

XXII. Inteligencia artificial, propaganda electoral y libertad de expresión: problemas constitucionales de la propuesta de reforma electoral

Sergio LÓPEZ AYLLÓN*
Pedro SALAZAR UGARTE**

SUMARIO: I. *Introducción*. II. *El discurso político como núcleo de la libertad de expresión*. III. *La responsabilidad impuesta a plataformas y medios*. IV. *El problema técnico de identificar contenidos generados con IA*. V. *Una regulación conceptualmente incompleta*. VI. *Una secuencia imposible*. VII. *El prejuicio regulatorio frente a la inteligencia artificial*. VIII. *Conclusión*.

135

I. INTRODUCCIÓN

La iniciativa de reforma constitucional en materia electoral presentada por la presidenta Claudia Sheinbaum introduce un régimen de etiquetado obligatorio para todos los contenidos electorales modificados mediante inteligencia artificial o cualquier otra tecnología. Este régimen plantea un conjunto de problemas importantes, tanto desde la perspectiva constitucional como de carácter operativo. El texto propuesto, que sería un nuevo párrafo que se adiciona a la parte final del apartado C del artículo 41 constitucional que regula la propaganda política, establecería que:

Todo contenido relacionado a los procesos electorales que sea modificado o alterado mediante el uso de inteligencia artificial o cualquier otra tecnología, deberá estar etiquetado por su emisor para su identificación

* Investigador en el Instituto de Investigaciones Jurídicas de la UNAM. ORCID: 0000-0003-1204-4898.

** Investigador en el Instituto de Investigaciones Jurídicas de la UNAM. ORCID: 0000-0001-9562-5043.

en los términos que establezca la ley. Las personas concesionarias de radio o televisión, así como las plataformas de servicios digitales, serán responsables de identificar, evitar la difusión y advertir sobre contenido no etiquetado. La ley establecerá las responsabilidades, medidas y sanciones aplicables.

La exposición de motivos del proyecto de Decreto de reforma omite justificar la medida. Por su parte, el Proyecto de Dictamen de las Comisiones Unidas de Puntos Constitucionales y Reforma Político-Electoral del Senado de la República señala que el propósito es combatir “la manipulación informativa a través de tecnologías como la inteligencia artificial”.¹ Más adelante señala que diversos organismos internacionales han advertido que estas tecnologías pueden ser “utilizadas para alterar o manipular contenidos informativos durante procesos electorales”² lo que justificaría la “regulación del uso de la inteligencia artificial en la comunicación política”.³

Sin duda, este escueto diagnóstico refleja preocupaciones legítimas. En distintos países se ha abierto un debate similar sobre el impacto de los llamados *deepfakes* y de la propaganda política sintética en los procesos electorales.⁴ Lamentablemente, la manera en que la iniciativa plantea esta regulación presenta problemas normativos, constitucionales y técnicos significativos.

En efecto, existe un consenso muy amplio que la libertad de expresión ocupa una posición estructural en las democracias constitucionales, pues constituye “la matriz indispensable de casi todas las demás libertades”.⁵ En consecuencia, cualquier restricción que incida en el debate público

¹ Cámara de Diputados LXVI Legislatura, Proyecto de dictamen de las comisiones unidas de puntos constitucionales y de reforma política electoral por el que se reforman, adicionan y derogan diversas disposiciones de la constitución política de los estados unidos mexicanos, en materia electoral, Comisiones Unidas de Puntos Constitucionales y de Reforma Política-Electoral, p. 41, disponible en: <https://www.brujula.news/wp-content/uploads/2026/03/Proyecto-Dictamen-CU-Reforma-Electoral-1.pdf>.

² *Ibidem*, p. 70.

³ *Idem*.

⁴ Por ejemplo, China, India y Singapur. En los Estados Unidos algunos estados como Texas, California, Nueva York e Illinois han emitido regulación al respecto. La Unión Europea ha establecido obligaciones de transparencia relevantes en el *AI Act*.

⁵ Barendt, Eric, “Freedom of Expression”, en Rosenfeld, Michel y Sajó, Andrés (eds.), *The Oxford Handbook of Comparative Constitutional Law*, Oxford, Oxford University Press, 2010, pp. 891-908.

exige una justificación reforzada y un diseño que genere las menores restricciones posibles. Este no es el caso de la reforma que analizamos. A continuación exponemos los argumentos.

II. EL DISCURSO POLÍTICO COMO NÚCLEO DE LA LIBERTAD DE EXPRESIÓN

En el sistema interamericano, la Corte Interamericana ha señalado reiteradamente que la libertad de expresión tiene una relación estrecha, indisoluble y fundamental con la democracia,⁶ en tanto permite el debate público necesario para el gobierno colectivo.⁷

Esta función es particularmente intensa en el ámbito del debate político. La jurisprudencia constitucional comparada coincide en que la expresión política debe recibir el nivel más alto de protección, dado que constituye la condición de posibilidad de un gobierno democrático. Esto implica que las restricciones al discurso político deben analizarse con el estándar más estricto de idoneidad, necesidad y proporcionalidad.

Junto con lo anterior, se debe considerar que el artículo 13.3 de la Convención Americana prohíbe explícitamente las restricciones indirectas a la libertad de expresión, incluidas aquellas que se ejercen mediante controles sobre los medios de comunicación o las infraestructuras de difusión.

III. LA RESPONSABILIDAD IMPUESTA A PLATAFORMAS Y MEDIOS

Uno de los elementos más problemáticos de la iniciativa es la obligación impuesta a concesionarios de radiodifusión y plataformas digitales de identificar y evitar la difusión de contenidos no etiquetados.

⁶ Comisión Interamericana de Derechos Humanos-Relatoría Especial para la Libertad, *Marco jurídico interamericano sobre el derecho a la libertad de expresión*, Washington, OEA, 2010. Comisión Interamericana de Derechos Humanos-Relatoría Especial para la Libertad.

⁷ Álvarez Ugarte, Ramiro, “Freedom of Expression and Populist Politics”, en Hubner Mendes, Conrado, Gargarella, Roberto y Guidi, Sebastian (eds.), *The Oxford Handbook of Constitutional Law in Latin America*, Oxford, Oxford University Press, 2022, pp. 598-616.

Este enfoque implica delegar en los agentes privados funciones de control del discurso público, cuando esta obligación debe recaer fundamentalmente en la autoridad. La literatura comparada advierte que este tipo de mecanismos puede derivar en censura privada o remoción excesiva de contenidos, pues los intermediarios (concesionarios de radio y TV y plataformas) tienden a eliminar aquéllos potencialmente problemáticos para evitar responsabilidades legales, lo que genera efectos de *overblocking* o bloqueo excesivo.⁸

El problema es más serio respecto a las plataformas. Un principio central en la arquitectura jurídica de Internet es la no responsabilidad o responsabilidad limitada de los intermediarios digitales por los contenidos generados por terceros. Este principio, desarrollado originalmente en el derecho comparado para preservar la apertura del ecosistema digital, parte de la premisa de que las plataformas que alojan o transmiten información —redes sociales, servicios de alojamiento o motores de búsqueda— no actúan como editoras del contenido producido por sus usuarios. En consecuencia, no pueden ser consideradas responsables de manera automática por su difusión, salvo en supuestos específicos en los que, una vez notificadas de la ilicitud del contenido, omitan retirarlo o actuar diligentemente.⁹

Las plataformas han desarrollado códigos, criterios y mecanismos de moderación, que han sido objeto de amplios debates, pero no se les ha exigido una responsabilidad directa sobre los contenidos que difunden sus usuarios.¹⁰

⁸ Benedek, Wolfgang y Kettemann, Mathias C., *Freedom of Expression and the Internet*, Estrasburgo, Council of Europe Publishing, 2020.

⁹ Este modelo fue consolidado en la legislación estadounidense mediante la sección 230 del *Communications Decency Act* de 1996, que establece que ningún proveedor o usuario de un servicio informático interactivo será tratado como el editor o autor de información proporcionada por otro proveedor de contenido. En Europa, un principio similar fue adoptado en la Directiva sobre comercio electrónico de 2000 y posteriormente reiterado por el *Digital Services Act*, que mantiene el modelo de responsabilidad condicionada y rechaza la imposición de obligaciones generales de monitoreo de contenidos. La lógica subyacente es que exigir a los intermediarios una supervisión previa del flujo informativo equivaldría, en la práctica, a instaurar formas de censura privada preventiva y generaría incentivos para la sobre-remoción de contenidos, con efectos negativos sobre la libertad de expresión y el pluralismo informativo.

¹⁰ López Noriega, Saúl, *El futuro de la libertad de expresión. Internet, plataformas y algoritmos*, México, Grano de Sal, 2023.

Existen tres riesgos identificables derivados de la propuesta que analizamos:

- 1) El riesgo de restricciones indirectas, pues cuando el Estado utiliza intermediarios para controlar el flujo de información, puede generar restricciones indirectas incompatibles con la libertad de expresión. Imponer a plataformas y concesionarios la obligación de detectar contenido manipulado mediante IA puede funcionar como un mecanismo de filtrado previo, incompatible con ese principio.
- 2) Incentivos a la censura preventiva. Al enfrentar posibles sanciones, los intermediarios tenderán a eliminar cualquier contenido sospechoso. Esto genera incentivos estructurales para la censura preventiva. Como advierte la literatura sobre intermediarios digitales, cuando las plataformas enfrentan responsabilidad por contenido de terceros, suelen adoptar estrategias de moderación excesiva para minimizar riesgos regulatorios. El resultado puede ser un empobrecimiento del debate público, especialmente durante los procesos electorales, cuando resulta más necesario.
- 3) Distorsión del sistema de responsabilidades. La iniciativa asigna responsabilidades a intermediarios, pero no a los actores políticos que eventualmente generen propaganda manipulada mediante IA. Desde el punto de vista normativo, esta distribución es problemática: quienes producen la propaganda deberían ser los primeros responsables de su veracidad y transparencia. La ausencia de esta responsabilidad directa genera un esquema regulatorio asimétrico e ineficaz.

IV. EL PROBLEMA TÉCNICO DE IDENTIFICAR CONTENIDOS GENERADOS CON IA

Un segundo problema central es la premisa tecnológica sobre la que descansa la norma. La iniciativa asume que es posible identificar con certeza los contenidos generados o manipulados mediante inteligencia artificial. Sin embargo, esta premisa es discutible.

Los sistemas actuales de detección de contenido sintético presentan tasas significativas de error. Además, los modelos generativos evolucionan

rápidamente, lo que dificulta la creación de herramientas confiables de identificación.

En consecuencia, exigir a plataformas y medios que identifiquen contenidos generados con IA supone imponerles una obligación técnica muy difícil o incluso imposible de cumplir.

Este problema puede generar dos efectos indeseables: la responsabilidad por fallas inevitables, debido a la imposibilidad técnica de detectar ciertos contenidos o bien caer en un problema de censura excesiva, al eliminar contenidos legítimos ante la duda.

V. UNA REGULACIÓN CONCEPTUALMENTE INCOMPLETA

La iniciativa también presenta un problema conceptual relevante. El texto obliga a etiquetar contenidos “modificados o alterados mediante inteligencia artificial”, pero no incluye contenidos generados completamente por IA.

Esta distinción carece de sentido tanto desde el punto de vista tecnológico como normativo. En efecto, gran parte de la propaganda potencialmente engañosa puede consistir en contenidos enteramente generados por IA (por ejemplo, discursos sintéticos o imágenes ficticias), que quedarían fuera del supuesto normativo.

El resultado es una regulación que no cubre muchos de los riesgos que pretende evitar, pero sí genera cargas regulatorias importantes para intermediarios. Por otro lado, los conceptos de “modificación” o “alteración” son vahos y elásticos, ¿editar un mensaje equivale a modificarlo?, ¿qué significa en términos de contenidos una alteración?

VI. UNA SECUENCIA IMPOSIBLE

La propuesta tiene un problema de secuencia que genera un resultado imposible. En efecto, por un lado establece que los concesionarios de radio y televisión y las plataformas deben “identificar” y “evitar” la difusión de contenido no etiquetado. Evitar supone un mandato claro: no deben transmitir o difundir contenidos que no estén etiquetados cuando éstos hayan sido modificados o alterados con IA o con “cualquier otra

tecnología”. Al mismo tiempo, los obliga a “advertir sobre el contenido no etiquetado”. La falta de sentido es obvia, no puedo “advertir” sobre un contenido que no pueden “difundir”.

VII. EL PREJUICIO REGULATORIO FRENTE A LA INTELIGENCIA ARTIFICIAL

Finalmente, la propuesta parte de una visión simplificada de la inteligencia artificial —no siquiera diferencia los diferentes tipos de IA existentes— y le añade además un enunciado —“y cualquier otra tecnología”— de tal generalidad que todo puede caber.

La tecnología no es intrínsecamente engañosa ni incompatible con el debate democrático. Como ocurre con cualquier herramienta de comunicación, su impacto depende de cómo se utilice. Timothy Garton Ash ha advertido que el desafío contemporáneo no consiste en restringir la comunicación tecnológica, sino en crear marcos normativos que permitan más y mejor libertad de expresión en un entorno digital global.¹¹

Regular desde el prejuicio tecnológico puede conducir a soluciones normativas ineficaces o contraproducentes.

VIII. CONCLUSIÓN

La propuesta analizada tiene severos problemas desde su concepción hasta su implementación. En efecto:

- 1) Impone responsabilidades indebidas a plataformas y concesionarios.
- 2) Genera riesgos de censura indirecta y sobre-remoción de contenidos.
- 3) Parte de una premisa técnica discutible sobre la detección de contenido generado con IA.
- 4) Omite atribuir responsabilidades directas a partidos y candidatos.
- 5) Diseña una regulación conceptualmente incompleta.

¹¹ Ash, Timothy Garton, *Free Speech. Ten principles for a Connected World*, New Haven-Londres, Yale University Press, 2016.

La defensa de la integridad del debate electoral frente a la desinformación es un objetivo legítimo. Sin embargo, las medidas deben ser compatibles con la libertad de expresión y con la arquitectura del ecosistema digital.

Una regulación constitucionalmente adecuada debería centrarse en la transparencia de los actores políticos, responsabilidades directas de quienes producen propaganda indebida y mecanismos de verificación pública, evitando trasladar funciones de control del discurso a intermediarios privados. De lo contrario, el remedio regulatorio puede resultar más dañino que el problema que pretende resolver.