A JOURNAL BY SISSA MEDIALAB

JCOMAL

ENSAYO

Usos éticos de la Inteligencia Artificial para la comunicación de la ciencia, con una perspectiva decolonial y un enfoque de Derechos Humanos

Gabriela Frías-Villegas

Resumen

Este ensayo examina los usos éticos de la inteligencia artificial (IA) en la comunicación pública de la ciencia, entendida como un modo fundamental de observar el derecho humano a la ciencia. Propone entender a la IA no solo como una herramienta eficiente para los comunicadores de la ciencia, capaz de redactar, de recrear la ciencia o de generar imágenes científicas, sino como un sistema sociotecnológico que puede reproducir sesgos de género, lingüísticos y geopolíticos, amplificando desigualdades epistémicas ya existentes. El ensayo subraya la necesidad de integrar marcos éticos y decoloniales en las prácticas de comunicación de la ciencia para garantizar el acceso democrático al conocimiento científico en la era digital.

Palabras clave

Comprensión pública de la ciencia y la tecnología; Estudios sociales de la ciencia y la tecnología; Popularización de la ciencia y la tecnología

Fecha de recepción: 21 de octubre de 2025
Fecha de aceptación: 7 de julio de 2026
Fecha de publicación: 12 de noviembre de 2026

1 ▪ Introducción

El derecho humano a la ciencia — reconocido desde hace más de medio siglo en la Declaración Universal de Derechos Humanos [Naciones Unidas, 1948] y en el Pacto Internacional de Derechos Económicos, Sociales y Culturales [Naciones Unidas, 1966] — exige que el conocimiento científico llegue a todas las personas de manera accesible, comprensible y libre de discriminación. La comunicación pública de la ciencia es un vehículo crucial para hacer efectivo ese derecho. En las últimas décadas, los estudios sobre comunicación de la ciencia han experimentado una transformación profunda: de los museos, escuelas y centros culturales como espacios centrales de divulgación, hemos pasado a entornos digitales donde internet, las redes sociodigitales y los buscadores en línea median el acceso cotidiano al conocimiento científico.

En ese contexto digital, la irrupción de las herramientas de inteligencia artificial (IA) representa un punto de inflexión. Si bien la IA existe como campo desde mediados del siglo XX — el científico computacional John McCarthy acuñó el término en 1956 [Stanford News, 2011] —, fue a partir de noviembre de 2022, con el lanzamiento masivo de ChatGPT [OpenAI, 2022], cuando los modelos generativos se volvieron accesibles al público general, desencadenando una expansión acelerada de su uso en la creación y circulación de contenidos. Desde entonces, la IA ha comenzado a transformar tanto la búsqueda de información como la producción de textos e imágenes, con consecuencias directas para la comunicación de la ciencia.

Las personas comunicadoras de la ciencia pueden apoyarse en estas herramientas para redactar y simplificar contenidos, adaptar explicaciones a distintos públicos, generar imágenes científicas o realizar traducciones; sin embargo, su uso plantea desafíos éticos que este ensayo busca examinar desde una perspectiva decolonial y de derechos humanos [Kessler et al., 2025].

Las problemáticas asociadas a estas dimensiones son considerables. Como documentan Afroogh et al. [2024], la confianza — o su ausencia — actúa como regulador central de la adopción de la IA: depende no solo de que los sistemas funcionen correctamente, sino de que la ciudadanía comprenda qué hacen, cómo lo hacen y qué valores incorporan. La comunicación de la ciencia ocupa aquí un lugar estratégico: es el puente entre el desarrollo técnico de la IA y la capacidad de la sociedad para evaluarla, cuestionarla y participar en las decisiones sobre su uso.

Por otra parte, como advierten Hendriks et al. [2025], la IA generativa produce respuestas a partir de patrones estadísticos en sus datos de entrenamiento, sin comprender realmente su contenido; esto la hace propensa a generar resultados incompletos, sesgados o erróneos con aparente seguridad, lo que puede distorsionar gravemente la comprensión pública de la ciencia. Además, dado que estos modelos se entrenan con la información disponible en internet, replican y amplifican sesgos estructurales preexistentes — de género, lingüísticos y geopolíticos — que ya afectaban la representación del conocimiento científico antes de la IA.

Por ello, este ensayo busca contribuir al debate latinoamericano sobre el papel de la IA en la democratización del conocimiento científico, argumentando que su uso en la comunicación de la ciencia debe estar guiado por marcos éticos, decoloniales y de derechos humanos. Se analizará cómo las dinámicas del extractivismo de datos replican lógicas colonialistas, invisibilizan saberes del Sur Global y perpetúan formas de injusticia epistémica que los comunicadores de la ciencia tienen la responsabilidad de identificar y contrarrestar.

2 ▪ Contexto

Aunque el científico computacional John McCarthy acuñó el término *inteligencia artificial* en 1956, fue a partir de 2022 cuando los modelos generativos se volvieron accesibles al público general. De acuerdo con Villegas Servín [2025], la IA es "un campo de la computación enfocado en desarrollar sistemas o máquinas capaces de realizar tareas que normalmente requieren inteligencia humana", incluyendo el aprendizaje automático y la toma de decisiones.

Entre las herramientas más relevantes para la comunicación de la ciencia se encuentran los *large language models* (LLM), entrenados con miles de millones de palabras para predecir y generar texto coherente. Es fundamental subrayar que estos modelos no "entienden" en el sentido humano: imitan el lenguaje a partir de patrones en sus datos de entrenamiento, lo que los hace propensos a generar resultados incompletos, sesgados o erróneos con aparente seguridad — las llamadas "alucinaciones" — con consecuencias potencialmente graves para la comprensión pública de la ciencia [Hendriks et al., 2025]. Entre los más difundidos se cuentan ChatGPT (OpenAI), Llama (Meta), Gemini (Google DeepMind), DeepSeek y Manus. En la comunicación de la ciencia, estas herramientas pueden apoyar la redacción, la simplificación de temas complejos, la traducción y la creación de materiales para distintos públicos; sin embargo, dado que se entrenan con la información disponible en internet, replican y amplifican sesgos estructurales preexistentes — de género, lingüísticos y geopolíticos — que ya afectaban la representación del conocimiento científico.

Algunas de estas IA generadoras de texto, como ChatGPT, han tenido tanto éxito en estas tareas que incluso han superado variantes restringidas del test de Turing — propuesto por el matemático Alan Turing en 1950 para evaluar si una máquina puede imitar de forma convincente el comportamiento humano en una conversación —, hasta el punto de que sus respuestas resultan prácticamente indistinguibles de las producidas por personas [Mei et al., 2024; Moskvichev et al., 2023]. Esta capacidad constituye un arma de doble filo para la comunicación de la ciencia: si bien permite generar textos de alta calidad, accesibles y adaptados a distintos públicos, también entraña el riesgo de que información incorrecta o sesgada sea presentada con la misma fluidez y aparente autoridad que un texto bien fundamentado, dificultando que las audiencias distingan entre contenido confiable y contenido engañoso.

Las IA generadoras de imágenes y video constituyen otro conjunto de herramientas de creciente uso en la divulgación científica. Modelos como GPT Image 1.5 (OpenAI), Midjourney o Veo 3 (Google DeepMind) producen resultados frecuentemente indistinguibles de la fotografía o el video convencionales, como ilustran la imagen *The Electrician* de Boris Eldagsen — ganadora del Sony World Photography Award 2023 hasta que su autor reveló su origen artificial [Barbera, 2024] — y la proliferación de imágenes al estilo Studio Ghibli que se viralizaron globalmente en marzo de 2025, reabriendo el debate sobre derechos de autor y estilos artísticos protegidos [Fortune, 2025]. Más allá de su potencial creativo, cada consulta a estos sistemas conlleva una huella energética e hídrica considerable, aspecto que la comunicación de la ciencia tiene la responsabilidad de visibilizar [A. S. Luccioni et al., 2022; S. Luccioni et al., 2024].

Todas estas herramientas ofrecen oportunidades reales para democratizar la comunicación del conocimiento científico, pero su implementación acrítica puede reproducir desinformación, sesgos y exclusiones. De ahí la necesidad de abordarlas con marcos éticos, decoloniales y de derechos humanos, que son el eje central de este ensayo.

3 ▪ La Inteligencia Artificial y los Derechos Humanos

Los riesgos descritos en la sección anterior no son meramente teóricos, pues tienen una dimensión de derechos humanos que es central para quienes nos dedicamos a la comunicación de la ciencia. El acceso al conocimiento científico está reconocido como derecho humano desde hace más de medio siglo, tanto en la Declaración Universal de Derechos Humanos [Naciones Unidas, 1948] como en el Pacto Internacional de Derechos Económicos, Sociales y Culturales [Naciones Unidas, 1966]. De este mismo tronco deriva el derecho a la tecnología, que Haochen Sun [2020] caracteriza como una "bella durmiente": un derecho al que instituciones, gobiernos y empresas tecnológicas han prestado insuficiente atención, pese a que la tecnología permea todos los ámbitos de la vida — incluida la producción y circulación del conocimiento científico.

En ese marco, la comunicación de la ciencia no es un complemento opcional del quehacer científico, sino el vehículo indispensable para hacer efectivos esos derechos. Como proponen Frías-Villegas et al. [2024], la comunicación de la ciencia debe ser "un vehículo para dar cumplimiento al derecho humano de beneficio, comprensión, información y participación en la construcción de la ciencia", entendida como un proceso de cocreación y diálogo mediante el cual se generan entendimientos comunes y se transforman los marcos conceptuales de quienes participan. Esta definición implica que toda persona tiene derecho a acceder a información científica precisa, comprensible y libre de sesgos que excluyan a cualquier grupo por motivos de género, origen, idioma, religión o edad.

Cuando la comunicación de la ciencia se realiza con herramientas de IA, este imperativo ético se vuelve más urgente. La Recomendación sobre la Ética de la Inteligencia Artificial de la UNESCO [2021] reconoce las profundas repercusiones de la IA en el pensamiento, las interacciones y la toma de decisiones, e insta a los Estados a utilizarla para mejorar el acceso al conocimiento, garantizar la libertad de expresión y combatir la desinformación mediante la alfabetización mediática. Varios países latinoamericanos — entre ellos México, Argentina, Brasil, Colombia y Costa Rica — han adoptado esta recomendación. En el ámbito regional, la Red Iberoamericana de Protección de Datos ha emitido lineamientos sobre el tratamiento de datos en IA [2020] y sobre neurodatos [2023], ambos avalados por la OEA. Es importante mencionar también la *Ley Olimpia*, conjunto de reformas legislativas en México que reconocen la violencia digital como delito y sancionan la difusión no consentida de contenido íntimo [Congreso de la Unión, 2021]. Este marco normativo resulta especialmente relevante para los comunicadores de la ciencia ante el avance de los *deepfakes* generados con IA: contenidos audiovisuales manipulados que pueden suplantar voces, rostros o declaraciones de personas reales con el riesgo de que circulen como divulgación científica legítima. Conocer los marcos legales vigentes es, por tanto, parte de la responsabilidad ética de quienes producen contenidos científicos mediados por tecnologías generativas.

Sin embargo, estos marcos regulatorios son condición necesaria pero no suficiente: en la práctica cotidiana de producción de contenidos científicos con IA pueden producirse violaciones a los derechos humanos — por sesgos, desinformación o exclusión — si no se aplican criterios éticos y de responsabilidad. Por ello, el enfoque de derechos debe complementarse con una perspectiva decolonial e intercultural que identifique y desactive jerarquías epistémicas, lingüísticas y culturales, aspecto que se desarrolla en la siguiente sección.

4 ▪ La Inteligencia Artificial, el giro descolonial y la comunicación de la ciencia

Si los marcos de derechos humanos descritos en la sección anterior establecen el horizonte normativo para el uso ético de la IA en la comunicación de la ciencia, el pensamiento decolonial ofrece las herramientas críticas para identificar por qué ese horizonte está aún lejos de cumplirse. Comunicar ciencia con IA no es un acto neutro: es un acto situado en estructuras de poder que determinan qué conocimientos circulan, en qué idiomas, desde qué referentes culturales y para qué públicos. El concepto de "giro descolonial", acuñado por Maldonado-Torres [2008], señala que la modernidad no solo produjo avances técnicos sino también mecanismos sistemáticos de deshumanización, y que los saberes de los pueblos más afectados por ese proyecto son fundamentales para comprender las formas contemporáneas de poder. Aplicado a la IA, este enfoque permite desmitificar la aparente neutralidad de herramientas desarrolladas mayoritariamente por corporaciones del Norte Global, cuyos sistemas reproducen los sesgos — de género, raza, clase, idioma y cosmovisión — de quienes los diseñan [Tierra Común Network, 2023].

La evidencia empírica ha refutado con contundencia esa ilusión de imparcialidad. Los sesgos de los sistemas de IA reproducen lo que Lugones [2007] describe como la lógica colonial de dominación, determinando qué científicas y científicos son representados, qué tradiciones de conocimiento son reconocidas como válidas y qué comunidades pueden acceder a información científica en su propia lengua. Estos sesgos no emergen con la IA: son el reflejo amplificado de estructuras previas de exclusión en la producción y circulación del conocimiento, que la IA hereda al entrenarse con información disponible en internet, donde las mujeres científicas, los saberes indígenas y las lenguas minoritarias están sistemáticamente subrepresentados. Cuando una herramienta de IA invisibiliza a una científica, omite un conocimiento tradicional o no puede procesar una lengua originaria, no está cometiendo un error técnico aislado: está reproduciendo y legitimando, a escala masiva y con apariencia de objetividad, las mismas exclusiones que durante siglos han definido quién cuenta como productor válido de conocimiento. Por ello, la comunicación de la ciencia tiene la doble tarea de alfabetizar críticamente a sus audiencias sobre el funcionamiento y los límites de la IA, y de formar a sus propios practicantes para que reconozcan estos sesgos antes de incorporar acríticamente sus resultados.

Como señala Joana Varón, los sistemas de IA constituyen "regímenes de datos" que imponen un orden epistemológico occidental y configuran un epistemicidio basado en la aniquilación de la diversidad racial y lingüística [Tierra Común Network, 2023]. Esta lógica se materializa en el extractivismo de datos [Couldry & Mejias, 2019]: de manera análoga a la extracción histórica de recursos naturales del Sur Global, hoy se extraen masivamente datos, cultura, idiomas y saberes comunitarios para entrenar modelos de IA sin el consentimiento ni la retribución de las comunidades que los generaron, definiendo desde el Norte qué cuenta como conocimiento legítimo.

Esta dinámica genera formas específicas de injusticia epistémica [Ricaurte, 2019]: una injusticia testimonial, cuando la IA otorga menor credibilidad a saberes fuera del canon occidental, y una injusticia hermenéutica, cuando no procesa lenguas originarias, privando a sus hablantes de participar en el entorno digital desde sus propias cosmovisiones. Ambas se hacen visibles en lo que Ruha Benjamin [2020] denomina "el nuevo código de Jim": tecnologías que profundizan inequidades existentes precisamente porque se perciben como objetivas y progresistas.

Frente a este panorama, un marco decolonial resulta indispensable para la comunicación de la ciencia en América Latina. Como argumentan Mohamed et al. [2020], dicho marco permite vincular la opresión algorítmica con contextos sociopolíticos más amplios, ir más allá de las definiciones de daño propias de Norteamérica o Europa, y diseñar estrategias comunicativas con mecanismos inclusivos de supervisión y reparación.

5 ▪ Sesgos, alucinaciones e información falsa

Uno de los usos más relevantes de la inteligencia artificial (IA) en la comunicación de la ciencia es la búsqueda de información. No obstante, resulta indispensable revisar las respuestas que obtenemos, pues sistemas como Gemini, ChatGPT y DeepSeek pueden presentar diversos problemas, como proporcionar información falsa, incorrecta o sesgada [Aspray & Hayes, 2011].

Los sesgos en los sistemas de IA son tendencias sistemáticas e injustas que pueden manifestarse al procesar información o al tomar decisiones, lo que conduce a resultados perjudiciales para ciertos individuos o grupos. En esencia, la IA aprende de los datos con los que es entrenada y, si dichos datos reflejan prejuicios o desigualdades existentes en el mundo real — o en la forma en que fueron recopilados —, no solo los replicará, sino que con frecuencia los amplificará. Es importante subrayar que estos sesgos no nacen con la IA: son el reflejo de estructuras de exclusión preexistentes en la producción y circulación del conocimiento que la IA hereda al entrenarse principalmente con información disponible en internet, donde determinados grupos, voces y tradiciones de conocimiento están sistemáticamente subrepresentados. Un ejemplo de sesgo como política deliberada es la censura de contenidos, observada recientemente en DeepSeek [Yang, 2025]. Al preguntarle sobre la Plaza de Tiananmén, en Pekín, China — donde en 1989 tuvieron lugar manifestaciones estudiantiles reprimidas por el ejército —, la respuesta fue: “Lo siento, no puedo responder a esa pregunta. Si tienes otras preguntas, estaré encantado de ayudarte” [DeepSeek, 2026].

Otro sesgo frecuente se relaciona con la representación de la apariencia de las personas: cuando se solicita a una IA que dibuje a una “mujer hermosa”, suele producir imágenes con rasgos estereotípicamente europeos, maquillaje y peinados normativos, contribuyendo a estándares de belleza inalcanzables para niñas y jóvenes. Este impacto es tan significativo que la marca Dove lanzó en 2024 la campaña “El Código”, comprometiéndose a no utilizar imágenes generadas por IA para representar personas reales [Dove, 2024].

Más allá de este ejemplo, Ramírez Autrán [2023] identifica los principales problemas éticos y sociales del uso de IA: la opacidad de los algoritmos o funcionamiento de “caja negra”; los sesgos inherentes que reflejan las visiones del mundo de quienes los diseñan; las alucinaciones o respuestas plausibles pero no fundamentadas en datos reales; la profundización de divisiones sociales en países de ingresos bajos y medios; los sesgos de género que normalizan roles sumisos; y la “brecha robótica” que excluye a quienes carecen de acceso a dispositivos tecnológicos.

Todos estos sesgos pueden perpetuar desigualdades sociales, ampliar brechas digitales y vulnerar derechos humanos. Conviene detenerse en el sesgo geográfico: dado que la mayoría de los sistemas de IA han sido desarrollados por grandes corporaciones del norte global, tienden a priorizar la información de esas regiones. Un ejemplo claro es la limitada

capacidad de la IA para producir contenidos en lenguas originarias. Según el estudio “El desempeño de la inteligencia artificial en el uso de lenguas indígenas americanas”, realizado por LLYC en colaboración con BID Lab y Microsoft, los modelos actuales muestran baja eficacia en lenguas indígenas americanas: entregan respuestas incorrectas, incompletas o en otros idiomas, y evidencian una marcada parcialidad cultural hacia Occidente. Entre las lenguas evaluadas con bajo rendimiento se encuentran quechua, guaraní, aimara, náhuatl, quiché, mapuche y tupí-guaraní [Lucas et al., 2025]. Las principales causas son la escasez de datos digitales en estas lenguas, la falta de herramientas lingüísticas adecuadas y las diferencias en la familiaridad tecnológica de sus hablantes. Para la comunicación de la ciencia, esto significa que vastas comunidades quedan excluidas de los beneficios de estas herramientas en su propia lengua y desde su propia cosmovisión.

Los sesgos de género pueden observarse mediante experimentos simples. Al pedir a ChatGPT (GPT-3) los veinte matemáticos más importantes de la historia, solo una mujer apareció en la lista: Emmy Noether, en el lugar 14, y ninguna científica del Sur Global. Este sesgo no es una falla técnica menor, sino un ejemplo de injusticia testimonial algorítmica: las contribuciones de mujeres y figuras no occidentales son sistemáticamente devaluadas por un sistema entrenado con un corpus que refleja la matriz colonial de dominación. Obtener representaciones diversas exige redactar instrucciones muy específicas que indiquen, por ejemplo, género o procedencia geográfica. Ante este panorama, la comunicación de la ciencia tiene un rol que va más allá de la denuncia: divulgar cómo funcionan estos sesgos, por qué existen y cómo detectarlos es en sí mismo un acto de alfabetización científica y mediática, tanto para las audiencias como para quienes producen contenidos con apoyo de la IA.

Un tercer desafío lo constituyen las alucinaciones. Estas se refieren a resultados incorrectos, engañosos o inventados que generan, sobre todo, los modelos de lenguaje al producir respuestas plausibles pero no fundamentadas en datos reales ni justificadas por la información de entrenamiento. Surgen por múltiples razones: datos insuficientes, sesgos en los conjuntos de entrenamiento, errores de patrón o intentos del modelo por “rellenar” lagunas con contenido fabricado. Tanto en textos como en imágenes, las IA pueden mezclar datos verídicos con datos ficticios [Alkaissi & McFarlane, 2023]. Este problema resulta especialmente grave en la comunicación pública de la ciencia y en la búsqueda de referencias confiables, pues modelos como ChatGPT pueden generar bibliografías falsas.

En el ámbito visual también se presentan alucinaciones. En la literatura se han documentado varios casos, como el de la revista Frontiers donde se publicó el artículo “Cellular functions of spermatogonial stem cells in relation to JAK/STAT signaling pathway” [Guo et al., 2024], que incluía imágenes generadas por IA sin sustento científico. En dicho estudio se intentó representar una ruta metabólica, pero la IA produjo una alucinación con palabras incompletas, numeraciones erróneas y ciclos inexistentes. Debido a estos errores, el artículo fue retractado [Frías-Villegas & Guzmán-Vendrell, 2024].

El problema de las alucinaciones afecta directamente a la divulgación científica: cuando se emplean textos, materiales o imágenes alucinados, se difunde información errónea que perjudica al público. La comunicación de la ciencia exige que los contenidos compartidos sean verificables y veraces. Una dependencia excesiva de estas herramientas puede derivar en la circulación de información falsa y minar la confianza en los esfuerzos de divulgación [Frías-Villegas & Guzmán-Vendrell, 2024].

Finalmente, resulta imprescindible considerar el tema de los derechos de autor. Muchos sistemas de IA se entrenan con grandes volúmenes de libros, obras de arte, fotografías y música, a menudo protegidos por copyright. El uso de estas obras sin autorización ni compensación ha derivado en demandas colectivas de autores y artistas contra desarrolladores como OpenAI y Meta.

El debate actual gira en torno a si el entrenamiento con obras protegidas constituye "uso justo" (fair use). Las empresas tecnológicas argumentan que el aprendizaje automático es transformador y no sustituye el consumo original; en contraste, los creadores sostienen que desplaza sus obras y explota su valor sin autorización. En México, la Suprema Corte de Justicia ha determinado que la IA no puede ser reconocida como autora, aunque permite registrar obras generadas con IA a nombre del solicitante, siempre que exista un mínimo de intervención humana [González, 2025].

A ello se suma el uso de IA para producir *deepfakes* sin consentimiento. Un *deepfake* es un tipo de contenido manipulado digitalmente mediante inteligencia artificial, que permite alterar imágenes, audios o videos para hacer que una persona parezca decir o hacer algo que nunca ocurrió. Este fenómeno representa un riesgo directo para la comunicación de la ciencia: voces o rostros de investigadores pueden ser suplantados para difundir desinformación con aparente legitimidad científica.

6 ▪ Conclusiones

La inteligencia artificial está reconfigurando la comunicación pública de la ciencia. Sin embargo, su aparente neutralidad oculta dinámicas de poder que, desde una perspectiva decolonial, es urgente develar. Los contenidos que genera no solo reproducen, sino que amplifican desigualdades epistémicas, de género y geopolíticas a través de un modelo de extractivismo de datos. Por ello, la pregunta rectora no es únicamente cómo usar estas herramientas sin vulnerar derechos, sino cómo avanzar hacia una soberanía tecnológica que permita a las comunidades latinoamericanas desarrollar, adaptar y controlar sus propias herramientas digitales para comunicar ciencia desde sus propios referentes culturales y lingüísticos.

A la luz de los marcos normativos vigentes — la Declaración Universal de Derechos Humanos, el Pacto Internacional de Derechos Económicos, Sociales y Culturales, la Recomendación de la UNESCO sobre la Ética de la IA y los estándares europeos y de la OCDE — la legitimidad del uso de IA en comunicación científica exige garantizar veracidad y verificabilidad, no discriminación, transparencia, rendición de cuentas y mecanismos de reparación en caso de daño. En términos operativos, estos principios deben integrarse en todas las fases de los proyectos — diseño, producción, difusión y evaluación —, incluyendo lenguas originarias y saberes locales, verificación humana experta (*human-in-the-loop*), acciones afirmativas de equidad de género y diversidad cultural, respeto estricto a los derechos de autor, políticas firmes contra los *deepfakes*, y programas de alfabetización mediática y digital crítica para equipos y audiencias.

En este marco, los comunicadores y comunicadoras de la ciencia tienen un rol que va más allá de la producción de contenidos: son agentes clave en la construcción de una cultura científica crítica frente a la IA. Esto significa no solo usar estas herramientas de forma responsable, sino divulgar activamente cómo funcionan, qué sesgos contienen, por qué

alucinan y qué marcos legales las regulan. En América Latina, donde las brechas digitales, lingüísticas y epistémicas son más pronunciadas, esta labor adquiere una urgencia particular. Bajo estas condiciones, la IA puede potenciar la comunicación de la ciencia haciéndola más accesible, plural y justa — sin convertirse en una nueva frontera de exclusión —, siempre que quienes comunicamos la ciencia conozcamos sus usos éticos, exijamos que sus herramientas sean desarrolladas con y para nuestras comunidades, y aseguremos que los contenidos que producimos sean científicamente precisos, culturalmente pertinentes y respetuosos de la dignidad humana.

Agradecimientos

Investigación realizada gracias al Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT) de la UNAM, IA401525. Se usaron las inteligencias artificiales ChatGPT-5, Gemini 2.5 Pro y Claude Sonnet 4.6 para revisar ortografía, gramática y coherencia del texto, además de para generar los ejemplos que se citan en el cuerpo del artículo; DeepSeek se utilizó únicamente para obtener el ejemplo citado en la sección 5.

Referencias

Afroogh, S., Akbari, A., Malone, E., Kargar, M., & Alambeigi, H. (2024). Trust in AI: progress, challenges, and future directions. *Humanities and Social Sciences Communications*, *11*(1). https://doi.org/10.1057/s41599-024-04044-8

Alkaissi, H., & McFarlane, S. I. (2023). Artificial Hallucinations in ChatGPT: Implications in Scientific Writing. *Cureus*. https://doi.org/10.7759/cureus.35179

Aspray, W., & Hayes, B. M. (2011). *Everyday Information: The Evolution of Information Seeking in America*. The MIT Press. https://doi.org/10.7551/mitpress/8820.001.0001

Barbera, D. (2024). Esta foto creada por IA que se hizo viral por ganar un concurso se acaba de vender por miles de dólares. *WIRED*. https://es.wired.com/articulos/foto-creada-por-ia-que-se-hizo-viral-por-ganar-concurso-se-acaba-de-vender-por-miles-de-dolares

Benjamin, R. (2020). *Race after technology: Abolitionist tools for the New Jim Code*. Polity.

Congreso de la Unión. (2021, 1 de junio). *Decreto por el que se adicionan diversas disposiciones a la Ley General de Acceso de las Mujeres a una Vida Libre de Violencia y al Código Penal Federal ["Ley Olimpia"]*. Diario Oficial de la Federación. https://www.dof.gob.mx/nota_detalle.php?codigo=5619905&fecha=01/06/2021

Couldry, N., & Mejias, U. A. (2019). *The costs of connection: How data is colonizing human life and appropriating it for capitalism*. Stanford University Press. https://www.sup.org/books/sociology/costs-connection

DeepSeek. (2026, 22 de septiembre). *DeepSeek [Modelo de lenguaje]. Respuesta a la consulta "Qué pasó en la plaza de Tiananmén en 1989*. https://chat.deepseek.com

Dove. (2024). *El compromiso con la Belleza Real de Dove*. https://www.dove.com/es/stories/about-dove/dove-real-beauty-pledge.html

Fortune. (2025). *Inside the history of ChatGPT's viral Studio Ghibli-style images: Founder once said he was 'utterly disgusted' by AI animation*. https://fortune.com/2025/03/27/chatgpt-studio-ghibli-hayao-miyazaki-openai-altman-copyright-lawsuit/

Frías-Villegas, G., & Guzmán-Vendrell, A. (2024). Usos éticos de la IA en la comunicación de la ciencia, con un enfoque de derechos humanos. *XXIII Congreso Nacional de Divulgación de la Ciencia y la Técnica*. https://www.somedicyt.org.mx/images/divulgadores/congresos/23/ponencias-congreso-XXIII.pdf

Frías-Villegas, G., García-Gómez, K. E., Guzmán-Vendrell, A., Mendoza-Hernández, I. A., Tránsito-Santos, R., & Vázquez-Quiroz, F. (2024). Science Communication as a Human Right. *JCOM*, *23*(02). https://doi.org/10.22323/2.23020201

González, F. (2025, 3 de julio). *En México, las obras generadas con IA son de dominio público, según la SCJN*. WIRED. https://es.wired.com/articulos/en-mexico-las-obras-generadas-con-ia-son-de-dominio-publico-segun-la-scjn

Guo, X., Dong, L., & Hao, D. (2024). RETRACTED: Cellular functions of spermatogonial stem cells in relation to JAK/STAT signaling pathway. *Frontiers in Cell and Developmental Biology*, *11*. https://doi.org/10.3389/fcell.2023.1339390

Hendriks, F., Barel-Ben David, Y., Banse, L., Fick, J., Greussing, E., Klein-Avraham, I., Rakedzon, T., Taddicken, M., & Baram-Tsabari, A. (2025). Generative AI in Science Communication: Fostering Scientists' Good Working Habits for Ethical and Effective Use. *Science Communication*. https://doi.org/10.1177/10755470251343486

Kessler, S. H., Mahl, D., Schäfer, M. S., & Volk, S. C. (2025). Science Communication in the Age of Artificial Intelligence. *JCOM*, *24*(2). https://doi.org/10.22323/2.24020501

Lucas, M., Burgueño, A., Buenadicha Sánchez, C., Carazas, M., Ramirez Rufino, S., Rosales Torres, C. S., Korn, D., Tesfaye, H., & Deo, G. (2025). *El desempeño de la inteligencia artificial en el uso de lenguas indígenas americanas*. https://doi.org/10.18235/0013542

Luccioni, A. S., Viguier, S., & Ligozat, A.-L. (2022). *Estimating the Carbon Footprint of BLOOM, a 176B Parameter Language Model*. arXiv: 2211.02001. https://doi.org/10.48550/arXiv.2211.02001

Luccioni, S., Jernite, Y., & Strubell, E. (2024). Power Hungry Processing: Watts Driving the Cost of AI Deployment? *The 2024 ACM Conference on Fairness Accountability and Transparency*, 85-99. https://doi.org/10.1145/3630106.3658542

Lugones, M. (2007). Heterosexualism and the Colonial/Modern Gender System. *Hypatia*, *22*, 186-209. https://doi.org/10.1111/j.1527-2001.2007.tb01156.x

Maldonado-Torres, N. (2008). *La descolonización y el giro des-colonial*. https://www.redalyc.org/pdf/396/39600905.pdf

Mei, Q., Xie, Y., Yuan, W., & Jackson, M. O. (2024). A Turing test of whether AI chatbots are behaviorally similar to humans. *Proceedings of the National Academy of Sciences*, *121*(9). https://doi.org/10.1073/pnas.2313925121

Mohamed, S., Png, M.-T., & Isaac, W. (2020). Decolonial AI: Decolonial Theory as Sociotechnical Foresight in Artificial Intelligence. *Philosophy & Technology*, *33*(4), 659-684. https://doi.org/10.1007/s13347-020-00405-8

Moskvichev, A., Odouard, V. V., & Mitchell, M. (2023). *The ConceptARC Benchmark: Evaluating Understanding and Generalization in the ARC Domain*. arXiv: 2305.07141. https://doi.org/10.48550/ARXIV.2305.07141

Naciones Unidas. (1948). *La Declaración Universal de los Derechos Humanos*. https://www.un.org/es/about-us/universal-declaration-of-human-rights

Naciones Unidas. (1966). *Pacto Internacional de Derechos Económicos, Sociales y Culturales*. https://www.ohchr.org/es/instruments-mechanisms/instruments/international-covenant-economic-social-and-cultural-rights

OpenAI. (2022, 30 de noviembre). *Introducing ChatGPT [Presentamos ChatGPT]*. https://openai.com/index/chatgpt/

Ramírez Autrán, R. (2023). Sesgos y discriminaciones sociales de los algoritmos en Inteligencia Artificial: Una revisión documental. *Entretextos*, *15*(39), 1-17. https://doi.org/10.59057/iberoleon.20075316.202339664

Red Iberoamericana de Protección de Datos. (2020). *Recomendaciones generales sobre tratamiento de datos en IA*. https://www.redipd.org/sites/default/files/2020-02/guia-recomendaciones-generales-tratamiento-datos-ia.pdf

Red Iberoamericana de Protección de Datos. (2023). *Declaración sobre neurodatos*. https://www.redipd.org/documento/declaracion-neurodatos-ripd.pdf

Ricaurte, P. (2019). Data Epistemologies, The Coloniality of Power, and Resistance. *Television & New Media*, *20*(4), 350-365. https://doi.org/10.1177/1527476419831640

Stanford News. (2011). *Stanford's John McCarthy, seminal figure of artificial intelligence, dies at 84*. https://news.stanford.edu/stories/2011/10/stanfords-john-mccarthy-seminal-figure-artificial-intelligence-dies-84

Sun, H. (2020). Reinvigorating the Human Right to Technology. *Michigan Journal of International Law*, *41*(2), 279. https://doi.org/10.36642/mjil.41.2.reinvigorating

Tierra Común Network. (2023). *Resisting Data Colionalism: A Practical Intervention*. https://networkcultures.org/blog/publication/tod-50-resisting-data-colonialism-a-practical-intervention/

UNESCO. (2021). *Recomendación sobre la ética de la inteligencia artificial*. https://unesdoc.unesco.org/ark:/48223/pf0000381133/PDF/381133eng.pdf.multi.page=62

Villegas Servín, C. (2025). *Inteligencia Artificial con perspectiva de género: Entendiendo sesgos y construyendo un futuro Inclusivo*. Programa Universitario de Derechos Humanos, Universidad Nacional Autónoma de México. https://www.pudh.unam.mx/conferencia11f_IA_con_perspectiva_genero_entendiendo_sesgos_construyendo_un_futuro_inclusivo.html

Yang, Z. (2025, 31 de enero). *Here's How DeepSeek Censorship Actually Works — And How to Get Around It*. WIRED. https://www.wired.com/story/deepseek-censorship/

Sobre la autoría

La Dra. Gabriela Frías-Villegas estudió las licenciaturas en Matemáticas y Lengua y Literatura Inglesas Modernas, además de la maestría y el doctorado en Filosofía de la Ciencia, con especialización en Comunicación de la Ciencia, en la UNAM. De 2008 a 2021 se desempeñó como Coordinadora de Comunicación de la Ciencia en el Instituto de Ciencias Nucleares de la misma institución. Actualmente es investigadora dentro del Programa Universitario de Derechos Humanos de la UNAM, donde sus líneas de investigación son las siguientes: estudios ciencia, tecnología, sociedad y derechos humanos; comunicación de las ciencias y las humanidades; puntos de encuentro entre el arte, la ciencia y la literatura; educación STEAM en la era digital, estudios de género e Inteligencia Artificial. Entre sus reconocimientos se puede mencionar el Reconocimiento Sor Juana Inés de la Cruz que otorga la UNAM a las académicas distinguidas.

gabriela.frias@humanidades.unam.mx

Cómo citar

Frías-Villegas, G. (2026). 'Usos éticos de la Inteligencia Artificial para la comunicación de la ciencia, con una perspectiva decolonial y un enfoque de Derechos Humanos'. *JCOMAL* 09(02), Y02. https://doi.org/10.22323/3613202607071111107.

© El autor o autores.
Esta publicación está bajo los términos de la licencia Creative Commons Atribución 4.0.
Todos los derechos de minería de texto y datos, entrenamiento de IA y tecnologías similares para fines comerciales están reservados.
ISSN 2611-9986. Publicado por SISSA Medialab. jcomal.sissa.it