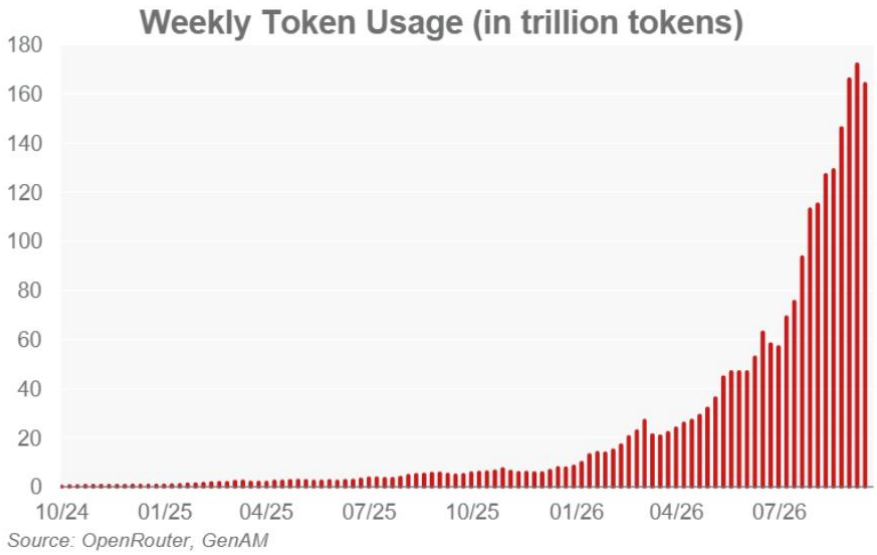
Demanda de IA: hasta ahora, justifica el gasto de capital

La IA está impulsando uno de los mayores ciclos de inversión de la historia de la tecnología moderna. Se prevé que la inversión global en infraestructuras relacionadas con la IA en 2026 ronde el billón de dólares, de los cuales la mayor parte corresponderá a EE. UU.[1]. La cuestión clave en materia de inversión es si la demanda de IA puede justificar la magnitud del gasto de capital que se está produciendo actualmente. Hasta ahora, los datos son alentadores.

Consumo de tokens

El consumo de tokens se ha convertido en la principal unidad de medida de la demanda de IA. A diferencia del software tradicional, que suele medirse en licencias o suscripciones, el uso de la IA se mide cada vez más en tokens: texto, código, imágenes y entradas multimodales procesadas por los modelos. A medida que la IA se expande hacia los copilotos, la programación, las búsquedas, el análisis, la atención al cliente y, sobre todo, los agentes, los volúmenes de tokens crecen mucho más rápido que la base de usuarios.

La utilización semanal de tokens en OpenRouter ha aumentado este año de unos pocos billones de tokens a más de 150 billones, y debería seguir creciendo a medida que se amplíe su adopción y se desarrolle la IA agentiva. En los últimos tres años, la caída de los precios unitarios de la IA se ha visto compensada por el rápido crecimiento del uso. En otras palabras, una IA más barata y eficiente no debilita necesariamente el argumento de inversión. Al estimular una demanda adicional y ampliar el mercado potencial, puede respaldar un gasto continuo en infraestructura, aunque la obtención de rendimientos financieros a corto plazo dependerá de la rapidez con la que la demanda se convierta en ingresos duraderos: incluido el uso combinado de modelos más baratos y de nueva generación.



Source: OpenRouter, GenAM



Source: Silicon Data, Bloomberg, GenAM

Adopción de la IA

Según el informe del Índice de IA de Stanford, el 88% de las organizaciones encuestadas afirmaron utilizar la IA en 2025, mientras que el 70% utilizaba IA generativa en al menos una función empresarial. Su difusión también ha sido más rápida que la de los ordenadores personales o Internet, alcanzando una tasa de adopción del 53% en los últimos tres años, tal y como también atestiguaron los directores generales de las empresas del S&P 500 durante la temporada de presentación de resultados del segundo trimestre de 2026.

La adopción de la IA es amplia pero superficial: la experimentación ya está muy extendida, mientras que el despliegue a gran escala, el rediseño de los flujos de trabajo y los beneficios financieros cuantificables se encuentran aún en una fase inicial. Una encuesta realizada por Mavvrik y Benchmarkit a 396 empresas reveló que, a pesar del seguimiento generalizado de los costes y de la existencia de presupuestos formales, solo el 11% podía prever el gasto en IA con un margen de precisión de ±10%. Los costes inesperados de la IA afectan cada vez más a los márgenes y a las decisiones empresariales.

Además, muchas empresas siguen en fase piloto, y los agentes de IA apenas están empezando a incorporarse a los flujos de trabajo empresariales, lo que sugiere que la demanda se encuentra aún en una fase temprana de su ciclo. La primera ola provino de la experimentación, la programación, los contenidos y el trabajo intelectual. La próxima ola traerá consigo una integración más profunda en los procesos, aplicaciones verticales y flujos de trabajo basados en agentes o autónomos. El KAIST —un instituto coreano— estima que, si las solicitudes diarias de los agentes igualaran los volúmenes de búsqueda de Google, la IA añadiría unos 200 GW de demanda energética, lo que supone más del 40% de la carga media de la red eléctrica de EE. UU. Se trata de un límite superior ilustrativo más que de una previsión: parte de la hipótesis de una eficiencia estática y de la adopción total de los agentes.

Expansión de los centros de datos

La expansión de los centros de datos constituye la columna vertebral física de las oportunidades que ofrece la IA. Además, pone de relieve los retos relacionados con la financiación, la refrigeración y las operaciones. McKinsey estima que la demanda mundial de capacidad de los centros de datos podría aumentar de aproximadamente 82 GW en 2025 a 219 GW en 2030. Se prevé que la capacidad relacionada con la IA aumente de unos 44 GW a 156 GW, lo que representará aproximadamente el 70% de la demanda total a finales de la década. La estimación de los analistas de que la utilización de tokens aumentará de casi 100 billones en 2026 a casi 250 billones en 2027 parece coincidir en líneas generales con el aumento previsto de la capacidad instalada en GW para 2030. Esta expansión también se observa en China[2], que cuenta con unos 24 GW de capacidad informática operativa en centros de datos y otros 50 GW en construcción o anunciados, lo que pone de relieve la escala cada vez más global de la carrera por las infraestructuras.

1 Según GS, la parte correspondiente a EE. UU. asciende a 581 000 millones de dólares; hasta 800 000 millones de dólares según el BIS y la Columbia Business School.

2 E. Olcott, «China se apresura a construir centros de datos en su apuesta por la supremacía en IA», Financial Times, 8 de octubre de 2026.

Michele Morganti, estratega sénior de renta variable, Generali AM (parte de Generali Investments)

Mattia Mammarella, analista de investigación, Generali AM (parte de Generali Investments)

Vladimir Oleinikov, analista cuantitativo sénior, Generali AM (parte de Generali Investments)

Federica Tartara, economista sénior, Generali AM (parte de Generali Investments)