ieee.es
Instituto Español de Estudios Estratégicos

Documento
Análisis

07/2025 28 de enero de 2025

David Ramírez Morán

La huella medioambiental de la IA



La huella medioambiental de la IA

Publicado originalmente en:

La inteligencia artificial en la geopolítica y los conflictos
Editado en junio de 2024

Resumen:

El impacto de la inteligencia artificial sobre el medioambiente es fruto de un amplio conjunto de factores bastante interrelacionados. Los requerimientos de energía, capacidad de procesado y refrigeración son los tres más directos, aunque hay otros factores, como la huella climática de las múltiples cadenas de suministro, las infraestructuras disponibles e incluso la regulación, que constituyen dimensiones relevantes en su desarrollo e implantación.

La sensibilidad existente actualmente en la población y en las organizaciones motiva que las empresas que proporcionan servicios estén haciendo constantes esfuerzos para reducir la huella climática de su actividad, a la vez que intentan minimizar los costes. Los esfuerzos se dirigen tanto a la optimización de sus propias operaciones como a la reducción de la huella en las cadenas de suministro. El objetivo es mostrar públicamente compromiso con el planeta, conseguir que el uso de sus servicios no afecte negativamente a la huella climática de la actividad de sus clientes y que incluso la reduzca, promoviendo así el uso intensivo.

Los escenarios futuros parecen augurar un crecimiento importante del uso de IA en muchos sectores, con soluciones que pueden imponer restricciones adicionales como latencia limitada, necesidad de confinar geográficamente los datos o medidas que aseguren el estricto cumplimiento de las normas de privacidad y no discriminación.

Todos estos factores configuran un entorno de competencia internacional en el que pueden estar en juego la autonomía y soberanía estratégica de los países.

Palabras clave:

Inteligencia artificial, Medioambiente, Huella de carbono, Competencia internacional.

NOTA:** Las ideas contenidas en los ***Documentos de Análisis son responsabilidad de sus autores, sin que reflejen necesariamente el pensamiento del IEEE o del Ministerio de Defensa.

Environmental footprint of artificial intelligence

Abstract:

The impact of artificial intelligence on the environment is the result of a wide range of interrelated factors. Energy requirements, processing capacity and cooling are the three most direct, although there are other factors, such as the climate footprint of multiple supply chains, available infrastructures and even regulation, which constitute relevant dimensions in its development and implementation.

The current sensitivity of the population and organisations means that companies providing services are making constant efforts to reduce the climate footprint of their activity, while at the same time trying to minimise costs. Efforts are directed both at optimising their own operations and reducing the footprint in supply chains. The aim is to publicly show commitment to the planet, to ensure that the use of their services does not negatively affect the climate footprint of their customers' activity and even reduce it, thus promoting intensive use.

Future scenarios seem to portend significant growth in the use of AI in many sectors, with solutions that may impose additional constraints such as limited latency, the need to geographically confine data, or measures to ensure strict compliance with privacy and non discriminatory regulations.

All these factors shape an environment of international competition in which autonomy and strategic sovereignty of the countries may be at stake.

Keywords:

Artificial intelligence, Environment, Carbón footprint, International competition.

Cómo citar este documento:

RAMÍREZ MORÁN, David. *La huella medioambiental de la IA*. Documento de Análisis IEEE 07/2025. enlace web IEEE y/o enlace bie[3] (consultado día/mes/año)

Introducción

El impacto sobre el medioambiente de las actividades humanas ha recibido una creciente atención durante las últimas décadas. En la actualidad, la emergencia climática se ubica con frecuencia en los primeros puestos de las encuestas a la población sobre los factores de riesgo para las sociedades, y es tema de discusión en la práctica totalidad de los foros internacionales.

La inteligencia artificial irrumpe en este contexto con unos antecedentes que, entre la población general, despiertan preocupación. El elevado consumo de energía necesario para el minado de criptomonedas basadas en pruebas de trabajo, como bitcoin o ethereum, ha generado alertas sobre las consecuencias que los elevados requerimientos de capacidad de cálculo de la inteligencia artificial pueden suponer para el medioambiente. Esto ha llevado a los principales actores a tomar medidas para reducir la huella medioambiental relacionada con los centros de datos en los que se alojan los dispositivos que realizan las operaciones necesarias. Se persigue reducir la cantidad de energía necesaria tanto para la computación de los resultados como para refrigerar los dispositivos que los evalúan. Esta reducción se traduce en una menor emisión de gases de efecto invernadero y también en una reducción de los costes asociados.

El problema no se puede circunscribir a los centros de datos y a actuar sobre lo que albergan, sino que se está dando un paso más. Las mejoras marginales que se obtienen actualmente con nuevos desarrollos en estas infraestructuras deben coordinarse con medidas adicionales relacionadas con las cadenas de suministro de energía y equipamiento, así como con la gestión del ciclo de vida de este.

En la búsqueda de soluciones, son muy diversas las aproximaciones adoptadas, sin ser exhaustivo, desde el desarrollo de procesadores más eficientes, el diseño de los centros de datos, el uso de refrigeración líquida, el origen de la energía utilizada, hasta la ubicación de los centros de datos para el aprovechamiento de fuentes de energía, naturales o artificiales, o de otros recursos o infraestructuras. Cada una de estas soluciones llega en muchos casos acompañada de problemas adicionales que limitan la viabilidad y el impacto sobre la huella de carbono que permiten alcanzar.

Por otro lado, la inteligencia artificial en sí misma también puede contribuir a la reducción de la huella climática de la actividad humana desde varios frentes. Como consumidor

importante de energía, puede jugar un papel estabilizador en el aprovechamiento de las fuentes renovables, pues su consumo es menos cíclico que el de la actividad humana y puede aprovechar el exceso de capacidad energética de una región cuando el resto de consumos se reduce. La conectividad global también permite trasladar el procesado de un lugar a otro con relativa facilidad, por lo que es posible aprovechar capacidad de proceso remota para conseguir una menor huella medioambiental.

La generalización del uso de la IA y la creciente dependencia de la disponibilidad de estas infraestructuras impone un requerimiento adicional en la resiliencia ante fallos y averías, que se traduce, a su vez, en necesidades de equipamiento adicional que también puede contribuir a la reducción de la huella energética de las instalaciones, entre otras ventajas que puede aportar la IA en la reducción de la huella climática de la actividad humana.

Los resultados proporcionados por la IA también contribuyen a la reducción de la huella medioambiental de la actividad humana. Sus capacidades de tratamiento de la información permiten identificar e implantar mecanismos para optimizar e incrementar la eficacia de las actividades y gestionar mejor la distribución y consumo, controlando cómo y de dónde se consume la energía.

La introducción paulatina de la IA en ámbitos de la sociedad alejados de la propia tecnología hace que la reducción de su huella medioambiental resulte, a su vez, de interés para aquellas empresas que hacen uso de ella. Cada vez hay una atribución más precisa de la huella de carbono de la aplicación de IA a lo largo de las cadenas de suministro y es necesario contabilizarla para evaluar la huella de carbono global de las organizaciones. Una huella nula o negativa del proceso de inteligencia artificial, junto con la optimización, fruto de su aplicación, de los procesos de las actividades principales de las organizaciones constituyen una ventaja competitiva e incrementa la sostenibilidad.

Sin embargo, no es posible alcanzar la eficiencia máxima porque otros factores también afectan a la viabilidad técnica o regulatoria de la implantación de soluciones que reducen la huella medioambiental. La menor latencia con la que hay que enviar resultados a los terminales está haciendo que la capacidad de cálculo se desplace desde los centros de datos al borde de la red, acercándola al terminal. La desconcentración de la capacidad de cálculo hace menos viables medidas de escala en la reducción de consumo u optimización de la refrigeración de los equipos.

Las regulaciones de protección de datos y privacidad también imponen limitaciones a la circulación de los datos. Pueden restringir la zona geográfica donde pueden ser almacenados y procesados. Así, deja de ser posible elegir dónde procesar la información sobre la base de criterios medioambientales y hay que tener en cuenta las limitaciones adicionales para ubicar el tratamiento de la información.

La preocupación creciente por la seguridad y privacidad de la información impone también restricciones de diversa naturaleza a la utilización de soluciones en la nube cuya implementación, por efectos de escala puede alcanzar mayores grados de eficiencia que la de centros de datos privados más pequeños.

Las soluciones aplicadas también dependen de la disponibilidad de equipamientos o infraestructuras que permitan implantarlas de forma viable en términos económicos.

El libre acceso a tecnologías, la disponibilidad de infraestructuras, las normativas de aplicación y las condiciones naturales se convierten así en dimensiones estratégicas que configuran los límites en los objetivos a los que pueden optar los diferentes países en la implantación y uso de la inteligencia artificial en su territorio.

Por último, no hay que olvidar las cuestiones políticas y geográficas que imponen condiciones de contorno infranqueables como el clima de la propia ubicación, la posibilidad de uso de energía nuclear o el acceso a agua o a fuentes renovables.

El consumo de la inteligencia artificial

La Agencia Internacional de la Energía cifra el consumo energético de los centros de datos, las criptomonedas y la inteligencia artificial, en 2022, en 460 TWh a nivel mundial, lo que supone un 2 % de la demanda global de energía (Agencia Internacional de la Energía, 2023). Esta cantidad se vería incrementada en un consumo equivalente al de toda Suiza o Alemania de cumplirse las previsiones de incremento para 2026.

A esta cantidad sería necesario incorporarle los consumos de los equipamientos involucrados en el ciclo de la información. La IA multiplica la cantidad de información que se puede procesar, lo que hace viable considerar un volumen de información mucho más amplio y diverso para la toma de decisiones. Pero, para poder actuar, es necesario recopilar la información para su procesado y uso. Es preciso establecer una cadena

completa formada por los sensores que captan la información, los dispositivos que la registran y transmiten a un nodo central cuando es pertinente y el nodo que recibe la información, la procesa y hace accesibles los resultados para que otros procesos los puedan utilizar. Ejemplo de estos equipamientos son las cámaras que recogen la actividad de las ciudades, sensores de variables meteorológicas o lumínicas, etc. así como toda la infraestructura necesaria para el transporte de la información que generan hasta donde es procesada.

Según un informe de PitchBook, para 2025, el 3,2 % de todas las emisiones de carbono del mundo provendrán de granjas de servidores de IA y «su costo ambiental no hace más que crecer a medida que la industria crece de una manera que prioriza la expansión en lugar de la eficiencia» (Bécares, 2023).

Las virtudes de la IA generativa se basan en procesados muy intensivos de la información, tanto en el proceso de entrenamiento como en la generación de resultados a petición de los usuarios. Ambas escalas no son comparables pues las dimensiones en las que generan cifras preocupantes son diferentes. En el caso del entrenamiento, requiere una elevada capacidad de procesado puntual, una vez, mientras que, para la generación de resultados, el riesgo proviene del número de peticiones realizadas por los usuarios. Por bajo que sea el consumo necesario para la generación de un resultado, cada vez serán más sistemas y cuestiones las que hagan uso de IA, resultando en un número de peticiones siempre creciente. Incluso las cuestiones más básicas podrían responderse recurriendo a IA, haciendo que cualquier actividad requiera gran número de consultas (Ramírez Morán, 2023).

Comparando las búsquedas tradicionales con consultas a modelos grandes de lenguaje como ChatGPT, un representante de Alphabet estimaba un coste diez veces superior con un consumo de hasta 3 Wh. Si cada búsqueda pasase a ser una consulta, la electricidad necesaria para dar las respuestas ascendería a 29,3 TWh al año, equivalente al consumo de un país como Irlanda (Robinson, 2023). Ante estos datos preocupantes, también hay motivo para albergar cierta esperanza pues entre 2010 y 2018 se quintuplicó la carga de trabajo de los centros de datos aunque el consumo de electricidad solo se incrementó en un 6 % (Chernicoff, 2023).

Respecto al consumo de energía de la utilización de sistemas ya entrenados, un estudio elaborado por profesores de universidades americanas concluye que las emisiones de

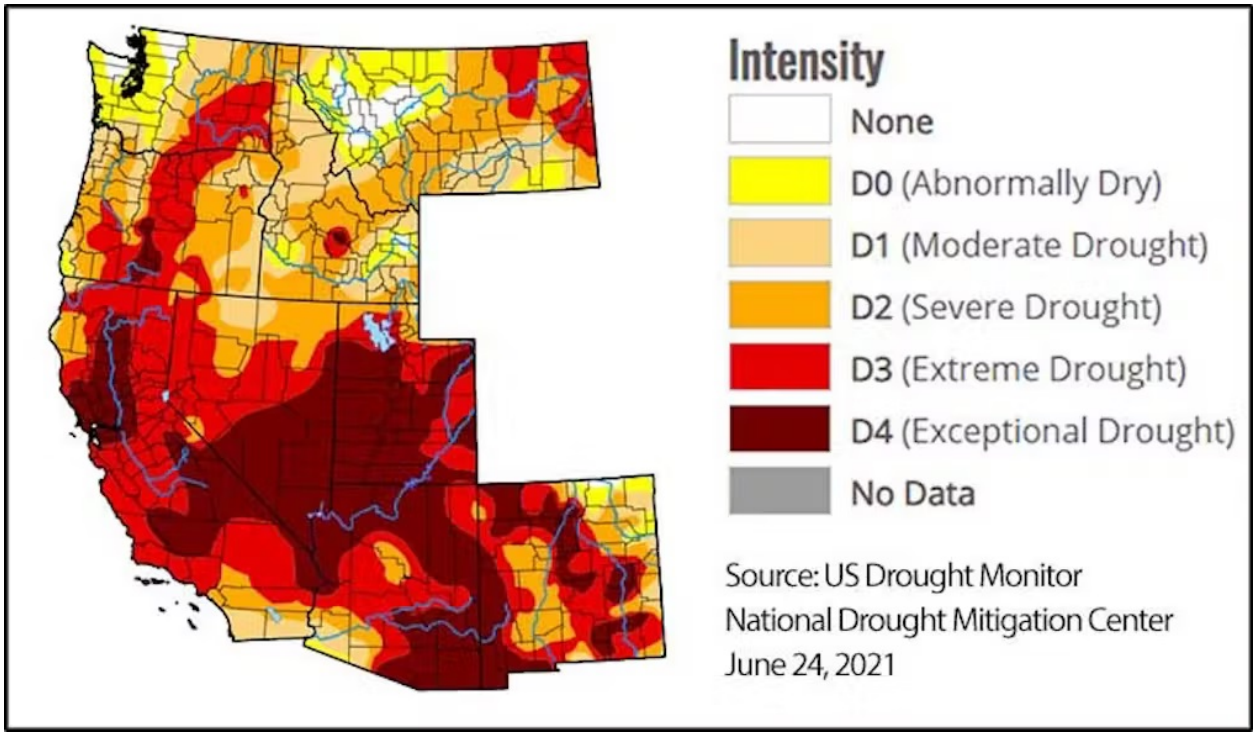
carbono asociadas a la escritura e ilustración son menores para la inteligencia artificial que para los seres humanos cuando se tiene en cuenta el consumo de los dispositivos utilizados (Tomlinson *et al.,* 2023). Una revisión crítica de este estudio destaca cómo se ha obviado la energía requerida por el entrenamiento de los sistemas de inteligencia artificial, los que generan el texto y las imágenes, que debe repercutirse en la elaboración de los contenidos (Mann, 2024).

Pero la energía no es el único consumo que está disparando las alertas medioambientales. El mercado de refrigeración de centros de datos en Europa alcanzó en 2020 los 3500 millones de dólares y se estimó un incremento anual de 2021 a 2027 de un 15 % adicional cada año (*Savills research*, 2023). El uso de agua para la refrigeración de los centros de datos es cada vez con más frecuencia un recurso escaso, especialmente cuando debe cumplir ciertos requisitos de calidad (Hidalgo García). En 2018, los centros de datos figuraban entre los diez principales consumidores de agua de los Estados Unidos, con un consumo de 513 millones de m^3, de los que una cuarta parte se utilizaban directamente para labores de enfriamiento directo (Siddik *et al.,* 2021).

CNDCP, una iniciativa autoregulatoria, presentó a la Comisión Europea un objetivo para 2040 de reducir el consumo de agua a 400 ml por MWh de energía de computación, objetivo muy ambicioso, pero que permite dirigir el trabajo y hacer un seguimiento de la evolución del sector (*Savills research*, 2023).

El problema del consumo de agua no se ha limitado a los ámbitos profesionales, sino que está llegando a la prensa general con noticias sobre los impactos[1] de centros de datos ubicados en zonas donde se producen sequías con frecuencia[2]. Tomando como ejemplo la costa oeste norteamericana, se observa cómo cerca de los polos tecnológicos de California, el problema de sequias es muy grave y puede afectar directamente a la instalación de centros de datos en esas zonas.

[1] Disponible en: https://www.newsweek.com/google-building-2-data-centers-which-require-heavy-water-use-drought-stricken-town-1647727.
[2] Véase en: https://time.com/5814276/google-data-centers-water/.



Tanto Meta como Intel declaraban que perseguían hacerse positivos en agua, proporcionando más agua de la que consumían, para 2030 (Robinson, 2022).

El informe del *think tank* Center for Data Innovation alerta sobre los riesgos que un exceso de regulación sobre la inteligencia artificial puede suponer para el compromiso entre los requerimientos de energía del modelo y la seguridad del sistema, con la implementación de técnicas de eliminación del sesgo y de salvaguardas que comprueben que los contenidos generados por los LLM no sean dañinos o contrarios a la ley (Castro, 2024: 13). Estos mecanismos imponen un consumo adicional de recursos que no contribuye directamente a la obtención de los resultados.

Reduciendo la huella medioambiental de la IA

La reducción de la huella de carbono de los centros de datos es un fin de los proveedores de servicios para que la utilización de sus productos no suponga un incremento de la huella de carbono de aquellas empresas que los incorporen a sus procesos de producción. Supondría una barrera a salvar para la incorporación de la tecnología a los procesos, mientras que el aún incipiente catálogo de funcionalidades que llevará a cabo

la inteligencia artificial permite prever un uso cada vez más intensivo y extensivo de la tecnología. Los esfuerzos de los proveedores constituyen, a su vez, una barrera por cuanto conseguir los elevados niveles de eficiencia energética fruto de su creciente experiencia resulta en muchos casos inalcanzable para equipos que desarrollaran las funcionalidades de inteligencia artificial en sus propios centros de datos. Se fomenta así la migración a la nube frente al desarrollo e instalación local de la tecnología.

A la hora de evaluar la huella medioambiental de la actividad de una organización, se definen tras ámbitos diferenciados, *scopes,* en inglés, correspondientes a la huella de la propia actividad de la compañía, *scope* 1, a la huella ambiental de los proveedores de recursos energéticos, o *scope* 2, y a la huella ambiental asociada a los proveedores de la empresa, *scope* 3, que aglutina la huella de fabricación y transporte de los productos que la empresa adquiere para realizar sus actividades. A través de estos ámbitos, es posible atribuir con precisión el origen de la huella medioambiental a lo largo de la cadena de valor de los productos y servicios.

Una actuación concienciada respecto al respeto al medioambiente es necesaria en todos los puntos de la cadena de valor. Desde la fabricación y distribución a la operación y posterior deshecho, es preciso mantener, en todo momento, una aproximación sostenible, tendente a la economía circular, en la que se minimizan los residuos, se favorece el reciclado y la recuperación de materiales y se persigue la minimización de los residuos que finalmente no pueden ser aprovechados o tratados.

Reducción de la huella de las fuentes de energía

El origen de la energía que alimenta el centro de datos también está tomando una importancia sustancial ante las crecientes imposiciones legales o voluntarias en relación con la huella de carbono. La generación de gases de efecto invernadero se evalúa en distintas dimensiones del modelo de negocio. La primera contribución proviene de la energía estrictamente necesaria para llevar a cabo las operaciones requeridas para la prestación del servicio.

Equipos más eficientes energéticamente y algoritmos más optimizados permitirán reducir el consumo exigido para el cálculo de los resultados de las peticiones. En un segundo escalón se sitúa el consumo auxiliar imprescindible para el funcionamiento de los

sistemas. Dentro de esta categoría se incluirían consumos como el necesario para la refrigeración del centro de datos, así como la eficiencia energética de la infraestructura en términos de pérdida de energía por su transformación, su aseguramiento con sistemas de alimentación ininterrumpida o las pérdidas propias de la distribución a los múltiples equipos existentes en el centro de datos. En un tercer escalón entra en juego la huella climática asociada a los proveedores de equipos, de construcción del centro de datos, de los suministros y el reciclaje de los residuos, etc.

En Estados Unidos se está planteando la construcción de los centros de datos en las proximidades de centrales energéticas ya existentes para reducir la carga sobre la red de distribución general que supone un centro de consumo tan intenso, mediante la construcción de conexiones dedicadas entre el centro de datos y la central. Se reducen así las pérdidas de energía debidas a su transporte a larga distancia, y se consigue también un incremento de la fiabilidad de la fuente de energía al utilizar infraestructuras dedicadas con las que se eliminan riesgos como la demanda variable y, por su menor complejidad, los asociados al mantenimiento.

Dos son las aproximaciones nucleares, por su reducida huella de carbono, que se están considerando. La primera es ubicar los centros de datos en las proximidades de centrales cuya capacidad no está plenamente explotada, de forma que la construcción de una línea dedicada que conecte la central eléctrica al centro de datos sea viable económicamente.

Por otro lado, las propuestas de nuevas centrales nucleares de un tamaño menor, bajo el modelo de reactores modulares pequeños, plantean la posibilidad de ubicar la central energética nuclear junto al propio centro de datos, en lo que se denomina colocalización (Chernicoff, 2023), con las ventajas antes reseñadas de emisión nula de carbono, menores pérdidas de transmisión, bajo coste y fiabilidad. Actualmente, el concepto de SMR (*Small Modular Reactors*), como los que alimentaban buques de guerra americanos desde los años cincuenta, no está completamente desarrollado, aunque, de resultar de interés en sectores como el del procesado de información, con las expectativas de crecimiento y estabilidad de demanda que presenta, podría acelerar el desarrollo de soluciones. En la actualidad hay varios SMR en construcción o esperando licencia en Argentina, Canada, China, Francia y Corea del Sur, aunque se encuentran todavía a años vista de su pleno rendimiento (Mann, 2022).

En todo caso, la contratación de suministros de energía de fuentes renovables a los proveedores tradicionales ya contribuye a limitar el impacto sobre el medioambiente de los centros de datos. La instalación de fuentes renovables en las propias instalaciones o en las proximidades también contribuye a la reducción de la huella energética de los centros de datos.

Mejora de las prestaciones de los equipos

La carrera tecnológica que se está produciendo en las dimensiones de los transistores que forman los circuitos integrados que procesan la información también tiene un trasfondo energético. Un menor tamaño del dispositivo reduce su capacidad eléctrica, por lo que la carga que es necesario inyectar o extraer del dispositivo cada vez que conmuta su valor lógico es más pequeña. La reducción de la carga que debe moverse conlleva una reducción de las corrientes que deben circular por los conductores de los circuitos integrados y, por tanto, una reducción de las pérdidas de energía que se traducen en calor. En definitiva, una tecnología de menor tamaño permitirá hacer las mismas operaciones con un menor consumo de energía o, equivalentemente, realizar más operaciones ante la misma disipación de calor, ser más eficientes. Pero no es posible reducir sin límite el tamaño de los dispositivos porque ya se está alcanzando el límite físico por el que los transistores, que deben funcionar como interruptores abiertos o cerrados, dejan de realizar correctamente su función debido a efectos cuánticos asociados al reducido número de átomos que forman los elementos semiconductores y aislantes de los dispositivos.

Reducción del impacto de fabricación

En un solo trimestre, una fábrica de chips produce casi 15 000 t de residuos, el 60 % peligrosos, así como un consumo de 927 millones de galones de agua, suficientes para llenar 1400 piscinas olímpicas (Belton, 2021). Cifras como estas justifican el importante esfuerzo realizado por las compañías para reducir el impacto. En 2022, Intel informaba de una reducción de un 4 % de generación de gases de efecto invernadero, en la senda de reducirlos un 10 % de 2019 a 2030, mientras que la energía que consumía provenía

ya en un 93 % de fuentes renovables. También para 2030 se había fijado como objetivo reducir a cero la generación de gases de efecto invernadero de su cadena de suministro.

Optimizar la refrigeración de los centros de datos

Incluso evaluar el consumo instantáneo de los equipos dentro del centro de datos se convierte en un problema cuando se imponen las condiciones necesarias para asegurar la privacidad de la información. Los equipos actuales incorporan numerosos mecanismos para poder evaluar su consumo instantáneo. El problema surge ante la necesidad de comunicar esta información hacia los dispositivos de control ambiental y de consumo eléctrico en un contexto donde la privacidad de los datos limita las conexiones que pueden realizar los equipos hacia el exterior.

Para hacer frente al problema que supone optimizar la refrigeración del centro de datos sin tener información sobre el consumo instantáneo de cada uno de los equipos que están operando en su interior, se recurre nuevamente a la IA con proyectos como el llevado a cabo por IBM en colaboración con la empresa japonesa NTT (Sharwood, 2024). Consiste en estimar el consumo de los equipos del centro de datos a partir de medidas de temperatura, humedad y flujo del aire en el centro de datos externas a los propios equipos de procesado de información. De esta forma, se conserva el aislamiento de los datos de los equipos a efectos de cumplimiento, desde la no conexión de dispositivos ajenos, la preservación del cifrado y la desconexión de redes ajenas a aquellas por las que circulan los datos. También se elimina la necesidad de que los equipos determinen su temperatura y la transmitan a los equipos de control de refrigeración del centro de datos.

En esta búsqueda de soluciones óptimas, Meta ha estado probando a elevar la temperatura de sus centros de datos hasta alrededor de 32° para reducir la necesidad de enfriamiento y, con ello, el consumo de agua y energía (Robinson, 2022).

Reducción de la emisión directa

Los centros de datos requieren la generación de energía en sus propias instalaciones para poder hacer frente a una pérdida del suministro de energía eléctrica. Para ello, deben contar con generadores que tradicionalmente han utilizado como combustible

gasóleo. Se están probando nuevas soluciones para reducir las emisiones de estos equipos, que van desde la utilización de biocombustibles (Miller, 2022) hasta la instalación de pilas de combustible alimentadas por hidrógeno (Roach, 2020).

Deshacerse del calor del centro de datos

En lugar de disipar el calor residual del centro de datos, emitiéndolo sin más al ambiente, se consideró aprovechar este calor para otros fines. Son varios los proyectos que trabajan en esta línea proporcionando agua caliente y o calefacción a ciudades cercanas a la ubicación del centro de datos. Sin embargo, no se trata de una solución aplicable de forma general pues requiere una infraestructura importante para llevar el calor desde el centro de datos a los destinatarios[3]. Esta infraestructura existe en algunas ciudades del norte de Europa y es allí donde se han implantado soluciones.

En esta línea, la Unión Europea, en la última revisión de la Directiva de Eficiencia Energética[4], estipula que los centros de datos con capacidad de procesado igual o superior a 1 MW deberán «tener en cuenta las mejores prácticas del código de conducta europeo en eficiencia energética de centros de datos», lo que significa implementar sistemas de recuperación de calor o demostrar que no era viable.

El Instituto Uptime estimaba (Smolaks, s.f.) en sesenta el número de proyectos de recuperación de calor en territorio europeo, mientras que en Estados Unidos solo hay seis proyectos actualmente y once se encuentran en desarrollo.

En Corea del Sur también se ha adoptado una aproximación novedosa por la que se colocalizan centros de datos y plantas de tratamiento de aguas residuales para reducir la energía necesaria para el tratamiento y aprovechar parte de las aguas tratadas para el enfriamiento del centro de datos (Robinson, 2022).

Inmersión de centros de datos en masas de agua

La inmersión del equipamiento de computación en grandes masas de agua es otra de las alternativas que se ha estudiado para eliminar el consumo y los costes de la

[3] Disponible en: https://www.theregister.com/2023/12/08/reusing_datacenter_heat_is_tricky/.
[4] Véase en: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ%3AJOL_2023_231_R_0001&qid=1695186598766.

refrigeración. Microsoft lideró una iniciativa, el proyecto Nattick[5], para evaluar las ventajas y la viabilidad de una solución técnica en este sentido. Los resultados fueron bastante buenos pues, no solo se consiguió el correcto funcionamiento del sistema durante todo el experimento, sino que se comprobó cómo los equipos incluidos en el contenedor habían tenido un número de averías inferior al de los ubicados en centros de datos tradicionales.

El deshecho del material de computación

No puede olvidarse que todo sistema tiene un ciclo de vida desde su incepción hasta su eliminación. Los equipos e instrumentos que posibilitan y dan soporte a una tecnología tienen un ciclo de vida que requiere una evaluación de su impacto medioambiental desde la obtención de las materias primas, a partir de las que se construirán los equipamientos, hasta la eliminación de los residuos que se generan.

La basura tecnológica es un problema identificado. Frente a este se plantea la solución del reciclaje, aunque, debido a la escala y complejidad que conlleva, es una solución que, aunque factible, no está siendo explotada actualmente en su máxima expresión. De hecho, son muchos los problemas que este tipo de basura está generando a nivel internacional porque aquellas tareas asociadas a la separación y reciclado de los materiales están siendo desdeñadas por los países más desarrollados. Se están estableciendo procedimientos para el deshecho de este tipo de materiales sin efectuar su reciclado en origen, enviándolos por diferentes procedimientos a otras zonas del globo en las que se tratan de manera menos eficiente. En estas zonas se genera una extracción menos respetuosa con el medio ambiente y con menor eficiencia. Sin embargo, la diferencia de costes de mano de obra, dado que se recurre a personal de muy baja cualificación en condiciones de trabajo insalubres, en muchos casos, se utiliza como argumento para justificar estas medidas que, orquestadas alrededor del reciclaje, son realmente malas soluciones a un problema complejo.

Sería necesaria una normativa que introduzca la simplificación del proceso de reciclado en el desarrollo de productos con tiempos de vida relativamente limitados, como son los

[5] Disponible en: https://natick.research.microsoft.com/.

aparatos electrónicos. Además de promoverse el reciclado de acuerdo con normativas que eviten la generación de basura tecnológica descontrolada, debería simplificarse el propio proceso de reciclaje y que sea posible incluso construir una industria a su alrededor que, de forma respetuosa con el medio ambiente, se aproveche de la reciclabilidad por diseño, para conseguir un elevado porcentaje de recuperación de materiales. Son materiales en muchos casos de gran valor, como son los metales preciosos o las tierras raras. Además, se encuentran en concentraciones muy elevadas y sometidos a procesos tanto físicos como químicos conocidos, lo que permitiría diseñar procesos de separación y purificación muy específicos y eficientes.

Los materiales recuperados resultantes podrían reintroducirse en las cadenas de producción para la fabricación de nuevos equipos sin necesidad de recurrir a las costosas operaciones de obtención en la naturaleza. Los plásticos con los que se fabrican las carcasas y soportes de los dispositivos son un ejemplo de cómo se está orientando la industria hacia estos objetivos, rescatando incluso residuos plásticos costeros para su fabricación[6].

Para las empresas, el coste asociado a la implantación de estas técnicas de diseño, donde deben quedar documentados los procesos aplicados a las materias primas, deben generar un beneficio que puede provenir por la vía de una reducción de la huella ambiental de la empresa, por la bajada de precios de las materias primas o por los beneficios que pueden obtener de encargarse de la gestión del reciclaje de sus propios equipos, añadiendo una actividad más al ciclo de prestación de servicios de sus productos. En caso de no poder alcanzarse este objetivo, podrían ser los Estados los encargados de implantar la normativa que obligara a la aplicación de estas medidas. Sin embargo, esta última vía se presta nuevamente a la aparición de contrastes por la aplicación de normativas de carácter nacional o regional diferentes en lugares distintos del globo.

Intereses y limitaciones nacionales

Ante el elevado incremento del consumo energético dedicado a tecnologías de la información, son necesarias medidas gubernamentales para dar soporte a la necesidad de generación y distribución de la energía eléctrica necesaria.

[6] El plástico reciclado llega a ordenadores y periféricos. Veáse en: https://plasticos-reciclados.es/2022/07/recycled-plastic-arrives-to-computers/.

Son varios los países que han implantado moratorias a la instalación de centros de datos en sus territorios, en atención al incremento del consumo energético que conllevan, así como a la mayor demanda que depositan sobre las redes de distribución. Los derechos de consumo industrial se están convirtiendo en objeto codiciado por las diferentes empresas dedicadas a los centros de datos.

Aquellos países con redes eléctricas más eficientes y actualizadas pueden ser polos de atracción para los nuevos centros de datos debido a la mayor fiabilidad que la provisión de energía proporciona. Por el contrario, los países que cuentan con infraestructuras de generación y distribución de la energía más precarias podrían ver mermado su atractivo para las empresas.

España se está haciendo hueco entre los principales anfitriones de centros de datos europeos, al lado de países como Holanda, Alemania, Francia, Italia… Los motivos para ello son su ubicación geográfica estratégica, que hacen del país un lugar propicio para la terminación de grandes cables submarinos, así como para la implantación de fuentes de energía renovables como paneles solares o generadores eólicos o en el mar.

La ubicación geográfica de España la sitúa en una excelente posición para competir internacionalmente. De hecho, se espera un crecimiento a altas tasas los próximos años debido a su ubicación, que motiva, a su vez, su conectividad a través de múltiples cables submarinos, la elevada capacidad de energía renovable y la creación reciente de regiones de nube por los principales proveedores. Todo ello con el respaldado gubernamentalmente con el plan España Digital (DatacenterDynamics).

La Comunidad de Madrid ha expresado su preocupación ante las limitaciones que la red nacional de distribución de energía puede suponer para el atractivo industrial de la región para la implantación de nuevos centros de datos en su territorio (Álvarez *et al.,* 2023). El consumo de este tipo de instalaciones requiere de una red de transporte que le de soporte y, en la actualidad, no hay suficiente provisión para hacer frente a la solicitud por parte de distintos operadores de autorizaciones de consumo de energía. Se trata de una muestra de la importancia que la red de distribución, de titularidad pública, en la mayor parte de los casos, supone para un sector tan eminentemente privado como es el de los centros de datos.

En el caso de los Países Bajos se han dado los primeros casos de rechazo social a la implantación de centros de datos que pueden afectar a la disponibilidad energética por el creciente consumo de energía. Los planes de Meta para construir allí un centro de datos se han trasladado a la localidad de Talavera de la Reina (Jiménez Arandia, 2024). Se beneficia así España de las restricciones de otros países por motivos energéticos o de otro tipo (DatacenterDynamics).

Contribución de la inteligencia artificial al medio ambiente

En contraste con el impacto que la inteligencia artificial tiene sobre el medioambiente, también proporciona mecanismos que permiten reducir la huella climática de la actividad humana. La aplicación de esta tecnología a cuestiones como el control ambiental y lumínico de edificios o la optimización de los consumos de energía para adaptarlos a las condiciones climáticas o de las infraestructuras resultan en beneficios para el medioambiente basados en dos premisas: la reducción del consumo total y la optimización del consumo para aprovechar al máximo las capacidades disponibles y ajustarlo a las condiciones específicas de cada momento.

La irrupción de la red inteligente o smart grid para la gestión de la energía supuso un punto de inflexión. El control tradicional de los equipamientos se hacía generalmente por franjas horarias definidas y mediante la aplicación de márgenes de temperatura controlados por termostatos. La introducción de la inteligencia artificial permite introducir nuevos paradigmas en la gestión de instalaciones como son el confort térmico y lumínico. En lugar de fijar una temperatura óptima general, esta se irá adaptando a las condiciones climáticas de forma automatizada, mediante algoritmos que determinan la temperatura y humedad óptimas.

Pero el consumo y la huella de carbono de los centros de datos no es la única relación de la inteligencia artificial con su huella. La propia IA se puede utilizar para la optimización de procesos industriales. Son muchas las soluciones que se están proponiendo en esta línea. El control del consumo de energía por intervalos se ha convertido en una medida de optimización de la generación eléctrica porque es posible controlar la carga de forma que, en los momentos de más demanda, algunos de los principales consumidores puedan reducir consumos instantáneos para descargar la red, o bien aprovechar

momentos de alta capacidad de generación pero bajo consumo para realizar tareas con altos requerimientos de energía a precios más reducidos y aprovechando la capacidad disponible.

Para las redes de suministro, tan importante como las fuentes de energía son los consumidores que la demandan. Por tratarse de un sistema que debe estar equilibrado, en el que el consumo instantáneo debe coincidir con la energía generada:

> «En algunos sistemas eléctricos, los centros de datos pueden ser capaces de ayudar a balancear es sistema o proveer otros servicios. En Irlanda, por ejemplo, la eólica ostenta un significativo y creciente porcentaje de la generación de energía (28 % en 2018). Sin embargo, gran parte de esta potencia eólica se genera por la noche, cuando la demanda de energía es baja en los sectores comercial y residencial. Con demanda nocturna sostenida, los centros de datos pueden absorber el exceso de suministro e incrementar la utilización de la electricidad de fuentes eólicas».

Además, los centros de datos podrían jugar un importante papel en la respuesta por el lado de la demanda. Aunque típicamente tienen un perfil de petición de energía estable, los grandes centros de datos están altamente automatizados y monitorizados, haciéndolos potencialmente más flexibles y adaptables en comparación con instalaciones industriales tradicionales. La regulación y las señales de precios pueden ayudar para sacar partido a este potencial (Kamiya *et al.,* 2019).

El Gobierno de Irlanda, dentro de los requisitos para la autorización de la construcción de un centro de datos, impone tres condicionantes relativos al suministro de energía de los que el tercero va precisamente en esta línea, requiriendo la flexibilidad de la demanda de energía, reduciéndola cuando así lo solicite un operador del sistema eléctrico (Agencia Internacional de la Energía, 2023).

Con la llegada de la IA y otras tecnologías habilitadoras aplicadas a la mejora de la eficiencia y sostenibilidad, se abren nuevos caminos como la denominada *Twin Transition,* clave para una industria digital y descarbonizada, o las ciudades inteligentes, que permitirán reducir la contaminación ambiental con medidas como la mejor gestión del tráfico y de los transportes públicos.

La economía de la empresa en la gestión de la inteligencia artificial

Cuestiones que se separan de la tecnología también tienen su impacto sobre la huella medioambiental de la IA. Detrás de las soluciones actuales se encuentran empresas privadas cuyo objetivo son los beneficios. Por ello, van a gestionar el negocio de forma que se maximicen sus intereses, lo que dará lugar a políticas que aumenten la salud financiera de la empresa, que reduzcan sus costes y de las que podrán hacer uso mediante una conveniente comunicación a la sociedad para respaldar su forma de actuar.

La preocupación por los riesgos en los que se incurre con la implantación de la IA es generalizada. Hay dos cuestiones especialmente relevantes en las que la regulación está incidiendo más con respecto a la IA generativa actual. Son la eliminación de sesgos por motivos como la raza, la religión o la ideología, y la prohibición de la generación de contenidos ilegales o fraudulentos.

Sin entrar en detalles tecnológicos sobre cómo abordar estos problemas, es necesario implantar en los sistemas de IA los mecanismos que aseguren el cumplimiento normativo a estos efectos, sin menospreciar el impacto social que podría suponer un flagrante incumplimiento de alguno de los principios. Los mecanismos requieren la aplicación de algoritmos que filtren las peticiones inapropiadas y que evalúen los resultados generados antes de remitirlos al usuario. La capacidad de cálculo que necesitan no contribuye directamente a la generación de resultados y suponen un uso adicional de recursos de computación para cada petición servida. Las empresas han destacado este último mensaje (Castro, 2024), incidiendo sobre el aumento del consumo de energía y de necesidad de refrigeración que supone, para abogar por un proceso legislativo que tenga en cuenta el impacto para el medioambiente que puede tener una legislación demasiado estricta.

El compromiso entre las restricciones impuestas y su viabilidad técnica queda también relacionado con los costes en necesidad de infraestructura y de consumo de energía para computación y refrigeración en que tienen que incurrir las empresas.

En términos puramente económicos, las empresas también están implantando políticas que responden a una maximización de la eficiencia de costes. La adquisición y amortización de los equipos informáticos constituyen dos de las principales cuentas en

los balances de las empresas. Pequeñas variaciones en cualquiera de estos ámbitos, debido al elevado volumen que manejan en su operación, suponen diferencias sustanciales en los resultados empresariales y, por tanto, en la evolución en los mercados de las empresas. Su madurez en el competitivo mercado actual contribuye a que hayan alcanzado un grado de eficiencia elevado, que reduce las opciones aún disponibles a las que pueden recurrir para mejorar su cuenta de resultados mediante medidas de fondo económico.

La extensión del plazo de amortización del equipamiento informático supone utilizar los mismos equipos durante más tiempo, asumiendo el creciente riesgo de mal funcionamiento debido a su mayor antigüedad antes de ser sustituidos. Permite retrasar la adquisición de nuevo equipamiento y, a lo largo de los años, reducir la cantidad de equipamiento adquirido para prestar los servicios. Son varios los hiperescalares[7] que han recurrido a estas medidas recientemente, entre los que se encuentran Amazon (Sharwood, 2024), Alphabet (Sharwood, 2024) o Cloudflare (Sharwood, 2024).

Esta medida también tiene sus efectos sobre el impacto ambiental debido a la rápida evolución tecnológica. Los nuevos dispositivos son más eficientes en términos de consumo de energía y de generación de calor gracias a nuevas arquitecturas más adaptadas a la resolución de algoritmos de inteligencia artificial y a la utilización de nuevos nodos tecnológicos (tamaño en nm) en la fabricación de los transistores, que reducen el consumo eléctrico y el calor generado.

Alargar la amortización del equipo supone su funcionamiento durante más años y que las ventajas de los nuevos equipos tarden más tiempo en poblar los centros de datos de las empresas. Durante este tiempo, los algoritmos estarán utilizando más energía de la que sería estrictamente necesaria si se compara con las prestaciones que los últimos dispositivos pueden proporcionar. La sustitución de un servidor antiguo por uno con las últimas tecnologías puede reducir hasta en un 30 % el consumo de energía necesario para realizar las mismas operaciones, lo que supondría un ahorro de alrededor de 480 $ durante un ciclo de vida de cuatro años, típico de este tipo de equipos[8].

[7] Proveedores de servicios de computación, comunicaciones y almacenamiento masivos.
[8] Disponible en: https://www.datacenterfrontier.com/voices-of-the-industry/article/11430647/waste-management-taking-out-the-trash-in-your-data-center.

Por el contrario, en favor de esta medida, los dispositivos fabricados estarán funcionando durante más tiempo, reduciendo el impacto de su fabricación en términos de consumo de materias primas y energía y de residuos generados durante su fabricación, así como el impacto que supone la retirada y reciclaje de los equipos decomisionados.

Conclusiones

El desarrollo de la IA no puede obviar el impacto medioambiental que su funcionamiento puede suponer. Los elevados consumos de energía y agua son factores muy importantes para la implantación de la IA porque afectan a dimensiones sensibles como el calentamiento global y la huella de carbono. Las tecnologías de la información, y dentro de ellas aquellas dirigidas a soportar la inteligencia artificial, suponen un consumo energético significativo con respecto al consumo energético global por lo que la optimización de estas infraestructuras y los algoritmos que en ellas se ejecutan son una obligación para conseguir la sostenibilidad del sistema en términos tanto económicos como ambientales. La creciente sensibilización sobre la sostenibilidad y el riesgo que la implantación de nuevas necesidades de consumo de energía supone para el calentamiento global es una cuestión que debe estar presente en toda nueva estrategia de trabajo.

Las empresas del sector tienen muy presente el impacto ambiental y muchas de sus decisiones técnicas y de negocio se encuentran guiadas tanto por la sostenibilidad como por la reducción de costes económicos y medioambientales de sus operaciones.

Las fuentes de generación de energía juegan un papel muy relevante en la huella climática de un centro de datos cuyo consumo, al menos, se va a mantener, si no crecer, a lo largo del tiempo. El uso de fuentes renovables como el viento o la luz solar permiten reducir la huella de carbono de las instalaciones. La energía nuclear también se considera como una fuente de energía sin huella de carbono, aunque no todos los países pueden recurrir a ella. Se están aplicando otras tecnologías para la reducción de la huella ambiental como la sustitución de combustibles fósiles por biocombustibles en los sistemas de respaldo dentro de la política de reducir al máximo la huella de carbono de los centros de datos

La geopolítica juega un papel importante en la huella climática de la inteligencia artificial, por cuanto afecta a qué soluciones resultan de aplicación y cómo de efectivas pueden resultar en cada territorio. Factores técnicos como las fuentes de generación disponibles, así como su tecnología —renovable, no renovable, nuclear…—, la disponibilidad y fiabilidad de la red de distribución o la distribución geográfica de las poblaciones en el territorio, no reducen la importancia de otros factores como la soberanía tecnológica, la autonomía o las regulaciones de protección de datos. Se genera así un mercado donde la competencia no se rige únicamente por los costes, sino que hay que incluir otros factores para determinar la viabilidad técnica y económica de las soluciones.

*David Ramírez Morán**
Analista del IEEE
Científico superior de la Defensa

Bibliografía

Agencia Internacional de la Energía. (2023). [Consulta: 2024]. Disponible en: Electricity 2024 - Analysis and forecast to 2026 (iea.blob.core.windows.net). Pp. 32-33.

Castro, D. (2024). *Rethinking Concerns About AI's Energy Use.*

DatacenterDynamics. (2022). Spain now a key data center hub for Southern Europe - Quark's Ricardo Abad. Disponible en: https://www.datacenterdynamics.com/en/marketwatch/spain-now-a-key-data-center-hub-for-southern-europe-quarks-ricardo-abad/

Hidalgo García, M. M. (2022). El consumo de energía y agua en los centros de datos: riesgos de sostenibilidad [en línea]. *Documento de Análisis IEEE 69/2022.* [Consulta: 2024]. Disponible en: El consumo de energía y agua en los centros de datos: riesgos de sostenibilidad (ieee.es)

Kamiya, G. y Kvarnström, O. (2019). *Data centres and energy – from global headlines to local headaches?*

Mann, T. (2024). *Think tank funded by Big Tech argues AI's climate impact is nothing to worry about.*

Ramírez Morán, D. (2013). Large Language Models: los nuevos actores de acceso al conocimiento [en línea]. *Documento de análisis 86/2023.* [Consulta: 2024]. Disponible en: Large Language Models: the new actors for knowledge access (ieee.es)

Robinson, D. (2023). AI processing could consume 'as much electricity as Ireland' [en línea]. *The Register.* [Consulta: 2024]. Disponible en: AI processing could consume 'as much electricity as Ireland' • The Register

Savills research. (2023). European Data Centres 2023: the watershed for data centre water usage.

Sharwood, S. (2024a). *Alphabet just banked $3B by stretching life of its servers.*

—. (2024b). *Amazon extends the life of its servers to six years, expects $900m benefit in 90 days.*

—. (2024c). *Cloudflare joins the 'we found ways to run our kit for longer' club.*

—. (2024d). IBM Japan and NTT think they can make datacenter aircon adjust to different workloads [en línea]. *The Register.* [Consulta: 2024]. Disponible en: https://www.theregister.com/2024/02/07/ntt_ibm_datacenter_heat_ai/.

Smolaks, M. (s.f.). *Heat Reuse: A Management Primer*. [Consulta: 2024]. Disponible en: https://uptimeinstitute.com/resources/research-and-reports/heat-reuse-a-management-primer

Tomlinson, B. *et al*. (2023). *The Carbon Emissions of Writing and Illustrating Are Lower for AI than for Humans.*