

PALABRAS PRELIMINARES

JUAN JOSÉ SOLOZABAL ECHAVARRIA

Catedrático de Derecho Constitucional UAM

Escribir sobre la libertad de expresión supone seguramente asumir tres cosas. En primer lugar, la relevancia de este derecho en el plano de la reflexión general sobre los derechos, pues buena parte de la problemática sobre los mismos se ha suscitado a propósito de esta libertad pública: así, la dimensión individual o institucional de los derechos, su titularidad individual o colectiva, su vigencia en las relaciones entre particulares, y eso sin mencionar la cuestión de sus límites o la resolución de los conflictos con otros derechos y bienes jurídicos constitucionales.

La radicalidad de la libertad de expresión, en segundo lugar, tiene que ver ciertamente con su trascendencia en el plano político, pero también con su relación con la dignidad de la persona, que no tolera su supresión o censura. De ahí la necesidad de una defensa firme de la libertad de expresión en un contexto en el que tienden a enfatizarse los peligros de su uso irrestricto. Como afirmaba un editorialista de *The Economist*, recogiendo el pensamiento de Timothy Garton Ash: «No intentemos jamás silenciar a aquellos con los que discrepamos. Contestemos al discurso de nuestro oponente con más debate. Hemos de discutir sin recurrir a la violencia, y simplemente dispongámonos a endurecer nuestra piel».

Sentado lo anterior, en tercer lugar, resulta comprensible que en la coyuntura actual el análisis de la libertad de expresión se centre en su carácter imprescindible para la formación de una opinión pública libre. En última instancia, la libertad de expresión tiene que ver con el *speech* político relativo al gobierno, ya sea para criticarlo o defenderlo, tanto en períodos electorales como durante el ejercicio ordinario del mandato. Lo esencial, desde esta perspectiva, no es que cada cual diga lo que quiera, sino que finalmente se diga aquello que merece la pena ser dicho. En esta línea de acentuación de la dimensión política de la libertad de expresión se inscriben las dos justificaciones clásicas ofrecidas por la

literatura académica estadounidense. En primer lugar, en la medida en que la acción gubernamental se convierte en objeto del debate público, la participación en la formación de la opinión pública legítima democráticamente al poder político, pues los ciudadanos se conciben a sí mismos como potenciales autores de las decisiones que les afectan. En segundo lugar, la existencia de una opinión pública rica y plural, a la que no se sustraigan ni temas ni participantes, es condición necesaria para una participación ciudadana consciente e informada, ya que el bienestar de la comunidad exige que quienes deciden conozcan adecuadamente los problemas sobre los que deliberan.

Es en este contexto teórico donde debe situarse la importante contribución de Isabel Giménez, que formula una propuesta descriptiva y normativa sobre el impacto de internet y de la inteligencia artificial en la sociedad contemporánea. A juicio de la autora, están disminuyendo las oportunidades para la formación de una opinión pública razonable, plural y libre. El panorama informativo tiende a caracterizarse por una creciente desorganización, generando un clima en el que resulta difícil distinguir entre la verdad y la invención, la mentira o la distorsión deliberada de la realidad. El resultado es un progresivo alejamiento de los ciudadanos respecto de las instituciones democráticas, acompañado de desafección y apatía política. Sin duda, la democracia —«régimen de ciudadanos informados», en expresión de Hesse— se resiente cuando no concurren las condiciones necesarias para una opinión pública saludable. En la denominada sociedad de la posverdad se fomenta una cultura predominantemente emocional, y no propiamente racional, basada en la manipulación de los hechos. Según nuestra autora la posverdad implica la distorsión intencionada de la realidad, la manipulación consciente de la percepción pública y la movilización de las emociones.

Como es sabido, uno de los principales factores en la generación de la desinformación son las *deepfakes* creadas mediante inteligencia artificial y difundidas a través de internet, especialmente en redes sociales. Nos hallamos ante un uso indebido de la IA con fines políticos, consistente en la propagación consciente de noticias falsas con el objetivo de obtener ventajas electorales. En este escenario, la proliferación de *deepfakes* pone en cuestión la autenticidad y la integridad de la información en el espacio público. La profesora Isabel Giménez subraya el propósito político de estas prácticas de manipulación, pero también analiza los instrumentos a través de los cuales se producen: los distintos tipos de contenidos generados por inteligencia artificial, las dificultades para verificar su veracidad y los contextos en los que opera la desinformación, como los sesgos de confirmación, la persistencia de las burbujas informativas o el uso de algoritmos para amplificar su impacto.

En el análisis de los problemas que plantea Internet para la formación de la opinión pública ocupa un lugar central la cuestión de la regulación de las plataformas digitales. Su protagonismo, especialmente en períodos electorales, es innegable, de ahí que resulte esencial garantizar tanto su neutralidad como mecanismos adecuados de moderación. En el trasfondo de este debate se encuentra el reconocimiento del papel decisivo que desempeñan estos nuevos actores privados en la configuración de la opinión pública y la consiguiente atribución de un rol específico al poder público. Se trata, en primer término, de admitir que las plataformas ejercen un control efectivo sobre la circulación de la información. No es descartable, incluso, que asuman funciones cuasi públicas mediante un control preventivo básico de los contenidos, control que, aun siendo provisional y susceptible de supervisión judicial, entraña el riesgo de una prepotencia potencialmente peligrosa.

Ante esta situación caben dos actitudes: o bien se afirma que sustancialmente la situación de los medios no ha sufrido con la irrupción de los nuevos actores digitales, más allá de la titularidad de los agentes operadores; o se suscribe la opinión de quienes mantienen que a las grandes empresas de la digitalización o plataformas les corresponde una naturaleza híbrida como foros a la vez públicos y privados lo que permitiría al Estado imponer sus políticas de moderación y contenidos. Hacerse eco de esta contraposición es muy importante no sólo para atribuir al poder público un determinado rol en el mercado de los medios ajeno a los planteamientos liberales, sino para justificar las exigencias a los operadores de la inteligencia artificial, aceptando determinadas políticas de control, y legitimando, con un propósito normativo, determinada dirección en la regulación del espacio público democrático.

Así, en relación con las redes sociales pueden distinguirse, en suma, dos modelos. El primero responde a una concepción libertaria o de irrestricción absoluta, que entiende la libertad de expresión como una libertad negativa frente al Estado. El segundo se articula en torno a una concepción moderada, que pone el acento en la relación entre derechos y deberes y en los límites derivados de la protección de la seguridad y el bienestar colectivos. Para el primer modelo, la defensa frente a la intervención estatal es la garantía única de un debate público «desinhibido, robusto y amplio» (*New York Times v. Sullivan*, 1964).

La concepción moderada de la libertad de expresión, en cambio, atiende a criterios como el daño, la proporcionalidad, las desigualdades de poder, la democracia y el bien común, y no exclusivamente a los derechos individuales o corporativos. Esta visión resulta especialmente pertinente en el contexto de las redes sociales, donde la concentración de plataformas confiere a actores priva-

dos funciones de regulación y control particularmente sensibles. Aunque el modelo de la desregulación parece imponerse en la doctrina reciente del Tribunal Supremo estadounidense y en la práctica política federal —con la excepción de ciertos ámbitos como la pornografía—, ello se justifica en parte por las expectativas de innovación tecnológica asociadas a la autorregulación. Sin embargo, desde la perspectiva de la participación democrática, los riesgos derivados de este enfoque no pueden ser ignorados.

El control estatal del contenido de los mensajes podría, ciertamente, reducir la utilidad y accesibilidad de las plataformas, afectando negativamente al caudal informativo y al pluralismo que estas favorecen. La Primera Enmienda protege a las plataformas frente a la imposición de contenidos por parte del Estado, sin impedir que adopten voluntariamente estándares éticos para reforzar su legitimidad. Ejemplo de ello es el *Facebook Oversight Board*, integrado por expertos independientes que revisan las decisiones de moderación de contenidos y formulan recomendaciones motivadas. Desde esta perspectiva, la solución pasaría por una autorregulación responsable de las propias plataformas.

No obstante, resulta igualmente razonable inclinarse por el segundo modelo —opción claramente asumida por Isabel Giménez—, que inspira los esfuerzos europeos de regulación frente a la relativa inacción estadounidense. En este marco adquieren especial relevancia normativas comunitarias como el Reglamento de Servicios Digitales (*Digital Services Act*) y el Reglamento de Mercados Digitales (*Digital Markets Act*), que imponen a las plataformas obligaciones de protección de los usuarios y del interés público bajo la amenaza de fuertes sanciones, así como la *Online Safety Act* británica de 2023.

Este segundo enfoque justifica asimismo determinadas regulaciones de la inteligencia artificial destinadas a prevenir abusos. Resulta esencial garantizar la transparencia de los procesos de IA, desde la recopilación de datos hasta la obtención de resultados, incluyendo los algoritmos utilizados y las decisiones adoptadas. Se asume, además, que la IA debe desempeñar una función meramente auxiliar en la adopción de decisiones públicas, sin sustituir el juicio humano cuando estén en juego conflictos de derechos, valoraciones subjetivas o potestades discrecionales. Estas regulaciones pueden incluir la creación de organismos independientes encargados de supervisar el impacto de la IA y asegurar el cumplimiento de estándares éticos y jurídicos.

La opción por el modelo ponderado no responde únicamente a un análisis teórico comparado, sino también al hecho de que este es el modelo efectivamente adoptado en el contexto europeo. Mientras en Estados Unidos las plataformas digitales operan con un margen de autonomía sin precedentes, Europa

avanza hacia un reforzamiento de la regulación de la IA y de la moderación de contenidos, consciente de que el reto jurídico no consiste en frenar el desarrollo tecnológico, sino en encauzarlo de modo compatible con los principios del Estado de derecho.

Este intervencionismo regulador, derivado además de las obligaciones impuestas por el derecho europeo y de la necesaria coordinación entre autoridades competentes, no debe temer el control jurisdiccional europeo. Muy al contrario, especialmente en el ámbito electoral y de la participación democrática, el Convenio Europeo de Derechos Humanos —en particular el artículo 3 del Protocolo n.º 1— impone a los Estados obligaciones positivas para garantizar procesos electorales democráticos. De ello se desprende un compromiso normativo claro para afrontar los problemas planteados por Internet. La regulación, aun compleja desde el punto de vista técnico y novedosa en su diseño, debe apoyarse en marcos normativos europeos, como el Plan de Acción para la Democracia de la Comisión Europea de 2020 o la Ley de Servicios Digitales. Sin duda el libro de la profesora Giménez pretende ayudar a desembarazar en nuestro ordenamiento el camino de la aplicación o el desarrollo del derecho europeo, en especial, como decimos, la norma sobre Servicios digitales.

Todo ello exige encontrar un equilibrio delicado entre el fomento de la innovación tecnológica y la protección efectiva de los derechos fundamentales y de los principios que sustentan los Estados democráticos. Se trata, en definitiva, concluye Isabel Giménez su valioso estudio, de establecer marcos regulatorios que impulsen el progreso tecnológico sin sacrificar valores esenciales como la privacidad, la libertad de expresión y la igualdad ante la ley.

PREFACIO

La génesis de este libro se remonta a apenas un par de años atrás, cuando inicié este trabajo como ejercicio para la obtención de la plaza de titular de universidad. En aquel momento, la relación entre la Inteligencia Artificial y la desinformación se antojaba un tema casi experimental; hoy, la realidad ha corrido tanto que lo que era novedad es ya una urgencia.

Para una constitucionalista casi siempre centrada en las estructuras del Estado (o de las Comunidades Autónomas, o incluso de la Unión Europea), este estudio me ha permitido una reconciliación inesperada. Siempre he mantenido una cierta distancia con los derechos fundamentales, por percibirlos a menudo como conceptos excesivamente filosóficos, abstractos y, en ocasiones, inaprensibles para el análisis jurídico más pragmático. Sin embargo, este tema me ha servido de puente, ya que he podido abordarlos desde la cuestión que verdaderamente siempre ha despertado mi curiosidad intelectual: el ejercicio del poder. Analizar cómo este poder se manifiesta ahora a través de nuevas tecnologías y es ejercido por agentes nuevos —actores privados y algoritmos que escapan a las categorías tradicionales— me ha permitido disfrutar de la investigación desde el folio en blanco, obligándome a leer y aprender como si fuera la primera vez.

En este proceso de aprendizaje, mi gratitud más profunda es para Juan Ignacio Godino. Nunca imaginé que terminaríamos compartiendo un objeto de investigación de forma tan intensa. A él le debo no solo la lectura crítica de las sucesivas versiones de este trabajo, sino la paciencia infinita de haberme explicado mil veces conceptos tecnológicos que me resultaban completamente ajenos y no siempre son fáciles de entender para un jurista. Si el lector encuentra solidez en la parte técnica de esta obra, es gracias a su generosidad y a su empeño en que yo los comprendiera. Saber que esa parte está bien es, sencillamente, saber que él está ahí.

Asimismo, agradezco a los evaluadores del manuscrito sus valiosas sugerencias. Su labor generosamente silenciosa es la que garantiza que los originales alcancen la madurez necesaria para ver la luz.

Es para mí un honor que este libro venga precedido por las palabras de Juan José Solozabal Echavarría. Contar con él para el prólogo es un privilegio que va más allá de lo académico. Juanjo Solozabal es uno de los grandes, tanto por su talla como jurista como por su calidad humana. De él no solo se aprende la materia; con su ejemplo, uno aprende a estar en la academia, a habitarla con rigor y con esa elegancia que solo da la integridad. Su manera de entender y ejercer nuestra profesión ha sido el mejor espejo en el que mirarse.

Por último, el recuerdo debido a José Tudela Aranda, a Pepe Tudela. Es incalculable lo que le debo en mi carrera profesional y lo que el Derecho Constitucional en España le debe a su figura. Pepe fue el alma que convirtió la Fundación Giménez Abad en un foro de encuentro absolutamente esencial para todos nosotros, un lugar de acogida y debate que no habría sido posible sin su entrega. Siento una pena profunda al saber que ya no podré comentar las tesis de este libro con él. Realmente, su ausencia deja un vacío tan grande como su legado.

INTRODUCCIÓN

En la última década, y de manera exponencial en el último lustro, el desarrollo de la inteligencia artificial (IA) ha venido adquiriendo un enorme impacto en diversos sectores de la sociedad. La explosión de la IA ha transformado radicalmente numerosos aspectos, tanto de la vida cotidiana como del funcionamiento de las organizaciones, impulsando la extracción de conocimientos a partir de los datos, la toma de decisiones automática y la realización de ciertas tareas repetitivas que requieren de conocimiento experto. Asimismo, ha abierto nuevas oportunidades para la innovación, la ciencia y la tecnología y, en el ámbito económico, ha dado lugar a la aparición de nuevos modelos de negocio y a la reconfiguración de las llamadas cadenas de valor, lo que ha supuesto la generación de infinidad de oportunidades, pero también de no pocas incertidumbres en términos de empleo y distribución de la riqueza¹. Por otro lado, en el ámbito de la seguridad y la privacidad, el uso generalizado de sistemas de IA plantea interrogantes sobre la protección de datos personales, la discriminación algorítmica y el potencial uso dañino de estas tecnologías.

Durante los primeros compases del auge de la inteligencia artificial generativa, desde cierta doctrina se mantuvo la esperanza de alcanzar consensos amplios en torno a un modelo global de gobernanza de la IA². Pese a la evidente carrera tecnológica entre Estados Unidos y China, el desarrollo de foros multilaterales —como el *AI for Good Global Summit* (2017), la *Global Part-*

¹ En palabras de Brad Smith, Presidente de *Microsoft*, la inteligencia artificial «es el invento más importante para la mente humana desde la invención de la imprenta». *El País*, 20 de febrero de 2024. <https://elpais.com/tecnologia/2024-02-20/brad-smith-presidente-de-microsoft-debemos-tener-una-manera-de-ralentizar-o-apagar-la-inteligencia-artificial.html>

² JOBIN, Anna, IENCA, Marcello y VAYENA, Effy: «The global landscape of AI ethics guidelines», en: *Nature Machine Intelligence*, 1, 2019, 389-399.

nership on AI (2018), los *Principios de la OCDE* (2019), o las recomendaciones éticas de la UNESCO (2021)— alimentaba una visión todavía cosmopolita de la IA como un bien común tecnológico. Esta expectativa se sustentaba, en parte, en la idea de que la IA, como fenómeno transversal y transfronterizo, requeriría de una arquitectura normativa compartida, basada en principios éticos universales, estándares técnicos interoperables y mecanismos globales de responsabilidad. Sin embargo, conforme se fue consolidando la centralidad estratégica de la IA en la transformación económica, militar y geopolítica del siglo XXI, se hizo cada vez más evidente que los Estados comenzarían a tratar la IA no como un problema técnico global, sino como un proyecto nacional irrenunciable³. Este giro —que puede calificarse de nacionalismo de la IA— ha introducido una nueva lógica en el campo de la gobernanza tecnológica, caracterizada por la asimilación de la IA a los intereses de seguridad nacional, la soberanía digital y la competitividad internacional⁴. En consecuencia, países tradicionalmente no rivales han comenzado a discrepar sobre cuestiones regulatorias, éticas y jurídicas, limitando las posibilidades de un consenso multilateral.

³ A pesar de algunas iniciativas aisladas, como el Órgano asesor de la ONU sobre IA, que nace en octubre de 2023 por iniciativa directa del Secretario General de Naciones Unidas, António Guterres, quien lo crea como parte de su agenda para reforzar la cooperación digital global. No lo establece la Asamblea General ni un tratado, sino la propia Secretaría General, que tiene la capacidad de convocar grupos de alto nivel para orientar debates internacionales. Su propósito es asesorar a la ONU y a los Estados miembros sobre cómo gobernar la IA, identificando riesgos, oportunidades y principios comunes, pero también servir como puente entre gobiernos, empresas tecnológicas, academia y sociedad civil para impulsar un marco global coherente. El objetivo de fondo era evitar que la tecnología avance sin coordinación internacional y garantizar que su desarrollo sea seguro, ético y compatible con los derechos humanos, además de apoyar el desarrollo sostenible. IGLESIAS VELASCO, Alfonso J., *La protección internacional de la libertad de expresión. Los retos planteados por Internet y las redes sociales*, Tirant lo Blanch, 2025, pp. 258-260.

⁴ BAZOOBANDI, Sara, et al.: «From global governance to nationalism: The future of AI», en: *GIGA Focus Global*, 2, 2025, <https://www.giga-hamburg.de/en/publications/giga-focus/from-global-governance-to-nationalism-the-future-of-ai>. Algunos autores, no obstante, siguen defendiendo que, pese a lo complejo del objetivo, es inevitable avanzar hacia una gobernanza internacional de la IA que garantice su uso ético, seguro y equitativo, para lo que proponen la creación de una autoridad reguladora global que establezca normas, supervise su cumplimiento y coordine a los distintos actores involucrados: gobiernos, empresas, sociedad civil y organizaciones internacionales. Vid. ZAIDAN, Esmat e IBRAHIM, Imad Antoine: «AI Governance in a complex and rapidly changing regulatory landscape: A global perspective», en: *Humanit Soc Sci Commun*, 11, 2024, pp. 1-18.

De otro lado, la irrupción de *DeepSeek R1* en el mercado de IA ha evidenciado que la competencia en este ámbito trasciende la mera capacidad de inversión de las grandes corporaciones, situándose como un escenario donde la eficiencia en el uso de los recursos redefine las reglas del juego⁵. Este fenómeno, lejos de circunscribirse al plano tecnológico, se proyecta como un factor determinante en la dinámica geopolítica contemporánea, donde Estados Unidos, China y Europa rivalizan por consolidar su supremacía en este terreno estratégico. En este contexto, el enfoque estadounidense privilegia la libertad de mercado y el liderazgo del sector privado, bajo la premisa de que la ausencia de regulación específica constituye una ventaja competitiva, motivo por el cual la administración Trump ha sostenido que las restricciones normativas podrían obstaculizar la innovación, otorgando así mayor autonomía a empresas como *OpenAI*, *Google* y *Anthropic*. No obstante, esta desregulación plantea serias dudas respecto del poder que dichas corporaciones ejercen sobre el desarrollo tecnológico y las implicaciones que ello conlleva en relación con los derechos fundamentales. Por su parte, China ha adoptado un modelo basado en la dirección estatal, donde la inversión pública y la supervisión gubernamental impulsan a compañías como *Alibaba*, *Baidu* y *DeepSeek*. A diferencia de Estados Unidos, cuya competitividad radica en la iniciativa privada, el país asiático apuesta por un control centralizado que le permite avanzar con mayor velocidad, consolidando su influencia global⁶. Así, según el *AI Index Report 2025* de *Stanford HAI*, Estados Unidos continúa liderando en la producción de modelos avanzados de IA, con una inversión privada que en 2024 superó los 109 mil millones de dólares, muy por encima de China y Europa. Esta ventaja competitiva se traduce ahora en una política de exportación estructurada, que pretende institucionalizar la influencia estadounidense en los incipientes marcos re-

⁵ En este contexto, la aparición de *DeepSeek R1* ha trastocado el modelo existente al demostrar que es posible competir con modelos como *ChatGPT* mediante el reciclaje de *GPUs* (procesadores especializados que han adquirido un papel esencial debido a su capacidad para entrenar modelos de *machine learning* y *deep learning*) reduciendo así los costes de infraestructura y democratizando el acceso a esta tecnología. Este avance supone una oportunidad para que países con recursos limitados puedan desarrollar sus propios sistemas de IA, cerrando la brecha tecnológica que históricamente los ha separado de las potencias económicas.

⁶ De hecho, la firma del manifiesto promovido en febrero de 2025 por la Unión Europea en favor de una IA ética e inclusiva ha sido interpretada como un movimiento estratégico de China para mejorar su imagen internacional, sin que ello implique una alteración sustancial de su modelo autoritario.

⁷ Disponible en: <https://hai.stanford.edu/ai-index/2025-ai-index-report>

gulatorios emergentes de la IA a escala global, a través de una nueva política regulatoria aprobada en julio de 2025.

Por último, la Unión Europea (UE), en cambio, procura equilibrar el desarrollo tecnológico con la protección de los derechos fundamentales, apostando por una regulación que garantice la transparencia y el respeto a los derechos fundamentales en el uso de la IA⁸. No obstante, ahora se discute si a este respecto se extenderá o no el llamado «efecto Bruselas»⁹, pues no son pocos los que consideran que el exceso de regulación vendría a provocar que Europa pudiera perder su posición económica en una futura cuarta revolución industrial, motivando que la generación de riqueza se trasladase a otros países. O, en otras palabras, que sin ser potencia consumidora ni productora no habría efecto Bruselas¹⁰. Sin embargo, como analizaremos en su momento, la rigidez normativa europea ha sido objeto de críticas por su potencial impacto en la innovación y la atracción de inversiones.

La inteligencia artificial trasciende así su dimensión técnica y económica para convertirse en un instrumento de poder global, cuya influencia se extiende a ámbitos tan diversos como la ciberseguridad, la automatización de la guerra y la manipulación de la opinión pública. En definitiva, la convergencia de estas capacidades sitúa a la inteligencia artificial como un factor determinante en las relaciones de poder globales, en relación con las cuales mientras Estados Unidos y China buscan consolidar su liderazgo, Europa intenta evitar su marginación mediante iniciativas como *Mistral*, concebida como la respuesta de la Unión Europea a las grandes tecnológicas extranjeras¹¹. La capacidad de adap-

⁸ Así, también la propuesta del presidente francés Emmanuel Macron de articular una tercera vía mediante alianzas con países del Sur Global en febrero de 2025, que pretendía posicionar al continente como un actor relevante en la carrera tecnológica.

⁹ En expresión afortunada acuñada por Anu BRADFORD (*The Brussels effect: How the European Union rules the world*, New York Oxford University Press, 2020).

¹⁰ En cierto sentido, con los modelos de lenguaje ya ha empezado a pasar. *Google*, *Microsoft* o *Anthropic* han lanzado algunos de sus mejores modelos en todo el mundo salvo en Europa, donde llegaron con meses de retraso. Por ello los críticos de la regulación sostienen que si la IA ayuda a que las empresas mejoren su productividad o aumenten sus ingresos, Europa estará perdiendo competitividad al no tener acceso a la última tecnología, con el agravante de que no hay tecnología producida en Europa en este caso.

¹¹ Un ejemplo paradigmático de este planteamiento de refuerzo tecnológico europeo es la iniciativa *Mistral AI*, una empresa emergente francesa fundada en 2023 que se ha convertido en una de las piezas centrales del proyecto europeo para alcanzar una soberanía tecnológica efectiva en el ámbito de la inteligencia artificial generativa.

Mistral AI fue creada por tres antiguos investigadores de compañías líderes del sector —*Meta* y *DeepMind*— y se ha distinguido desde sus inicios por su apuesta por democratizar el acceso a

tarse con inmediatez a esta nueva realidad determinará las posiciones relativas de los Estados en el sistema global, mientras que desde una perspectiva jurídico constitucional el problema radica en conciliar el desarrollo tecnológico con la preservación de los valores democráticos fundamentales, evitando que la concentración del poder en manos de unas pocas entidades derive en formas de control que socaven los derechos y libertades de la ciudadanía¹².

De manera paralela (o más bien secante), se viene produciendo en nuestras sociedades un fenómeno de intoxicación de la opinión pública con falsedades,

modelos de lenguaje de gran escala, mediante el uso de código abierto y licencias permisivas. Su posicionamiento ha sido respaldado de forma explícita por el Gobierno francés, en particular por el presidente Emmanuel Macron, quien ha defendido reiteradamente la necesidad de que Europa no quede subordinada tecnológicamente a actores privados de otras regiones geopolíticas, especialmente Estados Unidos y China (PRESCOTT, Katie, «Arthur Mensch, the Frenchman leading the AI charge against US domination», *The Times*, 17 de junio de 2025, <https://www.thetimes.com/business-money/technology/article/arthur-mensch-the-frenchman-leading-the-ai-charge-against-us-dominance-bj6r55mwf>; SCHECHNER, Sam, «Fears over U.S. AI dominance boost business for France's Mistral», en: *Wall Street Journal*, 11 de junio de 2025, <https://www.wsj.com/tech/ai/fears-over-u-s-ai-dominance-boost-business-for-frances-mistral-14405419>).

Más allá de su valor como empresa, *Mistral AI* se inscribe en una política de colaboración público-privada impulsada por Francia y adoptada por otras instituciones europeas. En 2024, la compañía anunció la creación de un «campus paneuropeo» de IA en los alrededores de París —con apoyo del fondo soberano francés *Bpifrance*, la empresa de semiconductores *NVIDIA*, *Microsoft*, y el grupo emiratí *MGX*— destinado a convertirse en la mayor infraestructura de entrenamiento y despliegue de IA de Europa. Este tipo de alianzas representa un cambio cualitativo en la estrategia europea, que pretende pasar de un planteamiento esencialmente regulador (como en el Reglamento de Inteligencia Artificial) a una estrategia de construcción de capacidades autónomas, inspirada en el concepto de «autonomía estratégica abierta» auspiciado desde la Comisión Europea (MILLÁN, Santiago, «La francesa Mistral, respaldada por Macron, invita a Abu Dabi a una gran ronda de financiación para la IA europea», en: *Cinco Días*, 10 de julio de 2025, <https://cincodias.elpais.com/companias/2025-07-10/la-francesa-mistral-respaldada-por-macron-invita-a-abu-dabi-a-una-gran-ronda-de-financiacion-para-la-ia-europea.html>).

El caso de Mistral AI ilustra con claridad cómo las instituciones europeas están comenzando a diseñar un modelo de soberanía digital que no se limita al ámbito normativo, sino que pretende garantizar el control sobre las infraestructuras, los datos, los modelos y las capacidades humanas necesarias para sostener un ecosistema tecnológico autónomo. Una política con implicaciones jurídicas relevantes, tanto en materia de regulación de la IA como en la reconfiguración del papel del Estado (o público, en general) en la promoción de tecnologías estratégicas.

¹² BALAGUER CALLEJÓN, FRANCISCO, «La constitución del algoritmo. El difícil encaje de la constitución analógica en el mundo digital», en BALAGUER CALLEJÓN, F.; COTINO HUESO, L. (coords.), *Derecho público de la inteligencia artificial*, Fundación Manuel Giménez Abad, Zaragoza, 2023, pp. 29-56.

mentiras, noticias falsas y bulos que se propagan a gran velocidad gracias a las herramientas de comunicación personal (como *WhatsApp* o *Telegram*) o las redes sociales (*Facebook*, *Instagram*, *TikTok*, *X* —antiguo *Twitter*—, etc.). Así, estos dos fenómenos, *a priori* independientes, se han acabado relacionando, de manera que el desarrollo de la IA ha servido para generar contenidos falsos de imagen, audio, vídeo o texto —o incluso multimodales—, además de para potenciar los efectos de las noticias falsas, si bien, como también veremos, puede coadyuvar en su control.

De otro lado, ya hace unos años se acuñó el término de democracia digital (también conocida en sus versiones anglosajonas como *e-democracy*, *digital democracy* o *Internet democracy*), respecto de la cual se ha ido asentando el protagonismo de las innovaciones tecnológicas y su influencia en la participación política¹³. Dicho concepto engloba una compleja intersección entre los procesos democráticos tradicionales y las dinámicas sociales propiciadas por el avance tecnológico. En este sentido, se observa un creciente reconocimiento del papel protagonista que las innovaciones tecnológicas desempeñan en la configuración y redefinición de las formas de participación política y de ejercicio del poder en las sociedades contemporáneas.

De entrada, la democracia digital, como construcción conceptual, conlleva un profundo impacto en la esfera política, al introducir nuevas modalidades y plataformas de interacción ciudadana, así como también al transformar los mecanismos tradicionales de representación y deliberación. De una parte, la proliferación de herramientas digitales, tales como herramientas de comunicación personal, redes sociales, plataformas de participación ciudadana en línea, sistemas de votación electrónica y portales gubernamentales interactivos, entre otros, ha ampliado significativamente el espectro de posibilidades para la expres-

¹³ La doctrina especializada (vid. GARCÍA-OROSA, Berta, «Disinformation, social media, bots, and astroturfing: the fourth wave of digital democracy», en: *Profesional de la información*, v. 30, n. 6, 2022, e300603) ha identificado cuatro grandes periodos en esta evolución: Una primera, en la década de los años 90 del siglo XX, se caracteriza por la presencia de los actores políticos en internet, pero sin modificaciones destacables en la comunicación. Una segunda, entre 2004 y 2008, en la que prima la consolidación de las redes sociales y el incremento de la mercantilización de la participación de la audiencia. A partir de la campaña electoral de Obama en 2008 con la microsegmentación y el uso de datos masivos (*big data*) se puede datar la tercera ola. Por último, la cuarta ola comenzaría en 2016, a partir de la campaña a favor del Brexit y las acciones de *Cambridge Analytica* y supone el auge de la inteligencia artificial. No obstante, es bastante probable que haya que añadir una quinta etapa que cubra en momento actual, con el despliegue de la IA generativa.

sión y el ejercicio de la ciudadanía en el entorno virtual. Además, la democracia digital ha generado un fenómeno de democratización de la información y el conocimiento, al facilitar el acceso a datos gubernamentales, documentación normativa, debates públicos y otros recursos relevantes para el ejercicio de la ciudadanía informada. Esta apertura y transparencia en la gestión pública contribuye a fortalecer la rendición de cuentas, a fomentar la participación activa de la ciudadanía en los procesos de toma de decisiones y a promover una mayor responsabilidad por parte de las autoridades en el ejercicio del poder. No obstante, cabe señalar que la democracia digital también suscita reparos jurídicos en relación con la privacidad, la seguridad de los datos, la desinformación y la polarización política en línea. Muy especialmente, la proliferación de noticias falsas —con independencia de su modalidad—, la manipulación de la opinión pública mediante algoritmos y la exposición a discursos antiliberales, representan algunas de las amenazas que pueden erosionar la calidad del debate democrático y socavar la legitimidad de las instituciones políticas.

Este escenario nos sitúa ante el reto de redefinir marcos, instrumentos de análisis y, sobre todo, elaborar propuestas normativas para abordar estos fenómenos. Todo ello, en el contexto de una era digital que se caracteriza, esencialmente, por dos notas: su complejidad y la velocidad a la que todo ocurre. Estos dos rasgos parecen además desarrollarse de manera exponencial y provocan una brecha creciente entre la realidad y el sistema institucional. En este sentido, como señala Tudela Aranda, la rapidez con la que la tecnología transforma la realidad constituye un factor preeminente, que da lugar a una intensidad de cambio que, objetivamente, obstaculiza la capacidad de respuesta del poder público. Este fenómeno, a su vez, plantea nuevos retos a los sistemas democráticos, concebidos originalmente para gestionar entornos más estables, por lo que ciertamente se encuentran desfasados al enfrentarse a una realidad que demanda un dinamismo y flexibilidad extraordinarios¹⁴.

En este contexto del entorno digital contemporáneo surge la IA como un nuevo protagonista, un agente dotado de una dualidad compleja, que presenta toda una gama de oportunidades, pero también de riesgos, que lo mismo genera contenido auténtico que fomenta la proliferación de engaños. Sin ir más lejos, durante las primeras horas del conflicto bélico entre Israel y Hamás, una avalancha de vídeos e imágenes relacionados con el conflicto, muchos de los cuales se revelaron posteriormente como falsos, invadió la plataforma X (*Twit-*

¹⁴ TUDELA ARANDA, José, «Gobierno, Parlamento, Democracia e inteligencia artificial», en: *Teoría y realidad constitucional*, 52, 2023, p. 305.

ter), evidenciando un ejemplo elocuente de esta dualidad¹⁵. Así, a medida que avanzamos en la utilización de la IA para la creación y manipulación de contenido digital, nos enfrentamos cada vez más al riesgo omnipresente de la creación de noticias falsas y la proliferación de *deepfakes*, es decir, contenidos de imagen, audio, vídeo o texto generados mediante modelos de inteligencia artificial que simulan la apariencia de autenticidad¹⁶. Este fenómeno plantea una seria amenaza para la integridad de la información y la opinión pública, como tendremos ocasión de desarrollar más adelante, ya que dichas herramientas pueden ser empleadas de manera fraudulenta para distorsionar la realidad y sembrar la confusión entre los receptores de dicho contenido. Para ello, los creadores de estos contenidos se aprovechan de la presunción de autenticidad de lo percibido por los usuarios, de la dificultad que entraña verificar la autenticidad de un contenido, de la necesidad del ser humano de confirmar sus propios prejuicios, o del deseo natural de asomarse al lado más morboso de la presunta noticia.

En 2024, a medida que se desarrollaban procesos electorales en diversas partes del mundo (a nivel estatal pero también internacional), se denunciaba con creciente inquietud la proliferación de contenido político engañoso y perjudicial impulsado por la llamada inteligencia artificial generativa —de la que más adelante nos ocuparemos—, una tendencia que ha dado lugar a una preocupante escalada en la difusión de *deepfakes* más que convincentes. Ejemplos relevantes en los Estados Unidos¹⁷ y

¹⁵ SEIBT, Sébastien, «Attaque du Hamas contre Israël: premier crash-test de Twitter sous l'ère Elon Musk», *France24*. 12 de octubre de 2023, <https://www.france24.com/fr/%C3%A9co-tech/20231011-attaque-du-hamas-contre-isra%C3%ABl-premier-crash-test-de-twitter-sous-l-%C3%A8re-elon-musk>

¹⁶ A veces la doctrina politológica distingue entre *deepfakes* (o ultrafalsedades) y falsificaciones superficiales (*cheapfakes* o *shallowfakes*), para denominar en este último caso a medios audiovisuales manipulados de forma sencilla para distorsionar información. Cfr. RUBIO NÚÑEZ, Rafael, *et al.*, *Inteligencia artificial y campañas electorales algorítmicas: disfunciones informativas y amenazas sistemáticas de la nueva comunicación política*. Centro de Estudios Políticos y Constitucionales, Madrid, 2024, 127-136. La realidad es que de estos últimos supuestos ni siquiera nos ocuparemos en este trabajo, ya que no suelen tener su origen en inteligencia artificial.

¹⁷ A finales de enero de 2024, varios electores demócratas en New Hampshire recibieron llamadas telefónicas presuntamente generadas por IA, haciéndose pasar por el presidente Biden, donde se les pedía no votar en las primarias con el objetivo de reservar su voto para las elecciones de noviembre. Además, en 2023, el Comité Nacional Republicano empleó esta tecnología para crear un vídeo que presentaba imágenes de escenarios apocalípticos tras el anuncio de la candidatura de Joe Biden a la reelección. El gobernador Ron DeSantis de Florida difundió imágenes

Eslovaquia¹⁸ proporcionan muestras tangibles de cómo imágenes y vídeos falsificados tienen la capacidad de influir en la opinión pública, desafiando, de este modo, la propia integridad de los procesos electorales¹⁹. Además, la sucesión de ejemplos reales demuestra de manera palpable la amenaza que supone la IA al introducir elementos distorsionadores que podrían alterar de manera significativa el discurso político y, en consecuencia, manipular la adopción de decisiones en el ámbito electoral y afectar sustancialmente al derecho a la participación política y la propia idea de representación democrática. Especialmente, este fenómeno adquiere particular relevancia en un año como fue 2024, donde concurría un contexto de elecciones en Estados Unidos y la Unión Europea (UE) (pero también en el Reino Unido o en la India), lo que suponía un incremento de la utilización indebida de la IA con fines políticos.

Asimismo, la reciente proliferación masiva de contenidos generados por inteligencia artificial en el contexto del conflicto con Irán en marzo de 2026 pone de manifiesto un cambio cualitativo en la naturaleza de la desinformación. No se trata ya de la difusión ocasional de informaciones falsas, sino de la creación industrial de simulaciones bélicas verosímiles —imágenes, vídeos y narrativas completas— que construyen una realidad paralela destinada a influir

falsas del expresidente Trump, su rival político en las primarias, junto al Dr. Anthony Fauci, el exdirector del Instituto Nacional de Alergia y Enfermedades Infecciosas (*National Institute of Allergy and Infectious Diseases, NIAID*) entre 1984 y 2022 (y principal asesor médico durante la gestión de la COVID-19). Cfr. HSU, Tiffany, «New Hampshire Officials Investigate A.I. robocalls mimicking Biden». *New York Times*. 22 de enero, 2024. <https://www.nytimes.com/2024/01/22/business/media/biden-robocall-ai-new-hampshire.html>

¹⁸ Apenas dos días antes de las elecciones en Eslovaquia, celebradas el 30 de septiembre de 2023, se publicó en *Facebook* una grabación de audio en la que se escuchaban supuestamente las voces del líder del partido liberal y de una periodista, discutiendo sobre la forma de manipular el proceso electoral mediante la compra de votos de la minoría romaní. Dicha grabación de audio se había realizado a través de inteligencia artificial, pero fue publicada en las redes durante la jornada de reflexión, lo que impedía a los medios tradicionales abordar cuestiones referentes a la campaña y, por tanto, desacreditar la grabación. Además, se aprovechó de una laguna en la política de *Meta* en materia de recursos manipulados, que solo hace referencia a los vídeos, pero no se refiere a los archivos de audio.

¹⁹ En realidad, aquí procede remitirse a la completísima obra de Óscar Sánchez Muñoz, publicada en esta misma colección, ya que en este aspecto las *deepfakes* constituyen un ejemplo más de los efectos de la desinformación en las campañas electorales. SÁNCHEZ MUÑOZ, Óscar, *La regulación de las campañas electorales en la era digital: desinformación y microsegmentación en las redes sociales con fines electorales*, Centro de Estudios Políticos y Constitucionales: Universidad de Valladolid, Ediciones Universidad de Valladolid (EdUVa), 2020.

en la percepción pública del conflicto. Este fenómeno se ve amplificado por los sistemas algorítmicos de recomendación de las plataformas digitales, que reconfiguran el consumo informativo de los usuarios y favorecen contenidos emocionales y espectaculares, debilitando los mecanismos tradicionales de contraste y contextualización. La combinación de generación automática de contenidos y amplificación algorítmica transforma la desinformación en una auténtica infraestructura tecnológica de guerra informativa, con efectos directos sobre la deliberación democrática, la confianza pública y la toma de decisiones políticas.

En definitiva, la intensificación del uso inadecuado de la IA en el ámbito político impone la necesidad y la urgencia de abordar eficazmente la propagación de información manipulada y la desinformación que amenazan la integridad de los procesos democráticos. De la especial trascendencia de esta amenaza da cuenta la celebración de un acuerdo entre las principales empresas desarrolladoras de esta tecnología en la Conferencia de Seguridad de Múnich, celebrada en febrero de 2024²⁰. Dicho acuerdo se centraba explícitamente en un conjunto concreto de abusos de las *deepfake*, el relativo al «contenido electoral engañoso generado mediante IA», para enfrentarse al cual se comprometían a la adopción de ocho medidas, de las que más adelante algo diremos. De este modo, incluso las grandes empresas tecnológicas parecían haber cobrado súbita conciencia de la necesidad de prevenir la repetición de fenómenos similares a los ocurridos con la propagación consciente de noticias falsas y la manipulación de información deformada mediante IA con el fin de obtener ventajas electorales, una situación que se ha visto alterada tras la llegada por segunda vez de Donald Trump a la Presidencia de los Estados Unidos (2025). En todo caso, esta percepción de urgencia regulatoria se ha reiterado posteriormente desde algunos de los principales responsables del desarrollo de la IA, que han advertido de la proximidad de importantes riesgos asociados a su uso, no limitados ya al ámbito electoral, sino también en relación con el empleo, la ciberseguridad y el crimen, reforzando la idea de que el debate normativo debe también responder a la creciente conciencia de los grandes impactos sistémicos que esta tecnología va a suponer²¹.

²⁰ Es el conocido como Pacto de Múnich (o «AI Elections Accord») en el que intervinieron veinte de las principales empresas tecnológicas, como *Microsoft*, *Meta*, *Amazon*, *Adobe*, *IBM*, *Google*, *OpenAI*, *TikTok* y otras 12 compañías, incluida *X (Twitter)*.

²¹ En una entrevista el 7 de abril de 2026, el Director Ejecutivo de *OpenAI*, Sam Altman al analizar la rápida llegada de la superinteligencia afirmaba que el mundo, especialmente Washington, aún no está preparado para la transformación que se avecina. Disponible en: <https://www.youtube.com/watch?v=B21KxGs8zDI>.

Desde esta perspectiva, probablemente el principal riesgo que entrañan las tecnologías de las *deepfakes* para el orden social es el de erosión de las bases sobre las cuales se sustenta la convivencia social. Así, la mayoría de la sociedad, en cierto momento, concuerda en la existencia de una forma de verdad mutuamente aceptada y ampliamente difundida, y las realidades sociales se sustentan en la confianza compartida en esa verdad común. Sin embargo, este fenómeno, caracterizado por la manipulación digital de imágenes y videos para generar contenido falso e hiperrealista, pone en entredicho los pilares sobre los cuales se edifican las interacciones sociales y la percepción colectiva de la realidad, que ahora ya no tiene una base común, ni una confianza compartida en la existencia de unos hechos reales e incuestionables. En este escenario, la proliferación de *deepfakes* nos lleva a cuestionar la autenticidad y la integridad de la información en el ámbito público. Como resultado, se difumina la línea divisoria entre lo genuino y lo simulado, y se mina la confianza en la veracidad de los medios de comunicación y otras fuentes informativas. Adicionalmente, la difusión masiva de *deepfakes* puede acarrear consecuencias devastadoras para la sociedad, al socavar la cohesión social y exacerbar divisiones preexistentes mediante la propagación de desinformación y el fomento de la desconfianza entre los individuos y las instituciones. En última instancia, pues, el surgimiento de las *deepfakes* subraya la urgente necesidad de reconsiderar los fundamentos de la verdad y la confianza en la era digital.

Es importante recalcar que no se trata simplemente del fin de la verdad en sí misma, sino más bien del colapso de la creencia en la verdad: Una sociedad fundamentada en la desconfianza es el peligro al que podría conducir el avance generalizado de las *deepfakes*; un escenario hipotético de desinformación masiva socavaría la confianza en las instituciones, erosionaría la cohesión social y minaría los fundamentos mismos de la democracia al limitar la capacidad de la sociedad para llegar a acuerdos compartidos basados en una comprensión común de la realidad²². Es precisamente en esa amenaza a la propia amenaza a la verdad en donde reside el principal riesgo que entraña esta tecnología. En consecuencia, podemos afirmar que la proliferación de *deepfakes* plantea no solo un problema técnico o de seguridad, sino también una cuestión fundamental sobre la naturaleza misma de la verdad y la confianza en la sociedad contemporánea.

²² O, como señalaba Michael SANDEL: «Este es el verdadero peligro, no solo que cada vez sea más difícil distinguir lo que es real de lo que es falso, sino que esa distinción deje de importarnos». SANDEL, Michael, Entrevista en *Telos: Cuadernos de comunicación e innovación*, 122 (Ejemplar dedicado a: Posverdad), 2023, p. 28.

No obstante los riesgos apenas señalados, las *deepfakes* de imagen, audio y/o vídeo no son el único problema asociado al uso de la IA en este contexto de la desinformación. Así, entre los problemas adicionales que plantea el uso de la IA en relación con la desinformación se encuentran otros como el de la generación automatizada de contenido falso —los conocidos como «textos autónomos» o «textos generativos»— con una apariencia de autenticidad, lo que dificulta la detección tanto por parte de los usuarios como por parte de los proveedores de herramientas para compartir contenidos digitales.

Además, potencialmente más peligroso aún que las *deepfakes*, la IA no solo sirve para la creación de contenidos falsos, sino también para llevar a cabo conductas manipulativas en las redes sociales. En primer lugar, por ejemplo, mediante la manipulación de algoritmos de recomendación: Con el objetivo de amplificar la difusión de información falsa y polarizante, mediante la utilización de IA para personalizar recomendaciones de contenido con base en el comportamiento del usuario, lo que puede dar lugar a la creación de las llamadas «burbujas de filtro», que refuerzan las creencias sesgadas y facilitan la propagación de desinformación. O también a través de la manipulación de opiniones y comportamientos, ya que la IA se utiliza en técnicas como la creación de perfiles falsos en redes sociales, la difusión (y redifusión) coordinada de mensajes y la utilización de asistentes conversacionales (*chatbots*) automatizados para amplificar determinadas narrativas. Como se verá seguidamente, todas estas tácticas pueden influir en la percepción pública, minar la confianza en las instituciones y polarizar aún más el debate político y social, con el consiguiente daño para nuestro Estado democrático de Derecho.

Asimismo, puede afirmarse que el auge de los sistemas de IA (como *ChatGPT*, *Gemini* o *Bard*) —de los que ya nos ocuparemos en su momento— ha iniciado una transformación estructural del modelo económico y funcional del Internet abierto que hasta ahora conocíamos. Ello se debe fundamentalmente a que la forma en que los usuarios acceden a la información ha mutado sustancialmente: ya no se produce una navegación mediante enlaces hacia las fuentes originales, sino que las respuestas se ofrecen de manera directa y cerrada en el propio entorno del modelo. Este desplazamiento del tráfico en Internet hacia los grandes intermediarios algorítmicos está alterando el equilibrio que sostenía el ecosistema informativo desde los albores de Internet, especialmente en lo que respecta a la lógica de la referencia, el enlace y la autoría. De entrada, en términos económicos, la consecuencia inmediata es la crisis del sistema de monetización basado en la publicidad y en el número de visitas, que sostenía a medios de comunicación, plataformas colaborativas y espacios de conocimiento abierto como *Wikipedia*.

Desde una perspectiva jurídico-constitucional, esta mutación supone además una alteración profunda del equilibrio entre el derecho de acceso a la información (art. 20 CE) y la garantía del pluralismo informativo como condición de la democracia deliberativa. La concentración del tráfico en manos de unos pocos operadores digitales implica que la circulación de la información pública se somete a lógicas de intermediación no transparentes, guiadas por intereses comerciales y no por criterios de relevancia o veracidad. No en vano, como se verá más adelante, el Tribunal Europeo de Derechos Humanos (TEDH), ha reiterado que el pluralismo informativo constituye un valor estructural del orden democrático²³, de modo que su deterioro afecta directamente a la calidad del debate público.

Frente a quienes sostienen —como *Google* o *Meta*— que la IA impulsa una expansión del entorno digital mediante la generación de nuevos contenidos y consultas, lo cierto es que se plantea ya un riesgo inverso: la erosión del Internet abierto como espacio descentralizado y plural. El tránsito de una «Internet de referencias» a una «Internet de respuestas» tiende a convertir a los ciudadanos en consumidores pasivos de información predigerida, minando las condiciones materiales de la deliberación democrática. En palabras de Harari, el verdadero peligro no es que las máquinas piensen por sí mismas, sino que los seres humanos dejen de pensar críticamente por sí mismos, delegando su juicio en sistemas opacos y autorreferenciales²⁴.

En este nuevo escenario surge un fenómeno que debería preocuparnos: la llamada «bazofIA» (*AI slop*), expresión que designa la proliferación masiva de contenido generado por inteligencia artificial caracterizado por su baja calidad, su falta de rigor y su naturaleza repetitiva o manipuladora. De hecho, un estudio de *Graphite* (2025) estima que más del 50% de los nuevos artículos en línea son ya producidos por IA²⁵. Este volumen inédito de información automatizada no pretende necesariamente engañar, sino saturar el espacio informativo con ruido, desplazando el contenido veraz y dificultando la verificación y viene provocado, fundamentalmente por la multiplicación de las llamadas granjas de contenido automatizadas que reciclan textos originales, explotan vacíos de dominio y ponen en peligro el trabajo de periodistas y

²³ Vid. *Manole c. Moldavia*, 2009.

²⁴ HARARI, Yuval Noah, *Nexus: Una breve historia de las redes de información desde la Edad de Piedra hasta la IA*, Penguin Random House, 2024, pp. 464ss..

²⁵ Disponible en: <https://graphite.io/five-percent/more-articles-are-now-created-by-ai-than-humans>

creadores independientes. Desde una perspectiva constitucional, la «bazofIA» compromete tres principios esenciales: el derecho a la información veraz, el pluralismo mediático y la formación de una opinión pública libre²⁶. La imposibilidad práctica de distinguir entre contenido auténtico y contenido artificial, junto con la opacidad de los sistemas de recomendación, amenaza con sustituir la noción de verdad por la de mera verosimilitud algorítmica. Si la información deja de ser verificable, el propio fundamento epistémico del Estado de Derecho se ve comprometido, en tanto que la democracia necesita hechos compartidos sobre los cuales deliberar.

Además, el uso reiterado de estos contenidos como material de entrenamiento de futuros modelos de IA genera un bucle de retroalimentación de baja calidad, donde los sistemas aprenden de su propio ruido y degradan progresivamente la fiabilidad del entorno digital²⁷. Este fenómeno, descrito por algunos investigadores como colapso del modelo, podría conducir a una pérdida estructural de la diversidad informativa y a un empobrecimiento irreversible del conocimiento colectivo. De este modo, el Internet abierto —concebido originalmente como un espacio de libertad, cooperación y acceso universal al saber— corre el riesgo de transformarse en un entorno cerrado, autorreferencial y contaminado, donde la confianza, la autoría y la verdad se diluyen en una marea de falsedades y contenidos sintéticos.

En resumen, si bien las *deepfakes* constituyen un aspecto preocupante de la desinformación en línea, es importante reconocer que existen otros problemas asociados al uso de la IA en este contexto, que serán analizados en las siguientes páginas²⁸.

²⁶ Quizás también habría añadir aquí el riesgo que implica para la propia integridad cognitiva de los ciudadanos, por su directa conexión con la dignidad de la persona.

²⁷ SHUMAILOV, Iliia, «AI models collapse when trained on recursively generated data», en: *Nature*, 631, 2024, pp. 755-759.

²⁸ La Declaración de Bletchley, publicada a finales de 2023 como resultado de la *AI Safety Summit* en Buckinghamshire, Inglaterra, incidía en el potencial de la IA para mejorar el bienestar humano, la paz y la prosperidad. Sin embargo, también subrayaba los riesgos asociados con su mal uso intencional o la falta de control, entre los que destacaba de manera especial el de amplificación de la desinformación, que se presenta como uno de los más relevantes en este ámbito. La cumbre fue histórica, no solo por el consenso logrado sobre el futuro de la inteligencia artificial, abogando por un acuerdo internacional para eliminar —o al menos paliar— estos riesgos y resaltando su importancia crítica y la urgencia que suscitan, sino también por la participación poco común de China en estos acuerdos. Entre los 28 participantes en la cumbre se incluían Reino Unido, Estados Unidos, la Unión Europea (así como algunos de sus Estados miembros individualmente, como España), Brasil, India, Kenia, Japón, Arabia Saudí, Nigeria y Emiratos

Por otro lado, para contrarrestar eficazmente la desinformación, es esencial encontrar un equilibrio entre regular el contenido dañino y proteger la libertad de expresión y el pluralismo informativo. Se tratará de estudiar, pues, si en la era digital no es necesario ya repensar y actualizar las garantías de la libertad de expresión para abordar los nuevos desafíos y peligros asociados con algunas medidas de limitación de contenidos, especialmente los referidos a las políticas de moderación.

Sin embargo, antes de demonizar la IA, es preciso reconocer que también se cuenta con ella para encontrar posibles herramientas dirigidas a combatir la desinformación. De hecho, la aplicación de la IA con el propósito de prevenir abusos inherentes a su propio funcionamiento constituye una estrategia clave que se está empezando a desplegar en los últimos tiempos. En este sentido, se están introduciendo importantes herramientas, tales como el *Fact Check Explorer* de *Google*²⁹ y los verificadores de imágenes de *Instagram*, concebidas como herramientas fundamentales en la contención de la diseminación de información falaz. Más aún, la lucha contra la desinformación se ha fortalecido precisamente con la implementación de proyectos financiados desde la Unión Europea. Estos proyectos, caracterizados por su envergadura y alcance, se han presentado desde la propia UE como esenciales para controlar la propagación de desinformación. La relevancia de estas iniciativas impulsadas por la UE radica en su capacidad para afrontar los retos éticos y prácticos asociados a la utilización de la IA, no solo como herramienta para combatir la desinformación, sino también como un medio para salvaguardar la integridad de los procesos informativos en la era digital. Otra cuestión será si estos instrumentos resultan o no suficientes, así como la valoración de los posibles problemas que dichas herramientas pueden plantear.

Esta obra se articula, por tanto, en torno a dos grandes bloques temáticos que, aunque pueden analizarse por separado, se entrecruzan en múltiples planos: de un lado, la problemática de la desinformación y su impacto sobre el debate público, la calidad democrática y el marco de la libertad de expresión; de otro, los retos normativos, éticos y epistémicos que plantea el desarrollo de

Árabes Unidos. AI Safety Summit, «Declaración de Bletchley de los países que asisten a la Cumbre de seguridad de la IA», *gov.uk*. 1 de noviembre de 2023. <https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/dbc58681-1b68-47e0-8e3f-f91435fdf8ce>

²⁹ Google. «Fact Check Explorer». Disponible *online* en: <https://toolbox.google.com/fact-check/explorer/search/list:recent:hl=es>

la inteligencia artificial. Lejos de considerar estas dos dimensiones como compartimentos estancos, el propósito es mostrar cómo se afectan mutuamente, y cómo ciertas transformaciones tecnológicas —particularmente las asociadas a la automatización del procesamiento y la distribución de información— están alterando de manera profunda las formas tradicionales de circulación del conocimiento, de formación de opinión y de deliberación política.

Para poder abordar esa intersección con la precisión necesaria, la estructura de la obra avanza primero mediante un tratamiento autónomo de cada problemática, exponiendo sus lógicas internas, sus dilemas y su evolución reciente. Solo a partir de ese desarrollo previo es posible delimitar con claridad los puntos donde ambas convergen —como ocurre en fenómenos como la microsegmentación, la generación automatizada de contenido o la manipulación algorítmica de la atención— y extraer de ahí conclusiones normativas fundadas. Este enfoque supone, por tanto, una progresión deliberada: el lector encontrará capítulos que se centran en la desinformación sin hacer referencia expresa a la IA, y otros que abordan riesgos propios de la inteligencia artificial sin tratar directamente la cuestión de la desinformación. Esa aparente bifurcación está pensada como una estrategia metodológica orientada a construir, paso a paso, un marco de análisis más sólido para comprender la mutua imbricación entre ambos fenómenos.

Para finalizar, los últimos capítulos de esta obra estarán dedicados a exponer algunos mecanismos para combatir la desinformación causada por la IA, así como a analizar la regulación aprobada en estos últimos años.

CAPÍTULO 1

DESAFÍOS CONTEMPORÁNEOS: *FAKE NEWS*, POSVERDAD Y DESINFORMACIÓN EN LA ERA DIGITAL

La proliferación de noticias falsas, también conocidas como *fake news* —entendidas como información no veraz, engañosa o falseada presentada como noticias verídicas—, ha experimentado un notorio incremento en la era digital, cobrando especial relevancia con la digitalización de los medios y la omnipresencia de las herramientas de comunicación personal y las redes sociales en las sociedades democráticas contemporáneas. Por ello, no es de extrañar que desde la doctrina filosófica, politológica, sociológica y, en general, la especializada en medios de comunicación se hayan multiplicado los estudios que analizan el alcance de conceptos como *fake news*, postverdad, infodemia o desinformación.

En este contexto, la discusión sobre el término «*fake news*» resulta una cuestión tan compleja que algunos expertos incluso han sugerido evitarlo por considerarlo demasiado genérico, e incluso políticamente arriesgado¹. Concretamente, se ha argumentado que este término no refleja adecuadamente el aspecto más relevante y peligroso del fenómeno, que radica precisamente en su capacidad de difusión. En consecuencia, ante la crítica de que las noticias falseadas son solo la punta del iceberg en el fenómeno informativo, se prefieren otros términos, como el de «posverdad»², por ser más inclusivo. Así, mientras las *fake news* se centran en la fabricación y difusión de información falsa, la

¹ RUBIO NÚÑEZ, Rafael, «Los efectos de la posverdad en la democracia», en: *Revista de Derecho Político*, 103, 2018, p. 194-5.

² Si bien el concepto fue acuñado por el escritor checo Steve Tesich en 1992, la definición más extendida de posverdad es la de «aquellas circunstancias en las cuales los hechos objetivos son menos influyentes para moldear la opinión pública que apelar a la emoción y las creencias personales» (*Oxford English Dictionary*, Oxford University Press, 2016). De hecho, el término posverdad se incluyó en la actualización del Diccionario de la lengua española (DLE), de diciembre de 2017, tras haber sido distinguida en 2016 con el título honorífico de palabra del año por el diccionario inglés de Oxford.

posverdad hace referencia a un contexto en el que las emociones y las creencias personales influyen desproporcionadamente en la formación de opiniones, independientemente de la veracidad de la información³.

La Real Academia Española define la posverdad como la «distorsión deliberada de una realidad, que manipula creencias y emociones con el fin de influir en la opinión pública y en actitudes sociales». En esta definición se destacan los tres elementos fundamentales que caracterizan la posverdad⁴. En primer lugar, la posverdad está intrínsecamente vinculada con la mentira, por lo que se entiende como una distorsión consciente de la realidad, convirtiéndola en una práctica engañosa y manipuladora. La esencia de la posverdad radica pues en la alteración intencionada de los hechos, creando una narrativa falsa que busca moldear la percepción pública⁵. En segundo lugar, la intencionalidad es un componente esencial de la posverdad, ya que esta práctica no surge de manera accidental o incidental, sino que implica la acción deliberada de un agente o grupo de agentes. La distorsión de la realidad se realiza con un propósito claro: influir en la opinión pública, por lo que no es de extrañar que este fenómeno se manifieste con mayor frecuencia en el ámbito político, donde las narrativas manipuladas pueden tener un impacto significativo en la percepción de los ciudadanos. Y en tercer y último lugar, la posverdad no se limita a afectar solo las creencias o pensamientos de las personas, sino que se dirige directamente a sus emociones. De este modo, al apelar a las emociones, la posverdad busca generar respuestas y reacciones intensas, creando vínculos emocionales que refuerzan la aceptación de la narrativa distorsionada.

Así pues, la posverdad representa una práctica compleja que implica la distorsión intencionada de la realidad, la manipulación consciente de la percepción pública y la movilización de las emociones. Este fenómeno, prevalente en el ámbito político y mediático, plantea importantes cuestiones respecto de la

³ APARICI, Roberto, *et al.*, «Noticias falsas, bulos y trending topics: Anatomía y estrategias de la desinformación en el conflicto catalán», en: *Profesional de la Información*, 28 (3), 2019.

⁴ RODRÍGUEZ ANDRÉS, Roberto, «¿De qué hablamos cuando hablamos de posverdad?», en: *The Conversation*, 3 de octubre de 2023, <https://theconversation.com/de-que-hablamos-cuando-hablamos-de-posverdad-213481>

⁵ El periodista británico Matthew D'ANCONA, columnista de diversos diarios es autor de una de las obras de referencia al respecto (*Post-Truth. The new war on truth and how to fight back*. Londres: Ebury Press, Penguin Random House, 2017). En su obra, D'ANCONA denuncia «cómo ha ido decayendo progresivamente el valor de la verdad como divisa de reserva de la sociedad, y el contagio epidémico de un pernicioso relativismo disfrazado de legítimo escepticismo» (p. 2, traducción propia).

búsqueda de una información veraz y la formación de una opinión pública informada, donde la verdad ya no compite con la mentira, sino con otras «verdades»⁶. Es quizá este el aspecto más interesante de la noción de posverdad, entendida como un fenómeno más amplio, que trasciende la mera falsedad, donde la verdad se ve eclipsada por la emocionalidad y las creencias personales⁷. La era de la información y la comunicación ha dado lugar a una cultura emocional, donde las narrativas que provocan emociones tienen un especial impacto. A su vez, la posverdad se manifiesta en el declinar de los medios tradicionales y el consiguiente auge de los medios digitales⁸. Y, como consecuencia de todo ello, la comunicación, fundamental en la sociedad actual, se acaba convirtiendo en un escenario político, y la posverdad en un desafío a la verdad objetiva, que orienta a la política hacia una guerra de información. Esto es, la posverdad consiste en la subordinación y reorganización de los hechos desde una voluntad ideológica y política concreta, lo que requiere de un mecanismo de legitimación en el que se persigue que aquello que se siente se convierta, además, en la verdad⁹.

Por su parte, la desinformación consiste en la elección intencional de datos parciales, incompletos, alterados... En otras palabras, la desinformación tiene una mayor complejidad, en sí misma y para su identificación, que las *fake news*, ya que para ser eficaz en su objetivo no necesita ser completamente falso. Así, a falta de una definición aceptada de manera general, la Comisión Europea ha definido la desinformación como aquel contenido fabricado y distribuido «falso, inexacto o engañoso [...] diseñado, presentado y promovido intencional-

⁶ RUBIO, Daniel, «La política de la posverdad», en: *Política exterior*, 176, 2017, pp. 58-67.

⁷ En el fondo, es una vuelta al concepto subjetivo de verdad que proponía Kierkegaard, por oposición a las concepciones epistemológicas y metafísicas de la verdad. KIERKEGAARD, Søren, *Migajas filosóficas. El concepto de angustia. Prólogos*, Trotta, 2023, pp. 28ss.

⁸ IBÁÑEZ FANÉS, Jordi, «La posverdad no es mentira», en *En la era de la posverdad: 14 ensayos*, IBÁÑEZ FANÉS, J.; ARIAS MALDONADO, M. (coords.), Calambur, Barcelona, 2017, pp. 37-48.

⁹ «Verdad y mentira son conceptos filosóficos y, por tanto, sujetos a múltiples niveles de debate (ontológico, lógico, semántico, metafísico, ético, teológico, epistemológico, etc.). Tal como son presentadas aquí se refieren a verdades o mentiras fácticas. Esto es, la existencia o no de correspondencia entre enunciados y hechos. Cuando no existe tal correspondencia, pero se sostiene de forma intencional que sí existe, estamos ante las *fake news*, que es conceptualmente lo más próximo a lo que el sentido común identifica como las mentiras de la propaganda. En todo caso, no estamos en un debate sobre la verdad o sus regímenes, en términos foucaultianos, sino sobre los desórdenes informativos, su identificación y sus consecuencias». DEL FRESNO GARCÍA, Miguel, «Desórdenes informativos: sobreexpuestos e infrainformados en la era de la posverdad», en: *El profesional de la información*, v. 28, n. 3, 2019, p. 2.

mente para causar daño público o beneficios» particulares¹⁰, siendo así que ese engaño también puede tener por objetivo, a los efectos que nos ocupan, alterar la percepción de forma intencional de grandes grupos de personas o sociedades e influir en su comportamiento (político, económico, ideológico...). En definitiva, la desinformación se presenta como un fenómeno complejo que va más allá de las noticias falsas o falseadas¹¹, de los bulos, la tendenciosidad informativa y la manipulación, e involucra a los medios de comunicación (sean estos los clásicos o los surgidos de las redes sociales), con el propósito de influir en el comportamiento y las percepciones de la ciudadanía. Este fenómeno, vinculado a conceptos como posverdad e infodemia, se convierte en un problema político con consecuencias hasta cierto punto impredecibles¹². La desinformación se manifiesta pues como un proceso con productos (noticias falsas), estrategias de distribución y respuestas ciudadanas, periodísticas o institucionales¹³. Pero si hay un ámbito donde quizá este fenómeno alcanza una mayor gravedad es en las campañas electorales, amenazando la seguridad electoral y poniendo en riesgo la integridad de los procesos democráticos¹⁴.

Por último, no son pocos los autores que prefieren crear una categoría común de todos estos fenómenos (desinformación, *fake news*, hechos alternativos, posverdad...), a la que denominan «desórdenes informativos» y que se definiría como «producciones intencionales cuya estrategia consiste en la fabricación de la duda y falsas controversias con el fin de conseguir beneficios económicos o

¹⁰ COM(2018) 236 final, «La lucha contra la desinformación online: un enfoque europeo», *Comisión Europea*, 26 de abril de 2018, 10.

¹¹ Algunos autores incluyen, dentro de la categoría de *fake news* a los llamados Sistemas de Recomendación de Contenidos (como *Taboola*, *Outbrain* o *RevContent...*), ya que al trivializar las noticias «blandas» falsas (aludiendo a la distinción anglosajona entre *hard news* —con un contenido puramente informativo— y *soft news* —con un tono más informal, que puede incluir elementos de entretenimiento), suponen una degradación del panorama informativo, pues acaban menoscabando también la confianza en los medios de información. FILLOUX, Frederic, «Fighting fake news is a losing battle, but there are other ways to win the war», en: *Monday Note*, 6 de agosto de 2018.

¹² LÓPEZ LÓPEZ, Paolo Carlos, *et al.*: «La desinformación en las democracias de América Latina y de la península ibérica: De las redes sociales a la inteligencia artificial (2015-2022)», en: *Uru: Revista de Comunicación y Cultura*, 75, 2023, pp. 69-89.

¹³ GUALLAR, Javier, *et al.*: «Desinformación, bulos, curación y verificación: revisión de estudios en Iberoamérica 2017- 2020», en: *Telos: Revista de Estudios Interdisciplinarios en Ciencias Sociales*, vol. 22, n. 3, 2020, pp. 595-613.

¹⁴ SÁNCHEZ MUÑOZ, Óscar. *La regulación de las campañas electorales en la era digital: desinformación y microsegmentación en las redes sociales con fines electorales*, *op. cit.*, pp. 81-92.

ideológicos»¹⁵. Dichos desórdenes informativos se caracterizarían por estar interrelacionados entre sí y muy especialmente por su necesaria dependencia de las tecnologías post Internet¹⁶. Si bien encuentran su origen en distintas causas que han facilitado su desarrollo, alcance e impacto actual, lo cierto es que estos desórdenes informativos han venido a modificar la naturaleza misma de la comunicación interpersonal colectiva.

En todo caso, a los efectos que aquí nos interesan, todos estos conceptos representan una auténtica amenaza para la calidad e integridad de la información en la sociedad actual. Por ello, no puede considerarse solo un problema cuantitativo, sino que también es claramente cualitativo, en cuanto que ello supone que «hemos entrado en una nueva fase del combate político e intelectual, donde las ortodoxias y las instituciones democráticas se ven sacudidas hasta sus cimientos por una oleada de alarmante populismo»¹⁷. Los riesgos que este fenómeno entraña para los sistemas democráticos han sido puestos de manifiesto ya desde hace unos años por la doctrina académica, pero también por numerosas instituciones nacionales e internacionales y, de manera muy especial, por la Unión Europea, desde que ya en el año 2018 la Comisión publicase la Comunicación «La lucha contra la desinformación online: un enfoque europeo».

Si bien es cierto que la generación deliberada de confusión informativa no constituye estrictamente una novedad histórica¹⁸, el elemento revolucionario que ha hecho sonar todas las alarmas, viene determinado, antes que por la IA,

¹⁵ Por todos, vid. WARDLE, Claire; DERAKHSHAN, Hossein, «Information disorder toward an interdisciplinary framework for research and policymaking», *Consejo de Europa*, 27 de septiembre de 2017. Dichos autores subdividen esta categoría general en tres subcategorías: *dis-information* (desinformación), *mis-information* (información errónea) y *mal-information* (información maliciosa). Por su parte, la propia Claire WARDLE, ha creado una imagen que ilustra de forma gráfica el ecosistema de la desinformación, vid. «Fake news. It's complicated», en: *First Draft News*, 16 de febrero.

¹⁶ La recopilación y análisis de falsas informaciones revela diversas estrategias utilizadas en su creación, como falsa atribución, exageración de hechos, manipulación de imágenes, fabricación de *fake news* y suplantación de identidad (*counterfeit*).

¹⁷ D'ANCONA, Matthew, *Post-Truth. The new war on truth and how to fight back*, op. cit., p. 7 (traducción propia).

¹⁸ «La producción de desórdenes informativos de forma intencional no es una novedad estricta de nuestro tiempo, por mucho que lo sean algunos conceptos, ya que la tecnología de cada época ha jugado un papel crítico a la hora de cambiar la escala e impacto de estos desórdenes en las sociedades. (...) Paul Virilio (1991, 1995) ya había evidenciado como toda tecnología lleva implícito, en sí misma, su propio accidente y, debido a la misma evolución tecnológica, un creciente *potencial catastrófico*». DEL FRESNO GARCÍA, Miguel, op. cit., p. 2.

por la introducción de Internet y las redes sociales. Ese nuevo elemento es, en definitiva, el de su volumen, su incremento exponencial provocado por las características del medio: por su extraordinario alcance, su velocidad y capacidad de distribución. Esas dimensiones de expansión de la información, mayor aún en el caso de las falsedades¹⁹, implica consecuentemente una multiplicación en los efectos y las consecuencias que son susceptibles de provocar, que ya se miden a escala global²⁰.

Por otro lado, si bien suele responsabilizarse a las redes sociales²¹ y a la clase política de esta situación, lo cierto es que los medios de comunicación «tradicionales» también han tenido su cuota de responsabilidad, en esta «crisis de la verdad»²², provocada fundamentalmente por la búsqueda incesante de contenidos sensacionalistas y la falta de rigor en la presentación de información²³. En

¹⁹ VOSOUGHI, Soroush, «The spread of true and false news online», en: *Science*, 359, 2018, pp. 1146-1151.

²⁰ El alcance de dicha amenaza es sin embargo bastante cuestionado, ya que algunos autores consideran que la trascendencia real de la amenaza de la desinformación digital en la opinión pública no se ve respaldada por los resultados empíricos, que vendrían a indicar que la desinformación es un problema muy relativo, con un alcance limitado entre el público. *Vid.*, JUNGHERR, Andreas; SCHROEDER, Ralph, «Disinformation and the structural transformations of the public arena: Addressing the actual challenges to Democracy», *Social Media+Society*, enero-marzo de 2021, p. 2.

²¹ Como señala Vallespín Oña: «El cambio tecnológico que recibe el nombre de *digital turn* ha provocado una reorganización profunda del mundo de la política o, más en particular, de la democracia. Por lo pronto, su efecto más inmediato ha consistido en producir una verdadera convulsión en el espacio público tradicional, constituido hasta muy recientemente por los medios de comunicación tradicionales como prensa, radio y televisión. A ellos se han superpuesto en nuestros días nuevos lugares de debate y comunicación política más o menos espontánea en blogs y redes sociales que están centrifugando la información y, sobre todo, la opinión». Cfr. «La gestión de la democracia en el entorno digital: política y nuevas redes», en *Telos: Cuadernos de comunicación e innovación*, 100 (Ejemplar dedicado a *La era digital: Balance y tendencias*), 2015, p. 45.

²² La revista *Time* le dedicó a esta cuestión una portada el 23 de marzo de 2017, en la que —sin imagen alguna— aparecía en letras rojas sobre fondo negro el siguiente titular: «Is truth dead?» (un homenaje tipográfico a su icónica portada «Is God Dead?» de 8 de abril de 1986. PINE, D. W., «Is Truth Dead? Behind the TIME Cover» *Time*, <https://time.com/4709920/donald-trump-truth-time-cover/>

²³ Así, podríamos destacar algunas de las actuaciones que han ido pavimentando el camino hacia la situación actual:

- i. Sustitución del rigor informativo por el equilibrio de versiones: Se ha producido un cambio de enfoque de los medios de comunicación, que dejaron de buscar y presentar hechos verificables para favorecer la objetividad y equilibrio entre diferentes versiones de la realidad. La búsqueda de imparcialidad llevó a la presentación de debates de

conjunto, pues, estas prácticas han afectado la credibilidad de los medios de comunicación, con las consiguientes consecuencias negativas para la sociedad al contribuir a la desinformación y la polarización. Por otro lado, tampoco el periodismo es ajeno a la IA y, de hecho, está librando su propia batalla para evitar desaparecer a manos de esta²⁴.

La doctrina destaca una serie de fenómenos que han venido coadyuvando a la actual situación de generalización de la desinformación y favoreciendo este fenómeno aparentemente imparable de las noticias falsas²⁵. Por un lado, la irrupción de las nuevas tecnologías ha traído consigo una difuminación entre el mundo real y el virtual, que se ve amplificada por un tercer componente, la ficción, que compite en su capacidad de influencia en la construcción de percepciones. Además, la cultura audiovisual ha alterado estructuralmente la manera en que se percibe la información, fundamentalmente porque la narración

opiniones y pseudo debates, generando el llamado *sesgo del equilibrio*, «cuando pasó de aplicarse de la presentación de opiniones divergentes a enfrentar los hechos científicos con el negacionismo intencional». DEL FRESNO GARCÍA, Miguel, «Desórdenes informativos...», *op. cit.*, 6.

- ii. Confrontación y falta de debate informado: La búsqueda de titulares impactantes y debates polarizados puede contribuir a una simplificación excesiva de problemas complejos, lo que no favorece una comprensión profunda de los temas.
- iii. Maximización de beneficios económicos: La priorización de la rentabilidad sobre la calidad informativa ha llevado a la difusión de noticias y titulares engañosos para atraer audiencia, con la consiguiente erosión en la confianza pública en los medios.
- iv. *Clickbait*: El *clickbait* (o ciberanzuelo, en la versión castellana), esto es, la práctica de utilizar titulares atractivos, pero a menudo engañosos para atraer clics, se ha convertido en una estrategia común en la prensa digital. Si bien puede aumentar el tráfico, ha sido al precio de socavar la integridad informativa y contribuir a la propagación de información errónea.
- v. *Infotainment* político: La *fusión* de información y entretenimiento en el ámbito político ha tenido como objetivo atraer a audiencias más amplias, pero a menudo a expensas de la calidad del debate político y la presentación precisa de los hechos.

²⁴ De hecho, son muchos los medios de comunicación tradicionales los que oficialmente admiten contar con la IA para la realización de distintos trabajos hasta ahora realizados por periodistas, hasta el punto de que no resulta exagerado afirmar que «[e]n breve, el público no será capaz de distinguir la autoría». MANFREDI SÁNCHEZ, Juan Luis; UFARTE RUIZ, M^a José, «Inteligencia artificial y periodismo», en: *Revista CIDOB d'Àfers Internacionals*, 124 Desinformación y poder: la crisis de los intermediarios, 2020, pp. 49-72.

²⁵ Vid., especialmente, ARIAS MALDONADO, Manuel, «Understanding Fake News: Technology, Affects, and the Politics of the Untruth», en: *Historia y comunicación social*, 24 (2), 2019, 533-546; y RUBIO NÚÑEZ, Rafael, «Los efectos de la posverdad en la democracia», *op. cit.*, 204ss.

visual rodea el texto con connotaciones emocionales, lo cual a su vez dificulta la explicación de conceptos complejos. Esta transformación ha convertido la comunicación en un espectáculo que premia la simplicidad y los titulares engañosos, al tiempo que la difusión masiva de imágenes ha contribuido al éxito de las *fake news*, otorgándoles una apariencia de infalibilidad.

A ello, asimismo, hay que añadir los tiempos vertiginosos en los que se desarrolla la nueva realidad digital. Esta nueva concepción del tiempo contribuye al éxito de las *fake news*, ya que la exigencia de actualización constante en la búsqueda de atención y tráfico en línea provoca que la información se elabore y distribuya de manera automática, sin las comprobaciones y reflexiones necesarias, primando la velocidad sobre la veracidad. Pero es que, además, la inmediatez también presenta una segunda derivada en cuanto que la capacidad de almacenamiento infinita y la accesibilidad instantánea permiten que cualquier declaración sea expuesta y, en ausencia de contexto, se convierta en objeto de *fake news*. Por otro lado, se produce la paradoja de que, pese a vivir en un momento donde todo es público, se genera un exceso de información que dificulta su comprensión, por no hablar de que temas aparentemente más llamativos pueden eclipsar los realmente importantes.

Volviendo a otra cuestión que ya se ha mencionado con anterioridad, la confusión de opinión e información, que distorsiona la transmisión de información, es otro fenómeno que contribuye a la propagación de la posverdad. La opinión disfrazada de información y la falta de claridad entre diferentes tipos de verdades también contribuyen a la confusión informativa. Además, como también se ha mencionado, las tecnologías de la información han priorizado el pensamiento emocional respecto del lógico, con el problema añadido de que este tipo de pensamiento es inherente a atajos lógicos —conocidos como sesgos cognitivos²⁶— que hacen que el sentido de percepción se vea influenciado por las circunstancias individuales del receptor, lo cual condiciona a su vez la selección de estímulos recibidos. En definitiva, por mor del sesgo de confirmación²⁷,

²⁶ Concepto introducido por Daniel Kahneman y Amos Tversky para hacer referencia a aquellos efectos psicológicos que producen una desviación en el procesamiento mental, conduciendo a una distorsión, o una interpretación ilógica. TVERSKY, AMOS; KAHNEMAN, DANIEL, «Availability: a heuristic for judging frequency and probability», en: *Cognitive Psychology*, 5(2), 1973, pp. 207-232.

(Vid. <https://www.elnotario.es/hemeroteca/revista-42?id=487:los-sesgos-cognitivos-y-el-derecho-el-influjo-de-lo-irracional-0-53842293707507>)

²⁷ Según el cual las personas tienden a favorecer, buscar, interpretar y recordar información de manera que confirme sus creencias preexistentes. Concepto acuñado por WASON, Peter C.;