de conflictos (Vargas et al., 2026). La lección dogmática que se desprende de estos procesos de integración radica en que toda articulación entre autoridades administrativas, plataformas tecnológicas y usuarios debe respetar las garantías esenciales del debido proceso para asegurar su validez material. En Colombia, cualquier modelo de correlación o moderación de la desinformación política digital debe incorporar estándares procedimentales rigurosos que impidan decisiones sumarias o lesivas de los derechos de defensa y contradicción.

Así, la protección a la democracia frente a la desinformación política no debe plantearse como un dilema entre libertad de expresión y regulación. El reto es encontrar los mecanismos para hacer frente a las prácticas de manipulación evitando de forma injustificada afectar el pluralismo y el derecho a la participación en el debate público. Los resultados apuntan hacia un modelo de protección fundamentalmente orientado hacia la transparencia, la supervisión humana y la responsabilidad posterior a las publicaciones y no hacia la censura previa o restricciones generales de los contenidos políticos.

Finalmente, para que las democracias se fortalezcan en relación con la potencial utilización política de la inteligencia artificial es preciso hacer frente a la problemática de forma conjunta mediante una articulación de herramientas jurídicas, institucionales y educativas, que no se limita a la fijación de nuevas obligaciones para quien interviene en el proceso de creación o difusión de contenidos, sino que también implica un fortalecimiento de la capacidad de la institución para identificar este tipo de prácticas y una promoción de la ciudadanía con mayor capacidad para reconocer información manipulada. De esta forma, la protección del pluralismo informativo, la autenticidad del debate público y los derechos

fundamentales pueden avanzarse de manera conjunta sin que la regulación de la inteligencia artificial signifique una limitación injustificable de la libertad de expresión.

CONCLUSIONES Y RECOMENDACIONES

La investigación permitió establecer que la inteligencia artificial representa un riesgo relevante para la libertad de expresión y el orden democrático cuando se utiliza para difundirles información política mediante deepfakes, bots, propaganda computacional y sistemas de segmentación algoritmo. Como tal, el análisis evidenció vacío regulatorio hacia el ordenamiento jurídico colombiano frente al uso político de estas tecnologías, principalmente en aspectos como la transparencia algorítmica, la identificación de contenido generados artificialmente y la determinación de responsabilidades en campañas automatizadas.

Sin embargo, estos riesgos no justifican la imposición de censura previa ni de restricciones desproporcionadas, ya que la libertad de expresión, especialmente en asuntos políticos y públicos, cuenta con una protección constitucional e interamericana que debe ser respetada al momento de establecer cualquier medida de control. Por otro lado, el principal desafío consiste en encontrar un equilibrio entre la protección de la democracia y la garantía y la libertad de expresión en los entornos digitales.

En ese sentido, cualquier regulación futura debería apoyarse en criterios de proporcionalidad, transparencia algorítmica, supervisión humana y responsabilidad ulterior, ya que el objetivo no debe ser controlar las ideas o limitar el contenido del discurso político, sino establecer mecanismos que permitan identificar y responsabilizar a quienes utilicen estas tecnologías para manipular o engañar de manera deliberada.

Por último, la respuesta frente a la desinformación no puede depender únicamente de nuevas formas, ya que es necesario fortalecer la alfabetización digital, la educación ciudadana y las herramientas de verificación de información, de modo que la población pueda reconocer contenidos manipulados y participar en el debate público, sin que corresponda al Estado determinar qué opiniones pueden circular.

LISTA DE REFERENCIA O BIBLIOGRAFÍA

- Acosta, J., Vargas, I., & Blanco, C. (2026). La seguridad jurídica: su resignificación en el Estado de Derecho y en el derecho procesal constitucional. *Revista Eletrônica de Direito Processual*, 27(3), 319–340. <https://civilistica.emnuvens.com.br/redc/article/view/1088>
- Alexy, R. (2002). *Teoría de los derechos fundamentales* (2.^a ed.). Centro de Estudios Constitucionales.
<https://www.pensamientopenal.com.ar/system/files/2014/12/doctrina37294.pdf>
- Alexy, R. (2010). *A theory of constitutional rights* (J. Rivers, Trans.). Oxford University Press. (Trabajo original publicado en 1985).
<https://doi.org/10.1093/acprof:oso/9780199582864.001.0001>
- Almache, J. C. (2025). *Inteligencia Artificial y Derecho: Incidencia en los Derechos Humanos* [Tesis de grado, Universidad de Córdoba]. Helvia.
<https://helvia.uco.es/handle/10396/32648>
- Álvarez, S. P. (2021). *Guía del ciudadano digital contra las fakes news o noticias falsas en los procesos electorales en Colombia* [Tesis de pregrado, Universidad de los Andes]. Repositorio Institucional Uniandes.
<https://repositorio.uniandes.edu.co/items/aadbc0d8-6811-4e10-b6a6-7643fef4bd01>
- Anstead, N., Edwards, L., Livingstone, S., & Stoilova, M. (2025). The potential for media literacy to combat misinformation: Results of a rapid evidence assessment. *International Journal of Communication*, 19, Article 23549.
<https://ijoc.org/index.php/ijoc/article/view/23549>
- Aral, S. (2021). *The hype machine: How social media disrupts our elections, our economy, and our health—and how we must adapt*. Crown Currency.

- Arango, A. V. (2021). *Memoria vs. Fake News: Un análisis del ejercicio de la libertad de información en el escenario del Pos-acuerdo en Colombia* [Tesis de maestría, Universidad Carlos III de Madrid]. e-Archivo. <https://e-archivo.uc3m.es/items/f700b0d8-1c82-46f4-ba32-8c2fafc6e6c5>
- Arendt, H. (1972). *Crises of the republic: Lying in politics, civil disobedience, violence, thoughts on politics, and revolution*. Harcourt Brace Jovanovich.
- Asamblea Nacional Constituyente. (1991). Constitución Política de Colombia. *Gaceta Constitucional* No. 116. <https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=4125>
- Atencio, R. E., de Jesús, J. C., Grunauer, G. R., & Tolozano, M. S. (2021). Metodología de la Investigación para el Estudiante de Derecho. *Fondo Editorial Fundación Koinonía*, 3(1), 132-132. <https://www.fondoeditorialfk.org/index.php/fondoeditorial/article/view/87>
- Azgin, B., & Kiralp, S. (2024). Surveillance, disinformation, and legislative measures in the 21st century: AI, social media, and the future of democracies. *Social Sciences*, 13(10), Artículo 510. <https://doi.org/10.3390/soc13100510>
- Azgin, B., & Kiralp, Ş. (2024). Surveillance, disinformation, and legislative measures in the 21st century: AI, social media, and the future of democracies. *Social Sciences*, 13(10), 510. <https://doi.org/10.3390/socsci13100510>
- Bailon, O. M. (2024). Disonancia entre la transformación tecnológica y ética periodística: Una reflexión crítica del impacto de la inteligencia artificial en los medios de comunicación. En *Inteligencia artificial y comunicación* (pp. 91-105). Universitat Jaume I. <https://repositori.uji.es/handle/10463/91>

- Balkin, J. M. (2018). Free speech in the algorithmic society: Big data, private governance, and new school speech regulation. *UC Davis Law Review*, 51(3), 1149–1210. https://heinonline.org/hol-cgi-bin/get_pdf.cgi?handle=hein.journals/davlr51§ion=40
- Baltezarević, R., Lović, V., & Baltezarević, I. (2025). Between Progress and Peril: The Role of Artificial Intelligence (AI) in Shaping Modern Political Communication. *International Journal of Cognitive Research in Science, Engineering and Education*, 13(3), 823-835. <https://doi.org/10.23947/2334-8496-2025-13-3-823-835>
- Bardelli, S. M. (2024). *El derecho y la tecnología en la era de la inteligencia artificial (IA)*. Editorial Autores de Argentina.
- Bassini, M. (2025). Speech without a speaker: Constitutional coverage for generative AI output? *European Constitutional Law Review*, 1–37. <https://doi.org/10.1017/S157401962500007X>
- Battista, D. (2024). Political communication in the age of artificial intelligence: An overview of deepfakes and their implications. *Society Register*, 8(2), 7–24. <https://doi.org/10.14746/sr.2024.8.2.01>
- Battista, D. (2025). When Democracy Meets AI: The Digital Battle for Political Truth. *HAPSc Policy Briefs Series*, 6(1), 45–52. <https://doi.org/10.12681/hapscpbs.39102>
- Beacken, G. D., & Woolley, S. (2026). Generative AI, Emerging Technology, and Political Communication: Examples, Methods of Study, and Implications for Democracy. En *Ethical AI and Data Science* (pp. 171–184). Auerbach Publications. <https://doi.org/10.1201/9781003657804-11>
- Borjas, J. E. (2020). Validez y confiabilidad en la recolección y análisis de datos bajo un enfoque cualitativo. *Trascender, Contabilidad y Gestión*, 5(15), 79–97. <https://doi.org/10.36791/tcg.v5i15.83>

- Brandao, P. R. (2025). The impact of artificial intelligence on modern society. *AI*, 6(8), Artículo 190. <https://doi.org/10.3390/ai6080190>
- Caballero, F. (2020). Ciberactivismo y nuevos movimientos urbanos: La producción del nuevo espacio público en la política contemporánea. *Perspectivas de la Comunicación*, 13(1), 177–202.
https://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-48672020000100177
- Carson, A., & Wright, S. (2022). Fake news and democracy: Definitions, impact, and response. *Australian Journal of Political Science*, 57(3), 221–230.
<https://doi.org/10.1080/10361146.2022.2122778>
- Carta Iberoamericana. (2023). *Carta Iberoamericana de Principios y Derechos en los Entornos Digitales*. Adoptada en la XXVIII Cumbre Iberoamericana de Jefas y Jefes de Estado y de Gobierno, Santo Domingo, República Dominicana, 25 de marzo de 2023.
https://www.segib.org/wp-content/uploads/2025/09/Carta_iberamericana_derechos_digitales_ESP_web.pdf?_gl=1*1qlcsi*_ga*MjAzMjUyMDczNS4xNzkwMTAzMjc2*_ga_MCLNSVDYMK*czE3OTAxMDMyNzYkbzEkZzAkDDE3OTAxMDMyNzYkajYwJGwwJGgw
- Castro, L. V. G., & Martínez, A. C. R. (2025). Análisis de la inteligencia artificial en las decisiones judiciales en Colombia: Estudio de caso de la sentencia T-323 y proyectos de Ley radicados. *Nuevos Desafíos del Derecho*, 5(1), 65–89.
<https://doi.org/10.48190/ndd.v5n1.4418>
- Cheong, I., Caliskan, A., & Kohno, T. (2024). Safeguarding human values: Rethinking US law for generative AI's societal impacts. *AI and Ethics*, 5(2), 1433–1459.
<https://doi.org/10.1007/s43681-024-00451-4>

- Cifuentes, J. E. S. (2025). *El algoritmo de la verdad: Un estudio de caso del impacto de las tecnologías digitales en la libertad de expresión, la desinformación y las relaciones de poder en el contexto democrático* [Tesis de maestría, Universidad de Brasilia]. Repositório Institucional da UnB. <https://repositorio.unb.br/handle/10482/52324>
- Ciusa, F., & Vargas, I. (2013). Considerazioni critiche nella dottrina giuridica italiana sul DRM. *Principia Iuris*, 19, 325–340. <https://revistas.santototunja.edu.co/index.php/piuris/article/view/779>
- Congreso de Colombia. (2012). Ley Estatutaria 1581 de 2012. Por la cual se dictan disposiciones generales para la protección de datos personales. *Diario Oficial No. 48.587*. <https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=49981>
- Congreso de Colombia. (2014). Ley 1712 de 2014. Por medio de la cual se crea la Ley de Transparencia y del Derecho de Acceso a la Información Pública Nacional y se dictan otras disposiciones. *Diario Oficial No. 49.084*. <https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=56882>
- Congreso de Colombia. (2023). Ley 2294 de 2023. Por la cual se expide el Plan Nacional de Desarrollo 2022–2026 “Colombia Potencia Mundial de la Vida”. *Diario Oficial No. 52.400*. <https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=209780>
- CONPES. (2025). *Política nacional de inteligencia artificial* (Documento CONPES 4144). Consejo Nacional de Política Económica y Social, Departamento Nacional de Planeación. <https://colaboracion.dnp.gov.co/CDT/Conpes/Econ%C3%B3micos/4144.pdf>
- Convención Americana sobre Derechos Humanos (1969). Organización de los Estados Americanos, Pacto de San José, Costa Rica.

https://www.oas.org/dil/esp/1969_Convenci%C3%B3n_Americana_sobre_Derechos_Humanos.pdf

Corte Constitucional de Colombia. (2007). *Sentencia T-391 de 2007* (M. P. Manuel Jose Cepeda Espinosa). <https://www.corteconstitucional.gov.co/relatoria/2007/t-391-07.htm>

Corte Constitucional de Colombia. (2015). *Sentencia C-150 de 2015* (M. P. Mauricio González Cuervo). <https://www.corteconstitucional.gov.co/relatoria/2015/C-150-15.htm>

Corte Interamericana de Derechos Humanos [Corte IDH]. (1985). *La colegiación obligatoria de periodistas (Art. 13 y 29 Convención Americana sobre Derechos Humanos)* (Opinión Consultiva OC-5/85). Serie A No. 5. http://www.oas.org/es/cidh/expresion/docs/publicaciones/OC5_ESP.PDF

Dan, V., Paris, B., Donovan, J., Hameleers, M., Roozenbeek, J., Van Der Linden, S., & Von Sikorski, C. (2021). Visual mis- and disinformation, social media, and democracy. *Journalism & Mass Communication Quarterly*, 98(3), 641–664. <https://doi.org/10.1177/10776990211035395>

Departamento Nacional de Planeación [DNP]. (2019). *Documento CONPES 3975: Política nacional para la transformación digital e inteligencia artificial*. Consejo Nacional de Política Económica y Social.

Departamento Nacional de Planeación [DNP]. (2023). *Estrategia Nacional Digital 2023–2026*. Gobierno de Colombia.

Downes, R. T. (2025). *Generative AI's Implications for Democratic Elections in a World at Risk: A Dahlian Perspective* [Ponencia]. Electoral Integrity Project, Technology and the Future of Democracy, Pre-APSA Workshop. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5988394

- Dryzek, J. S. (2002). *Deliberative democracy and beyond: Liberals, critics, contestations*. Oxford University Press. <https://doi.org/10.1093/019925043X.001.0001>
- Escobar, J. A. (2023). La Democracia en la Era Digital: El caso de Colombia. *Trayectorias Humanas Trascontinentales*, (16). <https://doi.org/10.25965/trahs.5682>
- Estupiñán, L. F., León, A. F., & Peña, L. J. (2020). *Fake news o los nuevos ecosistemas de comunicación personalizada: Comunicación política en las elecciones regionales de 2019 en Colombia*. Universidad Autónoma de Bucaramanga. <https://repository.unab.edu.co/handle/20.500.12749/7137>
- Ferrajoli, L. (2024). *Constitucionalismo más allá del Estado*. Trotta. <https://www.trotta.es/libros/constitucionalismo-mas-alla-del-estado/9788498797633/>
- Filho, I. N., Marrafon, M. A., & Medón, F. (2022). A Inteligência Artificial a serviço da desinformação: Como as deepfakes e as redes automatizadas abalam a liberdade de ideias no debate público e a democracia constitucional y deliberativa. *Economic Analysis of Law Review*, 13(3), 32–47. <https://doi.org/10.22456/2179-8969.123533>
- Galarza, C. A. R. (2020). Los alcances de una investigación. *CienciAmérica: Revista de Divulgación Científica de la Universidad Tecnológica Indoamérica*, 9(3), 1–6. <https://doi.org/10.33210/ca.v9i3.336>
- García, B. (2021). Disinformation, social media, bots, and astroturfing: The fourth wave of digital democracy. *El Profesional de la Información*, 30(6), e300602. <https://doi.org/10.3145/epi.2021.nov.02>
- García, J. M. P. (2025). Integración de la inteligencia artificial en democracias constitucionales: Desafíos conceptuales y propuestas normativas. *De Jure*, (2), 235–258. https://doi.org/10.29018/978-607-8919-48-2_10

- Guerrero, S. E. S. (2024). *La influencia de los medios de comunicación en el derecho penal en Colombia—Impacto en la opinión pública y las decisiones judiciales. Estudio de caso Yuliana Samboni-delito sexual y feminicidio* [Tesis de pregrado, Universidad Mariana]. Repositorio Institucional UMariana.
<https://repositorio.umariana.edu.co/server/api/core/bitstreams/c7f0d729-bc56-467a-97ef-840b98f3ac3a/content>
- Guzmán, L. F., Toledo, A., Vargas, I., Cortés, C., Botero, C., & Vargas, E. (2014). La función “Autocompletar” en los motores de búsqueda y la libertad de expresión en Internet. *Revista Informática Jurídica*, 11, 1–15.
https://www.informatica-juridica.com/wp-content/uploads/2014/01/La_funcion_autocompletar_motores_busqueda_libertad_expresion_Internet.pdf
- Habermas, J. (2015). *Between facts and norms: Contributions to a discourse theory of law and democracy* (W. Rehg, Trans.). John Wiley & Sons.
- Hernández, R., & Mendoza, C. (2020). *Metodología de la investigación: Las rutas cuantitativa, cualitativa y mixta*. McGraw-Hill Interamericana.
- Hunter, L. Y. (2023). Social media, disinformation, and democracy: How different types of social media usage affect democracy across-nationally. *Democratization*, 30(6), 1040–1072.
<https://doi.org/10.1080/13510347.2023.2208355>
- Ibáñez, S. J., Casadiegos, M. H., & Machado, M. B. (2024). Análisis integral del marco jurídico de los derechos fundamentales en Colombia. *Formación Estratégica*, 9(01), 1–14.
<https://doi.org/10.5281/zenodo.10811985>

- Jiménez, I. (2025). *La difícil convivencia entre la inteligencia artificial y los derechos fundamentales: La tecnología deepfake* [Tesis de maestría, Universidad de Sevilla]. idus. <https://idus.us.es/handle/11441/157297>
- Kaminski, M. E., & Jones, M. L. (2024). Constructing AI speech. *Yale Law Journal Forum*, 133, 1212–1235. <https://www.yalelawjournal.org/forum/constructing-ai-speech>
- Kaplan, A. (2020). Artificial intelligence, social media, and fake news: Is this the end of democracy? En *Digital transformation in media & society* (pp. 149–161). Istanbul University Press. <https://nek.istanbul.edu.tr/ekos/KITAP/ekt0000026.pdf>
- Kelani, Z. A., & Lepadatu, D. (2026). Generative artificial intelligence, misinformation, and political polarization: A content analysis of deepfake propaganda during the 2024 US presidential election. *Journal of Information Technology & Politics*, 1–19. <https://doi.org/10.1080/19331681.2026.2652897>
- Lassi, A. (2025). Inteligencia Artificial Generativa integrada al ecosistema digital. Un marco de situación para la gubernamentalidad algorítmica. *Inmediaciones de la Comunicación*, 20(1), 1–25. <https://doi.org/10.18861/ic.2025.20.1.3789>
- Lessig, L. (2009). *Code: And other laws of cyberspace* (Ver. 2.0). ReadHowYouWant.
- Ley 1341. (2009). Por la cual se definen principios y conceptos sobre la sociedad de la información y la organización de las tecnologías de la información y las comunicaciones. *Diario Oficial* No. 47.426. <https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=36913>
- Ley 1581. (2012). *Ley Estatutaria 1581 de 2012: Por la cual se dictan disposiciones generales para la protección de datos personales*. *Diario Oficial* No. 48.587. <https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=49981>

- Ley 1978. (2019). Por la cual se moderniza el sector de las tecnologías de la información y las comunicaciones. *Diario Oficial* No. 51.025.
<https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=98210>
- Limia, M. (2025). Inteligencia Artificial y Desinformación: Tensiones Éticas y Estrategias Sociotécnicas en la Era Digital. *Estudos em Comunicação*, 2(41), 86–108.
<https://doi.org/10.25768/2025.ec41.05>
- López, M. (2025). *La difusión de información falsa en redes sociales y sus límites en la libertad de expresión. Análisis jurídico de la problemática y soluciones al respecto* [Tesis de maestría, Universidad Pontificia Comillas]. Repositorio Comillas.
<https://repositorio.comillas.edu/xmlui/handle/11531/91095>
- Loretti, D., & Lozano, L. (2019). *El derecho a comunicar: Los conflictos en torno a la libertad de expresión en las sociedades contemporáneas*. Siglo XXI Editores.
- Maldonado, F. L. M., Yáñez, K. A. Y., & Salgado, J. D. M. (2021). Una aproximación a la metodología de la investigación jurídica. *Revista Pedagogía Universitaria y Didáctica del Derecho*, 8(2), 81–96. <https://doi.org/10.5354/0719-5885.2021.61113>
- Manfredi, L., Reboyedo, I. R., & Lasparrilla, M. P. U. (2021). Identificación, impacto y tratamiento de Fake News en campañas políticas en el contexto colombiano. *Escribanía*, 19(2), Article e4481. <https://doi.org/10.30554/escribania.19.2.2021.4481>
- Martínez, B., Vargas, I., & Salgado, E. (2018). El contrato de transferencia de tecnología: Caracterización e importancia estratégica. *Revista Brasileira de Direito*, 14(2), 22–39.
<https://doi.org/10.18256/2238-0604.2018.v14i2.2942>
- McCarthy, j. (2007). *What is artificial intelligence?* Stanford University.
<http://jmc.stanford.edu/articles/whatisai/whatisai.pdf>

- McKay, S., & Tenove, C. (2021). Disinformation is a threat to deliberative democracy. *Political Research Quarterly*, 74(3), 703–717. <https://doi.org/10.1177/1065912920938143>
- Ministerio de Ciencia, Tecnología e Innovación. (2024). *Hoja de ruta para el desarrollo y aplicación de la inteligencia artificial en Colombia*. Dirección de Desarrollo Tecnológico e Innovación. https://minciencias.gov.co/sites/default/files/upload/noticias/hoja_de_ruta_adopcion_etica_y_sostenible_de_inteligencia_artificial_colombia_0.pdf
- Ministerio de Tecnologías de la Información y las Comunicaciones. (2022a). *Decreto 1263 de 2022. Por el cual se adiciona el Título 22 a la Parte 2 del Libro 2 del Decreto 1078 de 2015, relacionado con transformación digital pública*. MinTIC.
- Ministerio de Tecnologías de la Información y las Comunicaciones. (2022b). *Resolución 460 de 2022. Plan Nacional de Infraestructura de Datos (PNID)*. MinTIC.
- Muñoz, J. A. (2024). Desinformación: Aproximación conceptual, riesgos y remedios. *Derecho PUCP*, (93), 55–97. <https://doi.org/10.18800/derechopucp.202402.003>
- OEA. (2019). *Guía para garantizar la libertad de expresión frente a la desinformación en contextos electorales*. Relatoría Especial para la Libertad de Expresión, Organización de los Estados Americanos. <https://www.oas.org/es/cidh/expresion/showarticle.asp?artID=1156&IID=2>
- ONU. (2025). *Relator Especial sobre la promoción y protección del derecho a la libertad de opinión y de expresión*. Naciones Unidas. [Página oficial del ACNUDH https://www.ohchr.org/es/special-procedures/sr-freedom-of-opinion-and-expression](https://www.ohchr.org/es/special-procedures/sr-freedom-of-opinion-and-expression)

Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura [UNESCO].

(2021a). *Recomendación sobre la ética de la inteligencia artificial*. UNESCO.

https://unesdoc.unesco.org/ark:/48223/pf0000381137_spa

Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura [UNESCO].

(2021b). *Disinformation and artificial intelligence: Challenges and responses*. UNESCO.

<https://www.unesco.org/en/articles/artificial-intelligence-and-disinformation-examining-challenges-and-solutions>

Organización de los Estados Americanos [OEA]. Relatoría Especial para la Libertad de

Expresión. (2017). *Declaración conjunta sobre libertad de expresión y “noticias falsas” (“fake news”), desinformación y propaganda*.

<https://www.oas.org/es/cidh/expresion/showarticle.asp?artID=1056&IID=2>

Organización para la Cooperación y el Desarrollo Económicos [OCDE]. (2019).

Recommendation of the Council on Artificial Intelligence (OECD/LEGAL/0449). OECD Legal Instruments. <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>

Ortiz, L. F., & Agudelo, L. (2023). La influencia de las redes sociales digitales en la política.

Análisis de las elecciones presidenciales de Colombia en 2022 y algunas lecciones globales. *Pluriverso*, 18(18), 57–86. <https://doi.org/10.24142/pluriverso.n18a4>

OSCE. (2025). Organization for Security and Co-operation in Europe *Representative on*

Freedom of the Media. <https://rfom.osce.org/>

Palma, J. C. S., Andrade, J. J. R., & Pinoargote, J. E. M. (2025). Sociedad digital y

desinformación: Un análisis del impacto en la comunicación política contemporánea. *Polo del Conocimiento*, 10(7), 2906–2924. <https://doi.org/10.23857/pc.v10i7.8930>

- Pawelec, M., & Labuz, M. (2026). Are softfakes “soft”? Semantic, ethical, and legal perspectives on deep-fake and AI-generated visualizations in political campaigning. *International Journal of Politics, Culture, and Society*, 1–26. <https://doi.org/10.1007/s10767-026-09579-z>
- Presidencia de la República de Colombia. (2021). *Directiva Presidencial No. 03 de 2021: Lineamientos para el uso de servicios en la nube, inteligencia artificial, seguridad digital y gestión de datos*. Presidencia de la República.
- Pulido, C. B. (2014). *El principio de proporcionalidad y los derechos fundamentales. El principio de proporcionalidad como criterio para determinar el contenido de los derechos fundamentales vinculante para el Legislador*, 4. U. Externado de Colombia. [https://books.google.com/books?hl=es&lr=&id=-4_aAwAAQBAJ&oi=fnd&pg=PA13&dq=Bernal+Pulido,+C.+\(2014\).+El+principio+de+proporcionalidad+y+los+derechos+fundamentales:+El+principio+de+proporcionalidad+como+criterio+para+determinar+el+contenido+de+los+derechos+fundamentales+vinculante+para+el+legislador.+Universidad+Externado+de+Colombia.&ots=o0-RSP6fL4&sig=xDEYG9jwUdZpqfewJchL66nDtq0](https://books.google.com/books?hl=es&lr=&id=-4_aAwAAQBAJ&oi=fnd&pg=PA13&dq=Bernal+Pulido,+C.+(2014).+El+principio+de+proporcionalidad+y+los+derechos+fundamentales:+El+principio+de+proporcionalidad+como+criterio+para+determinar+el+contenido+de+los+derechos+fundamentales+vinculante+para+el+legislador.+Universidad+Externado+de+Colombia.&ots=o0-RSP6fL4&sig=xDEYG9jwUdZpqfewJchL66nDtq0)
- Relatoría Especial para la Libertad de Expresión de la Comisión Interamericana de Derechos Humanos. (2013). *Libertad de expresión e Internet*. Organización de los Estados Americanos. https://www.oas.org/es/cidh/expresion/docs/publicaciones/internet_2013_esp.pdf
- Ros, D., Tamayo, R. D., & Robles, G. K. (2025). Influencia de la manipulación informativa en la democracia y la opinión pública. *Revista Metropolitana de Ciencias Aplicadas*, 8(3), 224–231. <https://doi.org/10.62452/3n5n7325>

- Rossel, M. (2025). *Regular el odio digital: Entre la libertad de expresión y la protección de colectivos vulnerables en Chile* [Tesis de pregrado, Universidad Politécnica Salesiana]. DSpace. <https://dspace.ups.edu.ec/handle/123456789/31014>
- Sampieri, R. (2023). *Glosario de términos y conceptos de investigación cualitativa* (Documento SSRN No. 4574353). SSRN. <https://doi.org/10.2139/ssrn.4574353>
- Sanz, M. (2025). *El papel de la inteligencia artificial en la sociedad: Retos y posibilidades* [Tesis de pregrado, Universidad de Valladolid]. UVaDoc. <https://uvadoc.uva.es/handle/10324/81698>
- Sera, F. A. (2024). *Inteligencia artificial en publicidad multimedia: Optimización de estrategias y métricas publicitarias* [Tesis de maestría, Universidad Internacional de Andalucía]. DSpace UNIA. <https://dspace.unia.es/handle/10334/9436>
- Sierra, C., Lemus, L., Ausín, T., Hernández, J., Cerquides, J., & Ribeiro, Á. (2025). *Inteligencia artificial. Transforming la gestión de políticas y bienes públicos en la era digital*. CSIC. <https://digital.csic.es/handle/10261/376024>
- Solar, J. I. (2020). *Dimensiones éticas y jurídicas de la inteligencia artificial en el marco del Estado de Derecho*. Aranzadi / Torrossa. <https://doi.org/10.33258/01.3789>
- Starbird, K. (2019). Disinformation spreaders: Bots, trolls, and all of us. *Nature*, 571(7766), 449–450. <https://doi.org/10.1038/d41586-019-02235-x>
- Vaccari, C., & Chadwick, A. (2020). Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. *Social Media + Society*, 6(1), 1–13. <https://doi.org/10.1177/2056305120903408>
- Valladolid, M. N., & Chávez, L. M. N. (2020). El enfoque cualitativo en la investigación jurídica, proyecto de investigación cualitativa y seminario de tesis. *Vox juris*, 38(2), 69-90.

- Vargas, I. (2013). Elementos doctrinales para el estudio de la argumentación como eje del control judicial. *Prolegómenos*, 16(32), 235–246. <https://doi.org/10.18359/prole.2241>
- Vargas, I. (2016). The ISP's role in improving intellectual property protection on the digital economy. *Vestigium Ire*, 10(1), 112–117. <https://doi.org/10.19053/2011608.5283>
- Vargas, I. (2025a). Take it Down Act: Un análisis crítico de su rol contra la difusión no consentida de imágenes y deepfakes. *Revista Criminalidad*, 67(3), 39–56. <https://doi.org/10.47741/17943108.729>
- Vargas, I. (2025b). Métodos de investigación jurídica. Un análisis desde la perspectiva tridimensional del derecho. *Verba Iuris*, 54, 128–150. <https://doi.org/10.18041/0121-3474/verbaiuris.54.12948>
- Vargas, I. (2025c). *Derecho internacional*. Grupo Editorial Ibáñez.
- Vargas, I. (2026). El bloqueo dinámico de retransmisiones deportivas: Una nueva frontera para el Derecho de telecomunicaciones y la propiedad intelectual. *Law, State and Telecommunications Review*, 18(1), 94–118. <https://doi.org/10.26512/lstr.v18i1.59884>
- Vargas, I., & Álzate, D. (2017). El Derecho como un instrumento generacional y dinámico al servicio de la sociedad: Un acercamiento desde la óptica holmesiana. *Vestigium Ire*, 11(1), 80–92. <https://dialnet.unirioja.es/descarga/articulo/7883806.pdf>
- Vargas, I., Roa, M., & Téllez, R. (2026). Las garantías procesales en el Derecho Comunitario Europeo y su armonización como condición de legitimidad para la justicia negociada. *Revista Brasileira de Alternative Dispute Resolution*, 8(15), 197–227. <https://doi.org/10.52028/rbadr.v8.i15.ART09.COL>

- Vásquez, E. D. C., Loza, R. F. N., Cherrez, A. M. F., & Montes, R. E. T. (2024). Uso de la inteligencia artificial en la personalización del aprendizaje. *Conocimiento Global*, 9(1), 75–83. <https://doi.org/10.47280/Rev.Con.Global.v9.n1.2024.07>
- Vásquez, J. A. (2024). Metodología de investigación en Derecho y Literatura para la educación jurídica y derechos humanos. *Justicia*, 29(45), Artículo e7342. <https://doi.org/10.17081/just.29.45.7342>
- Vélez, J. E. V. (2026). Comunicación Digital en la sociedad ecuatoriana: ¿Libertad de expresión y desinformación en la era de la Inteligencia Artificial durante las elecciones presidenciales 2025? *Maestro y Sociedad*, 23(1), 87–100. <https://maestroysociedad.uo.edu.cu/index.php/MyS/article/view/7367>
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146–1151. <https://doi.org/10.1126/science.aap9559>
- Wardle, C., & Derakhshan, H. (2017). *Information disorder: Toward an interdisciplinary framework for research and policy making* (DGI(2017)09). Council of Europe. <https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-research/168076277c>
- Yepes, R., Newman, V., & Güiza, D. I. (2015). *Concepto sobre el proceso de tutela T-4296509: Respuesta a oficio OPT-A-646/2014. Accionados: Casa Editorial El Tiempo S.A. y Google Colombia*. Centro de Estudios de Derecho, Justicia y Sociedad (Dejusticia).