

TRANSFORMACIÓN TECNOLÓGICA Un nuevo acelerador

La IA multimodal lleva el impacto de los algoritmos en las empresas a otro nivel

Los modelos que combinan varios tipos de datos a la vez son una especie de cerebro holístico que ya está impactando en la eficiencia de las organizaciones

CHARO BARROSO

La IA multimodal, una nueva variante perfeccionada de la revolución de los algoritmos, ya está impactando a fondo en la gestión de las organizaciones. «Por inteligencia artificial multimodal se entienden aquellos modelos de IA que tienen varios tipos de datos a la vez», comenta Jaime Pizarroso, profesor de Aprendizaje profundo y Explicabilidad de inteligencia artificial de la Universidad Pontificia Comillas (Madrid).

«Estamos acostumbrados –asegura este experto– a que se pregunte a una IA y devuelva un texto o a lo mejor genere una imagen. Una IA multimodal podría recibir a la vez un texto, una imagen y un audio de algo y devolver también texto, imágenes, audio... Ahora es más capaz de entender sobre aquello que le preguntas».

En competencia

La propuesta más famosa de IA multimodal es ChatGPT en su versión 4, señala Pizarroso: «Puede recibir directamente audio, texto, imagen, fichero PDF, incluso buscar información por internet. O generar imágenes, darnos texto, tiene un sintetizador que permite la respuesta a través de voz», precisa.

Otras grandes empresas también disponen de sus modelos: Google con Gemini o Meta con Llama 4. Hay muchas empresas, incluso compañías 'open source', con sus propios modelos que entienden sobre todo imágenes y texto, que es lo que más se está usando ahora.

Las posibilidades abarcan todos los sectores, detalla: «Para diagnóstico, en medicina, por ejemplo, podrías darle todos los libros que necesi-

te y una radiografía y preguntarle si un paciente está sufriendo algún tumor en los pulmones o qué tipo de rotura tiene en un hueso, lo que sea. En abogacía, le podrías entregar la legislación de un país, darle un caso y la sentencia y consultarle si cumple con la normativa de ese país. En cualquier sitio donde se estén utilizando datos de cualquier tipo va a tener un gran impacto la IA», asegura el profesor de la Universidad Pontificia Comillas.

Entre los peligros de la IA multimodal, según Pizarroso, figura la protección de la privacidad de datos y los costes: «De los modelos más grandes, son propietarios las empresas. Normalmente tienes que pasar a través de algún tipo de formulario que concedan los datos o pagar sumas de dinero bastante importantes. Y si lo quieres montar tú en tu propia empresa para no llevar los datos a un servidor de terceros, es una infraestructura muy cara, no solo de instalar, sino de mantener».

Además consume mucha electricidad. También recuerda el docente que estos modelos no son perfectos. «Tienen errores,

alucinan, te pueden dar respuestas equivocadas. Precisan supervisión humana», apunta. Y en el mundo laboral, a juicio de Pizarroso, deben actualizarse los perfiles: «Hay que pensar en la formación de profesionales en lo que se denomina 'lifelong learning'».

No hay escape

El impacto de la IA multimodal en la gestión de empresas y administraciones se siente en todas las áreas de negocio, según Juan Carlos Gutiérrez, CTO y socio de la consultora Setesca: «Los modelos multimodales y de razonamiento avanzado están detrás, por ejemplo, de chatbots o agentes conversacionales para atención al cliente o para asistentes de productividad del día a día. No hay ningún caso de uso que se escape a la inteligencia artificial».

Afecta a todos los ámbitos, incluso sin el conocimiento del usuario, sostiene: «Los modelos multimodales están inmersos en herramientas como Office. En en-

señanza, se aplica a plataformas tipo Moodle, Blackboard, de learning, plataformas de enseñanza administrada. En sanidad, el 'machine learning' permite la detección de enfermedades como cardiopatías o insuficiencias renales analizando imágenes».

ChatGPT o Gemini también son ejemplos de uso de IA multimodal. «Todas las empresas apuestan por este modelo, afirma, también el mundo de los abogados, las compañías que prestan servicios profesionales y aquellas que tengan en el 'customer service' su principal punto de acceso. La inteligencia artificial puede hacer informes, memorandos, ensayos, artículos de investigación que antes llevaban una semana y ahora pueden hacerse en una hora».

Pero cuidado con los riesgos, advierte Gutiérrez: «El modelo de adopción de despliegue de IA generativa, multimodal, monomodal o con modelo de razonamiento se tiene que basar siempre en un modelo de 'compliance', es de-

cir, en el gobierno del dato, en la privacidad. Antes de optar por un proyecto de inteligencia artificial generativa, la compañía debe definir la seguridad del dato», defiende este experto.

Con respecto a los países de nuestro entorno, el socio de Setesca estima que «aún estamos un poco a rebufo de Reino Unido o Francia, que tiene su propia IA, que es Le Chat». «España está haciendo alguna innovación, aunque no tenemos modelo de generación propio. Pero el movimiento potente viene de las empresas norteamericanas como Microsoft, OpenAI, Google, y también de las chinas, no únicamente DeepSeek o Manus AI», explica.

Falta de formación

Sobre los perfiles necesarios para acometer las tareas para liderar los procesos de IA multimodal, Gutiérrez ha detectado, también en los procesos de selección dentro de Setesca, una falta de formación, mientras que las grandes, «como Microsoft, Google y Amazon tienen sus programas de capacitación».

Juan Manuel Cigarrán, experto en IA, considera que la inteligencia



Publicación	ABC Empresa, 14
Soporte	Prensa Escrita
Circulación	133 666
Difusión	121 283
Audiencia	341 000

Fecha	15/06/2025
País	España
V. Comunicación	66 408 EUR (76,736 USD)
Tamaño	125,80 cm² (20,2%)
V.Publicitario	9020 EUR (10 423 USD)



artificial multimodal va a suponer un cambio total de paradigma en la gestión de las organizaciones. Y compara sus posibilidades con la configuración de un Tesla: «Es un coche con muchos inputs, cámaras por todos los sitios, gran cantidad de algoritmos y modelos que se unen, como la localización por satélite o el reconocimiento con tecnología Lidar. Gracias a toda esa información, el automóvil es capaz de detectar a una persona cruzando o que viene una curva y así ir adaptando dinámicamente el volante».

«No es una mejora y ya está. Lo que reciben las empresas es la capacidad para procesar los datos de una manera mucho más rica, un modelo más complejo que va a permitir sacar más provecho. Ahora mismo las empresas están automatizando tareas aisladas, pero podrán hacerlo con más cobertura», subraya. Pero la multimodalidad no es solo combinar tipos de datos. Se trata de crear una especie de inteligencia con visión periférica, que no depende de un solo ángulo, una especie de cerebro holístico. «Se mete todo en la coctelera para que el sistema tenga más contexto y sea más preciso», argumenta Cigarrán.

Comprensión integral

El gran salto no consiste solo en realizar tareas más rápidamente, sino en entender contextos completos. En lugar de pasar datos entre herramientas, los sistemas empiezan a entender y decidir con inputs multimodales (texto, imágenes, voz, datos estructurados). «Vamos a pasar de una automatización de tareas aisladas a una comprensión integral del negocio», insiste.

Uno de los aspectos donde más incide la revolución de la IA multimodal es en el 'machine learning'. «Estamos en el proceso de un aprendizaje sinérgico, más unificado –aclara el experto–. Kip, de OpenAI, aprende conceptos visuales y textuales simultáneamente. Flamingo, de DeepMind, junta todo desde el principio. Vamos a necesitar bastantes menos datos para llegar al mismo rendimiento, porque el modelo ya está relacionando los distintos formatos cuando se está entrenando. El aprendizaje automático antiguo estaba mucho más limitado. Gracias a la explotación de manera eficiente de redes neuronales como transformers o Stable diffusion estamos en el siguiente nivel».