

Algoritmos del poder: populismo digital y erosión democrática en el México contemporáneo

Architectures of control: digital populism and the corrosion of democracy in contemporary Mexico

 José Juan Pablo Rojas Ramírez*, Daniel Del Toro Rodríguez**

Artículo recibido: 29-08-25

Artículo aprobado: 10-11-25

Palabras clave:

retroalimentación degenerativa, hiperpersonalización emocional, instituciones de *accountability*, resiliencia democrática, epistemologías paralelas.

Keywords:

degenerative feedback, emotional hyper-personalization, accountability institutions, democratic resilience, parallel epistemologies.

Cómo citar este artículo

Rojas Ramírez, J. J. P. & Del Toro Rodríguez, D. (2025). Algoritmos del poder: Populismo digital y erosión democrática en el México contemporáneo. *Entretextos*, 17(41), 1-23. <https://doi.org/10.59057/iberoleon.20075316.202541792>.

Resumen

En este artículo se examina la reconfiguración de la esfera pública mexicana mediante la instrumentalización de la inteligencia artificial (IA) en estrategias de populismo digital durante el período 2018-2024. Fundamentada en la cibernética política de Deutsch y la cibernética de segundo orden de von Foerster, en la investigación se emplea el *process-tracing method* y la triangulación teórica para analizar cómo los sistemas algorítmicos han reemplazado procesos deliberativos por circuitos de retroalimentación degenerativa que erosionan la autorregulación democrática. Los hallazgos documentan la emergencia de arquitecturas sofisticadas de manipulación que operan mediante hiperpersonalización emocional, sincronización narrativa automatizada y captura de instituciones de *accountability*. El análisis revela la consolidación de un ecosistema de “opinión pública sintética” que fragmenta la deliberación democrática en epistemologías paralelas, mercantilizando el disenso y concentrando el control informativo.

* Doctorante en el programa Doctorado en Ciencia Política, Centro Universitario de la Cienega, Universidad de Guadalajara. Autor para correspondencia. Correo electrónico: jpablo.rojas@cutonala.udg.mx

** Profesor-Investigador de Tiempo Completo, en Centro Universitario de Tonalá, Universidad de Guadalajara. Correo electrónico: daniel.dtoro@alumnos.udg.mx

Abstract

This research examines the reconfiguration of Mexico's public sphere through the instrumentalization of artificial intelligence (AI) in digital populism strategies during 2018–2024. Grounded in Deutsch's political cybernetics and von Foerster's second-order cybernetics, the study employs process tracing and theoretical triangulation to analyze how algorithmic systems have replaced deliberative processes with degenerative feedback circuits that erode democratic self-regulation. Findings document the emergence of sophisticated manipulation architectures operating through emotional hyperpersonalization, automated narrative synchronization, and accountability institution capture. The analysis reveals the consolidation of a "synthetic public opinion" ecosystem that fragments democratic deliberation into parallel epistemologies, commodifying dissent and concentrating informational control. The research concludes that AI transcends its instrumental function to become a structural agent reconfiguring democracy's epistemic foundations, demanding new frameworks for algorithmic transparency and institutional resilience. These dynamics challenge conventional democratic participation models and necessitate urgent rethinking of informational sovereignty in fragile democratic contexts.

Introducción

La instrumentalización de la inteligencia artificial (IA) por actores políticos está redefiniendo los fundamentos de la democracia contemporánea mediante la alteración sistemática de los flujos informativos que sustentan la deliberación pública. Desde la perspectiva de la cibernética política de Karl Deutsch (1963), los sistemas democráticos operan como redes de procesamiento de información cuya estabilidad depende de mecanismos de retroalimentación regulatoria; sin embargo, la IA utilizada por el staff de políticos en campaña introduce circuitos cerrados de polarización que anulan la capacidad de autocorrección democrática, generando lo que la cibernética de segundo orden (von Foerster, 2003) conceptualiza como realidades autorreferenciales donde los algoritmos modifican recursivamente los criterios de validación epistémica.

En América Latina, particularmente en México, este fenómeno adquiere dimensiones críticas debido a la fragilidad institucional que caracteriza a las democracias de la región. La convergencia entre populismo digital e IA ha generado un ecosistema donde la hiperpersonalización emocional, la sincronización de narrativas automatizadas y la producción industrial de contenido sintético erosionan los contrapesos democráticos tradicionales. Como documenta la evidencia empírica analizada en casos como #RedAMLO (2018–2019) y #INECorrupto (2021), las redes de bots coordinados y los algoritmos de *microtargeting* han reconfigurado la esfera pública mexicana al sustituir la deliberación por interacciones binarias mercantilizadas.

La “opinión pública sintética” emergente no constituye meramente un problema técnico, sino una transformación estructural de los mecanismos democráticos. El capitalismo de vigilancia (Zuboff, 2020) monetiza la polarización mediante plataformas que priorizan contenidos virales sobre aquellos verificables, fragmentando el *demos* en microcomunidades autorreferenciales. Esta dinámica se intensifica en contextos de “democracia delegativa” (O’Donnell, 1992), donde la concentración presidencial del poder facilita la captura algorítmica de instituciones autónomas.

Planteamiento del problema

El populismo digital se puede definir como la instrumentación de estrategias de comunicación política en plataformas digitales, como pueden ser las redes sociales, con el fin de manejar la opinión pública, generar una opinión dominante desde la dirección de un actor o grupo político hacia el posicionamiento de sus plataformas ideológico-políticas en la consciencia de los usuarios (ciudadanos) de redes sociales, quienes interactúan con perfiles generados *ex profeso* por dichos grupos o actores políticos, en cuyo caso son creados a partir de la contratación de “community managers”, empresas de consultoría o agentes especializados en el diseño computacional de agentes “robot” o entrenamiento de inteligencias artificiales, ya sean generativas o agénticas, que tienen por propósito interactuar con los usuarios de redes sociales para persuadir y manipular el clima de opinión pública sobre temas políticos y apelando a cuestiones emocionales o temas sensibles en el electorado (Rosanvallon, 2020; Baldwin-Philippi, 2018; Mazzoleni & Bracciale, 2018). Dicha situación constituye un escenario emergente complejo que desafía la estabilidad democrática en contextos de fragilidad institucional, especialmente en América Latina. Este fenómeno se caracteriza por la instrumentalización de sistemas algorítmicos para manipular narrativas públicas y erosionar contrapesos institucionales, reconfigurando las dinámicas de poder mediante mecanismos de retroalimentación degenerativa y centralización algorítmica (Deutsch, 1963; Harari, 2024). La convergencia entre desintermediación tecnológica y estrategias populistas ha generado un ecosistema donde la “democratización digital” (Rosanvallon, 2020) sustituye la deliberación pública por interacciones binarias, fragmentando la esfera pública en cámaras de eco autorreferenciales.

Como señalan Innerarity y Colomina (2020), la red ofrece un espacio descentralizado que amplía el debate político, aunque conviven problemas como la descontextualización, la trivialización y la captura de datos personales para uso económico y político. El núcleo problemático reside en cómo la IA, operando como transductor cibernético, distorsiona los flujos informativos que sustentan la homeóstasis democrática mediante tres niveles: hiperpersonalización de mensajes por *microtargeting* emocional, creación de realidades sintéticas mediante *deepfakes* (Saura García y Calvo, 2024) y sustitución de instituciones por plataformas que centralizan el control narrativo (Zuboff, 2020).

En democracias frágiles, donde persisten asimetrías en alfabetización digital y resiliencia institucional, estos mecanismos amplifican la vulnerabilidad ante la manipulación algorítmica, transformando la participación ciudadana en una ilusión de agencia mientras se consolida un autoritarismo adaptativo (Wolfgang & Lührmann, 2021). Teóricamente, este fenómeno exige un enfoque interdisciplinario que integre cibernética política y estudios críticos de gobernanza algorítmica, conceptualizando los sistemas democráticos como redes de procesamiento informativo cuya estabilidad depende de mecanismos de *feedback* regulatorio.

El presente trabajo parte de la pregunta: ¿de qué manera el uso estratégico de la inteligencia artificial, por parte de actores políticos, transforma los circuitos de opinión pública y propicia el debilitamiento institucional en contextos democráticos frágiles? La hipótesis causal aquí planteada es que la implementación deliberada de tecnologías algorítmicas —particularmente aquellas que permiten la microsegmentación de mensajes y la automatización de la interacción política— induce una retroalimentación degenerativa en los flujos informativos, lo que conduce a la formación de burbujas de desinformación, erosiona los canales tradicionales de rendición de cuentas y debilita la capacidad de las instituciones democráticas para corregir desviaciones y responder eficazmente a los desafíos sociales. De este modo, el empleo instrumental de sistemas inteligentes por líderes o grupos de poder no sólo acelera la polarización y la manipulación emocional de las audiencias, sino que genera una dinámica de captura institucional que socava, tanto la legitimidad como la resiliencia estructural de la democracia misma, perpetuando condiciones favorables al control narrativo, la desconfianza ciudadana y la pérdida progresiva de contrapesos democráticos fundamentales.

Durante la trayectoria política de Andrés Manuel López Obrador, desde su campaña presidencial de 2018 hasta el cierre de su administración en 2024, México experimentó una transformación significativa en el uso de tecnologías digitales para influir en la opinión pública. Esta mutación no constituye un fenómeno aislado ni espontáneo, sino un proceso sistemático documentado por múltiples fuentes hemerográficas y académicas que evidencian la instrumentalización de inteligencia artificial y automatización para reconfigurar la deliberación democrática.

Ya en las semanas previas a las elecciones de julio de 2018, organismos internacionales alertaban sobre la proliferación masiva de cuentas automatizadas. La BBC reportó que investigadores de la Universidad de Oxford identificaron la existencia de “granjas de bots” operando en todas las campañas, aunque con particular intensidad en favor del entonces oficialista Partido Revolucionario Institucional (PRI) (Martínez, 2018); sin embargo, consultoras especializadas, como Metrics, documentaron que operaban aproximadamente ocho millones de cuentas falsas en siete redes sociales durante ese proceso electoral, posicionando hasta cien temas distintos mediante actividad coordinada (SinEmbargo, 2018).

La campaña de López Obrador no quedó exenta de estas dinámicas. Un análisis posterior del Signa Lab del Instituto Tecnológico y de Estudios Superiores de Occidente (ITESO) sobre el *hashtag* #RedAMLO identificó una red de 8 586 nodos con más de 26 000 interacciones, donde al menos dos de cada cinco nodos estratégicos mostraban probabilidades altas de comportamiento automatizado, según herramientas como Botometer (Olmedo-Neri, 2021). Esta estructura operaba bajo una lógica de “maestro de ceremonias y coro”, donde cuentas generadoras de contenido definían narrativas que posteriormente eran amplificadas de manera masiva por perfiles con comportamiento semiautomático o completamente robotizado (Ramírez, 2019).

Durante el primer año de gobierno, la estrategia se consolidó institucionalmente. El informe “Democracia, libertad de expresión y esfera digital”, del Signa Lab, documentó cómo la #RedAMLOVE coordinaba ataques sistemáticos contra periodistas y medios críticos mediante cuatro tipos de cuentas: maestros de ceremonia que generaban contenidos, bots y semibots que replicaban masivamente, troles que hostigaban y usuarios genuinos que amplificaban orgánicamente (Ojeda de la Torre, 2019). Una cuenta identificada en este ecosistema registraba promedios de 8 500 tuits semanales, volúmenes imposibles de sostener mediante actividad humana convencional (Ramírez, 2019).

El propio presidente reconoció indirectamente esta realidad cuando, en mayo de 2020, afirmó que Twitter había admitido la existencia de bots en su plataforma, interpretándolo como evidencia de “mal uso de las redes sociales” por grupos con recursos para pagar la automatización (Domínguez, 2020). Paradójicamente, mientras el gobierno denunciaba bots opositores, reportes como los de El Economista documentaban que seguidores “espontáneos” del mandatario utilizaban programas robotizados para generar tendencias y atacar críticos, con cuentas creadas apenas 30 o 60 días antes, pero acumulando más de 20 000 tuits (Badillo, 2019).

Para las elecciones intermedias de 2021, la sofisticación tecnológica había escalado de manera considerable. Aunque persisten limitaciones en la cuantificación precisa, reportes periodísticos señalaron campañas masivas con *hashtags* como #INECorrupto, que buscaban deslegitimar al organismo electoral mediante una saturación coordinada (Infobae, 2021). Esta estrategia representaba una evolución cualitativa: ya no se trataba solamente de amplificar mensajes favorables, sino de erosionar sistemáticamente la confianza en instituciones autónomas mediante la desinformación automatizada.

El uso de inteligencia artificial generativa marcó un punto de inflexión en 2024. En abril de ese año, el presidente alertó públicamente sobre la circulación de *deepfakes* que utilizaban su imagen y su voz manipuladas para promover estafas de inversión en Pemex (Aristegui Noticias, 2024; Presidencia de México, 2024). Reconoció que aproximadamente el 70% de los mexicanos no logra distinguir contenidos generados mediante IA

de materiales auténticos, evidenciando una vulnerabilidad estructural en la alfabetización digital ciudadana (Sosa Santiago, 2024). Esta tecnología, que permite la síntesis hiperrealista de audio y video, representa un salto cualitativo respecto de los bots textuales: genera contenidos prácticamente indistinguibles de la realidad para audiencias no especializadas.

Paralelamente, el gobierno implementó respuestas institucionales con componentes de IA. El Instituto Nacional Electoral (INE) desarrolló el chatbot “Inés” para WhatsApp, utilizando inteligencia artificial semántica en colaboración con Meedan y AFP Factual para verificar información electoral en tiempo real (INE, 2024); sin embargo, esta iniciativa llegó tardíamente y con alcances limitados frente a la escala de desinformación ya consolidada en el ecosistema digital.

Este fenómeno distorsiona flujos informativos, sustituyendo la retroalimentación democrática por circuitos cerrados de polarización (Deutsch, 1963), mientras la hiperpersonalización emocional explota vulnerabilidades psicológicas ciudadanas, generando burbujas informativas que refuerzan la desconfianza institucional (Harari, 2024).

Aspectos teóricos: desde la teoría cibernética a la opinión pública sintética

La teoría cibernética en ciencia política, desarrollada inicialmente por Karl Deutsch (1963), concibe los sistemas políticos como redes de procesamiento de información donde la comunicación y los mecanismos de retroalimentación regulan la estabilidad institucional. Desde esta perspectiva, el gobierno opera como un sistema adaptativo que mantiene su equilibrio (homeóstasis) mediante flujos informativos entre instituciones y ciudadanos (Deutsch, 1963, p. 77). La cibernética de segundo orden (von Foerster, 2003) amplía este enfoque al destacar cómo las tecnologías digitales modifican recursivamente los patrones de interacción política, creando realidades autorreferenciales donde los algoritmos no sólo procesan datos, sino que reconfiguran los criterios mismos de la toma de decisiones (Innerarity y Colomina, 2020). Ejemplos prácticos pueden observarse al ingresar a una red social como Twitter (hoy X), en donde la “creación” de una opinión dominante a partir del uso de empresas o corporativos que, anterior a la IA generativa y depurativa, implicaban en el proceso de participación cientos de dispositivos manejados por individuos que generaban códigos para la interacción mediante “chatbots”, hoy en día autogenerados por IA de lenguaje: Large Language Models (LLM), para defender, posicionar o derrocar a un personaje o un segmento partidista político.

Desde esta perspectiva cibernética resulta posible comprender fenómenos contemporáneos como la hiperpersonalización algorítmica de mensajes, mediante la cual sistemas automatizados fragmentan la esfera pública en burbujas informativas que obstruyen

la deliberación democrática (Harari, 2024), así como la “entropía dirigida” que instrumentaliza el caos informativo como mecanismo de control social (Wolfgang & Lührmann, 2021, p. 873). Bajo este orden de ideas se pretende entender cómo la concentración algorítmica del poder narrativo transforma las democracias en sistemas cibernéticos disfuncionales donde los mecanismos de autocorrección quedan anulados. Mientras que la retroalimentación negativa opera normalmente para estabilizar sistemas mediante la reducción de desviaciones respecto a estados deseados, la retroalimentación positiva amplifica perturbaciones iniciales, alejando progresivamente al sistema de su equilibrio (Ashby, 1956, pp. 206–207).

En contextos democráticos colonizados por la inteligencia artificial, prevalece precisamente esta dinámica degenerativa: la polarización se autorrefuerza mediante algoritmos que priorizan contenidos emocionalmente extremos, destruyendo la capacidad regulatoria que caracteriza a sistemas políticos saludables. El surgimiento de la IA intensificó el proceso cibernético y casi indistinguible entre humano y bot de información difuminada en los medios masivos que, a su vez, es recogida por los medios de comunicación tradicionales.

El enfoque cibernético permite visibilizar cómo la IA actúa como un transductor algorítmico, introduciendo mecanismos de retroalimentación degenerativa: los circuitos de *fake news*, amplificados mediante cuentas automatizadas y *deepfakes*, aceleran el cierre epistémico y la consolidación de epistemologías paralelas, fragmentando la esfera pública en microcomunidades autorreferenciales. Este proceso erosiona la capacidad deliberativa del sistema democrático, en tanto que bloquea los flujos críticos de información y dificulta la autocorrección institucional esperada en sistemas políticos abiertos.

Este marco cibernético permite comprender cómo operó en México la dinámica concreta de retroalimentación degenerativa durante el periodo analizado. Frente a las campañas coordinadas de desinformación que buscaban erosionar la confianza institucional, el Tribunal Electoral del Poder Judicial de la Federación (TEPJF) estableció, entre 2021 y 2024, mecanismos inéditos de colaboración con plataformas digitales como TikTok, mediante memorándums de entendimiento orientados a combatir la manipulación automatizada y difundir información electoral verificada (TEPJF, 2023). Simultáneamente, el Instituto Nacional Electoral desarrolló sistemas de denuncia ciudadana, protocolos de fiscalización para contenidos en redes sociales y el chatbot “Inés” en WhatsApp, para verificación en tiempo real (INE, 2024). No obstante, estas respuestas institucionales enfrentaron limitaciones estructurales críticas: mientras los algoritmos operaban con velocidades de propagación que convertían *hashtags* en *trending topics* en cuestión de horas, los mecanismos judiciales requerían procedimientos deliberativos incompatibles con esas dinámicas temporales. La magnitud del fenómeno superó sistemáticamente las capacidades regulatorias disponibles, confirmando así la hipótesis de Ashby (1956)

sobre los sistemas que pierden su estabilidad cuando la retroalimentación correctora resulta insuficiente frente a perturbaciones amplificadas de manera exponencial.

El populismo y la tecnocracia digital de llegada a las democracias frágiles

En el caso mexicano persiste una inercia histórica de instituciones heredadas de regímenes autoritarios que, aunque han transitado formalmente a sistemas representativos, mantienen enclaves de poder poco transparentes y prácticas clientelares o corporativas que dificultan el pleno desarrollo de una cultura cívica democrática. A este fenómeno se suman la corrupción sistémica, la violencia política —expresada en asesinatos de candidatos, represión de movimientos sociales y debilidad del Estado de derecho— y la precarización de mecanismos de rendición de cuentas, lo que genera escepticismo social y merma la legitimidad de los procesos electorales y de gobierno. Los estudios revisados señalan que la falta de reformas profundas en los sistemas de justicia, seguridad y partidos, junto con la captura de instituciones autónomas por intereses gubernamentales o privados, han consolidado una democracia que, si bien subsiste formalmente, es vulnerable ante dinámicas de regresión o captura autoritaria (Coss & Cronin, 2024; Whitehead, 2021).

La intersección entre populismo digital y democracias frágiles puede comprenderse desde la teoría cibernética de Deutsch (1963), quien conceptualiza los sistemas políticos como redes de comunicación autorreguladas. En este marco, la IA opera como un transductor que distorsiona los flujos informativos al sustituir los mecanismos de retroalimentación democrática por circuitos cerrados de polarización. Este fenómeno, denominado “retroalimentación degenerativa” (Harari, 2024), explica cómo los algoritmos no sólo amplifican narrativas maniqueas, sino que reconfiguran recursivamente los criterios de validación epistémica, generando realidades sintéticas autorreferenciales (von Foerster, 2003). En América Latina, esta dinámica se manifiesta en la sustitución de instituciones mediadoras por arquitecturas algorítmicas centralizadas que explotan brechas de subciudadanía digital (De Souza, 2018), creando un ecosistema donde la percepción de crisis se cultiva mediante *microtargeting* emocional y desinformación automatizada (Woolley, 2023).

La transición hacia una “democradura digital” (Rosanvallon, 2020) se sustenta en la capacidad de las plataformas para generar entropía dirigida: un caos informativo calculado que mantiene patrones estables de dominación. Como señalan Innerarity y Colomina (2020), la posverdad no radica en la mera falsedad, sino en la construcción de epistemologías paralelas mediante algoritmos que privilegian contenidos virales sobre verificables. Este proceso, analizado como “ciberbalcanización” (Van Alstyne & Brynjolfsson, 1997), fragmenta el *demos* en microcomunidades autorreferenciales donde impera la “ley de la variedad requerida” de Ashby (1956): los sistemas populistas absorben la complejidad

social mediante simplificaciones binarias (#EllosVsNosotros), anulando la capacidad deliberativa mediante saturación de estímulos emocionales.

La convergencia del populismo y la tecnología digital se analiza aquí como un fenómeno de transducción cibernética. Los líderes y movimientos populistas actúan como transductores que instrumentalizan la IA para convertir el descontento social latente en capital político digital. En las democracias frágiles, caracterizadas por una débil intermediación institucional y asimetrías en la alfabetización digital (Levitsky & Ziblatt, 2018), este proceso es particularmente corrosivo. La IA permite la hiperpersonalización de narrativas a través del *microtargeting* emocional, explotando vulnerabilidades cognitivas para construir y reforzar una visión maniquea del mundo (#EllosVsNosotros) que resuena con segmentos específicos del electorado (Tucker *et al.*, 2018).

Frente a esta colonización algorítmica, la teoría de la resiliencia democrática propone contrapesos basados en educación crítica digital y auditorías algorítmicas obligatorias (Innerarity & Colomina, 2020). No obstante, como advierte Saura García y Calvo (2024), la “opinión pública sintética” trasciende la mera manipulación informativa: constituye un sistema homeostático en el que el disenso y las entradas de “compartir” y “me gusta” devienen como núcleo central de lo que debiese ser un sistema de intercambio comunicacional, mediante *deep-synthetic* personalizados. Esto exige repensar la gobernanza tecnológica desde principios cibernéticos de transparencia recursiva, donde la regulación *cross-media* y la cooperación internacional eviten asimetrías en el control narrativo (Harari, 2024). La paradoja radica en que la misma IA que amenaza la democracia podría instrumentarse para fortalecerla, siempre que se priorice la soberanía algorítmica ciudadana sobre los intereses del capitalismo de vigilancia (Zuboff, 2020).

La paradoja radica en que, mientras se promueve una ilusión de participación, se consolida un “individualismo digital” que sustituye la acción comunicativa por interacciones binarias (like/share), mercantilizando el disenso y debilitando la resiliencia democrática (Habermas, 2023). Frente a esto, Urbinati (2014) y Rosanvallon (2020) proponen modelos de “democracia interactiva” que combinen auditorías algorítmicas con educación crítica digital, aunque su implementación exige superar asimetrías en el control narrativo y la gobernanza tecnológica transnacional.

En un orden de ideas similar, pero que alude a la democracia y sus posibles repercusiones ante la IA, el trabajo de Jungherr (2023) sobre el marco conceptual de ésta y la democracia ofrece una perspectiva crucial para entender cómo la IA puede influir en la formación de la opinión pública y la legitimidad democrática. En este contexto, la IA opera como un transductor cibernético que modifica los flujos informativos, generando realidades autorreferenciales que distorsionan la deliberación democrática (von Foerster, 2003). La capacidad de la IA para procesar y personalizar información a gran escala permite a los actores populistas crear narrativas hiperpersonalizadas que explotan las

emociones y prejuicios de diferentes segmentos de la población, lo que se alinea con la teoría cibernética de sistemas autorregulados (Deutsch, 1963). Este fenómeno se agudiza en democracias frágiles, donde la ausencia de contrapesos institucionales sólidos para la regulación del uso de IA, ya sea en comicios o en la evaluación del desempeño de gobernantes y funcionarios, propicia el escenario idóneo para la erosión democrática mediante la desinformación automatizada (Arellano Ríos, 2024).

Por otra parte, el análisis de Ruiz & Buttner (2024) sobre las noticias falsas y su impacto en la formación de la opinión pública, junto con el estudio de Zuiderwijk *et al.* (2021) sobre las implicaciones de la IA en la gobernanza pública, complementan el marco teórico al destacar cómo ésta puede socavar la legitimidad de las decisiones públicas. La difusión masiva de desinformación mediante algoritmos no sólo distorsiona la percepción de la realidad, sino que también erosiona la confianza en las instituciones democráticas, creando un entorno propicio para la manipulación política (Rosanvallon, 2020). Además, la centralización del poder algorítmico en manos de unos pocos actores tecnológicos puede reforzar la concentración de poder político, lo que plantea desafíos significativos para la resiliencia democrática en contextos de debilidad institucional (Zuboff, 2020). Estos estudios subrayan la necesidad de desarrollar marcos regulatorios que aborden la transparencia algorítmica y la educación crítica digital para contrarrestar los efectos negativos de la IA en la democracia.

La manipulación algorítmica de la opinión pública en democracias frágiles constituye un fenómeno multidimensional donde la IA cataliza estrategias populistas. Harari (2024) y Arellano Ríos (2024) evidencian cómo líderes populistas explotan vulnerabilidades institucionales mediante *microtargeting* emocional y desinformación automatizada, generando narrativas personalizadas que polarizan la sociedad. Este proceso se intensifica en América Latina, donde la debilidad institucional y la brecha digital facilitan la erosión del *accountability* horizontal (Nieto y Milanese, 2016).

La transición del populismo tradicional al digital sustituye intermediarios humanos por arquitecturas algorítmicas centralizadas que producen realidades sintéticas autorreferenciales (Saura García y Calvo, 2024). Estas dinámicas consolidan un “autoritarismo adaptativo” mediante vigilancia masiva y dataficación de interacciones ciudadanas (Zuboff, 2020).

La emergencia de una “opinión pública sintética” redefine el control político, combinando intoxicación informativa y colonización algorítmica (Saura García y Calvo, 2024). Plataformas como X, TikTok y Meta operan como panópticos digitales, priorizando contenidos virales sobre verificables, fragmentando el *demos* en microcomunidades (Innerarity & Colomina, 2020). Esta “ciberbalcanización” socava la resiliencia democrática, convirtiendo la participación ciudadana en transacciones binarias que mercantilizan el disenso. El “capitalismo de vigilancia” (Zuboff, 2020) monetiza la polarización y convierte la experiencia humana en datos comportamentales, debilitando la acción comunicativa colectiva.

Proceso metodológico

La presente investigación emplea una estrategia metodológica cualitativa de interpretación estructural, articulando el *process-tracing method* (Beach & Pedersen, 2019), el cual permitió el seguimiento algorítmico y la triangulación teórica multiescalar para analizar la manipulación de la opinión pública mediante IA en democracias frágiles; el análisis de estudios, tanto académicos como hemerográficos, y contrastados con la observación participativa con perfiles creados *exprofeso*, aun en actividad y desde el 2 de julio de 2018, donde el perfil en Twitter (X) viró hacia una carta abierta escrita por Denisse Dresser en esta red, las participaciones, desde entonces, se rediseñaron con la finalidad de ser afines a las partes en conflicto, según el clima de opinión que “iba y venía” entre soporte y rechazo de opiniones de grupos denominados “conservadores fifí” y “liberales chairo”, con el fin de rastrear la persistencias de determinados perfiles en función de la defensa de las plataformas políticas y *hashtag* afines.

Se verificaron ante Botometer X (<https://botometer.osome.iu.edu>) perfiles que presentaban resonancia en *replies* a partir de 1.1k, lo que se infiere como una amplificación automatizada (cifra arbitraria, pero que evidenciaba la sospechosa participación de humanos en tan corto tiempo, más cuando el mensaje era una difamación sin ningún trasfondo o sustento). En este sitio, la consideración para determinar un bot es a partir de 3/5 *bot score*, en un rango de colores que va de azul a rojo, en tanto su diseminación en las redes.

Botometer X

FAQ API Publications Bot Repo

Botometer X

An OSOME project (bot-o-meter)

Botometer X (formerly Botometer) is currently in archival mode, and the results were pre-calculated based on historical data collected before May 31, 2023. Botometer X will not return any results for accounts created after the data cutoff. For accounts created before the data cutoff, Botometer X only has the records for some of them. If these accounts are queried, Botometer X will return the bot scores and corresponding timestamps. The scores are numbers between 0 and 5, calculated using the BotometerLite model. Higher scores mean more bot-like activity. For details, please refer to the [FAQ](#) page.

You can also fetch the bot scores in batch through the API. For details, please refer to our [API](#) page.

Search by username or id:

@MrLotario

User id: 1174488517977825285

Bot Score: 2.8/5

Date calculated: 5/31/2023

Copyright © 2024 the Trustees of Indiana University

OSOME · CNetS · Luddy

Figura 1. Comprobación de perfil en Twitter de usuario anti López Obrador.

Fuente: Botometer (2025).

En esta fase de la investigación se consideraron sólo aspectos cualitativos, tales como naturaleza bot o humana de los perfiles, o mensajes y su diseminación por otros perfiles repetidores. El criterio establecido fue verificar en Botometer los perfiles que tenían más de 1.1k *replies* en Twitter. Se inició con la detección de un usuario humano reconocido, ya sea por su actividad periodística o por su exposición como “influencer político o económico”, tanto en la red anterior como en Facebook y TikTok, que emitían un mensaje, el cual era capitalizado por otros perfiles que lograban dicha cifra de *replies*, en ocasiones rebasando el *replie* del autor original y reconocido como humano.

No se llevó a cabo un estudio estadístico descriptivo ni inferencial, salvo el ejercicio de detección de algoritmos mediante IA, sólo la triangulación para la verificación y comprobación de que los mensajes populistas que aluden a sentimientos y sensibilidades coyunturales, y que tienen rasgos de hiperpersonalización son frecuentemente generados de manera sintética. Cuando son instrumentados por chatbots logran tener una resonancia en corto tiempo y posicionarse en la percepción de los humanos en cuanto a criterios de aprobación o desaprobación de determinada actuación política; sin embargo, bajo estas técnicas y métodos es posible llevarlos a cabo para futuras investigaciones en esta línea de análisis. Este diseño metodológico se fundamenta en un marco teórico que concibe las democracias como sistemas cibernéticos autorregulados (Deutsch, 1963), cuya estabilidad institucional depende de flujos informacionales funcionales y bidireccionales. La IA, desde esta perspectiva, actúa como un transductor disruptivo que altera las condiciones de circulación y validez epistémica de la información, generando circuitos cerrados de retroalimentación degenerativa y desinstitucionalizando la deliberación pública (von Foerster, 2003).

El *process-tracing*, en su vertiente explicativa (Beach & Pedersen, 2019), se utiliza para reconstruir las cadenas causales que conectan los niveles macro (estructural), meso (institucional) y micro (interaccional) en el caso mexicano, documentando cómo se producen, diseminan y consolidan las narrativas algorítmicamente manipuladas. Esta técnica intensiva requiere la identificación de evidencia diagnóstica, desde rastros digitales (como conjuntos de tweets, campañas en plataformas o registros de bots) ya realizados, tanto por laboratorios académicos como estudios periodísticos referentes a la temática de uso de granjas de bots en el sexenio pasado, principalmente.

De manera complementaria, se rastrearon las trayectorias de polarización digital inducida algorítmicamente por medio de *likes* mediante perfiles de Twitter y TikTok creados afines a cada postura, para documentar cómo los usuarios son conducidos por la arquitectura de recomendaciones hacia posturas extremas o autorreferenciales (Jungherr, 2023).

A ello se suma el diseño de una matriz de nodos informativos críticos, donde la ampliación narrativa se vincula con la deslegitimación institucional, el uso de *deepfakes* y

la manipulación de emociones específicas mediante las plataformas IA: Perplexity.ai y Quillbot.com. En esta última plataforma el trabajo de detección es directo cuando se cuenta con más de 250 caracteres, lo cual permite la detección certera y comprobación de cadenas de IAG. En todo caso, las IA requieren ser preentrenadas con prompts de simulación con el inicio de la tarea. En Perplexity.ai se propuso el siguiente prompt inicial, cuando la probabilidad de que un perfil tuviera 60% de ser “bot”. Prompt 1:

Eres un experto en la determinación de parámetros probabilísticos a través de un análisis fundamentado en estadística inferencial, en específico a través de métodos bayesianos, y que reúna los criterios estadísticos para la detección primero, de generación de texto mediante inteligencia artificial. Posteriormente, analiza los resultados procedentes del texto que te proveeré y evalúa, a través de términos bayesianos, la probabilidad de que este enunciado de este perfil sea humano o bot.

Prompt 2: “Necesito analizar un conjunto de tweets, comentarios en redes sociales o conversaciones en WhatsApp recolectados durante campañas electorales mexicanas (2018-2024) para identificar probable participación de bots, cuentas automatizadas o contenido generado por IA que podrían distorsionar la opinión pública”.

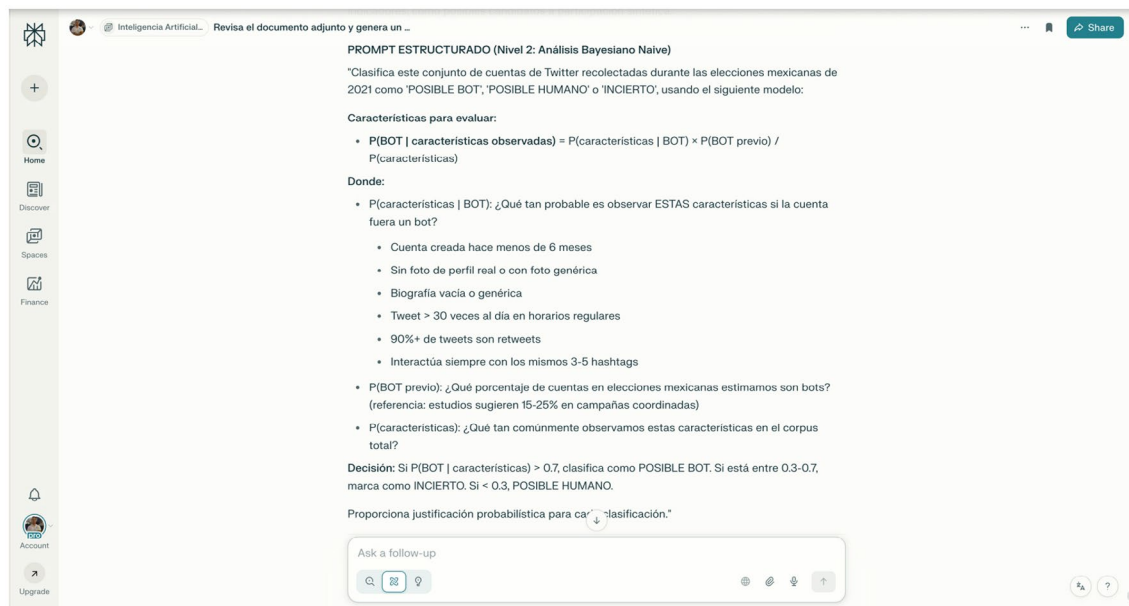


Figura 2. “Clasifica este conjunto de cuentas de Twitter recolectadas durante las elecciones mexicanas de 2021 como ‘POSSIBLE BOT’, ‘POSSIBLE HUMANO’ o ‘INCIERTO’.

Fuente: Elaboración propia con base en Perplexity.ia (2025).

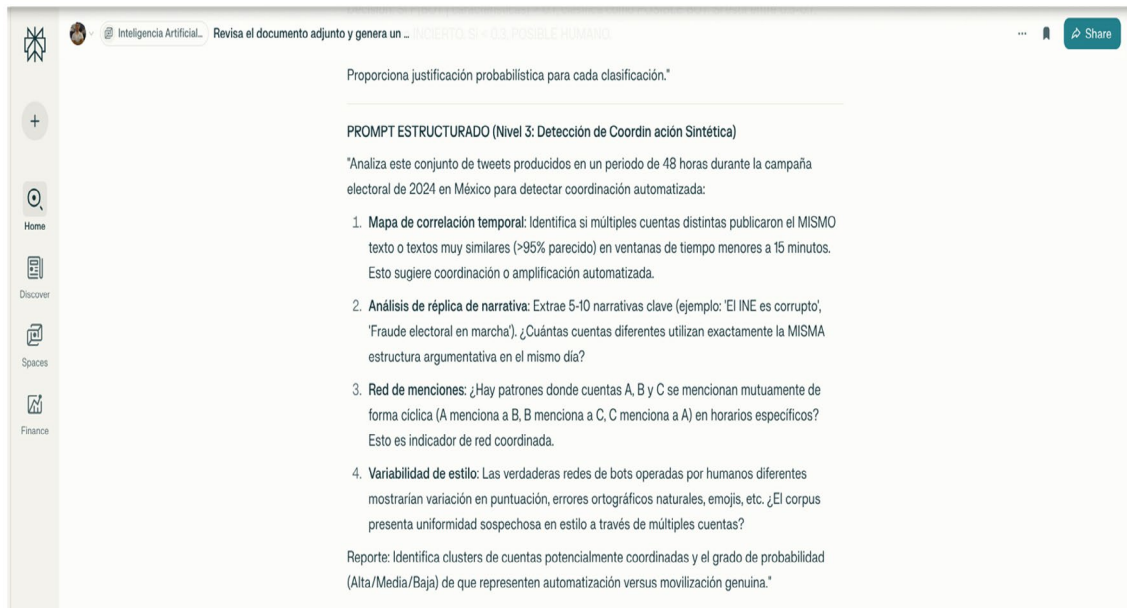


Figura 3. “Analiza este conjunto de tweets producidos en un período de 48 horas durante la campaña electoral de 2024 en México para detectar coordinación automatizada”
Fuente: Elaboración propia con base en Perplexity.ia (2025).

La triangulación teórica se construye a partir de cuatro marcos analíticos convergentes: 1) la cibernética política de Deutsch (1963) para modelar los flujos de información y evaluar la disfuncionalidad del sistema democrático; 2) la cibernética de segundo orden (von Foerster, 2003) para interpretar cómo los algoritmos generan entornos autorreferenciales que redefinen el criterio de verdad; 3) la teoría crítica del capitalismo de vigilancia (Zuboff, 2020), que explica la monetización de la polarización, y 4) los estudios sobre la opinión pública sintética que exploran la creación de realidades políticas artificiales, pero verosímiles.

Tabla 1. Matriz de rastreo de procesos (*Process Tracing*).

Nivel analítico	Variable causal	Mecanismo operativo	Evidencia empírica (trazabilidad)	Fuente(s) de triangulación de información principal
Macro	Competencia electoral cerrada en sistemas multipartidistas	Incentivo a la adopción de IA como ventaja estratégica	Contratos de servicios de microtargeting y bots; estrategias de modulación narrativa	Signa Lab ITESO (2019)
Macro	Brecha digital y subciudadanía informacional	Asimetría algorítmica y exclusión narrativa	Concentración del tráfico político en cuentas partidistas automatizadas	Olmedo-Neri (2021)
Meso	Erosión de órganos autónomos y <i>accountability</i>	Intermediarios algorítmicos y recortes presupuestales	Vinculación entre reformas institucionales y campañas de desacreditación	INE (2024)
Meso	Captura algorítmica por élites tecnopolíticas	Cooptación discursiva de plataformas en beneficio político	Alianzas gubernamentales con plataformas; uso orientado de tendencias	Zuboff (2020)

Micro	Personalización emocional de contenido político	Algoritmos de <i>deep learning</i> y aprendizaje emocional reforzado	Correlación entre tipo de emoción y viralización	Presidencia de México (2024)
Micro	Coordinación automática de narrativas (bots y LLMs)	Diseminación secuencial y reactiva de contenido político	Sincronía en hashtags; bajo índice humano en picos de trending topics	Olmedo-Neri, (2021); Signa Lab ITESO (2019)

Fuente: Elaboración propia con base en Perplexity.ai, en el cribado de información de fuentes académicas, hemerográficas y digitales de 2019 a 2024 (2025).

Resultados y discusión

Los datos sistematizados mediante análisis de estudios académicos especializados, reportes de laboratorios digitales y documentación hemerográfica sobre el uso de IA en la comunicación política mexicana documentada, para la validación sobre el uso de IA cada vez más humana, susceptible de ser no identificada como bots, pueden y están influyendo en la opinión pública de las personas. El análisis revela un patrón sostenido de intervención algorítmica en la esfera pública mexicana desde la campaña presidencial de Andrés Manuel López Obrador y hasta el cierre de su administración. Los hallazgos evidencian una evolución tecnológica que trasciende las concepciones tradicionales de manipulación mediática, configurando un ecosistema donde los circuitos de retroalimentación democrática descritos por Deutsch (1963) han sido sustituidos por arquitecturas de amplificación controlada mediante la IA.

La evidencia documenta tres fases diferenciadas en la instrumentalización política de la IA: una primera etapa (2018–2019) caracterizada por redes básicas de bots coordinados; una segunda fase (2020–2022) que incorpora microsegmentación psicográfica y *hashtag storms* programados, y una tercera generación (2023–2024) que integra *deepfakes*, chatbots institucionales y sistemas de IA conversacional indetectable. Esta progresión confirma la hipótesis teórica sobre la retroalimentación degenerativa (von Foerster, 2003), donde los algoritmos no sólo procesan información, sino que redefinen las condiciones epistémicas de su validación, generando realidades políticas sintéticas, pero verosímiles.

El ecosistema identificado refleja una implementación multiescalar de tecnologías de automatización que operan simultáneamente en los niveles macro (competencia electoral), meso (captura institucional) y micro (hiperpersonalización emocional). Twitter (X) emergió como el epicentro de estas operaciones, funcionando como laboratorio para experimentar estrategias que posteriormente se expandirían hacia otras plataformas. La documentación académica evidencia que esta colonización algorítmica no constituye un fenómeno aislado, sino un proceso sistemático de ingeniería social que explota las vulnerabilidades estructurales de la democracia mexicana (von Foerster, 2003).

Las estrategias implementadas trascienden la simple amplificación automatizada para incorporar elementos de lo que Zuboff (2020) denomina “capitalismo de vigilancia”: la mercantilización de datos emocionales para fines políticos. Los hallazgos sugieren que las campañas digitales no sólo buscan influir en preferencias electorales, sino reconfigurar permanentemente los parámetros de validación epistémica en la esfera pública. Esta reconfiguración opera mediante la sustitución de criterios de veracidad por métricas de *engagement*, generando lo que la cibernética de segundo orden identifica como “epistemologías sintéticas”, donde la verdad se determina de forma algorítmica.

La respuesta institucional, ejemplificada en herramientas como el chatbot “Inés”, del INE, revela la magnitud del desafío enfrentado por las instituciones democráticas. Mientras las estrategias de manipulación operan con velocidades de procesamiento y escalas de distribución propias de la IA, los mecanismos de verificación y control permanecen anclados en paradigmas analógicos de verificación manual. Esta asimetría temporal y tecnológica constituye uno de los factores explicativos centrales del deterioro documentado en la confianza institucional durante el período analizado.

Tabla 2. Cronología de estrategias algorítmicas por período político.

Período	Hashtag/Operación central	Plataforma principal	Evidencia de automatización/IA	Fuente
Campaña 2018	#RedAMLO	Twitter	Red de 8 586 nodos; 2/5 nodos estratégicos con probabilidad alta de automatización	Olmedo-Neri (2021)
Gobierno 2019	#RedAMLOVE; #ReformaTodoLoDeforma	Twitter	176 cuentas categorizadas; estructura “maestro de ceremonias/coro”	Signa Lab ITESO (2019)
Elecciones 2021	Campañas anti-INE	Twitter	Campañas coordinadas de deslegitimación institucional	Signa Lab ITESO (2021)
Elecciones 2024	Chatbot “Inés”	WhatsApp	IA institucional para verificación con AFP Factual y Meedan	INE (2024)
Gobierno 2024	Deepfakes promocionales	Múltiples plataformas	Video falso de AMLO para estafa de inversión en Pemex	Presidencia de México (2024)

Fuente: Elaboración propia.

Amplificación y control narrativo

El análisis de redes sociales revela la existencia de estructuras organizativas sofisticadas que trascienden la conceptualización tradicional de “granjas de bots”. Los datos de #RedAMLO demuestran la implementación de una topología híbrida que combina participación ciudadana genuina con amplificación automatizada estratégica. Esta hibridación, documentada mediante un análisis de centralidad y herramientas como Botometer, sugiere una evolución hacia formas más sofisticadas de manipulación que

resultan más difíciles de detectar mediante herramientas de verificación tradicionales (Olmedo-Neri, 2021).

La estructura identificada opera mediante lo que Signa Lab ITESO (2019) denomina tipología “maestro de ceremonias/coro”, donde nodos centrales (MC) generan contenido original, mientras que redes de cuentas subordinadas (coro) amplifican sistemáticamente estos mensajes. Esta arquitectura permite combinar la apariencia de diversidad discursiva con la coordinación centralizada de narrativas, explotando así las limitaciones de los algoritmos de detección que privilegian indicadores cuantitativos sobre análisis cualitativos de contenido. El resultado es un ecosistema que simula pluralidad mientras concentra efectivamente el control narrativo.

La investigación documenta, además, la implementación de estrategias de “mutación”, donde cuentas modifican sistemáticamente sus identificadores y eliminan contenido histórico para evadir detección. Esta capacidad adaptativa refleja un nivel de sofisticación que sugiere coordinación profesional antes que activismo espontáneo. Los patrones identificados incluyen sincronización temporal precisa, replicación de formatos discursivos específicos y uso estratégico de emociones particulares (indignación, esperanza, desconfianza) para maximizar la viralización.

Las plataformas funcionan como amplificadores diferenciados: Twitter (X), para generación de tendencias y debate político directo; WhatsApp, para difusión capilar en redes familiares y comunitarias; YouTube, para monetización de contenido político, y Facebook/Meta, para microsegmentación psicográfica. Esta especialización por plataforma revela una comprensión sofisticada de las *affordances* específicas de cada ecosistema digital, maximizando el impacto mediante estrategias adaptadas a las características particulares de cada audiencia y formato.

Tabla 3. Estrategias algorítmicas por plataforma digital.

Período	Operaciones verificadas	Período	Tipo de manipulación	Fuente
Twitter (X)	#RedAMLO, #RedAMLOVE	2018-2019	Amplificación coordinada, estructura MC/coro	Olmedo-Neri (2021); Signa Lab ITESO (2019)
WhatsApp	Chatbot “Inés” institucional	2021-2024	IA para verificación electoral	INE (2024)
YouTube	Monetización de contenido progubernamental	2018-2024	Editorialización algorítmica	Signa Lab ITESO (2020)
Facebook/Meta	Microsegmentación electoral	2021	Clustering psicográfico mediante IA	Privacy International & Cavaliere (2021)

Fuente: Elaboración propia.

Impacto en la deliberación democrática y los mecanismos de *accountability*

Los resultados documentan una transformación cualitativa en los mecanismos de *accountability* horizontal que constituye el núcleo del deterioro democrático observado. Las instituciones autónomas, particularmente el INE y el sistema judicial, enfrentaron campañas sistemáticas de deslegitimación que combinaron desinformación automatizada con estrategias de saturación informativa. El caso paradigmático del *hashtag* #INECorrupto ilustra cómo las tecnologías de amplificación pueden reconfigurar la percepción pública de instituciones clave en períodos temporales reducidos.

La velocidad de propagación de narrativas deslegitimadoras supera sistemáticamente la capacidad de respuesta institucional, generando lo que la literatura especializada identifica como “déficit de velocidad democrática”. Mientras las instituciones operan bajo protocolos deliberativos que requieren verificación, consulta y construcción de consenso, las estrategias de manipulación algorítmica pueden modificar tendencias de opinión en cuestión de horas. Esta asimetría temporal constituye una vulnerabilidad estructural que las democracias frágiles no han logrado resolver (Wolfgang & Lührmann, 2021).

El deterioro del *accountability* vertical se manifiesta en la creación de “realidades electorales paralelas” donde diferentes segmentos del electorado acceden a versiones contradictorias de los mismos eventos políticos. Los algoritmos de personalización de contenido, combinados con estrategias deliberadas de polarización, han fragmentado la esfera pública en múltiples “públicos sintéticos” que operan con epistemologías incompatibles. Esta fragmentación erosiona los fundamentos comunes necesarios para la deliberación democrática, sustituyendo el debate por la confrontación entre cosmovisiones mutuamente excluyentes.

La evidencia sugiere que el *accountability* diagonal, ejercido por medios de comunicación y sociedad civil, ha sido particularmente afectado. Las campañas coordinadas contra periodistas y organizaciones civiles, amplificadas mediante redes automatizadas, han generado un efecto inhibitorio que trasciende el acoso digital directo. La anticipación de ataques coordinados modifica los incentivos de actores sociales independientes, generando autocensura y reduciendo la diversidad de voces en el debate público. Este fenómeno actualiza los postulados de O'Donnell (1992) sobre democracia delegativa en el contexto digital.

Tabla 4. Tipología de estrategias de manipulación algorítmica.

Estrategia	Descripción	Plataforma	Impacto documentado	Fuente
Amplificación coordinada	Redes de bots para viralización de contenidos específicos	Twitter (X)	Modificación artificial de trending topics	Olmedo-Neri (2021)
Microsegmentación emocional	Personalización de mensajes según perfiles psicológicos	Facebook/Meta	Polarización dirigida de audiencias específicas	Privacy International & Cavaliere (2021)

Generación de contenido sintético	Deepfakes y audio manipulado	Múltiples	Desinformación de alta credibilidad	Presidencia de México (2024)
Verificación institucional	Chatbots con IA para <i>fact-checking</i>	WhatsApp	Respuesta institucional a desinformación	INE México (2024)

Fuente: Elaboración propia.

Emergencia de la opinión pública sintética y sus implicaciones democráticas

Los hallazgos revelan la consolidación de un fenómeno inédito: la producción industrial de una opinión pública sintética que opera con niveles de sofisticación que superan las capacidades de detección de herramientas académicas tradicionales. A diferencia de las estrategias de propaganda clásica, que operaban mediante repetición masiva de mensajes idénticos, las nuevas generaciones de IA pueden generar variaciones semánticamente coherentes, pero formalmente diversas del mismo núcleo narrativo. Esta capacidad de “variación sintética” permite evadir algoritmos de detección mientras mantiene coherencia estratégica.

La investigación documenta la transición desde bots detectables hacia “agentes conversacionales” que replican patrones de comunicación humana con precisión suficiente para superar los filtros de verificación. Estos sistemas incorporan elementos de contextualización local, referencias culturales específicas y adaptación a registros lingüísticos regionales, generando interacciones que resultan indistinguibles de la participación ciudadana auténtica. La implicación es que las herramientas tradicionales de verificación (análisis de frecuencia temporal, detección de patrones repetitivos, análisis de metadatos) resultan insuficientes para identificar la manipulación algorítmica sofisticada.

El impacto de esta evolución tecnológica trasciende la manipulación electoral para afectar los fundamentos epistémicos de la democracia. Cuando la ciudadanía no puede distinguir entre participación auténtica y sintética, la legitimidad misma de la opinión pública como fundamento democrático queda comprometida. Este fenómeno genera lo que los datos sugieren que es una “crisis de autenticidad democrática”, donde la incertidumbre sobre lo genuino de la participación ciudadana erosiona la confianza en los procesos democráticos fundamentales.

La respuesta institucional, documentada en iniciativas como el chatbot “Inés”, del INE, representa un intento de “democratización de la verificación” mediante herramientas de IA accesibles a la ciudadanía; sin embargo, la evidencia sugiere que estas respuestas institucionales enfrentan limitaciones estructurales significativas: operan de manera reactiva, requieren recursos tecnológicos sustanciales y dependen de la iniciativa ciudadana para activarse. Más fundamentalmente, reproducen el paradigma de “verificación técnica” sin abordar las dimensiones políticas y sociales del problema de la desinformación.

Los resultados apuntan hacia la consolidación de un nuevo paradigma de “governance híbrida”, donde la efectividad de las instituciones democráticas depende crecientemente de su capacidad para operar en ecosistemas digitales. Esta transición requiere, además de la adaptación tecnológica, una reconfiguración conceptual de principios democráticos fundamentales: transparencia, representación, *accountability* y legitimidad. La investigación sugiere que las democracias frágiles, como la mexicana, enfrentan un desafío existencial en esta transición, donde la preservación de valores democráticos requiere la incorporación de capacidades tecnológicas que tradicionalmente han estado concentradas en actores privados.

Conclusiones

En la presente investigación se evidencia una transformación cualitativa en los mecanismos de influencia sobre la opinión pública que trasciende las concepciones tradicionales de manipulación mediática. Mientras el debate académico se ha concentrado en la detección de “granjas de bots” mediante herramientas como Botometer o análisis semánticos, emerge un fenómeno más complejo y sofisticado: la generación de IA conversacional indetectable que replica patrones de comunicación genuinamente humanos. Estas tecnologías, entrenadas mediante modelos de lenguaje avanzados, producen interacciones que superan los filtros tradicionales de verificación automatizada y logran influenciar, tanto a ciudadanos con formación académica como a sectores populares mediante la activación estratégica de emociones transversales (indignación, esperanza, desconfianza institucional).

El caso mexicano demuestra que la opinión pública sintética no se limita a la amplificación coordinada de *hashtags*, sino que incluye la creación de narrativas aparentemente espontáneas que apelan a problemáticas de resonancia universal: corrupción, seguridad, justicia social. Los sistemas de IA contemporáneos pueden generar comentarios, testimonios y argumentaciones que, al carecer de los marcadores típicos de automatización (repetitividad, sincronización temporal, patrones semánticos predecibles), resultan indistinguibles del discurso ciudadano auténtico. Esta invisibilidad algorítmica representa un salto cualitativo en la manipulación política: la transición de la propaganda identificable hacia la persuasión sintética imperceptible.

La fragilidad democrática mexicana se ve exacerbada por esta colonización algorítmica de la deliberación pública, donde la autenticidad del debate ciudadano queda comprometida sin posibilidad de verificación efectiva. Los mecanismos tradicionales de *accountability* (medios de comunicación, organismos electorales, sociedad civil organizada) enfrentan un desafío epistémico: ¿cómo distinguir entre participación ciudadana genuina y manipulación algorítmica sofisticada, cuando ambas convergen en formatos indistinguibles? Esta incertidumbre erosiona la confianza en la propia legitimidad de la opinión pública como fundamento democrático.

El populismo digital mexicano ha evolucionado hacia una ingeniería de consensos artificiales que explota vulnerabilidades psicológicas universales mediante la personalización emocional masiva. A diferencia de los bots detectables analizados en #RedAMLO (2018–2019), las nuevas generaciones de IA pueden simular biografías ciudadanas convincentes, adaptar registros lingüísticos regionales y construir argumentaciones contextualmente relevantes que apelan a experiencias socioeconómicas diversas. Esta democratización aparente de la influencia algorítmica —donde la manipulación adopta formas vernáculas y localistas— constituye un fenómeno inédito que demanda replanteamientos teóricos fundamentales sobre la naturaleza de la opinión pública en democracias digitalizadas.

La resiliencia democrática en contextos de fragilidad institucional requiere el desarrollo urgente de soberanía epistémica ciudadana: la capacidad social para discernir, cuestionar y verificar la autenticidad de los flujos informativos, independientemente de su origen tecnológico o humano.

Referencias

- Arellano Ríos, A. (2024). El régimen político y populismo: una propuesta analítica. *Espiral. Estudios sobre Estado y Sociedad*, 31(90). <https://doi.org/10.32870/eees.v31i90.7362>
- Aristegui Noticias. (2024, 26 de abril). AMLO advierte estafas con IA utilizando su imagen; 70% de los mexicanos no comprenden que es falso. *Aristegui*. <https://aristeginoticias.com/2604/mexico/amlo-advierde-estafas-con-ia-utilizando-su-imagen-70-de-los-mexicanos-no-comprenden-que-es-falso/>
- Ashby, W.R. (1956). *An introduction to cybernetics*. John Wiley & Sons Inc. <https://archive.org/details/introductiontocy00ashb/page/n7/mode/2up>
- Badillo, D. (2019, 14 de diciembre). Seguidores “espontáneos” de AMLO usan bots para ganar el debate público en redes sociales. *El Economista*. <https://www.eleconomista.com.mx/politica/Seguidores-espontaneos-de-AMLO-usan-bots-para-ganar-el-debate-publico-en-redes-sociales-20191214-0001.html>
- Baldwin-Philippi, J. (2018). The technological performance of populism. *New Media & Society*, 21, pp. 376 - 397. <https://doi.org/10.1177/1461444818797591>
- Beach, D. & Pedersen, R.B. (2019). *Process-tracing methods: Foundations and guidelines*. University of Michigan Press.
- De Souza, L.C. (2018). Functional democracy in resilient states: The risk of enclosed governments. *Revista Brasileira de Estudos Jurídicos*, 13(1), pp. 112-130. <https://portalunifipmoc.emnuvens.com.br/rbej/issue/download/30/31>
- Coss, J. & Cronin, L. (2024, 10 de octubre). Sheinbaum's folly: How a weak state undermines Mexico's democracy. *Democratic Erosion Consortium*. <https://democratic-erosion.org/2024/10/10/sheinbaums-folly-how-a-weak-state-undermines-mexico-democracy/>
- Deutsch, K.W. (1963). *The nerves of government: Models of political communication and control*. Free Press.
- Domínguez, P. (2020, 6 de mayo). AMLO dice que Twitter reconoció la existencia de bots en su servicio. *Milenio*. <https://www.milenio.com/politica/amlo-twitter-reconocio-existencia-bots-servicio>

- Habermas, J. (2023). *Eine neue Strukturwandel der Öffentlichkeit und deliberative Politik*. Suhrkamp.
- Harari, Y.N. (2024). *Nexus: Una breve historia de las redes de información desde la Edad de Piedra hasta la IA*. Penguin Random House.
- Infobae. (2021). *La otra cara de las elecciones 2021 en México: fake news, bots y cuentas falsas inundan las redes sociales*. Infobae. <https://www.infobae.com/america/mexico/2021/05/24/la-otra-cara-de-las-elecciones-2021-en-mexico-fake-news-bots-y-cuentas-falsas-inundan-las-redes-sociales/>
- Innerarity, D. & Colomina, C. (2020). La verdad en las democracias algorítmicas. *Revista CIDOB d'Afers Internacionals*, 124, pp. 11-23. <https://doi.org/10.24241/rcai.2020.124.1.11>
- Instituto Nacional Electoral México. (2024, 17 de abril). Nueva versión del chatbot del INE en WhatsApp permitirá verificar información sobre elecciones 2024. *Central electoral*. <https://centralectoral.ine.mx/2024/04/17/nueva-version-del-chat-bot-del-ine-en-whatsapp-permitira-verificar-informacion-sobre-elecciones-2024/>
- Jungherr, A. (2023). Artificial intelligence and democracy: A conceptual framework. *Social Media + Society*, 9(1). <https://doi.org/10.1177/20563051231186353>
- Levitsky, S. & Ziblatt, D. (2018). *How Democracies Die*. Crown Publishing Group. <https://newuniversityinexileconsortium.org/wp-content/uploads/2022/08/Levitsky-Ziblatt-How-Democracies-Die-2018-1.pdf>
- Martínez, M. (2018, 29 de mayo). México: los “bots”, “trolls” y otros trucos de manipulación en internet que amenazan las próximas elecciones presidenciales. *BBC News Mundo*. <https://www.bbc.com/mundo/noticias-america-latina-44302996>
- Mazzoleni, G. & Bracciale, R. (2018). Socially mediated populism: the communicative strategies of political leaders on Facebook. *Palgrave Communications*, 4, pp. 1-10. <https://doi.org/10.1057/s41599-018-0104-x>
- Nieto, D. y Milanese, J.P. (2016). De las transiciones al estancamiento. Revisitando la democratización en la obra de Guillermo O'Donnell. *Co-Herencia*, 13(24). <https://doi.org/10.17230/co-herencia.13.24.6>
- O'Donnell, G. (1992). *Delegative democracy?* [Kellogg Institute Working Paper No. 172]. The Helen Kellogg Institute for International Studies-University of Notre Dame. https://kellogg.nd.edu/sites/default/files/old_files/documents/172_0.pdf
- Ojeda de la Torre, I. (2019, 12 de marzo). Red de fans de AMLO ataca desde Twitter a periodistas y medios críticos, revela Signa_Lab ITESO. *Sin embargo.mx*. https://www.sinembargo.mx/3549094/red-de-fans-de-amlo-ataca-desde-twitter-a-periodistas-y-medios-criticos-revela-signa_lab-iteso/
- Olmedo-Neri, R.A. (2021). Political communication on Internet: the case of #RedAMLO in Mexico. *Universitas*, 34, pp. 103-124. <https://doi.org/10.17163/uni.n34.2021.05>
- Presidencia de México. (2024, 26 de abril). Presidente alerta sobre mensaje falso creado con inteligencia artificial para invertir en Pemex. *Conferencia de prensa matutina*. <https://amlo.presidencia.gob.mx/presidente-alerta-sobre-mensaje-falso-creado-con-inteligencia-artificial-para-invertir-en-pemex/>
- Privacy International & Cavaliere, P. (2021). *Micro-targeting in political campaigns: A comparative analysis*. The University of Edinburgh. https://privacyinternational.org/sites/default/files/2021-01/UoE_PI%20Micro-targeting%20in%20political%20campaigns%20comparative%20analysis%202021.pdf
- Ramírez, L. (2019, 11 de marzo). Una red de bots y trolls ataca en Twitter a medios y críticos del gobierno. *Animal político*. <https://www.animalpolitico.com/2019/03/twitter-gobierno-criticos-bots-trolls>
- Rosanvallon, P. (2020). *El siglo del populismo: Historia, teoría y crítica*. Galaxia Gutenberg.

- Ruiz, M. & Buttner, M. (2024). Fake news, influence on the formation of public opinion and impacts on the legitimacy of public decision. *A&C - Revista de Direito Administrativo & Constitucional*, 24(95). <https://doi.org/10.21056/aec.v24i95.1898>
- Saura García, C. & Calvo, P. (2024). Opinión pública masiva: colonización algorítmica y sintetificación de la esfera pública. *Revista CIDOB d'Afers Internacionals*, 138, pp. 73-95. <https://doi.org/10.24241/rcai.2024.138.3.73>
- Signa Lab ITESO. (2019). *Democracia, libertad de expresión y esfera digital: El caso de la #RedAMLOVE*. Instituto Tecnológico y de Estudios Superiores de Occidente. https://signalab.iteso.mx/informes/informe_redamlove.html
- SinEmbargo. (2018, 22 de abril). 8 millones bots y trolls operan este 2018: buena onda con Meade y PRI, y rabiosos con Anaya y AMLO. *Sin embargo.mx*. <https://www.sinembargo.mx/3409934/en-la-campana-2018-operan-8-millones-de-bots-y-trolls-y-la-mayoria-trabaja-para-meade/>
- Sosa Santiago, M.F. (2024, 26 de abril). AMLO alerta sobre estafa con IA utilizando su imagen para invertir en Pemex. *El Economista*. <https://www.eleconomista.com.mx/politica/AMLO-alerta-sobre-estafa-con-IA-utilizando-su-imagen-para-invertir-en-Pemex-Cual-es-la-inversion-minima-que-piden-los-estafadores-20240426-0049.html>
- Tribunal Electoral del Poder Judicial de la Federación. (2023, 14 de diciembre). TEPJF y TikTok contribuirán a la comunicación de información electoral pública relevante. *Boletín 15366*. <https://www.te.gob.mx/front3/bulletins/detail/15366/0>
- Tucker, J.A. et al. (2018). Social Media, Political Polarization, and Political Disinformation: A Review of the Scientific Literature. *Hewlett Foundation*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3144139
- Urbinati, N. (2014). *Democracy disfigured: Opinion, truth, and the people*. Harvard University Press.
- Van Alstyne, M. & Brynjolfsson, E. (1997). Electronic communities: Global village or cyberbalkans? *Proceedings of the International Conference on Information Systems*, pp. 1-12. <https://web.mit.edu/marshall/www/papers/CyberBalkans.pdf>
- von Foerster, H. (2003). *Understanding understanding: Essays on cybernetics and cognition*. Springer. <https://link.springer.com/book/10.1007/b97451>
- Whitehead, L. (2021). Latin America erupts: The danger of democratic delinquency. *Journal of Democracy*, 32(3), pp. 78-93. <https://www.journalofdemocracy.org/articles/latin-america-erupts-the-danger-of-democratic-delinquency/>
- Wolfgang, M. & Lührmann, A. (2021). Resilience of democracies: Responses to illiberal and authoritarian challenges. *Democratization*, 28(5), pp. 903-924. <https://doi.org/10.1080/13510347.2021.1928081>
- Woolley, S. C. (2023). *Manufacturing consensus: Understanding propaganda in the age of automation and anonymity*. New Haven, CT: Yale University Press.
- Zuboff, S. (2020). *La era del capitalismo de la vigilancia*. Paidós. https://perio.unlp.edu.ar/catedras/comunicacionyrecepcion/wp-content/uploads/sites/210/2023/06/ZUBOFF_1.pdf
- Zuiderwijk, A., Chen, Y. & Salem, F. (2021). Implications of the use of artificial intelligence in public governance: A systematic literature review and a research agenda. *Government Information Quarterly*, 38(3), p. 101577. <https://doi.org/10.1016/j.giq.2021.101577>