

Inteligencia artificial para todos (parte 3): el costo invisible de cada conversación

VALERIA MAEDA GUTIÉRREZ Y JUAN JOSÉ OROPEZA VALDEZ

La Dra. Maeda-Gutiérrez es Docente-Investigadora en la Unidad Académica de Ingeniería Eléctrica de la Universidad Autónoma de Zacatecas. Investiga en el desarrollo y aplicación de modelos de Inteligencia Artificial, Procesamiento de Imágenes y Aprendizaje Profundo en biomedicina, con énfasis en el estudio de complicaciones microvasculares asociadas a la Diabetes Tipo 2. Es miembro del Sistema Nacional de Investigadoras e Investigadores, nivel C, en el área de Medicina y Ciencias de la Salud. El Dr. Oropeza-Valdez es Profesor de Tiempo Completo en la División de Investigación de la Facultad de Medicina de la Universidad Nacional Autónoma de México, adscrito a la Unidad de Investigación en Obesidad del Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán. Investiga la nefropatía diabética mediante enfoques de metabolómica, bioinformática e inteligencia artificial, para identificar biomarcadores, comprender los mecanismos metabólicos asociados a la progresión de la enfermedad y contribuir al desarrollo de estrategias diagnósticas y de medicina personalizada.

Esta publicación fue revisada por el comité editorial de la Academia de Ciencias de Morelos.

Introducción

En las entregas anteriores [1, 2] explicamos qué son los Modelos Grandes de Lenguaje, cómo aprenden y qué ocurre dentro de ellos cuando los usamos. Describimos un mundo aparentemente etéreo de *tokens*, vectores y mecanismos de atención. Esa descripción, sin embargo, deja fuera algo esencial. Cada conversación con un asistente conversacional, cada imagen generada por inteligencia artificial (IA), cada traducción automática y cada recomendación de música o video tienen un costo material y humano que rara vez aparece en pantalla. Detrás de la fluidez con la que un modelo responde a nuestras preguntas hay litros de agua que se evaporan en torres de enfriamiento, gigavatios de electricidad que vienen de plantas térmicas o hidroeléctricas, kilogramos de cobre y minerales raros extraídos en remotas regiones del planeta, y miles de personas que trabajan en condiciones precarias etiquetando datos y filtrando contenido. Esta tercera entrega busca hacer visible esa otra cara de la IA: lo que cuesta, en términos físicos y humanos, que un sistema responda con elegancia y rapidez. No se trata de demonizar la tecnología, sino de comprender su huella completa, porque solo desde esa comprensión es posible exigir un desarrollo más responsable y participar de manera informada en los debates públicos que ya tenemos enfrente.

La sed de los servidores: el agua que bebe la IA

Los procesadores que ejecutan los modelos de IA generan enormes cantidades de calor. Para evitar que se dañen, los centros de datos requieren sistemas de enfriamiento intensivos, y muchos de ellos utilizan agua dulce que se evapora durante el proceso. A esto se suma el agua consumida en la generación de la electricidad que alimenta esos mismos centros, lo que duplica o triplica la huella hídrica total. Investigadores de la Universidad de California en Riverside estimaron en 2023 que generar una respuesta de cien palabras con un modelo del tamaño de GPT-4 puede consumir alrededor de 519 mililitros de agua, el equivalente a una botella mediana [3]. Otras estimaciones sitúan el consumo entre 10 y 25 mililitros por consulta corta, dependiendo del tipo de tarea y de la matriz energética del país donde se aloja el centro de datos. Multiplicado por los miles de millones de consultas diarias, el resultado es asombroso: las Naciones Unidas estiman que la huella hídrica global de la IA alcanzó cerca de 9.3 billones de litros en 2025 [4], una cifra comparable al consumo anual de varios países europeos. El problema no es solo de magnitud, sino de localización. Cerca de la mitad de los centros de datos del mundo están instalados en regiones que ya enfrentan estrés hídrico. Esto significa que el agua que enfría a los servidores compete con el agua que las comunidades necesitan para beber, cultivar y sostener sus ecosistemas. En España, Amazon solicitó en 2024 un aumento del 48 por ciento en sus permisos hídricos para una instalación en Aragón. En el Reino Unido, una región designada como zona de crecimiento para la IA coincide con una de las áreas con mayor riesgo de escasez de agua del país. Y la tendencia se acelera: análisis recientes proyectan que el consumo hídrico de los centros de datos podría multiplicarse por once entre 2024 y 2028. Las empresas tecnológicas han comenzado a responder. Microsoft anunció en 2024 que sus nuevos centros de datos a partir de 2027 utilizarán enfriamiento sin agua. Google publica informes anuales con métricas de eficiencia hídrica. Algunas instalaciones de hiperescala ya operan con sistemas cerrados de líquido refrigerante que reducen drásticamente la evaporación. Sin embargo, estas mejoras avanzan más lentamente que el crecimiento exponencial de la demanda, y la pregunta de fondo persiste: ¿es razonable instalar infraestructura intensiva en agua precisamente en regiones que ya sufren sequías recurrentes?

El hambre de electricidad: del bit al megavatio

Si el agua es la primera dimensión del costo material de la IA, la electricidad es la segunda y la más estudiada. Cada consulta a un modelo de lenguaje moderno consume entre 0.3 y 19 vatios-hora de electricidad, dependiendo del tamaño del modelo y de la complejidad de la pre-

gunta [5]. Una consulta sencilla equivale a encender un foco LED durante unos cuantos segundos; una consulta compleja con razonamiento extendido puede consumir hasta cincuenta veces más. Una imagen generada por IA consume hasta mil cuatrocientas veces más energía que una clasificación básica de texto. Estos consumos individuales, multiplicados por miles de millones de consultas, dan cifras de escala nacional. ChatGPT procesa alrededor de 2,500 millones de consultas diarias. El consumo total de electricidad de los centros de datos del mundo se estimó en 448 teravatios-hora en 2025, una cantidad que, si se considerara como un país, lo colocaría en el undécimo lugar mundial, por encima de Arabia Saudita y por debajo de Francia [4]. Las proyecciones indican que para 2030 podría alcanzar 945 teravatios-hora, más del doble. Un aspecto poco conocido del problema es que la mayor parte de este consumo no proviene del entrenamiento de los modelos, sino de su uso cotidiano. Una vez que un modelo está entrenado, cada interacción con un usuario activa miles de millones de cálculos. Este proceso, llamado *inferencia*, absorbe entre el 80 y el 90 por ciento del consumo energético total a lo largo del ciclo de vida del sistema. Es decir, la mayor parte de la huella ambiental de la IA se genera mientras la usamos, no mientras se construye.

Esta demanda creciente está teniendo consecuencias palpables en las redes eléctricas. En el norte de Virginia, en Estados Unidos, sede de muchos centros de datos, los planificadores energéticos calculan que se necesitarían varias centrales nucleares adicionales para satisfacer la demanda proyectada. Algunas compañías tecnológicas han firmado contratos para comprar la totalidad de la producción de centrales nucleares completas. La paradoja es evidente: una tecnología que se presenta como aliada de la transición energética, está en realidad retrasándola, al absorber capacidad renovable que de otro modo descarbonizaría hogares, industrias y transporte.

Los obreros invisibles: el trabajo humano detrás del aprendizaje automático

Existe la imagen popular de que la IA aprende sola, alimentándose pasivamente de los datos que existen en internet. La realidad es más incómoda. Antes de que un modelo sea seguro, útil y socialmente aceptable, miles de personas, en su mayoría en países del Sur Global, deben revisar, clasificar y etiquetar enormes cantidades de texto, imágenes y video. Su trabajo es lo que separa a un generador de texto crudo de un asistente conversacional que evita producir contenido racista, sexualmente explícito o violento. Una investigación de la revista *Time* publicada en enero de 2023 reveló que para depurar los datos de entrenamiento de ChatGPT, OpenAI subcontrató a una empresa intermediaria que empleaba trabajadores en Kenia. Estos trabajadores recibían entre 1.32 y 2 dólares estadoun-

idenses por hora por leer y etiquetar fragmentos de texto que describían escenas de violencia extrema, abuso sexual infantil, suicidio, tortura, bestialidad e incesto [6]. La empresa contratante, según documentos revisados por reporteros, cobraba a OpenAI alrededor de 12.50 dólares la hora por cada trabajador, conservando una buena parte del pago como margen. Varios de los empleados entrevistados describieron daño psicológico persistente: visiones recurrentes, insomnio y deterioro de su vida familiar. El caso de Kenia no es excepcional. Empresas similares operan en India, Filipinas, Venezuela, Colombia y otras regiones donde existe mano de obra alfabetizada, frecuentemente con educación universitaria, y altos índices de desempleo. Cadenas de televisión como CBS y reportajes posteriores han documentado patrones consistentes: salarios bajos, exposición prolongada a contenido traumático, escaso apoyo psicológico y precariedad laboral [7]. En 2024, un grupo de trabajadores kenianos presentó una carta al Congreso de Estados Unidos en la que describían sus condiciones como “trabajo de esclavitud moderna” y solicitaban que cualquier acuerdo comercial bilateral considerara la protección de sus derechos.



FIGURA 1. EL trabajo humano detrás de los sistemas de IA (elaboración propia empleando NotebookLM).

desarrollo de su propia industria nacional y limitando la exportación de minerales críticos. El resultado es que la IA, presentada con frecuencia como una tecnología abierta y universal, depende en la práctica de cadenas de suministro frágiles, sujetas a decisiones políticas que pocos ciudadanos conocen y ninguno controla. La minería asociada a estos materiales tiene también sus propios costos ambientales y sociales. La extracción de cobalto en la República Democrática del Congo, indispensable para baterías y componentes electrónicos, está vinculada a denuncias documentadas de trabajo infantil y condiciones laborales peligrosas. La extracción de litio en el llamado triángulo sudamericano, que abarca

regiones de Argentina, Bolivia y Chile, demanda enormes cantidades de agua en zonas desérticas donde ya escasea. La huella material de la IA, por tanto, no comienza en el centro de datos: comienza mucho antes, en pozos, minas y procesos químicos que rara vez aparecen en las narrativas de innovación.

México en el mapa: el caso de Querétaro

Estos costos, que pueden sonar lejanos al lector mexicano, comienzan a tener expresiones locales muy concretas. En septiembre de 2025, el gobierno federal anunció la instalación de un campus de seis centros de datos de IA en Querétaro, a cargo de la empresa CloudHQ, con una inversión de 4 mil 800 millones de dólares [8]. El proyecto, parte del Plan México, contempla una superficie de 52 hectáreas y una demanda eléctrica final de 900 megavatios, una cantidad equivalente al consumo de una ciudad mediana. Querétaro se ha consolidado, junto con la Ciudad de México y Monterrey, como uno de los tres polos principales del país para esta industria. Solo en 2025 se anunciaron inversiones por más de 18 mil millones de dólares en centros de datos en territorio nacional. La Comisión

dustria automotriz y el crecimiento urbano? La empresa ha anunciado que su diseño utilizará sistemas de enfriamiento sin agua, lo cual es una buena noticia que merece verificación independiente y seguimiento público. ¿De dónde provendrá la electricidad? La meta nacional contempla que al menos un tercio de la nueva capacidad sea renovable, pero la velocidad de instalación de centros de datos puede ser mayor que la velocidad de despliegue de energías limpias. ¿Qué proporción del valor agregado se quedará en México? ¿Habrá transferencia tecnológica real o solo arrendamiento de territorio para infraestructura extranjera? Estas preguntas no descalifican la inversión, pero sí merecen estar en la conversación pública. Hay que tomar en cuenta las principales dimensiones materiales asociadas al uso y expansión de la IA (Figura 2).



FIGURA 2. EL costo material de la IA. (elaboración propia empleando NotebookLM).

¿Qué podemos hacer?

Como usuarios individuales, nuestra capacidad de incidir en estos procesos puede parecer mínima, y en cierto sentido lo es. Cada decisión personal de usar o no un asistente conversacional tiene un impacto marginal frente a los miles de millones de consultas diarias. Sin embargo, hay algunas prácticas razonables que disminuyen, aunque sea modestamente, la huella de nuestro uso. Conviene reservar las consultas a modelos pesados para tareas que realmente lo ameriten, y usar herramientas más simples para preguntas que un buscador tradicional puede resolver. Es prudente también evitar el uso recreativo intensivo de generación de imágenes y video, que es la categoría más costosa en energía. Y desde luego, desconfiar de la narrativa que presenta a la IA como mágica e inocua: cada interacción tiene un costo real, aunque no lo veamos. Como ciudadanos, lo más importante quizá no sea modificar el comportamiento individual sino exigir transparencia y regulación. Las empresas tecnológicas han comenzado a publicar reportes de sostenibilidad, pero la mayoría son selectivos y no permiten comparación rigurosa. Necesitamos métricas estandarizadas y verificables sobre consumo hídrico, energético y emisiones por unidad de servicio prestado. Necesitamos también marcos legales que exijan condiciones laborales dignas a lo largo de toda la cadena de valor, no solo en la oficina de Silicon Valley

donde se diseñan los productos, sino también en las salas donde se etiquetan datos en Nairobi, Bogotá o Manila. La Unión Europea ha avanzado con su Ley de IA: México y América Latina apenas inician estos debates. Como investigadores y comunidad científica, tenemos la responsabilidad adicional de generar evidencia sobre estos costos y de incorporar criterios de sostenibilidad en el diseño mismo de los modelos. Investigaciones recientes muestran que existen técnicas para reducir el consumo energético sin sacrificar de manera significativa el desempeño: modelos más pequeños especializados en tareas concretas, destilación de modelos grandes en versiones eficientes, planeación de cargas computacionales en horarios con mayor disponibilidad de energía limpia. La sostenibilidad no es

donde se diseñan los productos, sino también en las salas donde se etiquetan datos en Nairobi, Bogotá o Manila. La Unión Europea ha avanzado con su Ley de IA: México y América Latina apenas inician estos debates. Como investigadores y comunidad científica, tenemos la responsabilidad adicional de generar evidencia sobre estos costos y de incorporar criterios de sostenibilidad en el diseño mismo de los modelos. Investigaciones recientes muestran que existen técnicas para reducir el consumo energético sin sacrificar de manera significativa el desempeño: modelos más pequeños especializados en tareas concretas, destilación de modelos grandes en versiones eficientes, planeación de cargas computacionales en horarios con mayor disponibilidad de energía limpia. La sostenibilidad no es

Conclusión

La IA parece operar en una nube etérea, sin peso ni cuerpo, pero detrás de cada respuesta breve y elegante hay agua que se evapora, electricidad que se consume, minerales que se extraen y personas que trabajan, muchas de ellas en condiciones que no escogerían si tuvieran alternativa. Reconocer estos costos no debería llevarnos a rechazar la tecnología, sino a usarla con conciencia y a exigir que su desarrollo sea más equitativo, más transparente y más sostenible. En las dos entregas anteriores discutimos cómo funcionan los modelos de lenguaje y qué ocurre dentro de ellos. En esta tercera entrega quisimos ofrecer la dimensión que rara vez se discute: la materialidad y la humanidad sobre las que se sostiene esa fluidez conversacional. La IA es una de las tecnologías más fascinantes y poderosas de nuestro tiempo. Precisamente por eso, su desarrollo no puede dejarse exclusivamente en manos de las empresas que la producen ni de los países que la concentran. Nos corresponde a todos, como usuarios, ciudadanos y científicos, participar informadamente en el debate sobre cómo, dónde y para qué la queremos. La verdadera inteligencia, después de todo, no consiste en aceptar acríticamente las herramientas que se nos ofrecen, sino en entenderlas a fondo, incluidas sus partes incómo-

das, y elegir con lucidez qué hacemos con ellas.

Referencias

- [1] Maeda-Gutiérrez, V., y Oropeza-Valdez, J. J. (2025). Inteligencia artificial para todos: ¿Cómo funcionan los modelos que conversan contigo? *La Unión de Morelos*, 15 de septiembre de 2025. Disponible en: <https://acmor.org/publicaciones/inteligencia-artificial-para-todos-c-mo-funcionan-los-modelos-que-conversan-contigo>
- [2] Maeda-Gutiérrez, V., y Oropeza-Valdez, J. J. (2026). Inteligencia artificial para todos (parte 2): ¿qué ocurre dentro de los modelos y en qué debemos fijarnos? *La Unión de Morelos*, 4 de mayo de 2026. Disponible en: <https://acmor.org/publicaciones/inteligencia-artificial-para-todos-parte-2-qu-ocurre-dentro-de-los-modelos-y-en-que-debemos-fijarnos>
- [3] Li, P., Yang, J., Islam, M. A., y Ren, S. (2023). Making AI less “thirsty”: Uncovering and addressing the secret water footprint of AI models. *arXiv preprint arXiv:2304.03271*. Disponible en: <https://arxiv.org/abs/2304.03271>
- [4] United Nations University Institute for Water, Environment and Health (UNU-INWEH). (2026). *Environmental cost of AI’s energy use: carbon, water and land footprints*. Disponible en: <https://unu.edu/inweh/news/environmental-cost-of-AIs-Energy-use-carbon-water-and-land-footprints>
- [5] de Vries, A. (2023). The growing energy footprint of artificial intelligence. *Joule*, 7(10), 2191–2194. doi: 10.1016/j.joule.2023.09.004
- [6] Perrigo, B. (2023). OpenAI used Kenyan workers on less than \$2 per hour to make ChatGPT less toxic. *TIME*, 18 de enero de 2023. Disponible en: <https://time.com/6247678/openai-chatgpt-kenya-workers/>
- [7] Stahl, L. (2024). Kenyan workers with AI jobs thought they had tickets to the future until the grim reality set in. *CBS News, 60 Minutes*, 24 de noviembre de 2024. <https://www.cbsnews.com/news/ai-work-kenya-exploitation-60-minutes/>
- [8] Gobierno de México, Presidencia de la República. (2025). *Plan México avanza: Se anuncia inversión de 4 mil 800 mdd de CloudHQ para la construcción de 6 Centros de Datos en Querétaro*. Comunicado del 25 de septiembre de 2025. Disponible en: <https://www.gob.mx/presidencia/prensa/plan-mexico-avanza-se-anuncia-inversion-de-4-mil-800-mdd-de-cloudhq-para-la-construccion-de-6-centros-de-datos-en-queretaro>
- [9] Asociación Mexicana de Data Centers (MEXDC). (2025). *Estudio de Mercado del Sector 2025–2030*. <https://asmexdc.com/solicitud-estudios-mexdc>

Esta columna se prepara y edita semana con semana, en conjunto con investigadores morelenses convencidos del valor del conocimiento científico para el desarrollo social y económico de Morelos.