

Por que preocupa que a IA poida ameazar á humanidade e como de real é ese temor?

Nas últimas semanas, as advertencias sobre os perigos potenciais da IA impulsaron á acción coordinada para desacelerar o desenvolvemento desta tecnoloxía.

As afirmacións de investigadores actuais e anteriores de que a IA porá en perigo á humanidade -un deles mesmo suxire que hai máis dun 10% de probabilidades de que "mate todos os humanos"- fixeron saltar as alarmas.

Os principais desenvolvedores, OpenAI e Anthropic, instaron aos lexisladores para regulalo, aínda que algúns suxiren, talvez con certo cinismo, que esta medida pretende afianzar o seu dominio sobre os seus rivais. En calquera caso, as acusacións de gran repercusión causaron preocupación.

Isto é o que sabemos.

Por que algunhas persoas pensan que a IA podería ser unha ameaza?

Algúns investigadores de IA afirman que existen sinais preocupantes de que os seus sistemas están a desenvolver as súas capacidades con demasiada rapidez.

Evan Hubinger acaparou os titulares das noticias ao afirmar que existe unha probabilidade superior ao 10% de que a IA "poida acabar con todos os humanos" na próxima década.

Hubinger traballa no campo da aliñación, que busca incorporar ideas e principios éticos humanos á IA. Noutras palabras, pretende que se aliñe cos valores humanos. Outro investigador abandonou Anthropic e advertiu que quen desenvolve a tecnoloxía está a perder o control sobre ela.

Estas preocupacións son as máis recentes dunha longa serie de temores sobre o potencial da IA para ameazar á humanidade se chega a igualar ou superar as capacidades humanas.

É algo que se remonta a décadas atrás, cando Alan Turing escribiu na década de 1950 que, se as máquinas intelixentes chegasen a ser posibles, poderían tomar o control.

-Que entendiches? Fai un resumo

- Cal é a túa opinión sobre o tema?