

7-3-2026 elDiario.es:

Anthropic es uno de los mejores ejemplos de cómo funciona el autobombo en la industria de la inteligencia artificial. Esta startup suele ser una gran desconocida fuera del sector. ChatGPT, Google e incluso algunos modelos chinos acaparan más atención que ella. Algo que contrasta con lo que ocurre de puertas adentro en eventos como el Mobile World Congress celebrado esta semana en Barcelona, donde un alto ejecutivo de una gran teleco europea reconocía en privado que lo que está haciendo Anthropic es “de lo más impresionante” que ha visto en toda su carrera.

En esos círculos especializados, afirmar que dentro de 10 años Anthropic será considerada la gran ganadora de la revolución de la IA ni siquiera se considera una apuesta demasiado arriesgada. Al menos así era hasta esta semana, después de que Donald Trump se haya propuesto acabar con ella tras denunciar que está dirigida por “locos de izquierda” que intentan “dictar” cómo el “gran Ejército de Estados Unidos lucha y gana guerras”.

El motivo es que en el momento clave de la ofensiva contra Irán, **Anthropic se negó a ceder a las exigencias del Pentágono para utilizar sus modelos de inteligencia artificial sin restricciones**. El gobierno estadounidense exigía usar su tecnología para “todos los propósitos legales”, pero la startup se opuso al considerar que los “guardarraíles” establecidos en su contrato de 200 millones de dólares con el Departamento de Guerra eran fundamentales para impedir que la IA tome decisiones equivocadas con armas autónomas o se use para la vigilancia masiva de su propia población.

La reacción de la Casa Blanca ha sido furibunda y ejemplarizante. No solo ha impedido que Anthropic pueda volver a contratar con la administración, sino que quiere etiquetarla como un “riesgo para la cadena de suministro”. Esto implicaría que ninguna empresa que contrate los servicios de la startup podría tener relaciones con el Estado. En la práctica, es un asesinato corporativo...

Anthropic

Anthropic es una empresa estadounidense de inteligencia artificial fundada en 2021 por exmiembros de OpenAI, como los hermanos Dario y Daniela Amodei. Se especializa en el desarrollo de modelos de lenguaje grandes con un fuerte enfoque en la seguridad, la ética y la alineación de la IA, utilizando conceptos como la “IA constitucional” para garantizar comportamientos predecibles y alineados con valores humanos.

Origen y Fundación

La compañía surgió por preocupaciones sobre la dirección comercial de OpenAI, priorizando en cambio la **investigación responsable de IA**. Ha recibido inversiones masivas de Google y Amazon, posicionándose como rival clave de OpenAI y ChatGPT.

Producto Principal: Claude

Su modelo estrella es Claude, una familia de chatbots generativos capaces de generar texto, analizar documentos, programar y responder preguntas complejas. Versiones recientes como Claude Opus destacan en tareas complejas como análisis financiero, manejo de datos masivos y detección de errores propios.

Enfoque Ético y Actualidad

Anthropic enfatiza la interpretabilidad y seguridad de sus sistemas, operando incluso en entornos clasificados del gobierno de EE.UU. En 2026, ha enfrentado **tensiones con la administración Trump por negarse a relajar protecciones éticas**.

Dario Amodei, cofundador y CEO de Anthropic, representa el liderazgo enfocado en IA segura.