

El círculo vicioso de la Inteligencia Artificial Generativa

La teoría del Internet Muerto –*Dead Internet theory*–

Por Elian Facundo Yorlano¹

Resulta redundante ahondar en cuestiones que ya son cotidianas para la gran mayoría de las personas que realizan ciertas actividades. Así, podríamos concluir que a estas alturas la mayoría de los lectores de este artículo, han oído hablar de ChatGPT, e incluso se han aventurado, en menor o mayor medida a intentar colaborar por esta novedosa tecnología, en el afán de tomar algún atajo temporal sin dañar productividad. Ahora bien, ¿todo lo que produce una IA generativa es real? ¿Y comprobable? Hoy día se pueden producir textos, imágenes, videos, audios, ya sea *from scratch* (desde cero) o a partir de un recurso que le damos. Más aún, si utilizamos su información sin verificación, ¿contribuimos a un fenómeno que amenaza la autenticidad de la red? Así llegamos a la Teoría del Internet Muerto –*Dead Internet Theory*–, un concepto que vincula la proliferación de contenido artificial con la erosión de la confianza online, a la vez que analizaremos sus implicaciones legales y tecnológicas. El tradicional refrán de «No creas todo lo que te dicen» toma intensidad en estos tiempos de IA, y podemos concluir «No creas todo lo que ves, lees o escuchas».

teoría del internet muerto – inteligencia artificial – deepfakes – bots

a. ¿Qué es la Teoría del Internet Muerto?

i. Orígenes y Fundamentos

La *Teoría del Internet Muerto* postula que gran parte del contenido en línea ya no es creado por humanos, sino por sistemas automatizados (bots) e inteligencia artificial (IA)². De acuerdo a quienes apoyan esta teoría, esto generó un ecosistema digital donde las interacciones auténticas son

marginales, y la información está dominada por algoritmos con agendas ocultas, desde manipulación comercial hasta desinformación política, incluso con maniobras que podrían llegar a inducir ciertos comportamientos en el turismo.³

Si bien algunos aducen que la Teoría del Internet Muerto es solo una hipótesis conspirativa, muchos otros afirman que es un fenómeno respaldado por datos cuantificables. Su origen, según afirman (ya que no es totalmente confiable su génesis),

¹ Abogado egresado por la Universidad Nacional de La Plata. Autor del libro “Inteligencia Artificial para abogados y ChatGPT” (2024) y distintos artículos que relacionan el Derecho y la tecnología. Correo electrónico: efacundoyorlano@gmail.com

² Ahmed *et al*, 2024.

³ Ozgurel *et al*, 2025.

se dio en distintos foros como *Agora Road's Macintosh Café* en 2021, cuando un usuario argumentó que los gobiernos y distintas corporaciones usan bots y algoritmos para manipular narrativas.

Aunque la veracidad de este relato fundacional es cuestionable —algunos señalan que el usuario IlluminatiPirate podría ser un personaje ficticio creado para satirizar la paranoia digital—, investigaciones posteriores han identificado fenómenos paralelos que sustentan la teoría. Por ejemplo, El informe “*Bad Bot Report 2024*”⁴ de Imperva -una empresa enfocada en ciberseguridad- confirma que 49.6% del tráfico web no es humano (32% bots maliciosos + 17.6% bots benignos), respaldando la premisa central de la teoría. Desde 2013, el tráfico de bots maliciosos aumentó del 23.6% al 32% en 2023, mientras el tráfico humano cayó al 50.4%. Dentro de estos números, podríamos diferenciar entre los *bots* costosos y los *low cost*. Estos últimos en especial son utilizados para:

- Web *scraping* masivo: 40% del contenido en Reddit es generado por IA, contaminando entrenamientos de modelos futuros. El *scraping* de datos, de un modo general, se refiere a una técnica en la cual un programa informático extrae datos del resultado generado por otro programa.⁵
- Fabricación de perfiles falsos: 44.8% de los bots maliciosos se disfrazan de navegadores móviles para imitar usuarios reales.
- Reservas fraudulentas: Diversos bots usan IA para acaparar reservas en restaurantes u otros servicios y revenderlas, un fenómeno que ya afecta al 31.1% del sector entretenimiento.

Por ejemplo, una práctica frecuente es la del *seat-spinning*. De manera particular, esto afecta a la industria de aerolíneas, y prevalece en Asia. Los *bots* reservan los asientos sin efectuar un pago, para luego

revenderlos sin haber hecho una inversión. Por eso, muchas veces pareciera que un vuelo está totalmente reservado y se libera “mágicamente” algunas horas antes de su despegue. Esto, obviamente, produce daños económicos y una merma en la reputación de distintas aerolíneas.

ii. El caso Elon

En octubre de 2022, el empresario Elon Musk compró Twitter (ahora X). Por aquel entonces, el empresario cuestionó las cifras de bots en la plataforma, y una de sus primeras medidas fue quitar a los que producían spam. Antes, había postado “*If our twitter bid succeeds, we will defeat the spam bots or die trying!*” («Si nuestra oferta por twitter tiene éxito, derrotaremos a los bots de spam, ¡o moriremos en el intento!»)⁶. Según él, cumplió con su cometido.

No obstante esta purga, la limpieza masiva de Twitter/X no resuelve el trasfondo del problema, ya que la dinámica de crear contenido automatizado se mantiene en otras redes y foros, e incluso en el mismo X, aunque de manera más encubierta. Mientras no exista un protocolo global, y cada plataforma adopte sus propias normas internas —muchas veces opacas—, la tendencia a la automatización y manipulación informativa seguirá latente. Esto puede conducir a la manipulación o inducción de cierta información a todos los usuarios, como por ejemplo cuando -nuevamente- Musk hizo una encuesta sobre si Donald Trump debería ser elegido nuevamente como Presidente de los Estados Unidos -lo que finalmente ocurrió en enero de 2025-. En noviembre de 2022, la encuesta arrojó un 51,8% por el “Yes”, aunque se verificó que entre 8% y 14% de las interacciones con este posteo fueron bots.⁷ ¿Y si alguien se dejó llevar por el resultado de esta encuesta manipulada para reafirmar su voto unos años después? (En esta encuesta, hay más de 15 millones de votos).

Las redes sociales crean una opinión popular en el inconsciente colectivo, y esto

⁴ Imperva, 2024.

⁵ Ver [aquí](#).

⁶ Ver [aquí](#).

⁷ Tan *et al*, 2023.

conduce a tomar ciertas decisiones o no hacerlo. Algunos académicos plantean que los algoritmos de las redes sociales, al mostrarnos solamente opiniones favorables a nuestro punto de vista inicial, contribuyen a la polarización y radicalización y a que nos encerremos en cámaras de eco o burbujas de auto-confirmación.⁸

iii. El caso “Jesús Camarón”

El caso “Jesús Camarón” -por más extraño que suene-, se refiere a un conjunto de publicaciones, perfiles de redes sociales e interacciones online atribuidas a un sujeto del mismo nombre. El fenómeno generó cierta controversia en comunidades virtuales dedicadas a la ciberseguridad, donde se especuló que el contenido —artículos, comentarios y reseñas— estaba siendo gestionado de forma sistemática a través de bots o scripts automatizados.⁹

Para ilustrar esta presunta manipulación, cabe destacar algunos rasgos característicos:

- *Alta frecuencia de publicación:* Las cuentas vinculadas a “Jesús Camarón” publicaban artículos e interacciones con una cadencia muy superior a la de un usuario promedio.
- *Homogeneidad temática:* Los temas abordados mostraban un patrón casi idéntico en forma y fondo, sugiriendo un guion o plantilla de generación de contenido.
- *Ausencia de verificación personal:* No se ha logrado constatar la existencia tangible de un usuario que se haga llamar o se identifique con las imágenes difundidas, más allá de los entornos virtuales en los que opera.

Si se parte de la base de que “Jesús Camarón” es una construcción digital automatizada, el caso se torna ilustrativo de las premisas de la Teoría postulada. Un perfil supuestamente real, pero cuyas publicaciones e interacciones provendrían de herramientas automatizadas, encaja en la

idea de un ecosistema virtual nutrido por entidades no humanas. A partir de que algunas plataformas cambiaron la configuración de sus algoritmos de visualización de *timeline* (pantalla principal), la generación de estos *bots* fue favorecida.

iv. Deepfake en guerra

En marzo de 2022, en el marco del conflicto entre Rusia y Ucrania, circuló en redes sociales un video donde el presidente ucraniano Volodymyr Zelensky aparecía “pidiendo a sus soldados que se rindieran ante las tropas rusas”^{10 11}. Sin embargo, dicho video no era auténtico; se trataba de un *deepfake*, esto es, un contenido audiovisual manipulado con tecnologías de inteligencia artificial generativa para imitar la voz y la apariencia de una persona real. La difusión masiva del video se vio reforzada por redes coordinadas de bots que replicaron el contenido en múltiples plataformas.

Varios analistas de ciberseguridad destacaron cómo en cuestión de horas el video falseado se compartió miles de veces, multiplicándose gracias a cuentas automatizadas y perfiles falsos de redes sociales. El propósito era confundir a la población y minar la moral de las tropas ucranianas, explotando la credibilidad que la imagen de Zelensky tenía en ese momento.¹²

Pocas horas después de la publicación, distintos funcionarios gubernamentales ucranianos emitieron desmentidos oficiales, y hasta Zelensky mismo apareció en un video auténtico denunciando la manipulación, lo cual ayudó a disipar la confusión general.

El incidente fue controversial y advirtió sobre la urgencia de regular la creación y difusión de *deepfake* con fines de desinformación, y mostró el rol de campañas coordinadas (incluyendo *bots*) para amplificar contenido falso. Además de esto, puso en evidencia la dificultad de identificar a los responsables reales cuando

⁸ Colina, 2023.

⁹ DiResta, 2024.

¹⁰ Ver [aquí](#).

¹¹ Video disponible [aquí](#).

¹² Twomey, 2023.

estos operan desde múltiples jurisdicciones o bajo el amparo de actores estatales.

Si bien el *deepfake* no surgió únicamente de bots, estos fueron fundamentales para la proliferación de la desinformación, contribuyendo al clima de confusión que refuerza la idea de que parte del contenido en línea está manipulado o automatizado. Efectivamente, un Internet Muerto es el que se contraría con el fin que tradicionalmente siempre tuvo: la información. Esto llegó al punto que las principales redes sociales han quitado el contenido, sabiendo que este podría crear un caos generalizado.¹³

Otros ejemplos de manual de *deepfake* han salido a la luz en los últimos tiempos, ridiculizando situaciones poco probables, entre ellos: el Papa Francisco usando un abrigo que parecía salido de un desfile de Milán¹⁴, Barack Obama dando discursos que no son reales¹⁵, e incluso estos dos personajes junto con Donald Trump usando armas¹⁶. Un *deepfake* puede difundirse tan ágilmente debido a su grado de polémica y asombro que cause en la sociedad, pero el desmentir tal evento probablemente signifique un esfuerzo vastamente mayor.¹⁷

b. La perpetuación del círculo vicioso: IA generativa y retroalimentación de datos

Para comprender cómo la Teoría del Internet Muerto se ve amplificada por la Inteligencia Artificial Generativa, conviene observar el círculo de retroalimentación que se produce. Las IA, como ChatGPT, se nutren de abundantes cantidades de información disponible en la red, recopilando datos de foros, redes sociales, periódicos digitales, bibliotecas virtuales y otras fuentes. El problema surge cuando ese “combustible” de entrenamiento no está verificado o, peor aún, ya ha sido contaminado por información inexacta creada por bots u otros sistemas automatizados. Así, el contenido que generan estos modelos puede perpetuar las

falacias y errores presentes en el original. Luego, la información resultante —repetida, alterada o expandida— vuelve a circular en internet, siendo potencialmente utilizada para entrenar a futuras generaciones de IA, agravando el problema.

Por ejemplo, un blog automatizado publica supuestas “noticias” sobre una marca de autos eléctricos, citando estadísticas falsas que resalta la autonomía de un modelo que no existe. Dichas publicaciones aparecen en búsquedas y son tomadas por modelos de IA durante su etapa de entrenamiento. Con el tiempo, estos modelos replican datos erróneos cuando se les consulta sobre autos eléctricos, y terceros las publican en redes sociales o en nuevos blogs. El círculo se mantiene y se retroalimenta.

3. Desinformación y fake news: la amenaza de la manipulación masiva

La proliferación de contenido falso, generado o amplificado por bots e IA, repercute en la calidad de la información y, por ende, en la capacidad de los usuarios para discernir lo verdadero de lo que no lo es. Numerosos casos recientes ilustran cómo la desinformación, reforzada por algoritmos de recomendación y bots de amplificación, puede afectar procesos democráticos, decisiones de consumo o incluso la salud pública.

i. Ejemplos concretos de desinformación

Campañas electorales

Un informe de *The Oxford Internet Institute* (2021)¹⁸ demostró que en 81 países se han detectado campañas de manipulación digital organizadas, donde miles de cuentas automatizadas replican mensajes falsos o sesgados, buscando influir en la opinión pública.

¹³ Ver [aquí](#).

¹⁴ Ver [aquí](#).

¹⁵ Ver [aquí](#).

¹⁶ Ver [aquí](#).

¹⁷ Yorlano, 2023.

¹⁸ Bradshaw, 2021.

Salud y vacunas

Durante la pandemia de COVID-19, la Organización Mundial de la Salud (OMS) alertó sobre la “infodemia”¹⁹: la saturación de contenido falso o impreciso que provocó confusión y desconfianza. Bots automatizados y granjas de trolls fueron señalados como factores agravantes en la propagación de teorías conspirativas.

Economía y finanzas

Influencers falsos o automatizados en redes sociales financieras (como Twitter/X y foros de inversión) promueven criptomonedas dudosas o acciones “milagrosas”, generando burbujas especulativas o manipulación de precios. Un estudio del MIT Sloan School of Management²⁰ encontró que gran parte de los consejos sobre economía y salud en Twitter provenían de perfiles automatizados que no correspondían a una identidad humana real.

c. Reacciones y contramedidas legales

La proliferación de estos fenómenos no ha pasado inadvertida para los gobiernos y los organismos internacionales, que han iniciado planes reguladores y de control, intentando mitigar el impacto de la desinformación sistemática y del contenido automatizado malicioso.

i. Legislación y proyectos normativos

Unión Europea (UE): El Digital Services Act (Reglamento (UE) 2022/2065)

Establece obligaciones para las grandes plataformas digitales en cuanto a la transparencia de algoritmos y la rendición de cuentas sobre contenido ilegal o perjudicial. Aunque no se dirige exclusivamente a la IA generativa, sí introduce responsabilidades amplias para evitar la manipulación informativa.

Estados Unidos

Existen proyectos legislativos como la Algorithmic Accountability Act, que buscan

imponer mayores requisitos de auditoría y transparencia en los modelos de IA empleados por grandes empresas tecnológicas.

ii. Impacto en la industria y en los usuarios

Las compañías tecnológicas enfrentan una creciente presión para moderar el contenido automatizado, lo cual implica un coste operativo significativo y, a menudo, controversias sobre la censura y la libertad de expresión.

Los usuarios, por su parte, deben adaptarse a un entorno en el que la información puede ser creada sin intervención humana y con propósitos poco claros. En respuesta, han proliferado iniciativas de alfabetización digital y verificación de datos, como los talleres de MediaWise o las guías de FactCheck.org.

d. Implicaciones tecnológicas: ¿hacia un nuevo paradigma o una burbuja?

El auge de la Inteligencia Artificial Generativa y su potencial de crear contenido de manera ilimitada ha llevado a plantearse si estamos a las puertas de un nuevo paradigma tecnológico o frente a una burbuja similar a la de las puntocom de principios de los 2000. Numerosos académicos sostienen que la IA puede liberar a los creadores de tareas repetitivas y potenciar la innovación, pero existe el riesgo de que la sobreabundancia de información no verificada genere más confusión que progreso. Hacia donde vamos es algo incierto, pero algo es seguro, la innovación es cada vez más ágil y explosiva, ¿se puede hacer algo o solo acostumbrarnos a vivir con medios contaminados y un Internet cuasi muerto?

e. Corolario

En definitiva, la Teoría del Internet Muerto plantea una realidad inquietante

¹⁹ Ver [aquí](#).

²⁰ Stackpole, 2019.

sobre el estado actual de la red, donde la proliferación de contenido automatizado y la manipulación algorítmica amenazan la autenticidad y la confianza en la información digital. A través de casos concretos como el fenómeno de "Jesús Camarón" como hemos reseñado, o la difusión de *deepfakes* en contextos geopolíticos, se evidencia cómo los *bots* y la IA generativas han erosionado la veracidad de los contenidos en línea, generando un entorno virtual donde distinguir entre lo auténtico y lo artificial resulta cada vez más complejo. Esto no solo afecta la percepción pública, sino que también tiene implicancias jurídicas y éticas que requieren una respuesta regulatoria y técnica adecuada, para proteger los derechos de los usuarios y preservar la integridad informativa.

En este contexto, es mandatorio que tanto los gobiernos como las plataformas digitales implementen mecanismos de verificación de contenido y adopten normativas que regulen el uso de la IA generativa. Asimismo, la educación digital y la concientización sobre la manipulación informativa son clave para que los usuarios desarrollen una mirada crítica frente al contenido que consumen. Solo a través de un enfoque integral, que combine la regulación, la tecnología y la alfabetización digital, será posible mitigar el impacto del contenido automatizado y evitar que la red se convierta en un espacio dominado por la desinformación y la manipulación algorítmica.

f. Bibliografía

Ahmed, A and others. "The Dead Internet Theory: Investigating the Rise of AI-Generated Content and Bot Dominance

in Cyberspace" (2024). NED University of Engineering & Technology

Bradshaw, S and others (2021). "Industrialized Disinformation. 2020 Global Inventory of Organized Social Media Manipulation". University of Oxford.

Colina, C. . (2023). Manipulación algorítmica y sesgo psicosocial en redes sociales. *Temas De Comunicación*, (46), 6–26.

DiResta, R & Goldstein. J (2024); "How Spammers and Scammers Leverage AI-Generated Images on Facebook for Audience Growth"; Stanford Internet Observatory

Imperva, a Thales Company. "Bad Bot Report" (2024)

Ozgurel, G. y otros. "Dead Internet Theory in Theoretical Framework and its possible effects on Tourism" (2025). *Journal of Lifestyle and SDGs Review*

Stackpole, B (2019). "AI ain't for everyone — who trusts bots, and why". MIT Sloan School of Management

Tan, Z and others (2023). "BotPercent: Estimating Bot Populations in Twitter Communities". University of Notre Dame and others.

Twomey, J. and others (2023) "Do deepfake videos undermine our epistemic trust? A thematic analysis of tweets that discuss deepfakes in the Russian invasion of Ukraine"

Yorlano, Elian Facundo (2023). "Deepfake y responsabilidad", *Diario Judicial*. <https://www.diariojudicial.com/news-94656-deepfake-y-responsabilidad>