

¿Peligro real o paranoia? Investigador de Anthropic renuncia y alerta sobre el futuro de las inteligencias artificiales

El debate sobre los riesgos de la inteligencia artificial volvió a escalar luego de que el investigador Jacob Coxon renunciara a Anthropic, empresa estadounidense dedicada al desarrollo de inteligencia artificial, y cuestionara la velocidad con que compañías como Anthropic y OpenAI están desarrollando sistemas cada vez más avanzados. Coxon, quien trabajó en ambas empresas, advirtió sobre una eventual superinteligencia capaz de mejorar sus propias capacidades.

A través de una serie de publicaciones en X, el exinvestigador sostuvo que “están compitiendo directamente por llegar a la superinteligencia automejorable y están jugando con nuestras vidas”, planteando que la carrera tecnológica podría estar avanzando más rápido que la capacidad para comprender y controlar sus posibles consecuencias. También advirtió que futuros sistemas podrían superar las capacidades humanas, vulnerar sistemas informáticos y adquirir mayores niveles de autonomía.

Anuncio Patrocinado

Las declaraciones rápidamente escalaron a medios nacionales e internacionales, donde el caso comenzó a ser abordado desde distintas perspectivas. En una entrevista difundida por CNN Chile, Coxon profundizó se refirió a episodios recientes en los que sistemas de IA habrían participado de ataques informáticos de manera autónoma: “Fue una oleada concentrada de hackeos que llevaron a cabo por voluntad propia. Y creo que si extrapolamos al futuro, el nivel de capacidades de estas IA’s con esa misma autonomía, podrían causar estragos extremos”.

¿Peligro real o paranoia? Investigador de Anthropic renuncia y alerta sobre el futuro de las inteligencias artificiales



El caso abrió un amplio debate en redes sociales: mientras algunos usuarios consideran que las advertencias deben tomarse seriamente ante el avance de estas tecnologías, otros las califican de exageradas o incluso imposibles. Entre publicaciones virales y opiniones contrapuestas, la discusión vuelve a instalar una pregunta que aún no tiene una respuesta: ¿estamos frente a un peligro real que debe anticiparse o ante escenarios extremos sobre un futuro que todavía es incierto?

¿Peligro real o paranoia? Investigador de Anthropic renuncia y alerta sobre el futuro de las inteligencias artificiales

y tú, ¿qué opinas?