IA: producción social, apropiación privada

Carta a Lucho
6 de julio 2026

Este documento se distribuye bajo una licencia propia. Véanse las condiciones completas en la última página.

¡Querido amigo!

Tu pregunta es perfectamente legítima, pero el singular «la IA» oculta una dificultad decisiva. La inteligencia artificial no existe como una cosa única, homogénea y separable. Lo que llamamos «IA» es un complejo sociotécnico compuesto por modelos, infraestructuras, datos y formas de acceso sometidos a relaciones de propiedad diferentes. No existe una sola IA, sino una pluralidad de sistemas y modelos concretos, ubicados en diferentes espacios geográficos y sometidos a distintas formas de propiedad. Además, hay que analizar los diferentes niveles del funcionamiento de la IA y en cada nivel preguntarse quienes controlan esos distintos niveles.

Primero. La base material en forma de servidores

La primera base material de la IA está constituida por inmensos centros de datos. Para desarrollar y operar grandes modelos como los ofrecidos por OpenAI o Anthropic se necesitan miles y miles de computadoras que trabajan juntas. Esos servidores contienen especialmente procesadores de Nvidia, originalmente procesadores gráficos usados como aceleradores de cómputo para IA, cada uno de los cuales cuesta alrededor de 40.000$. Un servidor completo con ocho de estos procesadores puede costar entre 300.000 y 500.000 dólares, mientras que un sistema integrado en un rack con decenas de procesadores alcanza varios millones de dólares, sin contar todavía el edificio, la alimentación eléctrica, la refrigeración y las redes necesarias. Las inversiones necesarias se cuentan en decenas y, en el conjunto de la industria, en cientos de miles de millones de dólares. Esta concentración de capital se refleja también en el valor bursátil de las empresas y en las fortunas de sus principales accionistas. Esa inmensa concentración de capital se ilustra en una de las noticias más recientes, según la cual, que el señor Elon Musk es el primer ser humano que controla personalmente un capital de más de un billón de dólares. Y con ese capital controla una gran parte del desarrollo técnico de la IA (xAI), de la robótica (Optimus), de las redes sociales (X) y del uso comercial del espacio (SpaceX). También los demás grandes actores dependen de una concentración extraordinaria de capital, aunque no todos alcancen las dimensiones del conglomerado controlado por Musk.

La primera consecuencia de esta concentración es la siguiente: ¿se puede reproducir esta cantidad de capital? Sabemos muy bien que el capitalista no invierte su capital porque quiere hacer avanzar con la tecnología. El capitalista quiere convertir dinero en más dinero. En este momento, los grandes actores están perdiendo un montón de dólares cada día. Por ejemplo, según documentos comunicados a sus accionistas, OpenAI consumió 3.700 millones de dólares en efectivo durante el primer trimestre de 2026, es decir, unos 41 millones diarios. Mientras OpenAI consume decenas de millones de dólares cada día, Nvidia obtuvo un beneficio neto de casi 73.000 millones de dólares en el ejercicio fiscal que terminó el 26 de enero de 2025 y uno de unos 120.000 millones en el ejercicio fiscal que terminó el 25 de enero de 2026.

Y eso nos lleva a la segunda consecuencia. Con esa concentración de capital también se concentra cada vez más el poder. Esto se ha visto con especial claridad últimamente. El presidente de los EE. UU., Trump, considerado incorrectamente el hombre más poderoso, se reúne a cada rato con los grandes actores. Al leer esas noticias, el «hombre más poderoso» parece un perrito que se levanta sobre las patas traseras delante de su dueño. La imagen es cómica, pero la relación real es más compleja: el capital tecnológico necesita al Estado para proteger sus mercados, subvencionar sus infraestructuras y asegurar sus intereses, mientras que el Estado depende cada vez más de esas empresas para su estrategia económica y militar.

Segundo. La base técnica: el modelo, sus parámetros y su control

La segunda base de la IA es el propio modelo. Su funcionamiento no se parece al de un programa clásico, en el que una serie de reglas escritas de antemano conduce necesariamente a un resultado determinado. Un modelo lingüístico se construye mediante un proceso de entrenamiento en el que aprende relaciones estadísticas a partir de una enorme cantidad de textos.

Para hacerlo, los textos se dividen en pequeñas unidades llamadas *tokens*. Un token puede ser una palabra, una parte de una palabra o incluso un signo. Estas unidades se convierten en representaciones matemáticas, es decir, en vectores dentro de un espacio de muchas dimensiones. Durante el entrenamiento, el modelo modifica miles de millones de parámetros internos, también llamados pesos, para aprender las relaciones entre esos tokens y los contextos en los que aparecen.

Cuando le hacemos una pregunta, el modelo normalmente no busca una respuesta ya escrita en una especie de biblioteca interna. Procesa la secuencia de tokens que recibe y calcula, paso a paso, cuál es el siguiente token más probable:

$P(x_t \mid x_1, \ldots, x_{t-1})$

Después repite el mismo cálculo para el token siguiente hasta construir una respuesta completa.

Podemos ilustrarlo con un ejemplo muy simplificado. Pensemos en las palabras «Francia», «capital», «París» y «amor». Durante el entrenamiento, el modelo ha aprendido que «Francia», «capital» y «París» aparecen con frecuencia en relaciones muy estrechas. Por eso, ante la pregunta «¿Cuál es la capital de Francia?», la continuación «París» recibe una probabilidad muy alta.

Pero si preguntamos «¿Cuál es la capital del amor?», el contexto cambia. La palabra «Francia» ya no desempeña ningún papel, mientras que la relación cultural entre «amor» y «París» adquiere importancia. Por eso, el modelo puede responder que París es conocida como la capital del amor.

Naturalmente, esto no significa que las palabras se desplacen literalmente dentro de un espacio geométrico ni que los parámetros del modelo cambien cada vez que hacemos una pregunta. Los pesos aprendidos durante el entrenamiento permanecen normalmente fijos. Lo que cambia en cada consulta es la ponderación contextual: el sistema calcula de nuevo qué elementos de la frase son relevantes y qué relaciones aprendidas deben intervenir en la respuesta.

Aquí conviene distinguir, por tanto, dos niveles. Por una parte están los parámetros del modelo, que se forman durante el entrenamiento y condensan las regularidades aprendidas. Por otra parte están las relaciones y ponderaciones que el sistema calcula durante cada uso concreto. Los primeros son relativamente estables; las segundas dependen de la pregunta y del contexto.

Pero el comportamiento de una IA no depende solamente de esos dos niveles. Después del entrenamiento inicial se realiza un proceso de ajuste en el que se evalúan y corrigen las respuestas del modelo. Además, durante el uso intervienen instrucciones internas, filtros, sistemas de seguridad y reglas establecidas por la empresa. Estos mecanismos pueden favorecer ciertos tipos de respuesta, limitar otros o impedir que determinados contenidos aparezcan.

Por eso la cuestión política no consiste simplemente en saber si el código es público o secreto. Hay que preguntar quién decide la arquitectura del modelo, quién selecciona los datos, quién fija los objetivos del entrenamiento, quién controla los parámetros resultantes y quién establece posteriormente los filtros y las instrucciones que orientan las respuestas.

Algunas partes de un modelo pueden ser públicas, mientras otras permanecen en secreto. La arquitectura matemática puede ser conocida, el código puede estar parcialmente abierto y, sin embargo, los parámetros entrenados, los datos utilizados y el proceso de ajuste pueden seguir bajo el control exclusivo de una empresa.

La cuestión decisiva es, por tanto, quién domina el proceso en cada una de sus fases. Quien controla el entrenamiento, el ajuste y las condiciones de uso no determina mecánicamente cada frase producida por la IA, pero sí establece el campo dentro del cual esas respuestas pueden generarse.

Tercero. El entrenamiento de la IA. ¿De qué se trata?

Para construir y entrenar un modelo de IA se necesita una enorme cantidad de datos. Durante el entrenamiento se introducen millones de textos que se convierten, como hemos visto, en tokens. En general, una base de datos más amplia y mejor seleccionada puede mejorar los resultados, aunque la cantidad por sí sola no garantiza la calidad. Por eso se presentan tres preguntas más: ¿Quién controla los datos? ¿Quién decide qué es calidad? ¿Y quiénes son los dueños de esos datos?

En los grandes corpus públicamente conocidos predominan los textos en inglés y las fuentes procedentes del ámbito angloamericano. En el caso de los modelos cerrados, no conocemos la composición exacta, pero existen buenas razones para suponer una desigualdad semejante. La consecuencia es que las respuestas de la IA, especialmente cuando se trata de preguntas socioculturales, tienden a reproducir prejuicios y perspectivas dominantes en ese ámbito. Una crítica bien conocida es que la IA da respuestas sexistas, racistas o clasistas cuando se le pregunta sobre temas sociales. La reacción de los grandes actores a esa crítica fue filtrar las respuestas y eliminar las inadecuadas. Pero ¿quién decide que se filtra?

Otro problema relacionado con esa base de datos es que la calidad de las respuestas de la IA. Cuando se pregunta sobre temas socioculturales de Europa o los EE. UU. las respuestas son muy buenas y elaboradas. Pero cuando se trata, por ejemplo, del Perú, la IA empieza a fabular, no porque el sistema funcione mal, sino porque la información disponible sobre el tema es más escasa, fragmentaria o contradictoria. En esas condiciones, el modelo puede producir con mayor frecuencia respuestas plausibles en su forma, pero falsas en su contenido. Puedes imaginarte a la IA en ese momento como un estudiante que no conoce bien el tema y responde a su maestro. Si quieres estudiar ese fenómeno mi con mayor profundidad, lee mi ensayo «**La IA y la conciencia -** Una consideración desde una perspectiva materialista-dialéctica».

El tercer problema de los datos de entrenamiento se encuentra en la propiedad intelectual privada. Esta limita la base de los datos de entrenamiento. Cada vez más propietarios de los datos prohíben que se usen sus textos, dibujos, su música o sus vídeos para entrenar la IA. Los dueños de los grandes modelos se encuentran en la situación de tener que firmar contratos con las grandes organizaciones que representan a esos titulares de derechos de propiedad intelectual. Esta contradicción no se resuelve ampliando los derechos privados sobre el conocimiento. Al contrario: muestra por qué la propiedad intelectual privada se ha convertido en un obstáculo para la socialización del saber y para el desarrollo ulterior de la propia IA. Precisamente ahora estoy terminando un ensayo sobre este tema analizando la contradicción entre la necesidad de una base amplia de datos de la IA y los derechos de los dueños de los datos. Cuando lo terminé, te lo voy a mandar.

Cuarto. ¿Quién tiene el derecho de usar la IA?

Los distintos modelos de la IA están hoy en las manos de diferentes empresas. Actualmente no sólo existe el acceso de pago, sino que prácticamente todas las grandes plataformas ofrecen también alguna forma de acceso «gratuito». Este acceso está limitado por el número de consultas, la capacidad del modelo o las funciones disponibles. En el caso que se necesite hay que pagar una tarifa a la empresa, que puede consistir en una cuota mensual o en un precio por una determinada cantidad de tokens.

¿Es realmente gratuito ese acceso? No. Los usuarios pagan su acceso con sus datos. Cuando usan la IA, sus datos quedan registrados y esos datos se pueden usar para el próximo entrenamiento. Es decir, los

dueños de los modelos se apropian de los datos de los usuarios para ampliar la base de datos utilizada en el entrenamiento. Pero no se utilizan todos. Se usa solamente una parte. El resto se ignora. ¿Por qué? Dicen que es porque hay que fijarse en la calidad de los datos. Cuanta más basura entra en el entrenamiento, más basura sale en las respuestas. Hasta ahí, el razonamiento es correcto. La gran pregunta es: ¿quiénes deciden qué es basura y qué tiene calidad suficiente para entrar en la base de entrenamiento?

Para ser precisos, hay que añadir que actualmente, sobre todo en Europa, existe la posibilidad de desactivar el uso de las conversaciones para el entrenamiento o la mejora de los modelos. Sin embargo, en la gran mayoría de las ofertas gratuitas, la configuración predeterminada permite ese uso, y es el usuario quien tiene que buscar en los ajustes la posibilidad de oponerse. Sólo algunos proveedores, como Anthropic, en el caso de Claude, solicitan actualmente una decisión expresa del usuario.

Estas posibilidades de oposición no fueron concedidas voluntariamente por las empresas. Son el resultado de una lucha política y jurídica, especialmente intensa en Europa, que obligó a aplicar con mayor rigor la legislación sobre protección de datos y reforzó la intervención de las autoridades públicas.

Un ejemplo es ChatGPT. El 31 de marzo de 2023, la autoridad italiana de protección de datos impuso una limitación provisional al tratamiento de datos por parte de OpenAI, señalando, entre otras cosas, la falta de información suficiente y de una base jurídica clara para la recopilación masiva de datos personales destinados al entrenamiento de sus modelos. Menos de un mes después, el 25 de abril, OpenAI introdujo un acceso mucho más sencillo a los controles que permitían excluir determinadas conversaciones del entrenamiento. No podemos demostrar que esta modificación fuera causada exclusivamente por la intervención italiana, pero resulta difícil considerar esa proximidad temporal una simple coincidencia

Conclusión

Como has visto, la respuesta a tu pregunta sobre quiénes son los dueños de la IA es bastante más compleja de lo que parece a primera vista. No existe un único propietario porque la IA no constituye una cosa única. Hay que distinguir sus diferentes niveles, la infraestructura material, los modelos y sus parámetros, los datos de entrenamiento y las condiciones de acceso, y preguntarse en cada uno de ellos quién posee los medios necesarios, quién toma las decisiones y quién ejerce el control.

La personificación del problema en figuras como Altman o Musk conduce, como tantas veces, a respuestas equivocadas. El problema decisivo no es el carácter personal de estos individuos ni siquiera sus intenciones subjetivas. Es una cuestión social y, en último término, una cuestión de clase. La producción de la inteligencia artificial se apoya en conocimientos y datos que son producto de trabajo social, mientras que la capacidad de decidir sobre su desarrollo y su utilización se concentra en manos privadas.

Por eso, la pregunta «¿quién es el dueño de la IA?» resulta demasiado estrecha. Las preguntas decisivas son otras: ¿qué clase controla su desarrollo y su utilización?, ¿qué función cumple la IA dentro del modo de producción capitalista?, ¿qué intereses determinan sus objetivos? Y, finalmente: ¿cómo podemos poner esta nueva fuerza productiva bajo el control democrático de la sociedad?

Espero que te pude ayudar con ese pequeño ensayo.

Saludos cordiales y un fuerte abrazo

Erik

© 2026 Erik Peter Roderwald

Se permite descargar, copiar y redistribuir gratuitamente este documento, siempre que se distribuya de forma completa y sin modificaciones.

Deben conservarse el nombre del autor, el título, esta nota de licencia y todos los enlaces incluidos en el documento, especialmente el enlace para apoyar Tecnosofía.

No se permite vender el documento, incluirlo en productos de pago, modificarlo, traducirlo ni publicar versiones derivadas sin autorización previa y escrita del autor.

Se permite citar fragmentos breves con indicación del autor, el título y la fuente original.

Descarga original y apoyo al proyecto:

https://www.tecnosofia.es

https://ko-fi.com/erikroderwald