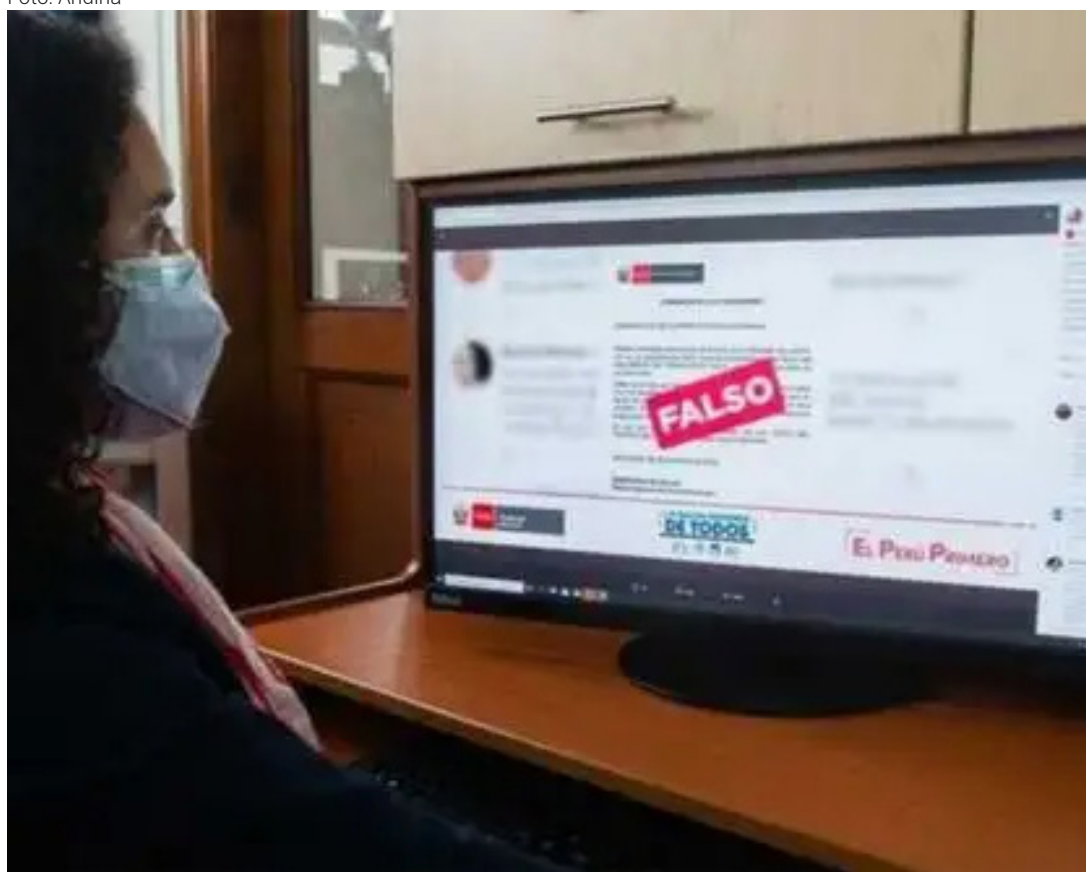


ÉTICA EN REDES SOCIALES E INTELIGENCIA ARTIFICIAL DURANTE CAMPAÑAS POLÍTICAS: LIBERTAD DE INFORMACIÓN VERSUS DESINFORMACIÓN Y BOTS

Foto: Andina



¿Cómo proteger la democracia sin censurar? El artículo analiza desinformación automatizada y propaganda segmentada en redes, y plantea responsabilidades compartidas entre ciudadanía, plataformas y Estado.

INTRODUCCIÓN

Vivimos una paradoja inquietante: las tecnologías de la información y la comunicación prometían democratizar el acceso al conocimiento, pero, en cambio, hemos terminado construyendo un ecosistema en el que proliferan las noticias falsas y en el que la manipulación informativa se ha convertido en una herramienta política cotidiana. Lo que debería ser nuestra plaza pública digital —ese espacio en el que intercambiamos ideas y debatimos libremente— está ahora poblado de robots, *trolls* pagados, algoritmos que nadie entiende y estrategias coordinadas para engañarnos. Todo esto nos obliga a preguntarnos cosas fundamentales: ¿qué significa realmente la verdad hoy en día? ¿Cuáles son nuestros compromisos éticos básicos? Y, quizá, lo más inquietante: ¿puede sobrevivir la democracia cuando nuestras conversaciones están controladas por la inteligencia artificial?

Piensen en lo que sucede durante las elecciones. Los ciudadanos queremos informarnos bien antes de votar, ¿verdad? Pero lo que encontramos es un bombardeo calculado de mentiras esparcidas por robots digitales y mensajes diseñados específicamente para manipular a cada tipo de votante. El resultado es grave: nos cuesta muchísimo formarnos una opinión basada en hechos reales, se distorsiona por completo la competencia política justa y, en casos extremos, estos mecanismos pueden incluso cambiar quién gana unas elecciones.

Es cierto que acceder libremente a la información es un derecho que nos protege del abuso del poder. Pero enfrentamos algo nuevo: una verdadera industria de la mentira que funciona en piloto automático y que destruye las condiciones básicas necesarias para tener debates democráticos significativos. Nos toca, entonces, hacernos preguntas difíciles sobre hasta dónde llega la libertad de expresión y cómo defendemos

el libre flujo de ideas sin permitir, al mismo tiempo, que nos manipulen de formas no contempladas por nuestros marcos éticos tradicionales.

LIBERTAD DE INFORMACIÓN COMO DERECHO FUNDAMENTAL

Para entender realmente qué está en juego hoy, necesitamos volver a lo básico: ¿por qué la libertad de expresión se considera un derecho tan fundamental? Dos tradiciones éticas ofrecen perspectivas complementarias: el utilitarismo liberal de John Stuart Mill y el universalismo deontológico de Immanuel Kant.

John Stuart Mill (1859/2012) formuló una de las defensas más influyentes de la libertad de expresión dentro del liberalismo. Desde un enfoque utilitarista, sostuvo que el libre intercambio de ideas maximiza el bienestar colectivo al crear condiciones óptimas para la búsqueda de la verdad. Su célebre tridente argumentativo señala que silenciar una opinión verdadera priva a la sociedad de la verdad, que refutar una opinión falsa fortalece la comprensión y que confrontar verdades parciales permite alcanzar una visión más completa. Además, Mill (1859/2012) advierte que una verdad incuestionada degenera en un dogma muerto, sostenido por costumbre y carente de vitalidad intelectual. Por su parte, Immanuel Kant (1784/s. f.) fundamentó la libertad de expresión en la dignidad humana y la autonomía moral. En *¿Qué es la Ilustración?*, Kant (1784/s. f.) definió la emancipación como la salida de la “minoría de edad autoimpuesta” y proclamó el lema *Sapere aude!*, como una exigencia moral. La libertad de pensamiento y expresión es, para Kant, una condición trascendental de la autonomía racional, y su restricción constituye una violación directa del imperativo categórico.

Ambas tradiciones coinciden en concebir la libertad informativa como un pilar civilizatorio. Con la irrupción de las redes sociales,

muchos interpretaron estas tecnologías como la realización de los ideales ilustrados: un ágora digital abierta y sin intermediarios. No obstante, esta promesa presenta diferencias sustanciales respecto del pasado: una escala masiva de usuarios, una difusión viral inmediata, una persistencia digital y una intermediación algorítmica opaca que condiciona el acceso a la información. Aunque fenómenos como la Primavera Árabe parecieron confirmar su potencial emancipador (Howard & Hussain, 2013), autores como Habermas (2023) advierten que las redes fragmentan la atención y privilegian respuestas emocionales. Casos como WikiLeaks, los Panama Papers o las revelaciones de Edward Snowden evidencian tanto el control ciudadano como nuevas formas de polarización. En el contexto peruano, las protestas del 2022 y 2023 mostraron simultáneamente el valor testimonial de las plataformas y su capacidad para amplificar discursos polarizantes y la desinformación. Así, los ideales de Mill y Kant enfrentan hoy desafíos inéditos derivados de la lógica algorítmica y la manipulación digital.

LA PARADOJA DE LA DESINFORMACIÓN AUTOMATIZADA

La forma en que están construidas las redes sociales hoy ha creado un ambiente vulnerable a la manipulación de la información mediante sistemas automatizados. Los robots digitales —programas informáticos que simulan el comportamiento humano— combinados con tecnologías de inteligencia artificial que pueden generar contenido político han cambiado profundamente cómo funciona la comunicación en el espacio público, especialmente cuando hay elecciones en curso. La automatización permite producir y distribuir desinformación a velocidades y volúmenes que superan la capacidad de verificación.

El escándalo de Cambridge Analytica, revelado en marzo del 2018 por Cadwalladr

Foto: Master1305/Shutterstock



Actualmente, las tecnologías de la información desafían la noción de libertad informativa

**LOS BOTS Y LA IA
PUEDEN ALTERAR EL
FUNCIONAMIENTO DE
LA COMUNICACIÓN
DE CONTENIDO
POLÍTICO Y
PRODUCIR
DESINFORMACIÓN**

CAMBRIDGE ANALYTICA ES UN CASO EMBLEMÁTICO RECIENTE DE MANIPULACIÓN POLÍTICA A GRAN ESCALA

y Graham-Harrison (2018), marcó un punto de inflexión. Hasta 87 millones de perfiles de Facebook fueron obtenidos sin consentimiento, lo que permitió desarrollar técnicas de perfilamiento psicográfico para segmentar votantes según rasgos de personalidad inferidos algorítmicamente. El *microtargeting* político representa una transformación cualitativa: cada votante recibe mensajes diseñados para activar respuestas emocionales, incluso mensajes contradictorios, sin saber que es objeto de manipulación psicológica. El problema ético va más allá del campo de la privacidad. Está en juego la autonomía racional que Kant (1784/s. f.) consideraba el fundamento de la dignidad humana. Cuando las preferencias políticas son moldeadas por técnicas algorítmicas que operan por debajo del umbral de conciencia, la formación autónoma de la opinión se vuelve conceptualmente frágil.

Los bots políticos intensifican este riesgo al producir la ilusión de consenso social en torno a narrativas fabricadas. Un informe de la Universidad de Oxford muestra que, en el 2019, 76 países emplearon la desinformación digital como estrategia política, mientras que 59 usaron *trolls* financiados por el Estado para atacar a adversarios (Bradshaw et al., 2021). La sofisticación de los bots —imitan con precisión al hombre mediante aprendizaje automático (Ferrara et al., 2016)— dificulta su detección. El Perú enfrenta dinámicas similares: estudios sobre campañas recientes identifican redes de

trolls que impulsan narrativas específicas, y que aprovechan la característica fragmentación y desconfianza institucional del entorno político nacional. Otros mecanismos son las cámaras de eco y las burbujas de filtro. Los algoritmos de recomendación priorizan contenidos que confirman creencias previas y que provocan reacciones emocionales intensas, lo que fragmenta la realidad compartida en múltiples versiones incompatibles entre sí (Pariser, 2011).

Este escenario plantea una paradoja ética central: ¿puede hablarse de libertad cuando se utiliza tecnología para mentir industrialmente mediante ejércitos de bots? Mill (1859/2012) confiaba en que las falsedades serían refutadas por el debate racional, pero ese mecanismo requiere condiciones que no existen en el ecosistema actual. La sobrea-bundancia de información falsa produce fatiga de verificación y erosiona la confianza epistémica (Giraldo, 2025), lo que transforma el debate político en una contienda tribal en la que importa más la lealtad grupal que la verdad verificable.

DILEMAS ÉTICOS DE LA REGULACIÓN

La expansión de la desinformación automatizada plantea una pregunta fundamental: ¿quién debe garantizar la verdad en el ecosistema digital y bajo qué criterios? La regulación estatal enfrenta objeciones liberales históricamente justificadas por abusos de poder y censura. El autoritarismo digital ya no es una hipótesis teórica, sino una realidad documentada.

El caso de China es revelador. El Gran Cortafuegos —parte del Proyecto Escudo Dorado— bloquea contenidos incómodos bajo el pretexto de mantener la armonía social, lo que impide el acceso a información sobre protestas, corrupción o represión a minorías. El sistema de crédito social ejemplifica cómo la tecnología se transforma en un sistema

de vigilancia masiva que monitorea y califica constantemente a los ciudadanos (Liang et al., 2018). En Rusia, leyes supuestamente dirigidas contra las *fake news* criminalizan el periodismo independiente sobre la guerra en Ucrania. En América Latina, existen propuestas regulatorias con letra pequeña peligrosa: reglas ambiguas que facilitan silenciar críticos bajo la excusa de combatir la desinformación (Bradshaw et al., 2019). El Perú es particularmente vulnerable: instituciones débiles y ausencia de marcos legales claros nos exponen a una manipulación sin defensas reales.

Delegar la moderación a las plataformas tampoco resuelve el problema. Estas corporaciones priorizan las ganancias sobre la democracia; su modelo de negocio prospera con el conflicto y la polarización, que generan más interacción. La suspensión de la cuenta de Trump en el 2021 ilustró cómo empresas privadas toman decisiones con implicancias constitucionales sin legitimidad democrática (Douek, 2020). Los sistemas de moderación presentan sesgos documentados y operan como cajas negras algorítmicas imposibles de auditar (Gillespie, 2018; Roberts, 2019). El desafío ético consiste en diseñar regulaciones que eviten dos peligros: gobiernos que censuran arbitrariamente y corporaciones que actúan sin responsabilidad. Necesitamos proteger la libertad de expresión sin permitir que la democracia sea destruida en el proceso.

HACIA UN ENFOQUE ÉTICO INTEGRADOR

La desinformación digital no admite soluciones mágicas ni tecnológicas únicas. Requiere responsabilidad compartida entre ciudadanos, plataformas, gobiernos, instituciones educativas y medios de comunicación. La democracia sobrevive solo mediante la coordinación efectiva de todos estos actores.

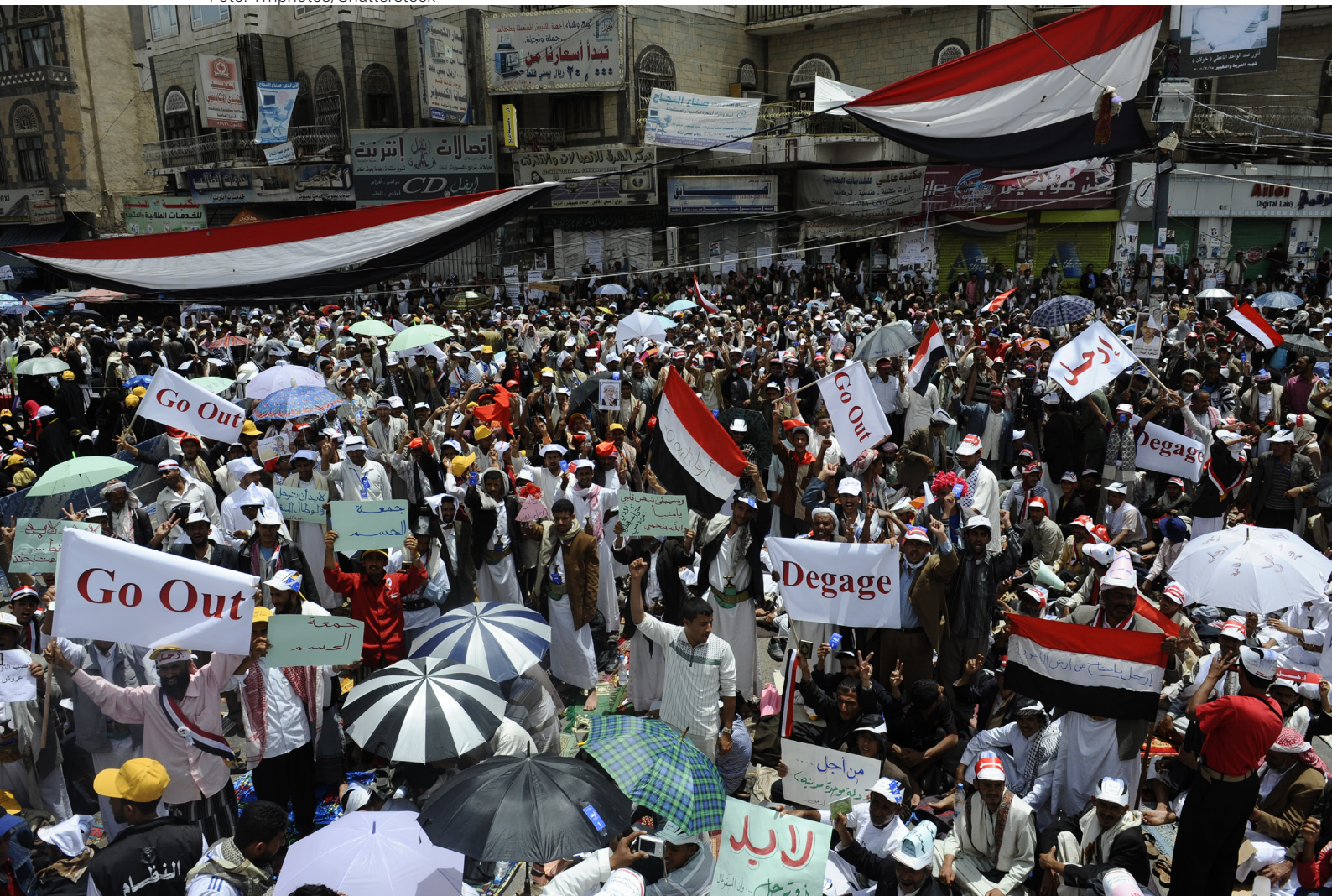
La educación digital constituye el fundamento. Sin habilidades críticas para evaluar

información *online*, ningún sistema de moderación externa será suficiente. Hobbs (2010) lo expresa claramente: el pensamiento crítico sobre las redes sociales es hoy tan esencial para la democracia como la alfabetización tradicional. Esto implica identificar fuentes confiables, distinguir hechos de opiniones, reconocer sesgos propios y comprender funcionamientos algorítmicos básicos. En el Perú, la situación es urgente: bajos niveles de educación digital y ausencia de políticas públicas sistemáticas nos dejan indefensos ante campañas sofisticadas de manipulación, lo cual profundiza la polarización y erosiona la confianza institucional.

Las empresas tecnológicas deben comprometerse con principios éticos verificables. El Código de Buenas Prácticas de la Unión Europea (Comisión Europea, s. f.) representa un avance, pero su efectividad depende de mecanismos reales de fiscalización y sanción. Sin consecuencias por incumplimiento, estos códigos son solo estrategias de relaciones públicas. De otro lado, las herramientas tecnológicas ofrecen un apoyo limitado. Sistemas automatizados de inteligencia artificial detectan patrones lingüísticos asociados a la desinformación, y etiquetar cuentas automatizadas aumenta la transparencia (Gorwa & Guilbeault, 2018). Sin embargo, estas herramientas presentan fallas inherentes: censuran expresiones legítimas por error o dejan pasar desinformación elaborada.

Finalmente, la libertad de expresión debe entenderse unida a la responsabilidad epistémica. Los ciudadanos tienen deberes informativos: verificar, contrastar y deliberar civilizadamente. Las plataformas deben diseñar sistemas que favorezcan la deliberación informada por encima de la polarización emocional, incluso reduciendo métricas de interacción. Solo así protegemos simultáneamente la libertad informativa y la calidad democrática.

Foto: Ymphotos/Shutterstock



La educación digital es una tarea urgente en el Perú.

JÓVENES COMO AGENTES DE CAMBIO EN LA ESFERA DIGITAL

Los adolescentes no son simples consumidores ni víctimas pasivas de la desinformación digital, sino los protagonistas en la transformación del ecosistema informativo. Como nativos digitales, poseen una familiaridad intuitiva con las plataformas, lo que puede convertirse en una ventaja estratégica para afrontar la manipulación informativa. Pero esta facilidad técnica no les dará la competencia crítica: estudios revelan que ellos, como otros grupos generacionales, son vulnerables a la desinformación, precisamente porque su

confianza en las redes sociales no siempre va acompañada de un escepticismo epistémico (Wineburg & McGrew, 2019). Por ello, la participación juvenil informada es una condición necesaria para la salud democrática futura. Esto exige desarrollar lo que Jenkins et al. (2016) denominan culturas participativas: comunidades en las que los jóvenes no solo consuman contenido, sino, también, en las que puedan crear, verificar y cuestionar la información política. Iniciativas como los *fact-checkers* estudiantiles, los observatorios ciudadanos de propaganda electoral y las redes de alfabetización mediática entre pares demuestran que, cuando se les proporcionan herramientas adecuadas,

pueden convertirse en vigilantes efectivos contra la desinformación y en promotores de una cultura digital más responsable y deliberativa.

CONCLUSIÓN

La tensión entre libertad de información y prevención de la desinformación automatizada durante campañas políticas no es un problema técnico transitorio, sino una contradicción inherente a la democracia digital. Los fundamentos filosóficos de Mill (1859/2012) y Kant (1784/s. f.) sobre la libertad de expresión conservan plena vigencia, pero el ecosistema actual de redes sociales algorítmicas y bots desafía los supuestos sobre los cuales descansaban aquellos argumentos clásicos.

No hay salidas simples. La regulación estatal robusta amenaza con un autoritarismo digital; la autorregulación corporativa carece de incentivos adecuados; y las soluciones puramente tecnológicas resultan insuficientes. Se requiere una ética de responsabilidad compartida en la que cada actor asuma obligaciones específicas: ciudadanos con alfabetización mediática crítica, plataformas comprometidas con la transparencia, Estados que regulen sin censurar e instituciones de verificación independientes.

En este contexto, los jóvenes emergen como actores estratégicos. Su familiaridad digital, combinada con una formación crítica apropiada, los posiciona como agentes transformadores capaces de construir culturas participativas más responsables. Sin embargo, esto exige dotarlos de herramientas para convertir su presencia digital en una fuerza democratizadora. La democracia deliberativa depende, paradójicamente, de reconocer que la libertad irrestricta puede destruir las condiciones de su propio ejercicio. Solo mediante un compromiso intergeneracional con la responsabilidad epistémica preservaremos

DIVERSAS INICIATIVAS CIUDADANAS, CON LAS HERRAMIENTAS ADECUADAS, PUEDEN SER VIGILANTES EFECTIVOS CONTRA LA DESINFORMACIÓN

tanto la libertad de expresión como la integridad democrática.

REFERENCIAS

- Bradshaw, S., Bailey, H. & Howard, P. N. (2021). *Industrialized disinformation: 2020 global inventory of organized social media manipulation*. Oxford Internet Institute. <https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/12/2021/01/CyberTroop-Report-2020-v.2.pdf>
- Bradshaw, S., Neudert, L.-M., & Howard, P. N. (2019). *Government responses to malicious use of social media*. NATO Strategic Communications Centre of Excellence. <https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/12/2019/01/Nato-Report.pdf>
- Cadwalladr, C., & Graham-Harrison, E. (2018, 17 de marzo). Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. *The Guardian*. <https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election>
- Comisión Europea. (s. f.). *Un código de buenas prácticas de la UE en materia de desinformación reforzado*. <https://>

- commission.europa.eu/topics/countering-information-manipulation/strengthened-eu-code-practice-disinformation_es
- Douek, E. (2020, 11 de febrero). *The rise of content cartels*. Knight First Amendment Institute. <https://knightcolumbia.org/content/the-rise-of-content-cartels>
- Ferrara, E., Varol, O., Davis, C. B., Menczer, F., & Flammini, A. (2016). The rise of social bots. *Communications of the ACM*, 59(7), 96-104.
- Gillespie, T. (2018). *Custodians of the internet. Platforms, content moderation and the hidden decisions that shape social media*. Yale University Press.
- Giraldo, I. (2025). *Las sirenas de Homero y las sirenas digitales: la Odisea como pedagogía de la resistencia crítica en la era de la información*. Ponencia presentada al I Congreso Internacional de las Humanidades, Universidad de Lima.
- Gorwa, R., & Guilbeault, D. (2018). Unpacking the social media bot: A typology to guide research and policy. *Policy & Internet*, 12(2), 225-248.
- Habermas, J. (2023). *A new structural transformation of the public sphere and deliberative politics*. Polity.
- Hobbs, R. (2010). *Digital and media literacy: A plan of action*. The Aspen Institute; The Knight Foundation. <https://files.eric.ed.gov/fulltext/ED523244.pdf>
- Howard, P. N., & Hussain, M. M. (2013). *Democracy's fourth wave? Digital media and the Arab Spring*. Oxford University Press.
- Jenkins, H., Ito, M., & Boyd, D. (2016). *Participatory culture in a networked era: A conversation on youth, learning, commerce, and politics*. Polity.
- Kant, I. (s. f.). *¿Qué es la Ilustración?* <https://educacion.uncuyo.edu.ar/upload/kant-que-es-la-ilustracion.pdf> (Obra original publicada en 1784)
- Liang, F., Das, V., Kostyuk, N., & Hussain, M. M. (2018). Constructing a data-driven society: China's social credit system as a state surveillance infrastructure. *Policy & Internet*, 10(4). <https://creatingfutures.org/wp-content/uploads/2021/10/LiangEtAl-2018-ChinaSocialCreditStateSurveillance.pdf>
- Mill, J. S. (2012). *Sobre la libertad*. Alianza Editorial. (Obra original publicada en 1859)
- Pariser, E. (2011). *The filter bubble: What the internet is hiding from you*. Penguin Press.
- Roberts, S. (2019). *Behind the screen: Content moderation in the shadows of social media*. Yale University Press.
- Wineburg, S., & McGrew, S. (2019). Lateral reading and the nature of expertise: Reading less and learning more when evaluating digital information. *Teachers College Record*, 121(11). <https://doi.org/10.1177/016146811912101102>