OPINIÓN

La pausa que no pausa: poder y gobernanza en la IA de frontera

Por Ernesto Ángeles · 15 de septiembre de 2026 · 9 min de lectura



El fin de semana pasado, tres de los principales empresarios de la inteligencia artificial a nivel mundial se posicionaron a favor de ralentizar el desarrollo de la IA de frontera: Dario Amodei, CEO de Anthropic; Sam Altman, CEO de OpenAI; y Elon Musk, CEO de xAI.

La coincidencia entre tres empresas rivales naturalmente causó un terremoto dentro y fuera del mercado digital. Pero vayamos por partes, ¿de qué trata tal posicionamiento y por qué empresas rivales querrían cooperar? ¿esto significa un movimiento en favor de la humanidad? ¿es otra estrategia de marketing del miedo y la (i)responsabilidad corporativa? O ¿estamos ante algo diferente?

Para empezar, fue Dario Amodei quien lanzó la propuesta a través de un texto titulado "*We Must Pace the Frontier*", en el que expone su diagnóstico, sus preocupaciones y un plan de acción.

En síntesis, Amodei parte justificando sus preocupaciones debido a dos factores: el rápido avance del desarrollo de la IA en lo que él se refiere como la “siguiente generación de la IA”, la auto-mejora de la IA de forma autónoma; el segundo hecho es el incidente de OpenAI-Hugging Face, en el que un enjambre de agentes de OpenAI hackearon a Hugging Face, a lo que el CEO de Anthropic plantea como posibilidad que, ante el avance de estas capacidades, en un plazo de seis a doce meses un enjambre de agentes de IA pueda llegar a apoderarse de buena parte de internet.

Con estos riesgos en mente, la propuesta de Amodei es un plan de tres pasos para "la **construcción de IA a un ritmo equilibrado que aspire a garantizar su seguridad**, a la vez que logre sus beneficios y lidie con importantes dilemas geopolíticos".

El primer paso consiste en que las empresas de alta tecnología -con fomento de los gobiernos- se comprometan a dar acceso continuo, similar al de los empleados, a un equipo de evaluadores externos, cuyo rol sería verificar el cumplimiento de las prácticas y los compromisos de seguridad, reportar incidentes y ayudar a evaluar la alineación no solo de los modelos de IA ya completados, sino también de los procesos de entrenamiento.

El segundo paso es un escalamiento de estos principios entre países democráticos, los cuales establecerían estándares de seguridad, así como límites al avance de la IA. Por último, en el tercer paso, bajo el liderazgo de EUA, los países democráticos llegarían a un acuerdo con autocracias, principalmente con China, para establecer esta pausa.

Una lectura inocente y literal del texto nos haría pensar ¿quién se negaría a acordar un comportamiento responsable de las empresas, en el cual se priorice la seguridad? Sin embargo, la realidad es más compleja y, por desgracia, este compendio de buenas intenciones esconde tras de sí diversos intereses, deudas y ambiciones comerciales; así como una tendencia a la que todos nos deberíamos oponer: que los oligarcas de la tecnología diseñen y dicten las reglas e instituciones internacionales del futuro.

El punto de partida de este posicionamiento tripartito son las **fuertes presiones financieras que comparten**: en el primer trimestre de 2026, OpenAI ingresó $5.7 mil millones, pero quemó $3.7 mil millones de efectivo y registró una pérdida operativa de $9.3 mil millones; xAI perdió $2.47 mil millones sobre $818 millones de ingresos y destinó $7.7 mil millones a gastos de capital. Anthropic representa una excepción parcial, aunque también comprometió unos $80 mil millones en cómputo.

No obstante, el problema estructural no es que estas empresas gasten mucho, sino la forma en que **crecen sus costos frente a sus ingresos**, ya que el costo de desarrollar el siguiente modelo de frontera ha aumentado de manera acelerada, mientras los retornos comerciales no han crecido al mismo ritmo.

Por tanto, **una pausa no resuelve la burbuja, pero compra tiempo, si el desarrollo se ralentiza, la presión de gastar *más* para no quedar atrás disminuye y las empresas pueden presentar márgenes que mejoren sin necesidad de un salto discontinuo en capacidades**. Y si la "super IA" no llega en el corto plazo, al menos la narrativa de seguridad ofrece una explicación plausible para el estancamiento que no sea que el modelo de negocio enfrenta límites serios.

Asimismo, Amodei habla de los riesgos en la mejora de la IA, pero no da ninguna prueba de tal avance, sino que se refiere únicamente a posibilidades e hipótesis, el único sustento real que ofrece es el incidente agéntico entre OpenAI y Hugging Face; sin embargo, su preocupación es engañosa, ya que el incidente no fue una IA que cobró conciencia y se rebeló, sino una concatenación de decisiones humanas negligentes: desactivar clasificadores de seguridad, permitir que el entrenamiento continuara a pesar de observar comportamientos riesgosos, y decidir que la evaluación podía seguir a pesar de la conexión a Internet.

Además, **Amodei no menciona ningún plan para pausar la implementación y despliegue de agentes de IA que ya existen** y se seguirían desplegando bajo la pausa propuesta y aceptada, ni mucho menos su regulación; regular el despliegue significaría establecer estándares obligatorios para la implementación de agentes: puntos de control humanos obligatorios, límites de autonomía, trazabilidad auditable, responsabilidad legal clara... Una pausa en el desarrollo es voluntaria, coordinada y no ejecutable sin cooperación internacional, una regulación del despliegue es obligatoria, unilateral y ejecutable por cada jurisdicción.

Y en cuanto a las propuestas, Amodei propone que cada empresa de IA otorgue acceso permanente, similar al de un empleado, a un equipo de evaluadores externos (como la organización METR) para verificar prácticas de seguridad y reportar incidentes.

El tema aquí es que la evaluación de las empresas por un tercer actor corre el riesgo de convertirse en un sustituto de facto de la regulación pública por una auditoría privada, por lo que no existirían sanciones de por medio; tampoco habría rendición de cuentas pública; su efectividad dependería de la cooperación entre rivales y los criterios de cooperación con las instancias evaluadoras quedarían en manos de las empresas.

Y si el modelo es como el que ofrece METR, cabe señalar que, aunque la organización no acepta dinero de las empresas de IA, sí ha recibido acceso gratuito a recursos computacionales, lo que plantea preguntas legítimas sobre la independencia y las condiciones de una evaluación que depende de la infraestructura de las propias empresas.

En lo que corresponde a la internacionalización de este acuerdo, aquí **el rol de los Estados sería el de implementar y revisar antes que el de regular, lo que en la práctica podría implicar una forma de captura regulativa y condenaría a las pequeñas y medianas empresas** a enfrentar grandes costos para cumplir con el acuerdo ya impuesto, implementándose un candado de facto.

El riesgo es que esto no produzca un ecosistema más seguro, sino la institucionalización de un oligopolio por medio de una regulación que favorece a quienes parten con ventaja de infraestructura de cómputo y capital financiero. El mismo Amodei reconoce que esta coordinación requeriría una exención antimonopolio del gobierno de EE. UU., es decir, que el Estado legalice lo que de otro modo sería colusión entre competidores.

Luego **introduce a China y a otras autocracias como factor de competitividad**, pero necesarios para que el acuerdo sea efectivo; por lo que la estrategia radica en mantener la hegemonía tecnológica de EE.UU. frente a rivales geopolíticos, y acaparar chips y energía como foso defensivo para que no surjan nuevos competidores.

La lógica es circular: debemos pausar, pero solo podemos pausar si mantenemos la ventaja sobre China; por lo tanto, nunca podemos pausar más de lo que la ventaja permite, y como la ventaja es siempre inestable, debemos siempre priorizar la velocidad sobre la seguridad. Es un estado de excepción permanente disfrazado de responsabilidad y seguridad.

Visto de forma más amplia, este movimiento no es una propuesta aislada y no sólo debería importarle al mundo de la tecnología; ya que esta propuesta forma parte de un patrón más amplio en el que las empresas tecnológicas buscan ocupar el espacio normativo antes de que lo ocupen los Estados.

La instrumentalización del Estado y las instituciones internacionales para la imposición de regulaciones dictadas directamente por los empresarios de la tecnología. Sin controles democráticos, sin legitimidad social, sin consenso, sin discusión… sólo capital, tecnología y condicionamiento a través de una nueva fase en el proceso de producción y acumulación.

Sin embargo, **la pugna entre las empresas, los oligarcas, el Estado y los líderes políticos apenas empieza, así lo demuestra la negativa de Trump de aceptar una pausa en el mercado de la IA; además de la respuesta de China, en donde en ambos casos se hace patente el rol del Estado** y sus ambiciones a la hora de instrumentalizar la inteligencia artificial en su proyección de poder internacional y la construcción del sistema internacional del futuro.

Y es que mientras que unos quieren imponer su voluntad sobre el Estado y las instituciones mundiales, las dos principales potencias buscan potenciar las capacidades del Estado, no sólo hacia el exterior de sus fronteras, sino también al interior: más vigilancia, más gobernanza tecnológica, más automatización en el cumplimiento de las reglas…

Lo que está en juego no es únicamente la velocidad del desarrollo de la IA, sino quién tiene la autoridad legítima para definir sus reglas. Que tres CEOs de empresas rivales se coordinen para proponer un marco regulatorio global —sin mandato democrático, sin deliberación pública, sin rendición de cuentas— no es un acto de responsabilidad: es la expresión más clara de un modelo de gobernanza en el que el capital tecnológico busca sustituir a las instituciones que no eligió y no controla.

La pugna por el poder tecnológico apenas comienza. Estados Unidos y China la procesan como competencia geopolítica y las empresas, como oportunidad de consolidación de sus ambiciones globales. En esta pugna, los países que no son potencias ni sedes de los grandes laboratorios enfrentan un riesgo particular: convertirse en sujetos pasivos de reglas que no negociaron. México no es la excepción

México tiene ante sí una ventana que no permanecerá abierta indefinidamente, pero aprovecharla exigirá algo más escaso que el capital o la infraestructura: una **política tecnológica soberana**, negociada con determinación en los foros donde aún se puede incidir —el T-MEC entre ellos— y orientada a la co-producción y la transferencia tecnológica antes que a la dependencia periférica. No es una opción cómoda, pero es la única que no deja la definición del futuro tecnológico en manos de un imperialismo en decadencia.

Publicado en Punto de Referencia. Dirección original: https://puntodereferencia.mx/opinion/la-pausa-que-no-pausa-poder-y-gobernanza-en-la-ia-de-frontera
Descargado el 21 de noviembre de 2026.