

La opinión pública en la era de la inteligencia artificial generativa: transformaciones en los mecanismos de intervención política

Public opinion in the age of generative artificial intelligence: transformations in the mechanisms of political intervention

José Antonio Martínez Díez Barroso¹

¹ Instituto de Investigaciones Bibliotecológicas y de la Información (IIBI), Universidad Nacional Autónoma de México (UNAM). Ciudad de México, México.

antonio.diez@iibi.unam.mx (Autor de correspondencia)

Resumen. Este artículo analiza las transformaciones en los mecanismos de incidencia sobre la opinión pública a partir de la irrupción de la inteligencia artificial (IA) generativa. La investigación se sustenta en una revisión bibliográfica de carácter interpretativo, centrada en publicaciones académicas recientes localizadas en bases de datos indexadas, con el objetivo de identificar continuidades y rupturas respecto de las formas tradicionales de influencia mediática. En una primera parte, se examinan los enfoques clásicos de la opinión pública desarrollados en el campo de las ciencias políticas y sociales, destacando el papel de los medios de comunicación, las encuestas y los procesos deliberativos en la formación de las percepciones colectivas. Posteriormente, se analizan las transformaciones introducidas por las redes sociales y, de manera más reciente, por la IA generativa y los grandes modelos de lenguaje (LLM). Los resultados muestran que estas tecnologías han intensificado fenómenos como la personalización del contenido, la formación de “cámaras de eco”, la simulación de consenso mediante bots y la generación de contenido sintético, lo que incide en la configuración de la opinión pública contemporánea. En este contexto, se discuten los riesgos asociados a la manipulación de la percepción colectiva, la desinformación y el debilitamiento de los procesos deliberativos, así como sus posibles implicaciones para las democracias actuales.

Palabras clave: Opinión pública; Inteligencia artificial; Redes sociales; Comunicación política; Desinformación; Democracia

Abstract. This article examines the transformations in the mechanisms influencing public opinion following the emergence of generative artificial intelligence (AI). The study is based on an interpretive literature review of recent academic publications indexed in scholarly databases, aiming to identify continuities and shifts in relation to traditional forms of media influence. The first section revisits classical approaches to public opinion within political and social sciences, emphasizing the role of mass media, surveys, and deliberative processes in shaping collective perceptions. It then explores the transformations introduced by social media and, more recently, by generative AI and large language models (LLM). The findings indicate that these technologies have intensified phenomena such as content personalization, “echo chambers”, the simulation of consensus through bots, and the generation of synthetic content, all of which influence the configuration of contemporary public opinion. In this context, the article discusses the risks associated with the manipulation of collective perception, disinformation, and the weakening of deliberative processes, as well as their potential implications for contemporary democracies.

Keywords: Public opinion; Artificial intelligence; Social media; Political communication; Disinformation; Democracy

Historial: Recibido: 30 de marzo de 2026 | Aceptado: 15 de mayo de 2026

Ética/Financiamiento: El autor declara no tener conflictos de interés. El artículo reconoce el apoyo del Programa de Becas Posdoctorales en la Universidad Nacional Autónoma de México (POSDOC).

Este documento está bajo una licencia
CC BY-NC-ND 4.0.

Cómo citar (APA): Martínez Díez Barroso, J. A. (2026). La opinión pública en la era de la inteligencia artificial generativa: transformaciones en los mecanismos de intervención política. *Conocimiento i Política*, 7(1), 1-36. DOI: 10.64480/cip.2026.n7.01.

1. Introducción

Este artículo expone una investigación teórica basada en la revisión bibliográfica de publicaciones localizadas en bases de datos indexadas. Se centra en las transformaciones de los mecanismos que inciden en la manipulación de la opinión pública tras la irrupción de la inteligencia artificial (IA).

En ese sentido, resulta pertinente señalar que la comprensión de los mecanismos a través de los cuales se influye en la opinión pública ha interesado históricamente a políticos, gobiernos y diversos actores sociales (Lippmann, 1922). Una percepción positiva se traduce en validación y reconocimiento por parte de la ciudadanía, mientras que una percepción negativa implica la pérdida de legitimidad y aceptación social.

Tradicionalmente, los medios de comunicación (como la prensa, la radio y la televisión) desempeñaban un papel central en la difusión de información con el propósito de posicionar las agendas de quienes financiaban los espacios informativos (Lippmann, 1922). Para ello, se recurría a figuras mediáticas que emitían opiniones a favor o en contra de determinados asuntos políticos (Lippmann, 1925). Asimismo, se realizaban encuestas orientadas a medir puntos de convergencia y divergencia, así como la percepción social alrededor de temas o acontecimientos específicos (Allport, 1940). El contexto sociohistórico de los ciudadanos constituía un factor determinante para comprender tendencias y comportamientos (Converse, 1970).

En este marco, los mass media saturaron la esfera pública con contenidos, lo que dificultó la capacidad de filtrar información y de construir opiniones fundamentadas en procesos deliberativos y en consensos democráticos (Zaller, 1992). Esta dinámica afectó la comunicación entre iguales como base para la toma de decisiones informadas (Habermas, 1987), una función que, en principio, se esperaba que fuera reivindicada por las redes sociales. En sus inicios, estas prometían democratizar la expresión de opiniones: si antes los medios actuaban como intermediarios, las redes sociales podrían hacer que cualquier persona con acceso a internet difundiera su punto de vista (Castells, 2009).

No obstante, las redes sociales evidenciaron rápidamente su carácter corporativo, al estar controladas por un número reducido de empresas que monetizan con los datos de sus usuarios. Su modelo de negocio depende de maximizar el tiempo de permanencia en línea, lo que les lleva a ofrecer publicidad focalizada y a intentar

predecir comportamientos. Para ello, se diseñan algoritmos que trazan gustos y actitudes, y que recomiendan contenido hiperpersonalizado (Saura García, 2025).

Esta dinámica se ha intensificado y complejizado con la irrupción de los grandes modelos de lenguaje (LLM) y de la IA generativa. Estas tecnologías son utilizadas de manera creciente para incidir en la opinión pública mediante estrategias sofisticadas de manipulación (Coeckelbergh, 2024). Los chatbots tienden a reforzar la exposición repetitiva a contenidos afines, limitando el contacto con perspectivas divergentes (Pariser, 2011; Bucher, 2018; Aral, 2021).

Lo mismo sucede con los perfiles falsos hechos con IA en redes sociales, los cuales simulan un apoyo inexistente (Elmas, 2023, Doshi et al., 2024). La IA facilita la construcción de falsos consensos, generando la percepción de que determinadas posturas cuentan con mayor aceptación de la que realmente poseen. Además, posibilita la emergencia de fenómenos políticos inéditos, como los “candidatos virtuales” o los “dobles digitales”, que inciden en la configuración de la opinión pública (Woolley, 2023).

Por otra parte, la generación de contenido sintético en redes sociales propicia la creación de realidades destinadas a influir deliberadamente en la percepción colectiva. En este contexto, los deepfakes (y su instrumentalización mediante fenómenos como el “liar’s dividend”) constituyen un ejemplo paradigmático del uso de tecnologías para incidir en la opinión pública (Meikle, 2023; Liu et al., 2025; Jackson et al., 2025). Esto se ha incrementado de manera sostenida y es cada vez más frecuente en procesos electorales y discusiones políticas en redes sociales: imágenes altamente verosímiles de candidatos en situaciones comprometedoras, audios con voces indistinguibles de las reales en los que se admiten actos ilícitos, o videos capaces de engañar incluso a observadores atentos. Lo que antes se manifestaba como una manipulación mediática más evidente se ha transformado en un entramado de mecanismos tecnológicos complejos.

En consecuencia, la IA puede entenderse como una herramienta con un alto potencial para la manipulación de la percepción colectiva (Coeckelbergh, 2024; Saura García, 2025). Por ello, resulta pertinente desarrollar una investigación cuyo objetivo sea analizar las transformaciones en los mecanismos de incidencia sobre la opinión pública a partir de la emergencia de la IA generativa. En este marco, la pregunta que orienta el presente trabajo es: ¿cómo han variado las formas de influencia sobre la

opinión pública en el contexto de la IA generativa?

Para responder a esta interrogante, se adopta una metodología interpretativa basada en el enfoque “Scoping Review” (McLeod, 2024), que permite sintetizar y analizar críticamente las fuentes consultadas. Los resultados se estructuran en dos ejes: por un lado, los mecanismos tradicionales de influencia sobre la opinión pública y, por otro, aquellos asociados a la IA generativa. Posteriormente, se presenta un cuadro comparativo que sintetiza los principales hallazgos.

Asimismo, se establece un diálogo con investigaciones previas en las que se abordan temas como la adopción de marcos regulatorios de la IA en América Latina, los datos utilizados para entrenar los modelos, los riesgos de asumir acríticamente las innovaciones tecnológicas y los posibles efectos adversos sobre las democracias de la región. Finalmente, se proponen líneas de investigación futuras y recomendaciones orientadas a mitigar la manipulación de la opinión pública mediante el uso de la IA generativa.

2. Metodología

El presente artículo se sustentó en una metodología cualitativa de carácter interpretativo. Se trata de una investigación teórica que examinó diversos artículos académicos, informes institucionales y libros centrados en la opinión pública para entender los mecanismos que sirven para incidir en ella, y sus variaciones a partir del uso de la IA generativa en redes sociales. Este procedimiento fue posible mediante una revisión bibliográfica sistemática sobre la relación entre opinión pública e IA.

Los antecedentes metodológicos de estudios similares pueden hallarse en trabajos que proponen análisis críticos de corte teórico para la comprensión de fenómenos propios de las ciencias políticas y sociales (Codina, 2020). En consecuencia, la presente investigación se distanció de métodos empíricos tradicionales, como la observación, las entrevistas, los estudios de caso o el diseño experimental. Esta decisión buscó contrastar las formas de incidir en la opinión pública a partir de la irrupción de la IA generativa; en ese sentido, la revisión bibliográfica sistemática permitió identificar las principales líneas de investigación, así como los temas de mayor interés para la comunidad académica.

La revisión se sustentó en procedimientos de selección, organización y análisis crítico de los autores, lo que dio lugar a un texto de carácter teórico presentado en la sección de resultados. Asimismo, se apoyó en criterios sistemáticos de búsqueda y selección, inspirados en el enfoque “Scoping Review” (McLeod, 2024), pero adaptados a un análisis teórico-interpretativo que preste atención a los modos en que se influye en la opinión pública en el contexto del uso emergente de la IA. Este enfoque permitió mapear la evolución de las formas en que se entiende la opinión pública en la producción académica reciente.

El procedimiento metodológico se desarrolló en cinco fases: (1) definición del propósito de la revisión bibliográfica; (2) identificación de las fuentes; (3) selección de las más pertinentes, con base en su calidad y relevancia conceptual; (4) extracción de la información útil para la investigación; y (5) síntesis crítica y sistemática de los hallazgos (Kitchenham, 2004; Okoli & Schabram, 2010). Las referencias bibliográficas se localizaron mediante el motor de búsqueda Google Scholar, dada su amplia utilización para la consulta de textos académicos; después, se incorporaron libros especializados, ampliamente citados en los artículos revisados.

El periodo de búsqueda abarcó de 2020 a 2025. Esta delimitación responde a que, durante dichos años, se consolidaron y popularizaron desarrollos tecnológicos que han impactado significativamente en la opinión pública, como los chatbots, la IA generativa y los LLM. Asimismo, en ese periodo, gobiernos, partidos políticos, organizaciones de la sociedad civil y diversos actores sociales comenzaron a implementar estas tecnologías con el propósito de influir en la percepción y el juicio de las personas, afectando los procesos deliberativos y de consenso democrático.

La búsqueda bibliográfica se realizó mediante combinaciones de palabras clave en español e inglés, entre ellas: opinión pública, IA generativa y redes sociales. Los textos se filtraron temáticamente a través de las herramientas proporcionadas por el propio motor de búsqueda. Para establecer los criterios de inclusión y exclusión, se consideró que los artículos estuvieran publicados en revistas indexadas y que se inscribieran en el ámbito de las ciencias políticas y sociales. Asimismo, se priorizaron aquellos trabajos que abordaran directamente las transformaciones de la opinión pública a partir del uso de la IA. En total, se recopilaron 87 artículos, los cuales, tras un proceso de revisión y sistematización, se redujeron a 68. Este proceso permitió evitar duplicidades y redundancias, así como alcanzar una saturación conceptual; es decir, se incorporaron publicaciones hasta que dejaron de aportar nuevos temas o

conceptos relevantes para la pregunta de investigación, lo que va en consonancia con metodologías empleadas en estudios previos (Kitchenham, 2004, Okoli & Schabram, 2010).

Los resultados se organizaron en torno a dos ejes principales: el primero corresponde a la etapa clásica de la opinión pública, y el segundo a sus transformaciones a partir de la irrupción de la IA generativa. De este modo, se ofrece una tabla comparativa que permite identificar un antes y un después en la configuración de los factores que inciden en la opinión pública. Finalmente, se establece un diálogo con investigaciones previas que han abordado esta problemática en América Latina, con el fin de profundizar en su análisis teórico.

3. Resultados

La revisión bibliográfica sobre los mecanismos para incidir en la opinión pública permitió identificar diferencias sustantivas entre las formas tradicionales de afectar la percepción social y lo que sucede con la llegada de los LLM y la IA generativa. Estas tecnologías alteran y reconfiguran los puntos de vista mediante estrategias de manipulación de la información desplegadas en redes sociales.

En este sentido, se presenta, en una primera instancia, la “Etapa clásica de la opinión pública” y, posteriormente, se la contrasta con “La opinión pública en la era de la IA generativa”. El análisis de estos dos ejes ayuda a comprender la forma en que se configura la opinión pública, los actores sociales que intervienen en este proceso, así como el papel de los medios de comunicación y la lógica de su funcionamiento. Asimismo, se examinan las particularidades de la deliberación pública, la producción de contenidos, la velocidad de circulación de la información y los mecanismos de influencia y persuasión, con el fin de identificar riesgos y fenómenos emergentes. Con base en esta comparación, se propone una tabla final que sintetiza los principales hallazgos.

4. Etapa clásica de la opinión pública

Uno de los primeros trabajos en abordar de manera detallada la opinión pública como objeto de estudio académico fue el de Bryce (1888). En esta obra temprana se advierte ya una preocupación por la estrecha relación entre los medios de comunicación y los intereses políticos. El argumento central de Bryce gira en torno a los procesos de simbolización mediante los cuales los sentimientos y pensamientos de las personas pueden influenciarse de manera deliberada con el fin de incidir en su subjetividad y, a la par, comercializar con la atención de los ciudadanos.

Desde su perspectiva, Bryce concebía a los medios de comunicación de la época, en particular los periódicos, como empresas que operan primordialmente en el ámbito del entretenimiento. Según su planteamiento, los partidos políticos recurren a ellos para captar el tiempo y la atención de las audiencias, a quienes se expone de manera reiterada a contenidos alineados con los intereses de los grupos con poder. De esta forma, los periódicos serían aparatos de difusión capaces de moldear la percepción social alrededor de temas de relevancia pública.

La perspectiva de Bryce fue contrastada y debatida por otros autores que, en las primeras décadas del siglo XX, profundizaron en el estudio de la opinión pública. Lowell (1913) sostuvo que, a pesar de la evidente conexión entre política y medios de comunicación, estos últimos tienen limitaciones para gestionar de manera adecuada la información relativa a asuntos de interés público. A su juicio, no pueden ser tratados como instrumentos pertinentes para el funcionamiento de los procesos deliberativos.

La razón de esta limitación radicaría en que los periódicos se orientan principalmente hacia la difusión de contenidos más que a la participación fundamentada en asuntos que conciernen a la mayoría. En otras palabras, su lógica prioriza la circulación de mensajes diseñados para atraer audiencias antes que la construcción de contextos informativos amplios y rigurosos. Como consecuencia, los ciudadanos carecerían de los elementos necesarios para deliberar de manera ponderada sobre, por ejemplo, el rumbo político de un país, las implicaciones de ciertas legislaciones o la conducta de sus gobernantes.

No obstante, la postura de Lowell también fue objeto de críticas debido a su marcada inclinación elitista en materia política. Sostenía que la participación ciudadana debía ser considerablemente limitada, pues era improbable e imposible que la totalidad de

la población dispusiera de los instrumentos informativos y formativos requeridos para comprender la complejidad de los procesos políticos y sociales. De ese modo, la elección de representantes tendería a sustentarse en simplificaciones, generalizaciones y estereotipos; por ello, Lowell defendía la conveniencia de que las decisiones más relevantes fueran delegadas en grupos mejor informados, a los que identificaba con sectores intelectuales o especializados.

En ese contexto de intensa producción académica en ciencias políticas y sociales apareció la obra de Lippmann (1922), que con el tiempo se convertiría en un referente sobre el tema. De acuerdo con él, los medios de comunicación contribuyen a que la información circule de manera fragmentaria. Desde su perspectiva, la imagen de una sociedad que toma decisiones políticas a partir de opiniones plenamente informadas constituye, en gran medida, un ideal difícil de alcanzar. La estratificación social característica de las sociedades modernas, junto con los accesos desiguales a los medios de comunicación y a los recursos educativos, obstaculiza una distribución equitativa de la información. Asimismo, el entorno social contemporáneo es tan vasto y complejo que la mayoría de las personas no dispone ni del tiempo ni de las herramientas cognitivas necesarias para comprenderlo cabalmente.

Al respecto, Lippmann argumentó que, en la práctica, la opinión pública se construye a partir de estereotipos. Estos funcionan como esquemas interpretativos que permiten a los individuos organizar y simplificar la realidad social. En otras palabras, operan como moldes mediante los cuales se percibe aquello que la cultura previamente ha delimitado, facilitando así la aprehensión conceptual de fenómenos que, en una primera instancia, resultan difíciles de comprender.

Por esta razón, al igual que Lowell, Lippmann manifestó una profunda desconfianza hacia las formas de democracia directa. Consideraba que la opinión pública podía variar considerablemente debido a la insuficiencia de información y a la susceptibilidad emocional de los ciudadanos. En este sentido, compartía la idea de que una posible solución consistía en limitar la toma de decisiones políticas y delegarla a una “clase de expertos”. Este grupo (integrado por individuos con formación especializada) tendría la capacidad de interpretar la complejidad del entorno social y orientar la percepción colectiva hacia decisiones más racionales enfocadas en el bienestar general. De ahí que el autor llegara a plantear la necesidad de una élite tecnócrata con la capacidad de articular los intereses de la ciudadanía.

No obstante, revisó críticamente esta postura en su obra posterior (Lippmann, 1925). En este texto reconoció que incluso los expertos podrían ignorar problemas particulares o ser incapaces de captar la totalidad de la vida democrática. En lugar de centrarse exclusivamente en la función de las élites especializadas, Lippmann dirigió entonces su atención hacia el papel de la propaganda en la formación de la opinión pública. A su juicio, los medios de comunicación producen contenidos diseñados deliberadamente para su difusión masiva; dichos contenidos son seleccionados y elaborados por los propietarios de los medios y por diversos líderes de opinión, quienes establecen jerarquías informativas que determinan qué temas reciben mayor visibilidad y cuáles permanecen marginados del debate público.

En relación con este proceso, Cohen formuló posteriormente una observación que se volvería ampliamente citada en los estudios de comunicación política: “puede ser que los medios no tengan éxito la mayor parte del tiempo en decirle a la gente qué pensar, pero tienen un éxito sorprendente en decirle a la gente sobre qué pensar” (1963, p. 13). Esta afirmación subraya el poder de los medios, entre los que ahora ya se encontraban también la televisión y la radio, para estructurar las agendas públicas y guiar la atención colectiva.

Frente a estas limitaciones, Lippmann sostuvo que la mejor vía para contrarrestar los efectos de los estereotipos, la propaganda y la influencia desproporcionada de la “clase de expertos” consistía en fortalecer la educación cívica y el pensamiento crítico de los ciudadanos. Desde esta perspectiva, debían existir herramientas que permitieran comprender los procesos políticos, evaluar críticamente la información disponible y participar de manera más reflexiva en la deliberación pública. En última instancia, Lippmann defendía la posibilidad de que los individuos desarrollaran juicios tanto personales como colectivos mejor fundamentados; para ello, consideraba indispensable reformular el papel de los medios de comunicación, llevándolos hacia la promoción de procesos deliberativos. Estas ideas serían posteriormente tomadas en cuenta para justificar la incorporación de debates políticos en medios durante las campañas electorales.

Tiempo después, los estudios sobre la opinión pública continuaron desarrollándose a lo largo de las primeras décadas del siglo XX, como lo evidencia el trabajo de Dewey (1927). En esta obra, Dewey concibe la opinión pública como un entramado de relaciones sociales que condiciona tanto el acceso a la información como el comportamiento individual. Dichas relaciones se encuentran arraigadas en tradiciones,

prácticas culturales e instituciones sociales, y desempeñan un papel fundamental en los procesos de formación cívica. Desde esta perspectiva, la opinión pública posee un carácter eminentemente social: las posturas respaldadas por una mayoría tienden, progresivamente, a influir en las minorías, generando un efecto que refuerza determinados consensos dentro del espacio público.

En ese mismo periodo se consolidaron diversas técnicas de muestreo, como las encuestas, que contribuyeron a profesionalizar el estudio empírico de la opinión pública en las ciencias políticas y sociales. Estas herramientas permitieron sistematizar la recolección de datos sobre percepciones y actitudes colectivas, ampliando las posibilidades para comprender la circulación social de la información y las formas en que esta incide en la formación de opiniones.

El desarrollo de estas técnicas también permitió reconocer que la opinión pública no constituye un fenómeno plenamente objetivo ni exento de ambigüedades, dado que en su configuración intervienen inevitablemente dimensiones subjetivas. Desde esta perspectiva, las encuestas buscaron medir preferencias individuales y comprender los contextos culturales e históricos que favorecen la emergencia de actitudes y comportamientos compartidos socialmente. En ese sentido, la opinión pública podía interpretarse como el resultado de patrones relativamente uniformes de percepción y valoración entre los ciudadanos, agrupados alrededor de marcos culturales comunes (Allport, 1940).

A partir de este enfoque, numerosos estudios comenzaron a analizar procesos sociohistóricos sustentados en bases culturales compartidas, con el propósito de relacionarlos con las actitudes de la ciudadanía frente a los gobiernos, las políticas públicas o los partidos políticos. Un ejemplo de esto fue el artículo de Converse (1964). Este trabajo marcó un desplazamiento significativo en la agenda de investigación: en lugar de centrarse exclusivamente en el papel de los medios de comunicación, comenzó a examinarse cómo los ciudadanos, mediados por sus trayectorias históricas y culturales, procesan la información disponible al momento de tomar decisiones políticas.

Sostuvo que el simple acceso a la información política no implica necesariamente un interés activo por parte de la ciudadanía. Para describir este fenómeno introdujo el concepto de “Non-Attitudes” (Converse, 1970), entendido como la ausencia de posturas consistentes frente a determinados asuntos públicos, motivada por la apatía o la desafección política. No obstante, esta interpretación fue objeto de críticas por

parte de diversos autores, quienes argumentaron que la aparente inestabilidad de las opiniones respondía a limitaciones metodológicas en el diseño de las encuestas. Así, se sostuvo que era posible identificar patrones relativamente estables de opinión pública e incluso anticipar comportamientos colectivos con cierto grado de precisión, siempre y cuando las preguntas fueran formuladas de manera rigurosa y contextualizada. En este sentido, la opinión pública podía exhibir rasgos de racionalidad, coherencia y estabilidad, aunque sus preferencias se ajustaran progresivamente a la información que circula en la sociedad (Page & Shapiro, 1992). Este periodo estuvo marcado por un intenso debate entre quienes defendían la validez analítica de las encuestas y quienes cuestionaban su capacidad para captar adecuadamente las actitudes políticas de la población.

Con el tiempo y la llegada de los medios masivos de comunicación, el campo de estudio de la opinión pública experimentó una revitalización significativa. El trabajo de Zaller (1992) abrió nuevas líneas de investigación. Propuso que la aparente inconsistencia de las opiniones ciudadanas no se explica por una carencia de ideas, sino por la sobreabundancia de información a la que están expuestos los individuos y por la ausencia de mecanismos eficaces para discriminar entre información verificada y contenidos basados en rumores o conjeturas.

Según este enfoque, gran parte de la información política se acumula en la memoria sin ser necesariamente filtrada o sistematizada. Cuando los individuos son llamados a emitir una opinión, tienden a recurrir a aquellas ideas que han tenido mayor visibilidad o repetición en el espacio público. De este modo, las respuestas no surgen de manera espontánea ni aleatoria: se apoyan en marcos interpretativos previamente disponibles. En consecuencia, la formación de la opinión pública se encuentra influida tanto por los mass media que jerarquizan la información como por la historia sociocultural de los individuos, que condiciona la manera en que la interpretan y recuerdan.

Por su parte, en las décadas de 1980 y 1990 circularon profusamente las ideas de Habermas (1987). El autor planteó que la opinión pública articula, por un lado, un juicio incierto y atravesado por la subjetividad, y, por otro, la reputación que un tema o persona goza a consideración de un grupo social. Dicha reputación puede verse afectada por estrategias de comunicación política, como el marketing, frente a las cuales Habermas subrayó la necesidad de marcos regulatorios que, sin menoscabar el derecho a la información, garanticen condiciones de equidad y transparencia.

Desde esta perspectiva, lo definitorio de la opinión pública no radica en su carácter representativo, entendido como un conjunto aceptado de preferencias, sino en un consenso alcanzado con el diálogo. Habermas rechazaba que lo relevante del término se encuentre en la cantidad de puntos de vista que una población tenga sobre determinado tema; en cambio, propone que se evalúe el entendimiento mutuo posibilitado por el lenguaje. En este sentido, define la opinión pública como una “red para la comunicación de contenidos y tomas de posición [...] que se reproducen a través de la acción comunicativa” (Habermas, 1998, p. 441). Esta red se configura y legitima en función de la calidad de la información disponible y de las condiciones en las que se desarrolla la acción comunicativa, la cual recae en los procesos deliberativos y en los diálogos libres de cualquier presión externa.

En consecuencia, la formación de la opinión pública depende en gran medida de las características institucionales que estructuran la vida social, de los niveles de confianza que estas instituciones suscitan y de la existencia de entornos que favorezcan la deliberación racional y el intercambio dialógico (Habermas, 1987). No obstante, esta concepción se sitúa en tensión con las dinámicas contemporáneas de circulación de la información, en las que las condiciones ideales de la acción comunicativa resultan cada vez más difíciles de sostener.

Con la expansión de internet a partir de la década de 2000, la opinión pública experimentó una transformación significativa: dejó de ser predominantemente receptiva para adquirir un carácter participativo, transitando de una lógica centrada en el consenso hacia otra basada en la interacción. En este contexto, Castells (2009) introdujo el concepto de “autocomunicación de masas” para describir la capacidad de los individuos de producir y difundir sus propios mensajes en entornos digitales, incidiendo potencialmente en la percepción de los otros. Desde esta óptica, la opinión pública habría experimentado un proceso de democratización, al posibilitar la circulación de múltiples perspectivas en espacios como foros, blogs y sitios web.

Estas transformaciones inauguraron el estudio sistemático de la opinión pública en entornos digitales; inicialmente marcado por un cierto optimismo respecto a sus potenciales efectos democratizadores. Se consideraba que la posibilidad de que los individuos difundieran sus opiniones de manera rápida y a gran escala podría fortalecer la deliberación pública y diversificar el debate. Sin embargo, esta visión fue rápidamente matizada por diversas críticas. Se señaló que no todos los discursos en internet poseen el mismo alcance ni visibilidad, ya que los sistemas de búsqueda tien-

den a privilegiar aquellos contenidos con mayores niveles de interacción, reforzando así determinadas narrativas dominantes (Couldry, 2012). También se subrayó que en internet hay poca deliberación; en su lugar, circulan contenidos polémicos o cargados de emociones, lo que va en detrimento de las opiniones basadas en información verificada y en la argumentación fundamentada (Dean, 2005).

En última instancia, la irrupción de las redes sociales terminó de consolidar este desplazamiento, trazando un punto de inflexión en las formas de construcción de la opinión pública. En estos espacios, la producción y circulación de información ya no depende única y exclusivamente de los medios de comunicación tradicionales, sino que se articula a través de múltiples actores y dispositivos. A ello se suman mecanismos como la personalización de contenidos basada en el análisis de datos y su gestión con algoritmos, así como la incorporación reciente del uso de IA para incidir en la percepción ciudadana, lo que ha complejizado aún más los procesos deliberativos y democráticos.

5. La opinión pública en la época de la IA generativa

El uso político y la creciente popularidad de los modelos de IA han introducido transformaciones significativas en la opinión pública de quienes interactúan y se informan en redes sociales. Estas tecnologías permiten la creación y difusión automatizada de contenidos en múltiples formatos, como texto, imagen, audio y video, al tiempo que facilitan altos niveles de personalización y una distribución prácticamente inmediata de los discursos.

En la actualidad, las redes sociales se han consolidado como un espacio central en la configuración de la opinión pública. Diversas estimaciones indican que una proporción significativa de la población en América Latina, cercana al 86 %, participa en dichos espacios (Larocca, 2025). Lejos de constituir meros canales de comunicación, las redes sociales operan como modelos de negocio sustentados en tres pilares interrelacionados: la hiperconectividad de los usuarios, el monitoreo sistemático de su actividad (dataficación) y la gestión de los contenidos a los que acceden (algoritmización) (Saura García, 2025).

Son controladas por un número reducido de grandes corporaciones tecnológicas, que han adquirido un papel central en la formación de la opinión pública contemporánea

(García-Marzá & Calvo, 2024). Su lógica de funcionamiento recae en la hiperconectividad: maximizar el tiempo de permanencia de los usuarios en línea, lo que exige una producción constante de contenido textual y audiovisual. Esta exposición continua ha propiciado fenómenos como la falsificación de identidades, la manipulación y la instrumentalización de la opinión pública (Moore, 2018).

Sostener esta dinámica requiere también de una administración personalizada del contenido, lo cual es posible gracias a la recolección masiva de datos generados por los propios usuarios. La dataficación, en este sentido, se refiere al proceso mediante el cual las empresas tecnológicas (y, en ocasiones, gobiernos que establecen acuerdos con ellas) recopilan, gestionan y analizan datos y metadatos derivados de la actividad digital (García-Marzá & Calvo, 2022; Splichal, 2022). La revisión e inspección de los datos permite conocer las actividades de los usuarios, sus comportamientos, gustos, información personal (dónde viven y con quién se relacionan) y anticipar las posibles acciones que tomará en el futuro (Feldstein, 2019; Snowden, 2019; Han, 2022; Farahany, 2023). El procesamiento de toda esa información sirve para la categorización y segmentación de los usuarios, lo que a su vez posibilita la algoritmización: la implementación de sistemas computacionales diseñados para perfilar a los individuos y recomendarles contenidos acordes con sus intereses, hábitos y preferencias. Todo esto para mantener a los usuarios conectados y expuestos a publicidad focalizada, monetización de datos, suscripciones a plataformas, y venta de servicios tecnológicos, entre otros fines.

Este proceso de algoritmización no solo determina qué información resulta visible, sino también qué contenidos permanecen ocultos, reduciendo la exposición a perspectivas divergentes y limitando la pluralidad del debate público (Hao, 2025). En este contexto, las redes sociales ejercen un control cuasi monopólico sobre la circulación de la información. Se trata, en su mayoría, de corporaciones estadounidenses y chinas que concentran tanto el acceso a los datos de millones de usuarios como el control sobre los algoritmos que organizan la visibilidad de los contenidos (Cave et al., 2019). En un entorno donde la desconexión implica, en la práctica, la exclusión de la esfera virtual, la aceptación de los términos y condiciones de uso se vuelve prácticamente ineludible (Zuboff, 2019). En consecuencia, la opinión pública tiende a reducirse a un conjunto de inferencias calculadas mediante sistemas computacionales que procesan y predicen comportamientos.

Al respecto, la irrupción de la IA terminó articulándose también con las dinámicas de hiperconectividad, dataficación y algoritmización propias de las redes sociales, e incidiendo en la formación de la opinión pública. El contenido producido por los modelos de IA suele reforzar dinámicas como la conformación de “cámaras de eco”, entendidas como entornos discursivos en los que los individuos se exponen predominantemente a opiniones afines (Pariser, 2011, Bucher, 2018, Aral, 2021). En tales condiciones, los procesos de formación ciudadana quedan circunscritos a marcos interpretativos homogéneos, lo que limita la confrontación con perspectivas divergentes y debilita las condiciones para una deliberación plural (Coeckelbergh, 2024).

En particular, el contenido que circula en estos entornos puede adoptar formas tanto textuales como audiovisuales, lo cual amplifica su capacidad de persuasión. Estos contenidos funcionan a partir de los LLM y de la IA generativa, cuya convergencia ha facilitado la manipulación de la opinión pública con mecanismos que alteran hechos, actores sociales y acontecimientos políticos.

Los LLM inciden en las redes sociales, en buena medida, a través de agentes conversacionales o chatbots, que han adquirido una presencia creciente en el espacio digital. Estos sistemas pueden ser programados para generar contenidos alineados con objetivos específicos, como posicionar candidaturas, legitimar políticas públicas o desacreditar adversarios, contribuyendo así a intensificar la polarización y a personalizar la exposición informativa. En este sentido, los chatbots operan como herramientas de intervención política, amplificando determinados discursos y modulando la interacción con los usuarios.

En América Latina, se han documentado diversos casos de utilización de “granjas de bots” con fines de influencia política. Por ejemplo, durante el proceso electoral mexicano de 2018, se estimó que una proporción significativa de las cuentas en la plataforma X que seguían a los principales candidatos no correspondía a usuarios reales, sino a sistemas automatizados diseñados para incidir en la percepción pública y generar interacción artificial (Redacción AN, 2018). Estas prácticas incluyen la compra de seguidores falsos y la reproducción automatizada de mensajes favorables, lo que crea una apariencia de apoyo social que no necesariamente se corresponde con la realidad (Pastor-Galindo et al., 2020).

Si bien estos sistemas también pueden desempeñar funciones legítimas, como la automatización de servicios informativos, en numerosos casos actúan como instrumentos de manipulación discursiva. Mediante técnicas de Machine Learning, los chatbots interactúan con usuarios y generan contenidos orientados a objetivos estratégicos, como la difusión de desinformación o la intensificación de la polarización (Ferrara et al., 2016; Yang et al., 2019). Además, su comercialización se inscribe en un mercado escasamente regulado y de carácter transnacional, en el que los operadores trabajan con infraestructuras distribuidas en distintos países y recurren a herramientas como redes privadas virtuales (VPN) para ocultar su localización (Gorodnichenko et al., 2021).

En esa misma línea, es relevante el uso de los LLM en la adecuación estratégica de mensajes políticos, mediante procesos de individualización basados en datos demográficos, conductuales e incluso psicométricos. Esta práctica, conocida como “microtargeting” asistido por IA, supone una transformación respecto de las formas tradicionales de segmentación (Simchon et al. 2024). Mientras que antes los discursos orientados a incidir en la opinión pública se estructuraban a partir de categorías relativamente estables —como el género, la edad o la clase social—, ahora se configuran perfiles altamente individualizados, contruidos a partir de rasgos psicológicos específicos y patrones de comportamiento inferidos por algoritmos.

Esto se aleja del marketing político tradicional y apunta hacia la consolidación de una nueva forma de intervención comunicativa que puede caracterizarse como una tecnología de persuasión automatizada, y que permite la generación masiva de mensajes ajustados a personalidades variadas (Hackenburg & Margetts, 2024; Simchon et al., 2024). Incluso con la evolución de los modelos de IA es posible que en un futuro se prescinda de equipos humanos dedicados a la comunicación política para que, en su lugar, sea la tecnología la que analice grandes cantidades de información, la procese y despliegue campañas de manera autónoma, apoyándose en los datos extraídos de las interacciones de los usuarios en redes sociales (Saura García, 2025).

Al respecto, un experimento basado en LLM mostró que los mensajes hiperpersonalizados son más persuasivos que los genéricos, observando que incluso pequeñas diferencias de estilo producen cambios cuantificables (Hackenburg & Margetts, 2024). Esto se debe, en parte, a que la persuasión aumenta cuando hay interacción continua. Debido a que la IA envía más de un mensaje y los chatbots pueden responder en tiempo real, modular el tono discursivo, insistir estratégicamente y reformular

argumentos en función de las reacciones del interlocutor, hay menor resistencia a cambiar de opinión (Argyle et al., 2025). En consecuencia, los LLM intensifican el “microtargeting”, transformándolo en un sistema de convencimiento personalizado y automatizado, apoyado en mensajes adaptados a perfiles de individuos y una continua exposición.

En conjunto, estas transformaciones dificultan la configuración de una opinión pública basada en procesos deliberativos y consensuales, y contribuyen a erosionar la confianza ciudadana en los sistemas democráticos. La deliberación tiende a ser sustituida por dinámicas que algunos autores han conceptualizado como una “tiranía de los algoritmos”, caracterizada por la fragmentación del tejido social, la polarización de las opiniones, la personalización extrema de la información y la exclusión de perspectivas disidentes. Este escenario, además, puede profundizar desigualdades estructurales y favorecer la emergencia de fenómenos como “tecnopopulismos” (Bickerton & Accetti, 2021) y “tecnofascismos” (Coeckelbergh, 2026). Los primeros operativizan la división maniquea entre pueblo y élite a través de una legitimación simbólica dada por el saber técnico: gobernantes que presentan a la tecnología como la solución a todos los problemas que dejaron élites corruptas. Mientras tanto, los segundos aprovechan la recolección de datos para vigilar a la ciudadanía y ejercer prácticas de control autoritario, en las cuales se monitorean de manera permanente las acciones y los discursos de las personas.

Ahora bien, además de los LLM, la configuración contemporánea de la opinión pública se encuentra crecientemente atravesada por IA generativa. Esta remite a un conjunto de tecnologías capaces de producir contenido audiovisual que no solo resulta indistinguible del real, sino que además es altamente verosímil (Goldstein et al. 2023). En ese marco, diversos actores políticos instrumentalizan dichas tecnologías con el propósito de intensificar el ruido informacional, mediante la producción masiva de tendencias prefabricadas orientadas a la saturación de las discusiones en redes sociales.

Un fenómeno particularmente ilustrativo es el denominado “astroturfing”, entendido como la simulación de apoyo ciudadano espontáneo que, en realidad, es coordinado mediante algoritmos. Esta práctica se materializa a través de la creación y gestión de cuentas que aparentan pertenecer a individuos reales, pero que son operadas mediante modelos de IA generativa. A diferencia de los chatbots convencionales, incluso aquellos desarrollados a partir de modelos de los LLM, estas entidades

exhiben un manejo lingüístico más sofisticado, así como una capacidad de interacción contextual, coherente y adaptativa que les permite responder de manera diferenciada a los argumentos y a las características de su interlocutor (Elmas, 2023).

En este sentido, el “astroturfing” se fundamenta en la producción de identidades sintéticas: “sujetos” artificiales cuya presencia en redes sociales se construye a partir de simular la identidad digital, adoptando una fotografía de perfil, un nombre propio, una biografía creíble, un historial de publicaciones y una red de contactos. Su rasgo distintivo no radica únicamente en la verosimilitud, sino en su carácter colectivo y coordinado. Lejos de ser usuarios aislados, estas identidades operan como parte de redes articuladas, integradas por cientos o miles de perfiles que amplifican, replican y sincronizan mensajes con el fin de aparentar dinámicas orgánicas de consenso. De este modo, se produce una ilusión de respaldo social que responde a una estrategia deliberada de manipulación de tendencias, visibilidad y alteración de la opinión pública (Doshi et al. 2024).

El objetivo central de estas prácticas es reconfigurar la percepción, haciendo que posiciones marginales o minoritarias adquieran la apariencia de ser ampliamente compartidas. Este simulacro de sobrerrepresentación puede desencadenar un “efecto de arrastre”, en el que individuos reales ajusten sus propias posiciones en función de lo que perciben como una opinión dominante. En México se observó esto cuando ciertos grupos de extrema derecha intentaron (sin éxito) replicar dinámicas de movilización juvenil similares a las registradas en Nepal, particularmente aquellas asociadas a la llamada generación Z, con la apropiación de símbolos provenientes del anime One Piece. Así, el “astroturfing” puede ser visto como una modalidad específica de intervención en la opinión pública mediada por la IA generativa, en la cual la participación ciudadana orgánica es simulada para fabricar consenso, manipular la visibilidad y, en última instancia, reconfigurar las percepciones colectivas (García-Estévez, 2025).

Por ello, más que persuadir en un sentido clásico, estas estrategias parecen orientarse a erosionar las condiciones mismas de posibilidad de la deliberación pública, configurando un entorno en el que la distinción entre verdad y falsedad parece ser irrelevante, mientras que la visibilidad y la capacidad de captar la atención de las personas se constituyen como los principales vectores de incidencia (Hao, 2025). Así, cuando un determinado tema adquiere centralidad en la agenda pública, puede neutralizarse mediante la generación automatizada de contenidos contrarios o polémicos, lo que

desplaza la atención colectiva de asuntos potencialmente perjudiciales para determinados intereses políticos, estatales o corporativos (Coeckelbergh, 2024; García-Marzá & Calvo, 2024).

En este contexto, el acelerado perfeccionamiento de los modelos de IA ha dado lugar a una reducción sustantiva de los costos asociados a la producción de contenido sintético hiperrealista, hasta el punto de que su generación puede depender únicamente de la formulación de instrucciones textuales mínimas (prompts) (Feldstein, 2023; Meikle, 2023; Woolley, 2023). Esta transformación tecnológica amplía el acceso a dichas herramientas, pero también reconfigura la opinión pública, en la medida en que habilita la producción, modulación y circulación de actores, eventos y contextos políticamente relevantes sin estar anclados en referentes reales. En otras palabras, la IA generativa no se limita a intervenir en la representación de la realidad, sino que participa activamente en su (re)construcción simbólica. Su rápida difusión se explica, en buena medida, por la democratización de su acceso, que ha multiplicado el número de agentes capaces de generar y distribuir contenidos audiovisuales a gran escala, incidiendo de manera directa en la formación de percepciones y disposiciones políticas. A su vez, el escalamiento técnico de estos modelos ha incrementado tanto la calidad como la coherencia de los contenidos producidos, al tiempo que ha reducido sus márgenes de error mediante el análisis eficiente de grandes volúmenes de datos y la identificación de patrones recurrentes. Como consecuencia, el contenido sintético se vuelve cada vez más difícil de detectar, especialmente para usuarios no especializados, lo que amplifica su potencial de incidencia en la opinión pública (Goldstein et al., 2023).

En redes sociales, esta dinámica es explotada por una pluralidad de actores (partidos políticos, élites económicas y diversos poderes fácticos) que encuentran en estas tecnologías un instrumento eficaz para intervenir en los procesos de formación de la opinión pública; en abierta tensión con la concepción clásica que la asocia con prácticas deliberativas sustentadas en hechos verificables, la opinión pública derivada de la IA se reduce a un cálculo estadístico. En este sentido, la bibliografía reciente sugiere que la capacidad de los individuos para distinguir entre rostros reales y generados artificialmente se ha deteriorado de manera significativa (Nightingale & Farid, 2022). Esta dificultad plantea un problema que introduce una fractura en la opinión pública clásica, al socavar los criterios mediante los cuales es posible discriminar entre información auténtica y contenido sintético: incluso los propios

modelos de IA presentan limitaciones para efectuar dicha distinción, lo que sugiere la emergencia de un entorno en redes sociales caracterizado por altos niveles de desinformación (Kreps & Kriner, 2023; Hyska, 2025).

Es en este escenario donde se vuelve plausible la proliferación de los llamados “candidatos virtuales”, esto es, entidades generadas por algoritmos que simulan la existencia de sujetos políticos. Se trata de avatares diseñados con atributos propios de la agencia humana (rostro, voz, discurso e interactividad), lo que les permite participar activamente en la comunicación política, emitir posicionamientos e incluso interactuar con audiencias en tiempo real. Desde una perspectiva teórica, ello supone una reconfiguración profunda de la noción de sujeto político, que deja de concebirse en términos antropocéntricos (o humanocéntricos) para desembocar en un ensamblaje sociotécnico. Desde otra arista, la IA sintética posibilita también la emergencia de los llamados “dobles digitales”, avatares artificiales de individuos reales que son utilizados estratégicamente para establecer interacciones personalizadas con el electorado y optimizar la gestión de la imagen pública. Ambos fenómenos, “candidatos virtuales” y “dobles digitales”, se inscriben en una lógica de producción de contenido altamente personalizado, basada en la explotación de datos de los usuarios y su procesamiento mediante sistemas algorítmicos. Este tipo de comunicación política no solo intensifica la segmentación de audiencias, sino que también abre la puerta a la amplificación estratégica de discursos falsos o distorsionados, orientados a la difusión de agendas ideológicas específicas (Woolley, 2023).

En este sentido, la creciente preeminencia del contenido sintético sobre el contenido real ha favorecido la expansión de un fenómeno comunicativo paradigmático de la era de la IA: los deepfakes, cuya capacidad para erosionar los regímenes tradicionales de verdad plantea desafíos para la noción de opinión pública y para el funcionamiento de las democracias contemporáneas.

El término deepfakes surge de la combinación de deep learning y fake, y se refiere a contenidos audiovisuales generados o manipulados mediante IA generativa con el propósito de aparentar autenticidad. Además de apoyarse en el deep learning, también lo hacen en arquitecturas más complejas como las redes generativas —en particular, las redes generativas antagónicas (GAN)— y los modelos multimodales, capaces de integrar y sincronizar información visual, auditiva y textual (Meikle, 2023). No obstante, resulta pertinente subrayar que no todas las deepfakes implican una falsificación total del contenido; más bien, operan a través de distintos grados de

intervención y manipulación, orientados estratégicamente a incidir en la opinión pública. Entre sus usos más recurrentes se encuentra la suplantación de identidad (mediante la modificación de rasgos faciales y vocales), así como la simulación de escenarios que representan eventos inexistentes (Liu et al. 2025).

Las deepfakes pueden clasificarse en función de los recursos técnicos que emplean y del tipo de contenido que producen. En el caso de las deepfakes de video, se ha avanzado significativamente en técnicas de sustitución facial y sincronización labial; sin embargo, persisten limitaciones en la recreación anatómica integral, lo que ocasionalmente se traduce en inconsistencias visuales detectables (manos con más dedos de los posibles, por ejemplo). Por su parte, las deepfakes de audio permiten la clonación de voces con altos niveles de fidelidad, lo que ha sido utilizado para difundir discursos confusos con potenciales implicaciones políticas y diplomáticas. En cuanto a las deepfakes de imágenes, estas suelen representar escenas visuales de eventos improbables o inexistentes, como la detención de líderes políticos, y circulan ampliamente en redes sociales como si se tratara de evidencia comprobada (Meikle, 2023, Liu et al. 2025).

Las implicaciones de estas tecnologías para la opinión pública son profundas. Las deepfakes contribuyen a la distorsión de la realidad, interviniendo en la percepción que conlleva la formación de criterios políticos. Ello se explica, en parte, debido a que los estímulos visuales y auditivos suelen tener mayor impacto que los textuales: en términos generales, los individuos tienden a interpretar lo que ven y escuchan como evidencia directa, mientras que lo que leen se analiza detenidamente y se duda de ello (Raj, 2025). Esto no implica que lo textual carezca de relevancia, sino que en redes sociales los formatos audiovisuales operan como dispositivos de validación simbólica que pueden ser usados para dañar reputaciones, fabricar escándalos o reconfigurar narrativas públicas (Meikle, 2023). En esta línea, diversos estudios han documentado el uso de deepfakes como herramienta de ataque contra periodistas, activistas y actores opositores, con la intención de desacreditarlos o silenciarlos (Pawelec, 2022).

El impacto de estas prácticas en la opinión pública se manifiesta, entre otros aspectos, en la intensificación de respuestas emocionales frente a eventos que, en muchos casos, nunca han ocurrido. Este fenómeno tiende a erosionar el diálogo y el consenso, al tiempo que incrementa el ruido informativo característico de las redes sociales (Liu et al. 2025). Asimismo, la producción y circulación de deepfakes se agrava en contextos electorales, donde su capacidad de incidir en la percepción de los votantes es más

significativa (Meikle, 2023). Dado que los procesos de verificación suelen requerir tiempo, los daños a la reputación persisten en la memoria de los votantes, influyendo en sus decisiones incluso después de que el contenido haya sido desmentido (Ranka et al., 2024).

A este escenario se suma un fenómeno particularmente relevante: el denominado “liar’s dividend”. Este concepto alude a la capacidad de los actores políticos para beneficiarse de la existencia de deepfakes con el fin de eludir mecanismos de rendición de cuentas, desestimando información verificada bajo el argumento de que se trata de contenido manipulado. En un entorno saturado de producciones sintéticas, donde la distinción entre hechos y ficciones se torna cada vez más difusa, se abre un espacio propicio para la negación sistemática de la evidencia. Así, la incertidumbre se convierte en un recurso estratégico: si todo puede ser falso, todo puede ser negado. En este contexto, la opinión pública se moviliza en beneficio de determinados actores políticos, cuyos seguidores interpretan las acusaciones o evidencias en su contra como parte de campañas negras, lo que incluso puede reforzar su apoyo, convirtiéndolos en víctimas. Diversos estudios sugieren que esta estrategia resulta más eficaz que las respuestas tradicionales como el silencio o la disculpa pública, en la medida en que contribuye a la polarización e intensifica las disputas en redes sociales (Jackson et al., 2025).

En conjunto, los LLM y la IA sintética no se limitan, como los medios de comunicación clásicos, a incidir en la opinión pública mediante la difusión de contenidos, sino que operan a un nivel más profundo, orientado a la saturación de información. Su lógica consiste en sustituir la información producida por actores humanos e inundar las redes sociales con contenidos sintéticos textuales y audiovisuales que favorecen la emergencia de realidades alternativas y discursos altamente personalizados y polarizantes. Es una estrategia en la cual se menoscaban los hechos y las realidades que sirven para contextualizar las decisiones políticas y apuntalar la opinión pública; en cambio, se presentan contenidos elaborados a partir de IA por gobiernos, actores políticos o poderes fácticos con la intermediación de las empresas tecnológicas. Esta saturación de información sirve para construir una falsa creencia en torno al desarrollo de un suceso o para trivializar la realidad política bajo la premisa de que todos los discursos son iguales y ameritan el mismo tratamiento. En este sentido, dinámicas como la hiperconectividad, la dataficación y la algoritmización configuran un escenario en el que la visibilidad de los contenidos (y, por ende, su capacidad

de incidencia) puede ser amplificada o reducida artificialmente, reconfigurando los procesos de formación de la opinión pública.

Al respecto, la comparación entre las formas tradicionales de incidencia en la opinión pública y sus cambios con la llegada de la IA generativa queda resumida en la Tabla 1.

Tabla 1: Comparación en las formas en que se incide en la opinión pública (elaboración propia, 2026)

	Etapa clásica de la opinión pública		Opinión pública en la era de la IA generativa
Formación de la opinión pública	Basada en estereotipos, marcos culturales y exposición mediática		Basada en perfiles algorítmicos, datos conductuales y modelado predictivo
Actores principales	Medios de comunicación tradicionales (prensa, radio, TV), élites políticas e intelectuales		Plataformas digitales, corporaciones tecnológicas, algoritmos, IA generativa, bots y actores automatizados
Rol de los medios	Intermediarios que seleccionan y jerarquizan la información		Infraestructura algorítmica que no solo distribuye, sino que produce, personaliza y oculta información
Lógica de funcionamiento	Difusión masiva relativamente homogénea		Hiperconectividad + dataficación + algoritmización = segmentación extrema y personalización
Deliberación pública	Idealmente, basado en el diálogo racional		Erosión de la deliberación: polarización, cámaras de eco y sustitución por interacción automatizada
Producción de contenido	Centralizada (medios y élites)		Descentralizada pero automatizada: IA genera texto, imagen, audio y video
Velocidad y escala	Relativamente lenta y limitada		Inmediata, masiva y continua, con producción automatizada 24/7
Mecanismos de influencia	Propaganda en medios		Microtargeting, astroturfing, bots, deepfakes, hiperpersonalización
Persuasión	Generalista (mensajes para audiencias amplias)		Individualizada y adaptativa (mensajes diseñados para perfiles psicométricos específicos)
Tipo de información dominante	Información mediada y parcialmente verificada		Contenido sintético y potencialmente indistinguible de lo real
Riesgos principales	Manipulación mediática, propaganda, falta de información		Desinformación masiva, simulación de consenso, polarización extrema, control algorítmico
Fenómenos emergentes clave	Estereotipos, agendas preestablecidas, nonattitudes		Deepfakes, astroturfing, candidatos virtuales, liar's dividend, tecnopopulismo, tecnofascismo
Posible definición de opinión pública	Resultado de interacción entre medios, cultura y deliberación		Resultado de cálculos algorítmicos, simulación discursiva y manipulación de visibilidad

Fuente: Elaboración propia.

6. Discusión

En los artículos académicos revisados para esta investigación, destaca la necesidad de que los gobiernos latinoamericanos adopten marcos regulatorios orientados a establecer lineamientos éticos para el uso de los LLM y de la IA generativa en redes sociales. Asimismo, se analizan modelos entrenados con datos de usuarios latinoamericanos con el propósito de impulsar proyectos regionales. De manera transversal, se advierte el riesgo de incurrir en un determinismo tecnológico en el que las innovaciones y los desarrollos terminen por suplantar o desplazar la agencia social. Finalmente, las investigaciones alertan sobre los posibles daños a las democracias latinoamericanas, las cuales, incluso antes de la incidencia de la IA en la opinión pública, ya enfrentaban problemas como liderazgos débiles, disputas por la representatividad legislativa y desconfianza en las instituciones.

En este contexto, la opinión pública en la era de la IA supone la coexistencia de múltiples contenidos que dificultan la conformación de procesos deliberativos y consensuados, a partir de los cuales las personas puedan construir criterios sustentados en información verificable y trazable. Fenómenos como las cámaras de eco, el astroturfing, los candidatos virtuales, los deepfakes y el liar's dividend hacen necesario que los gobiernos establezcan legislaciones que regulen el contenido generado por IA que circula en redes sociales. Dado el carácter confidencial de los algoritmos, protegidos como secretos corporativos, la IA puede derivar en sistemas opacos; en consecuencia, resulta imprescindible establecer mecanismos de rendición de cuentas y de participación ciudadana (Ariza & Ruiz, 2025). De este modo, se podrían mitigar los daños en la opinión pública mediante la catalogación de videos, audios, imágenes y textos bajo la leyenda "hecho con IA" y la identificación de los datos con los que se entrenaron los modelos, así como al permitir que los posibles afectados inicien procesos legales contra las corporaciones que faciliten la desinformación. Esto abre la puerta a una gobernanza de la IA en América Latina en la que destaquen la transparencia, la responsabilidad y la protección de los derechos de los usuarios (Estrada, 2025).

De acuerdo con el Centro de Estudios Estratégicos de Derecho de la Inteligencia Artificial (CEEDIA), en América Latina se han registrado intentos por regular la IA, cuyas orientaciones privilegian enfoques éticos basados en la protección de datos personales, en consonancia con los derechos humanos, y la supervisión humana de los

sistemas de IA, en oposición a la automatización total (Criado, 2024). En este sentido, en México aún no existe una política pública integral que regule la IA; no obstante, se han producido avances a partir de iniciativas como la “Agenda Nacional Mexicana de Inteligencia Artificial (2020)”, que contempla la protección de datos y se enfoca en los derechos digitales. En Brasil destaca la “Estratégia Brasileira de Inteligência Artificial”, orientada a fortalecer principios éticos de gobernanza para promover la innovación y el desarrollo en sectores estratégicos. Por su parte, en Bolivia, la CEEDIA no documenta iniciativas consolidadas de regulación; sin embargo, existe el Proyecto de Ley N.º 558/2024-2025, “Ley General de Inteligencia Artificial”, presentado a inicios de 2026, cuyo objetivo es proteger a la ciudadanía frente a la automatización y al uso indebido de datos personales.

No obstante, pese a estas iniciativas, persisten desafíos significativos, como la debilidad institucional frente al poder de las grandes empresas tecnológicas, la brecha digital que amplifica las desigualdades en el acceso a la información y la dependencia tecnológica respecto de modelos de IA desarrollados fuera de la región. Así, aunque el riesgo de menoscabar la opinión pública ha impulsado una voluntad generalizada en los países latinoamericanos por diseñar marcos normativos, todavía se observa una falta de consenso y de claridad en torno a los puntos de partida (OECD, 2022, CEPAL, 2023).

De manera paralela a la regulación de los modelos de IA, algunos gobiernos con rasgos progresistas, como el México de la presidenta Sheinbaum o el Chile de Boric, han impulsado iniciativas tecnológicas orientadas a contrarrestar el peso de los monopolios. Con este propósito, se han desarrollado modelos de IA entrenados con datos de usuarios latinoamericanos. Aunque estos proyectos no cuentan con el mismo nivel de financiamiento que sus contrapartes estadounidenses, ni procesan volúmenes comparables de información (Carrasco-Ahen et al., 2025), buscan limitar el predominio del inglés como lengua hegemónica en el entrenamiento de los modelos de IA. Al mismo tiempo, promueven el uso de datos generados en español y en lenguas originarias, bajo el supuesto de que ello permitirá obtener respuestas más plurales y representativas de las realidades latinoamericanas y de su expresividad cultural (Mora-Reyes et al., 2025). De este modo, el contenido generado por IA que incide en la opinión pública incorporaría matices propios de América Latina, incluyendo dialectos, culturas indígenas y diversas cosmovisiones. Asimismo, los modelos regionales constituyen un intento por alcanzar la soberanía tecnológica;

en este sentido, puede mencionarse el caso de Latam-GPT (Lagos, 2025), una IA de código abierto, susceptible de ser compartida y modificada libremente, que se desarrolla en Chile en el Centro Nacional de Inteligencia Artificial (CENIA).

Pese a lo anterior, todavía no existe un proyecto regional articulado; lo que hay son iniciativas nacionales que responden a condiciones estructurales específicas. Así, el Índice Latinoamericano de Inteligencia Artificial señala la presencia de una infraestructura tecnológica débil y rezagos en la inversión, lo cual dificulta la consolidación de una “IA latinoamericana”. A ello se suma la ausencia de una matriz tecnológica robusta, lo que limita la contratación de especialistas en programación de IA, quienes con frecuencia deben migrar a otros países. En contraste, prevalecen formas de trabajo precario centradas en el etiquetado de datos, lo que lleva a que la opinión pública mediada por los escasos modelos de IA en América Latina tienda a construirse sobre condiciones laborales desiguales y en el marco de esfuerzos aún incipientes por superar la dependencia tecnológica (Jung & Katz, 2025).

A esta dependencia debe sumarse el riesgo de incurrir en determinismos tecnológicos. Ningún modelo de IA opera en abstracto ni de manera neutral. Estos desarrollos se sustentan en una materialidad concreta: microprocesadores, centros de datos, recursos energéticos y trabajo humano, entre otros. La tecnología no se produce por sí misma, sino que es el resultado de decisiones humanas y de relaciones de poder situadas en contextos sociohistóricos específicos (López del Castillo, 2024). Su naturalización oculta la participación de gobiernos y empresas, y constituye, además, una estrategia comercial mediante la cual los errores se presentan como “alucinaciones” y los procesos estadísticos como “razonamientos”, alimentando narrativas mediáticas cuyo propósito es sobredimensionar las capacidades de los modelos de IA para posicionarlos en el mercado.

Las decisiones humanas involucradas en la programación de algoritmos inciden directamente en la opinión pública, en la medida en que persiguen una “pretensión de verdad” que hace pasar contenidos generados artificialmente, y atravesados por diversos sesgos, como verdades sustentadas en evidencia (Marda, 2019). Los artículos revisados para esta investigación señalan que el determinismo tecnológico promueve una visión acrítica que menoscaba la agencia social y conduce a concebir la automatización como un proceso inevitable e ineludible. Ello deriva en una despolitización que presenta a los usuarios como actores pasivos, desprovistos de capacidad de transformación (Guanche et al., 2025). Asimismo, atribuye a la tecnología una autonomía

de la que carece, encubriendo los entramados sociales, económicos y políticos que condicionan su accionar.

La dimensión política resulta fundamental para comprender los cambios que ha experimentado la opinión pública a raíz de la irrupción de la IA generativa. En América Latina, las democracias han registrado una intensificación de la desinformación durante las campañas electorales, lo que ha alterado el debate público y favorecido la circulación de contenidos no verificados (CEPAL, 2023; López, 2023). Un informe del International Institute for Democracy and Electoral Assistance (IDEA) señala que la IA exacerba la baja confianza institucional y la fragmentación en las percepciones ciudadanas sobre sus gobernantes, lo que deriva en una débil credibilidad de los líderes políticos y de las estructuras parlamentarias. Asimismo, la asimetría de poder entre las grandes empresas tecnológicas globales y gobiernos latinoamericanos con escasa cohesión institucional genera un entorno propicio para la acción de lobbies, incrementando la influencia política de las corporaciones y la ausencia de marcos normativos efectivamente vinculantes (Becker Castellaro et al., 2025).

En este contexto, los gobiernos de la región se enfrentan a un desfase estructural: por un lado, una acelerada evolución tecnológica y, por otro, capacidades institucionales limitadas para responder de manera ética y responsable (OECD, 2022). Finalmente, las democracias latinoamericanas pueden verse afectadas por prácticas de vigilancia y monitoreo social, en las que los gobiernos contratan servicios de IA (como el reconocimiento facial o el análisis de comportamiento) para el control de la ciudadanía (Ibarreche et al., 2025), así como por el acceso desigual a estas tecnologías, cuyo uso se concentra en sectores con mayores niveles de educación e ingreso.

Todo lo anterior pone de manifiesto los riesgos que enfrenta la región. La IA complejiza un escenario previamente marcado por estructuras de desigualdad, debilidad institucional y dependencia tecnológica. En este sentido, la bibliografía revisada evidencia que la estabilidad de los procesos electorales y la calidad del debate público, lejos de mejorar, se enfrentan actualmente a desafíos más profundos, que transforman de manera sustantiva la forma en que se ha construido la opinión pública.

7. Conclusiones

El presente artículo desarrolló una revisión bibliográfica sobre los principales mecanismos mediante los cuales se incide en la opinión pública, con especial énfasis en las tecnologías de IA. La manipulación de percepciones y posturas a través de técnicas aplicadas a contenidos textuales y audiovisuales entraña riesgos significativos para las democracias. En América Latina, en particular, los procesos deliberativos y de construcción de consensos enfrentan desafíos estructurales, como la desconfianza en las instituciones y la fragilidad de los liderazgos políticos. A estos problemas se suman ahora imágenes y videos generados mediante IA que, por su alto grado de verosimilitud, dificultan la distinción entre lo verdadero y lo falso. Asimismo, los chatbots inciden en la opinión pública en la medida en que permiten la creación de cuentas apócrifas o se utilizan para difundir mensajes políticos altamente personalizados.

Este escenario abre la posibilidad de una opinión pública configurada a partir de la mediación de algoritmos y la preponderancia de grandes corporaciones tecnológicas, las cuales desplazan progresivamente a los medios de comunicación tradicionales y se consolidan como actores centrales en un entorno crecientemente influido por la IA.

Lo anterior pone de relieve la necesidad de profundizar en futuras líneas de investigación sobre el papel de estas tecnologías en las ciencias políticas y sociales. Si bien este trabajo buscó esclarecer los mecanismos que configuran la opinión pública en el contexto actual, aún es necesario explorar con mayor detenimiento cómo los Estados pueden afrontar los desafíos de la IA, no solo desde una perspectiva regulatoria, sino también mediante su apropiación estratégica para mejorar su operación. En la actualidad, se observa un proceso de plataformización en el consumo de bienes culturales, como películas y música, en el que la IA desempeña un papel clave en la gestión y automatización de procesos. En este sentido, resulta pertinente preguntarse si los Estados adoptarán estas tecnologías para optimizar su desempeño, por ejemplo, en la gestión burocrática y administrativa.

De igual forma, cabe cuestionar si la IA podría contribuir a la construcción de democracias más inclusivas. Hoy en día, es técnicamente posible crear avatares digitales que tomen decisiones en nombre de los individuos, lo que implicaría una transferencia significativa de la agencia política en sistemas tecnológicos cuyo funcionamiento

aún presenta zonas de opacidad, como ocurre con las denominadas “cajas negras”. En este marco, resulta fundamental distinguir entre el tecnosolucionismo y el tecnopesimismo; es decir, entre una confianza irrestricta en la tecnología y su rechazo absoluto. La adopción de una postura intermedia, crítica y equilibrada permitiría reconocer tanto los beneficios como los riesgos de la IA en el ámbito político. Así, esta tecnología podría emplearse, por ejemplo, para difundir propuestas electorales, siempre que se eviten usos contrarios a principios éticos.

En última instancia, se recomienda abordar la IA desde un marco de responsabilidad que no busque manipular la opinión pública, sino fortalecerla mediante el acceso amplio a información que permita la formación de criterios propios. La implementación de marcos regulatorios sólidos constituye un primer paso indispensable; sin embargo, también resulta urgente promover una alfabetización en IA que permita a los ciudadanos identificar contenidos falsos o generados artificialmente y distinguirlos de información verídica. De este modo, la opinión pública podrá configurarse a partir del conocimiento de diversas posturas y favorecer la toma de decisiones conscientes.

Agradecimientos

Este artículo pudo escribirse con ayuda de una estancia posdoctoral realizada gracias al Programa de Becas Posdoctorales en la Universidad Nacional Autónoma de México (POSDOC).

Referencias

Allport, F. H. 1940. “Polls and the Science of Public Opinion.” *The Public Opinion Quarterly* 4 (2): 249-257.

Aral, S. 2021. *The hype machine*. Penguin Random House.

Argyle, L. P., Busby, E. C., Gubler, J. R., Lyman, A., Olcott, J., Pond, J., & Wingate, D. 2025. Testing theories of political persuasion using AI. *Proceedings of the National Academy of Sciences*, 122(18). <https://doi.org/10.1073/pnas.2412815122>

Ariza, L., & Ruiz, C. B. 2025. *Inteligencia artificial en el sector público: Enfoques*

teóricos y propuestas para una gestión innovadora para América Latina. *Gobierno y Administración Pública*, 10, 119–134. <https://doi.org/10.29393/GP10-9SPLA20009>

Becker Castellaro, S., Carvalho, M., Fernández Gibaja, A., Grassi, A., Hammar, C., Müller, J., Pereira, L., Piaia, V., & Ruediger, M. A. 2025. Inteligencia artificial e integridad de la información: Experiencias latinoamericanas (Documento de política N.º 34). International IDEA. <https://www.idea.int/publications/catalogue/html/inteligencia-artificial-e-integridad-de-la-informacion-experiencias>

Bickerton, C. J., & Invernizzi Accetti, C. 2021. *Technopopulism: The new logic of democratic politics*. Oxford University Press. <https://doi.org/10.1093/oso/9780198807766.001.0001>

Bryce, J. 1888. *The American Commonwealth*. New York, Macmillan.

Bucher, T. 2018. *If... then: Algorithmic power and politics*. Oxford University Press.

Carrasco-Ahen, C. J., Castillo Arrosquipa, G., Palacios Ticona, G. S., & Sánchez-Córdova, M. 2025. Producción científica sobre inteligencia artificial en América Latina: Un análisis bibliométrico. *EPISTEMUS*, 19(38). <https://doi.org/10.36790/epistemus.v19i38.457>

Castells, M. 2009. *Comunicación y Poder*. Alianza Editorial.

Cave, D., Hoffman, S., Joske, A., Ryan, F., & Thomas, E. 2019. *Mapping China's technology giants*. The Australian Strategic Policy Institute.

Centro de Estudios Estratégicos de Derecho de la Inteligencia Artificial (CEEDIA). 2026. Inicio. <https://ceedia.org/>

CEPAL. 2023. *Inteligencia artificial en América Latina y el Caribe: desafíos, oportunidades y políticas públicas*. <https://repositorio.cepal.org/handle/11362/69210>

Codina, L. 2020. Cómo hacer revisiones bibliográficas tradicionales o sistemáticas utilizando bases de datos académicas. *Revista ORL*, 11(2), 139–153. <https://doi.org/10.14201/orl.22977>

Coeckelbergh, M. 2024. *Why AI undermines democracy and what to do about it*. Polity Press.

Coeckelbergh, M. 2026. *Technofascism: AI, Big Tech, and the new authoritarianism*.

AI & Society. <https://doi.org/10.1007/s00146-026-02862-9>

Cohen, B. 1963. *The press and foreign policy*. Princeton University Press.

Converse, P. E. 1964. "The Nature of Belief Systems in Mass Publics." *Ideology and Discontent*. New York, The Free Press.

Converse, Philip E. 1970. "Attitudes and Non-Attitudes: Continuation of a Dialogue." *The Quantitative Analysis of Social Problems*. Reading, Addison-Wesley.

Couldry, N. 2012. *Media, Society, World: Social Theory and Digital Media Practice*. London, Polity.

Criado, J. I. 2024. Inteligencia artificial en el sector público latinoamericano: Estudio comparado a partir de la Carta Iberoamericana de Inteligencia Artificial en la Administración Pública. *Revista del CLAD Reforma y Democracia*, 88, 116–143. <https://doi.org/10.69733/clad.ryd.n88.a387>

Dean, J. 2005. Communicative capitalism: Circulation and the foreclosure of politics. *Cultural Politics*, 1(1), 51–74. <https://doi.org/10.2752/174321905778054845>

Dewey, John. 1927. *The Public and its Problems*. New York, Holt

Doshi, J. A., Novacic, I., Fletcher, C., Borges, M., Zhong, E., Marino, M., Gan, J., Mager, S., Sprague, D. N., & Xia, M. 2024. Sleeper social bots: A new generation of AI disinformation bots are already a political threat. *arXiv*. <https://doi.org/10.48550/arXiv.2408.12603>

Elmas, T. 2023. Analyzing activity and suspension patterns of Twitter bots attacking Turkish Twitter trends by a longitudinal dataset. *arXiv*. <https://doi.org/10.48550/arXiv.2304.07907>

Estrada Merino, A. 2025. Inteligencia artificial: hacia una gobernanza ética. *Acta Herediana*, 68(2), 33–40. <https://doi.org/10.20453/ah.v68i2.7505>

Farahany, N. A. 2023. *The battle for your brain*. St. Martin's Press.

Feldstein, S. 2019. *The Global Expansion of AI Surveillance*. Carnegie Endowment for International Peace.

Feldstein, S. 2023. The consequences of generative AI for democracy, governance and war. *Survival. Global Politics and Strategy*, 65(5), 117–142. <https://doi.org/10.108>

0/00396338.2023.2261260

Ferrara, E.; Varol, O.; Davis, C.; Menczer, F. y Flammini, A. 2016. "El auge de los bots sociales". *Comunicaciones de la ACM*, 59(7), 96-104.

García-Estévez, N., Ballesteros-Aguayo, L., & Colussi, J. 2025. Desinformación y manipulación de la opinión pública: Una revisión sistemática sobre astroturfing (2004–2024). *Revista de Comunicación*, 24(2), 159–181. <https://doi.org/10.26441/RC24.2-2025-3988>

García-Marzá, D., & Calvo, P. 2022. Democracia algorítmica: ¿un nuevo cambio estructural de la opinión pública? *Isegoría*, 67. <https://doi.org/10.3989/isegoria.2022.67.17>

García-Marzá, D., & Calvo, P. 2024. *Algorithmic democracy: A critical perspective based on deliberative democracy*. Springer.

Goldstein, J. A., Sastry, G., Musser, M., DiResta, R., Gentzel, M., & Sedova, K. 2023. Generative Language Models and Automated Influence Operations: Emerging Threats and Potential Mitigations. <https://doi.org/10.48550/arXiv.2301.04246>

Gorodnichenko, Y.; Pham, T. y Talavera, O. 2021. "Social Media, Sentiment and Public Opinions: Evidence from #Brexit and #USElection". *European Economic Review*, 136. <https://doi.org/10.3386/w24631>

Guanche, J. C., Reyes, N., & Guoliang, L. 2025. Inteligencia artificial (IA) y competencias mediáticas, digitales, informacionales y comunicacionales en América Latina y el Caribe: Una perspectiva ética y crítica en la era algorítmica. *Revista Latinoamericana de Comunicación*, 159, 17–36. <https://doi.org/10.16921/chasqui.v1i159.5219>

Habermas, J. 1998. *Facticidad y validez*. Madrid, Trotta.

Habermas, J. 1987. *Teoría de la acción comunicativa*. Madrid, Taurus.

Hackenburg, K., & Margetts, H. 2024. Evaluating the persuasive influence of political microtargeting with large language models. *Proceedings of the National Academy of Sciences*, 121(24). <https://doi.org/10.1073/pnas.2403116121>

Han, B. C. 2022. *Infocracy: Digitalization and the crisis of democracy*. Polity.

Hao, K. 2025. *Empire of AI. Inside the race for total domination*. Penguin.

Hyska, M. 2025. The politics of past and future: Synthetic media, showing, and telling. *Philosophical Studies*, 182(1), 137–158. <https://doi.org/10.1007/s11098-023-02062-x>

Ibarreche, X., Martínez Elebi, C., & Di Lorenzo, C. 2025. Reconocimiento facial y tecnologías de vigilancia en América Latina: Casos, proveedores y dinámicas comerciales. *AlSur*. <https://www.alsur.lat/sites/default/files/2025-07/Reporte%20reconocimiento%20facial.pdf>

Jung, J., & Katz, R. 2025. Impacto económico de la inteligencia artificial en América Latina. Transformación tecnológica y rezago en materia de inversión y capacidades laborales. CEPAL. <https://repositorio.cepal.org/server/api/core/bitstreams/954c64fe-197c-49f9-b118-0707cc6f8910/content>

Kitchenham, B. 2004. Procedures for performing systematic reviews. Keele University. <http://www.inf.ufsc.br/~aldo.vw/kitchenham.pdf>

Kreiss, D., Lawrence, R. G., & McGregor, S. C. 2023. In their own words: Political practitioner accounts of the democratic implications of artificial intelligence. *New Media & Society*. <https://doi.org/10.1177/146144482311605>

Lagos, A. 2025. Latam-GPT: The free, open source, and collaborative AI of Latin America. *WIRED*. <https://www.wired.com/story/latam-gpt-the-free-open-source-and-collaborative-ai-of-latin-america/>

Larocca, N. 2025. Redes sociales llegan a 86 de cada 100 personas en América Latina. *DPL News*. <https://dplnews.com/redes-sociales-llegan-a-86-de-cada-100-personas-en-america-latina/>

Lippmann, W. 1922. *Public Opinion*. New York, Macmillan.

Lippmann, W. 1925. *Phantom Public. A Sequel to "Public Opinion"*. Nueva York.

Liu, Y., Padmanabhan, B., & Viswanathan, S. 2025. What exactly is a deepfake? *arXiv*. <https://doi.org/10.48550/arXiv.2510.22128>

López del Castillo, F. L. 2024. Inteligencia artificial generativa: determinismo tecnológico o artefacto construido socialmente. *Palabra Clave*, 27(1). <https://doi.org/10.5294/pacla.2024.27.1.9>

López López, P. C., Maldonado, A. M., & Ribeiro, V. 2023. La desinformación en las

democracias de América Latina y la península ibérica. *Uru: Revista de Comunicación y Cultura*. <https://doi.org/10.32719/26312514.2023.8.5>

Lowell, A. L. 1913. *Public Opinion and Popular Government*. New York, Longmans Green.

Marda, V. 2019. *Inteligencia artificial: una perspectiva desde el Sur*. América Latina en Movimiento. <https://alainet.org/es/articulo/203490>

McLeod, S. 2024. *Doing a scoping review: A practical, step-by-step guide*. ResearchGate. <https://doi.org/10.13140/RG.2.2.23703.79528>

Meikle, G. 2023. *Deepfakes*. Polity Press.

Moore, M. 2018. *Democracy hacked: How technology is destabilising global politics*. Oneworld.

Mora-Reyes, B. A., Drewyor, J. A., & Reyes-Angulo, A. A. 2025. *Advancing equitable AI: Evaluating cultural expressiveness in LLMs for Latin American contexts*. arXiv. <https://doi.org/10.48550/arXiv.2511.04090>

Nightingale, S. J., & Farid, H. 2022. *AI-synthesized faces are indistinguishable from real faces and more trustworthy*. *Proceedings of the National Academy of Sciences*, 119(8), e2120481119. <https://doi.org/10.1073/pnas.2120481119>

OECD. 2022. *Uso estratégico y responsable de la inteligencia artificial en el sector público de América Latina y el Caribe*. https://www.oecd.org/es/publications/uso-estrategico-y-responsable-de-la-inteligencia-artificial-en-el-sector-publico-de-america-latina-y-el-caribe_5b189cb4-es/full-report/component-6.html

Okoli, C., & Schabram, K. 2010. *A Guide to Conducting a Systematic Literature Review of Information Systems Research*. *Communications of the Association for Information Systems*, 37(43), 879-910. <http://dx.doi.org/10.2139/ssrn.1954824>

Page, B. I., Robert Y. Shapiro. 1992. *The Rational Public. Fifty Years of Trends in Americans Policy Preferences*. Chicago, The University of Chicago Press.

Pariser, E. 2011. *The filter bubble: how the new personalized web is changing what we read and how we think*. Penguin Books.

Pastor-Galindo, J., Zago, M., Nespoli, P., López Bernal, S., Huertas Celdrán, A., Gil Pérez, M., Ruipérez-Valiente, J. A., Martínez Pérez, G., & Gómez Mármol, F. 2020. Spotting political social bots in Twitter: A use case of the 2019 Spanish general election. arXiv. <https://doi.org/10.48550/arXiv.2004.00931>

Pawelec, M. 2022. Deepfakes and democracy (theory): How synthetic audio-visual media for disinformation and hate speech threaten core democratic functions. *Digital Society*, 1(2), 19. <https://doi.org/10.1007/s44206-022-00010-6>

Raj, P. 2025. The impact of deepfake technology on elections and democratic trust. *Indian Journal of Law*, 3(6), 56–64. <https://doi.org/10.36676/ijl.v3.i6.137>

Ranka, H., Surana, M., Kothari, N., Pariawala, V., Banerjee, P., Surve, A., Sankepally, S. R., Jain, R., Lalwani, J., & Mehta, S. S. 2024. Examining the implications of deepfakes for election integrity. arXiv. <https://doi.org/10.48550/arXiv.2406.14290>

Redacción AN. 2018, junio 20. Los bots y su influencia en la elección presidencial mexicana: Albert-László Barabási. *Aristegui Noticias*. <https://aristeguinoicias.com/2006/mexico/los-bots-y-su-influencia-en-la-eleccion-presidencial-mexicana-albert-laszlo-barabasi/>

Saura García, C. 2025. Synthetification of public opinion: Impacts on deliberative democracies. *Ethics and Information Technology*, 27(4), 1–14. <https://doi.org/10.1007/s10676-025-09870-1>

Schiff, K. J., Schiff, D. S., & Bueno, N. S. 2025. The liar's dividend: Can politicians claim misinformation to evade accountability? *American Political Science Review*, 119(1), 71–90. <https://doi.org/10.1017/S0003055423001454>

Simchon, A., Edwards, M., & Lewandowsky, S. 2024. The persuasive effects of political microtargeting in the age of generative artificial intelligence. *PNAS Nexus*, 3(2). <https://doi.org/10.1093/pnasnexus/pgae035>

Snowden, E. 2019. *Permanent record*. Metropolitan Books.

Splichal, S. 2022. Datafication of public opinion and the public sphere: How extraction replaced expression of opinion. *Anthem*.

Woolley, S. 2023. *Manufacturing consensus: Understanding propaganda in the era of automation and anonymity*. Yale University Press.

Yang, K.-C.; Varol, O.; Davis, C. A.; Ferrara, E.; Flammini, A. y Menczer, F. 2019. "Arming the public with artificial intelligence to counter social bots". *Human Behavior & Emerging Technologies* 1(1), 48-61. <https://doi.org/10.1002/hbe2.115>

Zaller, J. R. 1992. *The Nature and Origins of Mass Opinion*. Cambridge: Cambridge University Press.

Zuboff, S. 2019. *The age of surveillance capitalism: The fight for The future at The new frontier of power*. Profile Books.