



RADICALITZEM

la vida

La mayoría de libros de Virus editorial es troben sota llicències lliures i per la seva lliure descàrrega. Però els projectes autogestionaris i alternatius, com Virus editorial, necessiten un important suport econòmic. En la mesura que oferim bona part del nostre treball pel comú, creiem important crear també formes de col·laboració en la sostenibilitat del projecte. **Subscriu-t'hi!!**

La mayoría de libros de Virus editorial se encuentran bajo licencias libres y para su libre descarga. Pero los proyectos autogestionarios y alternativos, como Virus editorial, necesitan de un importante apoyo económico. En la medida en que ofrecemos buena parte de nuestro trabajo para lo común, creemos importante crear también formas de colaboración en la sostenibilidad del proyecto. **¡Subscríbete!**



Satèl·lit
sense quota



Càpsida



Replicant

60€

- 5% de llibres i llibres en línia
- 5 compres en línia anual
- Punt de lectura "gestat"

60€

- 5% de llibres i llibres en línia
- 4 llibres Virus editorial sin límite de precio
- Pack bienvenida (libro + Multilibros + Virus + Punto de lectura "gestat")
- Descuentos En grupos de lectura y otras actividades

100€

- 5% de llibres i llibres en línia
- 8 novedades Virus editorial durante un año
- Pack bienvenida (libro + Multilibros + Virus + Punto de lectura "gestat")
- Descuentos En grupos de lectura y otras actividades

<https://viruseditorial.net/hazte-virica/>



Meredith Whittaker

La plantación y la computadora

**Tres ensayos sobre IA,
trabajo y esclavitud**

Prólogo de Proyecto UNA

Traducción de Elena Fernández-Renau Rozas





LICENCIA CREATIVE COMMONS
AUTORÍA - NO DERIVADOS -
NO COMERCIAL 1.0

Esta licencia permite copiar, distribuir, exhibir e interpretar este texto, siempre y cuando se cumplan las siguientes condiciones:

- ⓘ **Autoría-atribución:** deberá respetarse la autoría del texto y de su traducción. Siempre habrá de constar la autoría del texto y/o de la traducción.
- € **No comercial:** no puede utilizarse este trabajo con fines comerciales.
- ⊖ **No derivados:** no se puede alterar, transformar, modificar o reconstruir este texto.

Los términos de esta licencia deberán constar de una manera clara para cualquier uso o distribución del texto. Estas condiciones solo podrán alterarse con el permiso expreso del autor o la autora.

Este libro tiene una licencia Creative Commons Attribution-NoDerivs-NonCommercial.

Para consultar las condiciones de esta licencia, puede visitarse: creativecommons.org/licenses/by-nd-nc/1.0/ o enviar una carta a Creative Commons, 559 Nathan Abbot Way, Stanford, California 94305, EE.UU.

© 2025 del texto, Meredith Whittaker

© 2025 de la presente edición, Virus Editorial

Título: La plantación y la computadora. Tres ensayos sobre IA, trabajo y esclavitud

Diseño de colección: Pilar Sánchez Molina y Silvio García-Aguirre

Diseño de cubierta: Virus Editorial

Traducción del inglés: Elena Fernández-Renau Rozas

Edición y maquetación: Virus Editorial

Corrección ortotipográfica y de estilo: Carlos Marín Hernández (L'Entrellat)

Corrección de galeradas: Rita Soler Colín

Primera edición: marzo de 2025

ISBN: 978-84-17870-43-0

Depósito legal: B-17935-2024



VIRUS Editorial i Distribuïdora, SCCL

C/ Junta de Comerç, 18, baixos

08001 Barcelona

Tel. / Fax: 934 413 814

editorial@viruseditorial.net

www.viruseditorial.net

Índice

- 7 PRÓLOGO. PROYECTO UNA
- 17 El elevado coste de la captura
- 43 Plantaciones, ordenadores y control industrial. Historias de origen
 - 48 Buscar la disciplina del trabajo industrial en las plantaciones
 - 53 Babbage y la gestión de las plantaciones
 - 62 Esclavitud, imperio y la cuestión del trabajo
 - 68 Las plantaciones en el ordenador
 - 77 Redefinir la libertad
- 79 Abierto (para las empresas). *Big tech*, concentración de poder y economía política de la IA abierta
 - 81 Introducción
 - 87 Definir la IA «abierta»
 - 99 Qué está (y qué no está) abierto en la IA de código abierto
 - 132 Breve historia del código abierto
 - 139 El caso de OpenAI LP: la apertura como marketing
 - 146 Los argumentos a favor y en contra de la IA «abierta» y quién los defiende
 - 166 Conclusión
- 171 BIBLIOGRAFÍA

Prólogo

Proyecto UNA

Escasos meses antes de que nos enfrentásemos a la pandemia del covid, se viralizaba el meme de «Rechazar la humanidad y volver al mono». Al tiempo que nuestras relaciones sociales y económicas se sometían a la virtualidad y nos recluíamos (quienes podíamos) en casa, nuestra ya complicada relación con la tecnología digital se volvió aún más intensa y dependiente. El confinamiento, como tanto se ha comentado, aceleró muchos de los procesos sociales que venían dándose, y la *tecnobajona* fue uno de ellos. Cada quien se refugiaba del malestar reinante como buenamente podía. Ya fuera a través del consumo compulsivo de series, haciendo *doomscrolling* o idealizando tiempos pasados en los que podíamos salir a la calle y todo era más sencillo.

Es sabido que vivimos en tiempos marcados por la nostalgia. Este sentimiento suele ser cooptado tanto por el marketing como por la propaganda política (dos conceptos cada vez más indistinguibles). Quien navegue por cualquier red social se topará fácilmente con

referencias a la moda o la música del pasado, muchas veces canalizadas a través de productos comerciales o cuentas que pretenden captar nuestra atención. En medio de esta obsesión con pasados míticos, surgen mensajes reaccionarios acerca de «rechazar al hombre moderno» y «volver a la tradición», que instrumentalizan la melancolía y la confusión hacia el presente con objetivos políticos conservadores. «Rechazar la humanidad y volver al mono» fue la vuelta de tuerca de este fenómeno, que tomó la forma en que los mensajes meméticos se propagan mejor: bajo capas de ironía. Era una manera de validar y a la vez llevar a un extremo ridículo ese sentimiento de impotencia frente a visiones de un presente ciberpunk y un mañana apocalíptico.

En las fotos usadas para dicho meme podíamos observar como nuestros parientes cercanos del reino animal nos tendían la mano, invitándonos a darle la espalda a la modernidad. Nos convidaban a penetrar en la jungla, tirar nuestros dispositivos digitales a la basura, abrazar la naturaleza y perdernos en la espesura. Aunque nunca haríamos eso, por supuesto. Lo más probable es que les diésemos un *like* y continuáramos con nuestro *scroll* hasta encontrar el siguiente estímulo que despertase nuestros sentidos. Es difícil calcular cuánta seriedad había en la propuesta primitivista del «return to monkey», porque el sarcasmo memético nunca revela del todo sus cartas.

Cuánto habrán ayudado realmente los memes a radicalizar a alguien en el rechazo tecnológico es difícil

de medir. Razones para la radicalización no nos faltan. Cuanto más manipulan nuestra atención para robarnos datos personales, más crece nuestra dependencia de la oligarquía de Silicon Valley (la gente más rica de toda la historia reciente). Como crece también su alianza con la ultraderecha mundial. Por su parte, la «nación start-up», como se conoce en el mundillo tech a Israel, experimenta con la automatización como arma de guerra contra la población civil. Es normal que, frente a estas visiones de genocidio y aumento del fascismo algorítmico, la reacción de tanta gente sea la de desesperarse y desear bajarse de ese tren que llaman progreso tecnológico. Sea de forma irónica o cínica, se expande el terror hacia las pantallas y la pérdida de privacidad. Los gurús de la autoayuda y otros cantamañanas intentan engatusarnos con historias sobre la merma de dopamina que supone usar las redes sociales. El pánico moral y el miedo a que la IA nos quite el trabajo copan las conversaciones. El tecnooptimismo parece reservado para unos pocos millonarios con los ojos puestos en Marte, para cuando el planeta Tierra no dé para más.

La pregunta es: ¿podemos realmente dar marcha atrás, deshacernos de todo lo que supone la digitalización y volver a ser felices monetes comiendo plátanos entre las lianas? Mucho nos tememos que esa visión idílica de una arcadia primigenia, de antes de que a un maldito homínido le diese por ponerse a chocar piedras, no solo no es del todo real, sino completamente

inalcanzable. Las tecnologías computacionales, el procesamiento masivo de datos o la cibervigilancia son realidades que, aun con el rechazo moral y todos los desafíos ambientales y energéticos que tenemos por delante, no parece que vayan a irse a ninguna parte. Por lo menos de momento. En este contexto, más que un rechazo individual hacia estas tecnologías, parece esencial entender de dónde salen, cómo funcionan y, sobre todo, quién las controla.

Por todo ello, el trabajo de Meredith Whittaker es esencial. Porque nos ayuda a entender la tecnología que nos rodea al margen de alarmismos o de conspiranoias. Poniendo el foco en cuestiones fundamentales para impedir que nos vendan la moto con soluciones mágicas o palabras que ocultan realidades incómodas. Whittaker analiza cómo la inteligencia artificial se ha convertido en la enésima promesa especulativa de la industria tecnológica, sin dejar de abordar polémicas como la apropiación de contenido y de arte o el uso tendencioso de la etiqueta de «código abierto» para conseguir trabajo barato bajo apariencia virtuosa. Cuando desmenuzamos el funcionamiento de las nuevas tecnologías, lo que encontramos son viejos patrones, basados en la posesión de los medios de producción y la extracción de plusvalía. Bajo una aparente búsqueda del bien común, reducen nuestras vidas y todo el planeta a recursos que pueden explotarse, beneficiando únicamente a intereses privados.

Las tecnologías de automatización y computación que se presentan con el nombre de «inteligencia artificial» (denominación que se aplica indiscriminadamente y que en muchas ocasiones no es más que un gancho comercial) esconden los conflictos sociales de los que provienen y a los que dan forma bajo su apariencia de inevitabilidad. Muchas IA necesitan de ingente cantidad de horas de trabajo humano para poder funcionar. No solo en un sentido académico, sino también en términos de extracción de materiales, construcción y mantenimiento de infraestructura, así como clasificación, calibración y revisión de contenidos. Estas son tareas que, en distintas escalas, son invisibilizadas con el objetivo de pretender que las máquinas son más autónomas de lo que realmente pueden llegar a ser. Astra Taylor definía este fenómeno como falsomatización (*fauxtomation*), que no deja de ser una forma de alienar a la clase trabajadora del objetivo y resultado de sus tareas, a la vez que se deslocalizan empleos y se abaratan costes aprovechando las desigualdades Norte-Sur de origen colonial. En Barcelona hemos presenciado la lucha de la plantilla encargada de la moderación de contenido de Meta, sometida a pésimas condiciones laborales y afectada por graves secuelas emocionales y psicológicas, causadas por extenuantes jornadas en las que tienen que tragarse y eliminar el contenido más nocivo de las redes sociales.

Los avances tecnológicos que Whittaker critica en este libro presentan un problema todavía mayor y más

complejo. Se trata del modo en que se nos presentan estas tecnologías como ineludibles, como el paso lógico hacia un futuro mejor contra el cual no tenemos nada que hacer sino rendirnos y aceptarlo pasivamente. Esta forma de ver la tecnología, parcialmente mística, parcialmente utilitarista, presupone que la manera en la que hemos desarrollado estas herramientas es la única posible. Así, se descarta cualquier otro camino alternativo que podríamos haber tomado y se equipara el desarrollo tecnológico con el mito de la historia del progreso: lineal, imparable, inexorable. Este es el único mundo que puede existir, puesto que esta es la forma «natural» en que se desarrollan los acontecimientos. Vivimos en este mundo porque las mejoras técnicas nos han conducido irremediamente hasta él. Y, si bien es mejor que los anteriores, nos ayudará a avanzar hasta futuros aún más prometedores.

Frente a esta visión, debemos entender la tecnología en su contexto. Desentrañar los orígenes e intenciones de la optimización en la explotación industrial e incluso esclava, como hace Meredith en el texto «Plantaciones, ordenadores y control industrial». Como decían las leyes de la tecnología propuestas por el historiador Melvin Kranzberg Jr., «La tecnología no es ni buena ni mala, pero tampoco neutral», más bien está determinada por la sociedad donde nace, por sus estructuras y necesidades. Se desarrolla con una intención concreta, pero también puede ser hackeada. En una sociedad distinta nacerían otras tecnologías, y

algunas de las que tenemos ahora no existirían. Al margen de la fe ciega en una automatización y un progreso que nunca son tan absolutos como nos quieren hacer creer, nosotras siempre apostamos por tomar conciencia de dónde nos movemos y qué tecnologías usamos, cómo funcionan y a qué intereses sirven. Porque el primer paso para plantar cara al control corporativo sobre nuestras vidas es enfrentar el control tecnológico que lo posibilita. Abandonar la mistificación tecnológica y recuperar nuestra capacidad de acción colectiva se revela fundamental. Leer textos tan enriquecedores como los de Meredith Whittaker puede ser un buen comienzo.

El elevado coste de la captura¹

Nos encontramos en un momento peligroso. Los sistemas informáticos privados que se comercializan como inteligencia artificial (IA) atraviesan nuestra vida pública e instituciones: concentran poder industrial, agravan la marginalización y, sin hacer mucho ruido, definen el acceso a los recursos y a la información.

Para pensar en cómo abordar esta arremetida de la IA industrial, primero debemos reconocer que los «avances» del sector, tan celebrados durante la última década, no son fruto de decisivos adelantos científicos en

¹. Publicado originalmente en *Interactions*, noviembre-diciembre de 2021. Mi sincero agradecimiento a Theodora Dryer, Lilly Irani, Nantina Vgontzas, Sarah Myers West y Michael Decker por sus concienzudas y generosas lecturas y sugerencias; también a Nicole Weber por su excelente ayuda con la investigación. La rigurosa y atenta edición de J. Khadijah Abdurahman también resultó imprescindible para dar forma a esta pieza.

las técnicas de este campo. En su mayor parte, se han debido y se deben a la notable concentración de datos y recursos informáticos en manos de unos pocos gigantes tecnológicos. La IA moderna depende intrínsecamente de los recursos corporativos y de las prácticas empresariales, y nuestra creciente subordinación a ella otorga un poder desmedido sobre nuestra vida e instituciones a un puñado de empresas tecnológicas. Esto también provoca que estas compañías ejerzan una gran influencia a la hora de orientar el desarrollo de su trayectoria y de las instituciones académicas que se interesan por investigarla. Así, las empresas tecnológicas se encuentran en una posición privilegiada para definir lo que sabemos —y lo que no— sobre la IA y sobre el negocio que hay tras ella; al mismo tiempo que sus productos con IA integrada definen nuestra vida e instituciones.

Si repasamos la historia de la influencia que tuvo el Ejército de Estados Unidos en la investigación científica durante la Guerra Fría, encontramos paralelismos con la actual intervención de la industria tecnológica en el ámbito de la IA. Esta historia también ofrece inquietantes ejemplos de cómo el dominio militar estadounidense definió la producción de conocimiento académico y castigó a quienes se mostraron disconformes.

Hoy el sector tecnológico se enfrenta a una creciente presión normativa y redobla sus esfuerzos por crear un relato amable de las empresas tecnológicas,

así como por silenciar y marginar las críticas; en gran medida, como ya hicieran el Ejército estadounidense y sus aliados en el pasado. Si observamos la industria tecnológica en su conjunto, vemos que, desde su posición dominante en la investigación de la IA y en la producción de conocimiento, considera desleales las líneas de investigación y las opiniones críticas (dentro y fuera de la academia). Esto amenaza con privar a las comunidades más vanguardistas, a la clase política y al público en general de información indispensable acerca de los costes y las consecuencias de la IA, y de la industria que hay tras ella; precisamente cuando esa conciencia es más necesaria que nunca.

Para analizar hasta dónde llega la influencia que ejercen las grandes tecnológicas sobre la IA y la investigación en torno a ella, resulta útil empezar con un breve repaso histórico hasta la actual preeminencia de esta tecnología. Si tenemos en cuenta que este campo tiene cerca de setenta años y que ha conocido varios «inviernos de la IA», ¿por qué ha crecido tanto en la última década? Y ¿de qué hablamos cuando hablamos de IA? Responder a estas preguntas pone de relieve la mutabilidad del término «IA». También nos advierte de lo imprescindible de la concentración de recursos corporativos para el actual bum del sector y de cómo el control monopolizador de estos recursos le ha otorgado a un puñado de empresas tecnológicas la autoridad para (re)definir este campo, mientras restringe la información sobre los sistemas de IA tras el secretismo corporativo.

En 2012 un equipo de investigación con sede en Toronto creó un algoritmo llamado AlexNet que ganó el concurso Large Scale Visual Recognition Challenge de ImageNet. Aquello supuso un momento decisivo en la historia reciente de la IA y trajo consigo un gran reto para el sector tecnológico. Demostró que el aprendizaje automático supervisado era sorprendentemente efectivo para el reconocimiento predictivo de patrones si se entrenaba con una capacidad computacional considerable y con enormes cantidades de datos etiquetados.² El algoritmo AlexNet se basaba en técnicas de aprendizaje automático de casi dos décadas de antigüedad. Sin embargo, no fue el algoritmo lo que supuso un gran avance, sino lo que este era capaz de hacer al combinarlo con datos y recursos informáticos a gran escala.

AlexNet definía la senda que debían seguir las grandes tecnológicas que quisieran consolidar y ampliar su poder. Los recursos de los que dependía el éxito de este algoritmo no eran otros que los que esas grandes corporaciones ya controlaban: una ingente infraestructura computacional; enormes cantidades de datos junto con los sistemas capaces de procesarlos y almacenarlos; una posición consolidada en el mercado

². Alex Krizhevsky, Ilya Sutskever y Geoffrey E. Hinton, «ImageNet classification with deep convolutional neural networks», en Peter L. Bartlett, Fernando C.N. Pereira, Christopher J.C. Burges, Léon Bottou y Kilian Q. Weinberger (eds.), *Advances in neural information processing systems* 25, Curran Associates, 2013, lc.cx/x1s-Pg.

que garantiza su recolección constante, y el capital para contratar y retener la escasa mano de obra. Yoshua Bengio, uno de los precursores de la investigación sobre IA, lo explica de forma sencilla: «El poder [computacional], los conocimientos, los datos..., todo se concentra en manos de unas pocas empresas».³

El año 2012 demostró el potencial comercial del aprendizaje automático supervisado, así como el poder del término «IA» como gancho de marketing. Las empresas tecnológicas no tardaron en (re)bautizar el aprendizaje automático y otros enfoques basados en datos para referirse a ellos como «IA», presentándolos como el resultado de la innovación científica de vanguardia. Las empresas adquirieron laboratorios y *start-ups*, y se esforzaron por mostrar la IA como una multiherramienta eficaz y precisa, apropiada para casi cualquier propósito en infinidad de ámbitos. A esto nos referimos cuando decimos que la IA está por todas partes.

La retórica y el flujo de capital procedentes de estas empresas sirvieron para redefinir el ámbito de investigación en torno a la IA, inundándolo de financiación y concentrando la atención del sector en técnicas y cuestiones que requieren un uso intensivo de datos y computación. Los laboratorios universitarios y las *start-ups* interesadas en desarrollar e investigar

³ Madhumita Murgia, «AI academics under pressure to do commercial research», *Financial Times*, 13 de marzo de 2019, [lc.cx/J_Q5br](https://www.ft.com/content/1c.cx/J_Q5br).

IA se vieron obligadas a solicitar acceso a onerosos entornos computacionales basados en la nube y explotados por grandes tecnológicas, y a pelearse por acceder a los datos; una dinámica que no ha hecho más que intensificarse desde 2012. J. Nathan Matias, catedrático en la Universidad Cornell y líder de Citizens and Technology Lab, señala el alcance de esta dependencia cuando apunta que «ciertas áreas no existirían sin los estrechos lazos que mantienen con la industria».⁴

Esto no significa que quienes investigan en estas áreas tengan que hacer concesiones. Tampoco quiere decir que no haya directrices de investigación que puedan eludir esas dependencias. Lo que evidencia, por el contrario, es que las preguntas y los incentivos que impulsan el sector no siempre dependen de las personas que llevan a cabo la investigación. También revela que las condiciones de este ámbito del saber —por ejemplo, qué preguntas se consideran dignas de respuesta o qué respuestas recibirán becas, premios o puestos de trabajo— están excesivamente supeditadas al enfoque empresarial de una IA que hace un uso intensivo de recursos y a los incentivos de la industria tecnológica que lo promueve.

Una iniciativa reciente de la Universidad Stanford ilustra esta dinámica. En agosto de 2021, la institución

⁴ J. Nathan Matias, «Why we need industry-independent research on tech & society», *Citizens and Tech Lab*, enero de 2020, ic.cx/CrsIDq.

anunció el nuevo Centro para la Investigación de Modelos Fundacionales (CRFM, por sus siglas en inglés). Acompañando al lanzamiento se presentó un informe avalado por más de cien autores en el que se definían estos modelos como un «cambio de paradigma» de tal relevancia en la IA que justificaba la creación de un nuevo y costoso centro de investigación.⁵ A lo largo del informe, los «modelos fundacionales» se presentan como ineludibles, vanguardistas y fruto del progreso científico. Pero ¿qué son los modelos fundacionales? Perdonamos a quien no sepa la respuesta. Stanford acuñó esta denominación para su informe y para el material de promoción del CRFM; así, daba un giro a la nomenclatura de lo que antes se conocía como grandes modelos de lenguaje (LLM).⁶ Los sistemas LLM —pensemos, por ejemplo, en GPT-3 o BERT— son una de las técnicas de IA que hacen un uso de datos y computación más intensivo y, por lo tanto, una de las que más captura la industria. Estos modelos también han recibido mucha atención

⁵ Rishi Bommasani, Drew A. Hudson, Ehsan Adeli *et al.*, «On the opportunities and risks of foundation models», *arXiv*, 2021, [lc.cx/XaRcUZ](https://arxiv.org/abs/2108.07258).

⁶ Los LLM (*large language model*) son modelos de lenguaje capaces de procesar de manera masiva grandes cantidades de texto, ya sea con autosupervisión (el propio modelo aprende, procesa y supervisa) o con semisupervisión humana. Estos modelos son la base de máquinas como el conocido ChatGPT u otros sistemas como Copilot, PaLM o Gemini. (*N. de la E.*)

mediática reciente y han sido objeto de críticas continuadas por sus importantes sesgos o costes medioambientales, o a causa del poder que concentran.⁷

Más allá de limitarse a considerar innovadoras las técnicas que la industria captura, el cambio de nombre que hace Stanford sirve para distanciarse a los LLM de este legado crítico. Aunque el informe reconoce que «la investigación sobre el desarrollo de modelos fundacionales se ha llevado a cabo casi exclusivamente dentro de la industria», no plantea la concentración de poder como un problema que debería hacer que nos replanteáramos la confianza en estas tecnologías, sino como una cuestión que se puede resolver aligerando las restricciones de acceso, para que a instituciones como Stanford también les corresponda un pedazo del pastel: «En última instancia, la industria toma decisiones concretas sobre cómo se utilizarán los modelos fundacionales, pero también deberíamos ampararnos en la academia, con su diversidad de disciplinas y ajena a los incentivos comerciales».⁸

⁷ Emily M. Bender, Timnit Gebru, Angelina McMillan-Major y Shmargaret Shmitchell, «On the dangers of stochastic parrots. Can language models be too big?», en *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, Association for Computing Machinery, Nueva York, 2021, pp. 610-623, [lc.cx/N16AyA](https://doi.org/10.1145/3458612.3458623).

⁸ Bommasani, Hudson, Adeli *et al.*, «On the opportunities and risks of foundation models», *op. cit.*

Los esfuerzos por ampliar el acceso a la investigación sobre IA también siguen este patrón: asumen los mecanismos de uso intensivo de datos y de computación y se centran únicamente en cómo lograr que más personas accedan a estos recursos concentrados. Al analizar una de las «soluciones» propuestas para este problema, que se plantea de una forma muy limitada, nos encontramos frente a frente con la cuestión del alcance de la captura industrial.

En marzo de 2020, la Comisión de Seguridad Nacional sobre Inteligencia Artificial de Estados Unidos (NSCAI, por sus siglas en inglés), presidida por el antiguo director ejecutivo de Google, Eric Schmidt, y dirigida por otros perfiles tecnológicos, propuso que el Gobierno estadounidense financiara una infraestructura nacional de investigación sobre IA para «democratizar» su acceso. La recomendación fue recogida en la Ley de Autorización de Defensa Nacional (NDAA) de 2021 y ordenaba la creación de «un sistema que permita que quienes investigan y estudian desde los distintos ámbitos y disciplinas accedan a los recursos de computación, así como a los conjuntos de datos de acceso público listos para procesar con inteligencia artificial, propiedad del Estado o de otras entidades».⁹ En virtud de la directiva de la NDAA, la Oficina de Políticas de Ciencia y Tecnología de la Casa Blanca y la

⁹ [lc.cx/FvTqvz](https://www.fvtqvz.com).

Fundación Nacional de Ciencias de Estados Unidos presentaron recientemente el Recurso Nacional de Investigación sobre IA (NAIRR) y designaron un grupo de trabajo para diseñar sus políticas y su aplicación.

¿Por qué un organismo gubernamental con intereses contrapuestos, formado por perfiles ejecutivos tecnológicos, recomendaría «democratizar» el acceso a infraestructuras que están en el corazón del poder que concentra? Porque en realidad esa propuesta no reduce su poder. De hecho, en caso de salir adelante, casi sin ninguna duda consolidaría y ampliaría el dominio y el alcance de las grandes empresas tecnológicas. El control que ejercen estas corporaciones sobre la infraestructura de investigación y desarrollo de la IA va más allá de ofrecer «plataformas neutrales». Estas empresas controlan las herramientas, los entornos de desarrollo, los lenguajes y el software que definen el proceso de investigación de la IA; son el hábitat en el cual se desarrolla la investigación. Incluso si fuese deseable (algo que, a tenor de sus perjuicios y deficiencias, debería cuestionarse), no existe un escenario plausible en el que la infraestructura nacional de investigación pudiera desarrollarse de forma significativa al margen del actual ecosistema de la industria tecnológica. Algo así exigiría desarrollar software, desplegar una nueva plataforma y habituar a decenas de miles de investigadores a nuevas herramientas e interfaces, además de contratar a miles de ingenieros fiables, desarrolladores de software, controladores de calidad y a todo el

personal necesario para mantener a perpetuidad un sistema tan grande y caro.

En la práctica, estas propuestas para «democratizar» el acceso a las infraestructuras de investigación en torno a la IA son, en realidad, mecanismos para subvencionar aún más a los gigantes tecnológicos mediante la concesión de licencias de uso de infraestructuras habituales, propiedad de estas mismas empresas, de tal forma que puedan seguir definiendo las condiciones de acceso a la IA y a su investigación. Todo ello al tiempo que centros como el nuevo CRFM de Stanford se disponen a afianzar todavía más ese control presentando unas técnicas de IA dependientes de la industria situada como la vanguardia de la investigación.

Desde el patrocinio de programas doctorales hasta iniciativas que colocan las oficinas de empresas tecnológicas literalmente en el centro de las universidades, o la asociación de la Fundación Nacional de Ciencias con Amazon para definir los parámetros de la «justicia» en la IA y conceder becas a quienes cumplan sus criterios positivistas,¹⁰ vemos como una miríada de estructuras acercan la academia a las tecnológicas. Entre ellas, encontramos los acuerdos de doble adscripción, frecuentes en el campo de la IA y que permiten que las empresas contraten a catedráticos del ramo sin que estos pierdan sus cargos y compromisos

¹⁰ lc.cx/Fx8CH.

académicos. Los investigadores con doble adscripción perciben un salario de las tecnológicas, trabajan codo con codo con empleados del sector y se valen de la infraestructura de investigación corporativa; todo ello mientras publican investigaciones avaladas por el emblema de una universidad. Estos acuerdos protegen a las empresas ante las acusaciones de provocar una fuga de cerebros al contratar a investigadores y sacarlos de las universidades. También permiten que las corporaciones cuenten con profesionales que respondan a las preguntas que más interesan a las tecnológicas, mientras aparentan ser disciplinas académicas que se dedican de forma orgánica e independiente a esas mismas cuestiones.

Probablemente, el hecho de que estos conflictivos acuerdos sean considerados como práctica habitual esté relacionado con la transparencia con la que los investigadores sobre IA y las universidades reconocen su dependencia de las grandes empresas y de los recursos que estas controlan. Maja Pantic, catedrática de Aprendizaje Automático que trabaja para Samsung con doble adscripción en el Imperial College London, lo explicaba en el *Financial Times*: «No podía seguir trabajando exclusivamente en la academia, no contamos con los recursos computacionales, no podía pagar a gente para que trabajara conmigo y no tenía dinero para desarrollar motores de procesamiento».¹¹ Ella y muchos

¹¹. Murgia, «AI academics under pressure to do commercial research», *op. cit.*

otros se enfrentan a tener que elegir entre asociarse a una empresa, con todos los requisitos tácitos que exige dicha dependencia, o no poder desarrollar el trabajo que genera prestigio y éxito académico.

El alcance de la influencia de la industria tecnológica en el ámbito de la investigación en IA guarda ciertos paralelismos con el dominio del Ejército de Estados Unidos sobre la investigación científica durante la Guerra Fría. Hoy en día, las grandes tecnológicas recurren a unos mecanismos muy parecidos.

En 1946, justo después de que terminara la Segunda Guerra Mundial, el general Dwight D. Eisenhower redactó un memorando titulado «Recursos científicos y técnicos como activos militares» con el que se proponía acercar a científicos e investigadores a la planificación militar estadounidense. Argumentaba que aquello permitiría que el Ejército forjara una relación de confianza con la ciencia, le otorgaría una posición de vanguardia ante los nuevos desarrollos científicos y —a través de la financiación y la proximidad con las universidades— haría que la institución militar pudiera dirigir los ámbitos de investigación para garantizar que la comunidad científica estuviera «familiarizada con nuestros problemas fundamentales».¹² Tres años después de

¹². Seymour Melman, *Pentagon capitalism. The political economy of war*, McGraw-Hill, Nueva York, 1970. [Hay trad. cast.: *El capitalismo del Pentágono. La economía política de la guerra*, trad. Alejandro Licona, Siglo XXI, Ciudad de México, 1972].

aquel memorando, en 1949, Estados Unidos halló pruebas de que la Unión Soviética estaba probando armas nucleares. Aquello ayudó a poner en marcha el plan de Eisenhower y catalizó la creación de oficinas y agencias de investigación en las distintas dependencias militares, destinadas a financiar y definir la investigación.¹³

Resulta relevante para el caso que nos ocupa señalar el poder que aquello le otorgó al Ejército estadounidense sobre el avance de la investigación científica y las instituciones que la acogían. Esta influencia no solo fue utilizada para garantizar que las cuestiones militares estadounidenses impulsaran la investigación y las inquietudes académicas. También sirvió para perseguir a denunciantes, apaciguar disensos y promover la complacencia ante exageradas afirmaciones disfrazadas de autoridad científica. Es en estas historias más oscuras donde nos enfrentamos al elevado coste de la captura —ya sea militar o industrial— y a sus peligrosas implicaciones para la libertad académica y la producción de conocimiento capaces de exigir responsabilidades a los poderosos.

Aldric Saucier fue un científico que trabajó para el Ejército de Estados Unidos durante la polémica Iniciativa de Defensa Estratégica (SDI, por sus siglas

13. Robert Krinsky, «Swords and sheepskins. Militarization of higher education in the United States and prospects of its conversion», *Bulletin of Peace Proposals*, vol. 19, n.º 1, 1988, pp. 33-51, lc.cx/4_Mq4A.

en inglés) del entonces presidente Ronald Reagan. La SDI era un gigantesco proyecto militar que reclutó a científicos de todo el país en un intento por construir un escudo anti misiles balísticos. La propuesta era fantásica y fueron muchas las voces de la comunidad investigadora que la consideraron científicamente infundada y capaz de aumentar las posibilidades de una guerra nuclear. Cuando Saucier denunció el despilfarro, el fraude y lo exagerado del programa, el entonces secretario de defensa, Dick Cheney, dio luz verde para su despido, que se sumó a una campaña para desacreditar públicamente su experiencia científica.¹⁴ Fuera de los laboratorios controlados por el Ejército, también se amenazaba a quienes disientían. La comunidad científica de las distintas universidades organizó un boicot a la investigación y financiación procedentes de la SDI. En consecuencia, el congresista de Indiana Dan Burton amenazó con anular las subvenciones a aquellas universidades en las que el profesorado rechazara las becas relacionadas con la SDI. Mientras tanto, el responsable universitario del Laboratorio Nacional Lawrence Livermore se ocupó de apartar al físico Hugh DeWitt, crítico con el papel de la institución en la promoción de la carrera armamentística. Aunque DeWitt logró conservar su

¹⁴. George Lardner Jr., «Army accuses SDI critic of falsifying credentials», *The Washington Post*, 13 de abril de 1992, lc.cx/1sRJvV.

puesto, se le negaron aumentos y ascensos, y fue excluido del trabajo de mayor interés.¹⁵ El subsecretario de Defensa para Investigación e Ingeniería, Donald Hicks —quien por aquel entonces era responsable de los contratos de investigación del Pentágono—, llegó a amenazar públicamente a la comunidad investigadora. En una entrevista, Hicks declaró que, igual que los catedráticos podían expresarse en un «país libre», asimismo tenían «la libertad de mantener la boca cerrada... [pues] yo también soy libre de no pagarles».¹⁶ *The Wall Street Journal* publicó un editorial respaldando al subsecretario.

Con la perspectiva de los años, sabemos que las críticas a la SDI por sus defectos y por la peligrosa lógica que impulsaba el programa eran correctas en su mayoría. Sin embargo, los argumentos y análisis basados en pruebas no evitaron las represalias, la coacción financiera y las campañas de desprestigio, incluso en instituciones supuestamente dedicadas a la libertad académica.

El hecho de que hoy en día un puñado de grandes tecnológicas tengan la misma influencia en relación con la investigación sobre IA debería alarmarnos. Sobre todo si tenemos en cuenta el interés cada vez más evidente de estas empresas por construir un

¹⁵. Brian Martin, «Science: contemporary censorship», en Derek Jones (coord.), *Censorship. A world encyclopedia*, vol. 4, Fitzroy Dearborn, Londres, 2001, pp. 2167-2170, lc.cx/MOy1-7.

¹⁶. Fred Hiatt, «Official seeks like minds in “Star Wars”», *The Washington Post*, 12 de mayo de 1986, lc.cx/KX4mJm.

relato amable en respuesta a la creciente presión reguladora y de la opinión pública; lo que se suma a la clara voluntad de la industria por silenciar y castigar las críticas. Sobran los ejemplos: Facebook retiró el acceso a los datos a quienes desde la Universidad de Nueva York investigaban el papel de la compañía en los disturbios del 6 de enero en el Capitolio; Google impuso a sus investigadores en plantilla que «adoptaran un tono positivo» en sus hallazgos,¹⁷ mientras ordenaba a sus «aliados académicos» externos que cuestionaran la intervención normativa,¹⁸ y Amazon dirigió ataques capciosos contra jóvenes investigadores negros que identificaron lógicas racistas en sus productos, a la vez que represaliaba a trabajadores organizados contra los daños medioambientales causados por la compañía. Por otro lado, Google despidió a la trabajadora Timnit Gebru tras pedirle que retirara su nombre y el de sus colaboradores de un artículo que criticaba los LLM, imprescindibles en la hoja de ruta de los productos de Google (los mismos que Stanford ha reinventado y revalorizado recientemente). La lista es larga y nos ofrece un buen barómetro de dónde ponen los límites

17. Paresh Dave y Jeffrey Dastin, «Google told its scientists to “strike a positive tone” in AI research – documents», *Reuters*, 23 de diciembre de 2020, [lc.cx/OGCCY](https://www.reuters.com/technology/google-told-scientists-strike-positive-tone-ai-research-documents-2020-12-23/).

18. Adam Satariano y Matina Stevis-Gridneff, «Big tech turns its lobbyists loose on Europe, alarming regulators», *The New York Times*, 14 de diciembre de 2020, [lc.cx/uf3gnh](https://www.nytimes.com/2020/12/14/technology/google-ai-research-lobbyists.html).

estas empresas: en la investigación y el disenso que amenazan el crecimiento y los ingresos.

Además de castigar la discrepancia y desprestigiar la investigación que consideran amenazante, las empresas tecnológicas trabajan para cooptar y neutralizar las críticas. En parte lo hacen financiando y promoviendo aquellas que resultan más débiles, normalmente las planteadas por instituciones y alianzas centradas en lo que denominan «ética de la IA», que exponen los problemas de poder y dominio tecnológico como cuestiones abstractas de gobernanza que no cuestionan la actual estructura de la industria tecnológica ni la proliferación de la IA. Por su parte, las empresas tecnológicas también defienden soluciones tecnocráticas, como las «recompensas por encontrar sesgos en la IA» y otros remiendos equitativos que presentan la discriminación que promueve la tecnología como meros errores de código o fallos de ingeniería.¹⁹ Estos planteamientos funcionan muy bien de cara al público. También sirven para convertir a ingenieros de élite en árbitros de la «parcialidad», al tiempo que excluyen a académicos y perfiles críticos que no son informáticos, pero cuyo análisis de las asimetrías de poder racializadas y de la economía política de la IA resultan imprescindibles para comprender y abordar los perjuicios que esta causa.

¹⁹. Jonathan Vanian, «Why Microsoft and Twitter are turning to bug bounties to fix their A.I.», *Fortune*, 10 de agosto de 2021, [lc.cx/QWPFQ](https://www.fortune.com/story/2021/08/10/microsoft-twitter-ai-bug-bounties/391279).

Todo esto sucede ante un telón de fondo en el que a las instituciones académicas (gestionadas cada vez más como empresas en busca de grandes inversores) les cuesta ignorar las ventajas financieras y reputacionales que ofrecen las alianzas y la financiación a través de empresas tecnológicas. Esta dinámica se ve agravada por la creciente precariedad profesional de la Universidad, en la que cada vez menos trabajadores cuentan con la seguridad laboral o con la solidaridad sindical necesaria para enfrentarse con cierto amparo a las políticas que podrían comprometer la libertad de cátedra. Así, las empresas tecnológicas consiguen una mayor influencia no solo sobre la investigación que financian directamente, sino también sobre las decisiones de qué proyectos se incluyen o se excluyen de la Universidad en general.

Tampoco podemos ignorar los ataques que actualmente tienen lugar contra las investigaciones que señalan el racismo y la desigualdad estructurales. Los *think tanks* de extrema derecha y los *apparátchiks* republicanos presionan a las instituciones educativas para que eliminen la pedagogía y la investigación centradas en la justicia racial, que engloban con torpeza bajo el término *teoría crítica de la raza*. Este ataque a la libertad intelectual es importante por muchos motivos. El trabajo académico y activista, consciente del capitalismo racial y del racismo estructural, ha planteado muchos de los métodos y marcos fundamentales para el trabajo crítico que aborda las implicaciones

sociales de la tecnología. Ha ayudado a situar la crítica tecnológica más allá de las limitadas nociones de parcialidad, al analizar cómo estas tecnologías reproducen patrones de marginación racial y concentran el poder en manos de quienes poseen los escasos y costosos recursos para desarrollar e implantar la IA. Esta línea de crítica ya ha logrado una poderosa influencia en el discurso público y en las prioridades reguladoras internacionales a las que las empresas tecnológicas se resisten activamente.

Entonces ¿qué camino hay que seguir? Para empezar, la academia, el activismo y la clase política que desarrollan un trabajo crítico con la tecnología, y que confían en ella, deben confrontar y poner sobre la mesa las dinámicas de captura, cooptación y compromiso de la tecnología; y deben hacerlo de frente y pronto. Esto significa incorporar críticas reflexivas sobre las circunstancias y la creación de conocimiento, y también sobre los compromisos y contrapartidas a los que se enfrentan los trabajadores del sector, quienes se encuentran bajo la influencia de las instituciones interesadas. Si tenemos en cuenta las políticas de proximidad con las universidades, que interfieren en las redes de prestigio académico a la vez que se esfuerzan por difuminar las fronteras entre trabajadores académicos e industriales, no cabe duda de que resulta incómodo. No obstante, poner sobre la mesa estas dinámicas es la única forma de abordarlas y de plantear preguntas que nos permitan concebir y exigir futuros alternativos.

La captura que hace la industria de la investigación sobre IA amenaza precisamente este tipo de intervenciones. Por lo tanto, ¿cómo apoyar ese trabajo crítico y proteger a quienes lo desarrollan tanto dentro como fuera de la academia?

En este sentido, tenemos que fijarnos en el papel fundamental de la mano de obra organizada: la del sector tecnológico, que ha logrado avances en la industria durante los últimos cinco años, y la de la academia, que se organiza en un entorno en el que el mito del talento individual sirve para perpetuar la desigualdad, incluso en un mercado laboral que se desmorona. La lucha de los y las trabajadoras académicas contra la precariedad de la profesión también es una lucha por la libertad de cátedra. La posibilidad de desarrollar una carrera estable y de lograr un control más democrático de la universidad ayudaría a reducir la influencia de donantes adinerados y de grandes patrones industriales. Por su parte, los y las trabajadoras tecnológicas organizadas tienen un papel fundamental a la hora de revisar el poder de su industria desde dentro: han de pelear por un mayor control del trabajo que llevan a cabo y esforzarse por reducir la influencia de sus empleadores en la academia y más allá de ella. En un escenario como ese, podríamos imaginarnos a investigadores y científicos organizados que exigieran otros destinos para las generosas partidas del Congreso estadounidense que hoy avalan la infraestructura nacional de investigación de la IA. Podrían servirse de su

experiencia y alzar la voz para, en su lugar, exigir dotaciones que respalden universidades, academias y programas verdaderamente públicos y accesibles, que integren a las comunidades experimentadas en el panteón de aprendices y expertos que cuestionan la tecnología.²⁰ Por supuesto, hay que tener en cuenta que la industria tecnológica no solo redacta sus políticas en los consejos que se nombran en el Congreso, como la NSCAI, sino que también dedica partidas mayores que las de grandes petroleras y tabaqueras a sus grupos de presión; por lo tanto, queda claro que cualquier intervención de este estilo exigiría una lucha seria y organizada.

En un futuro próximo no se vislumbra un Congreso de Estados Unidos que defienda ampliamente el trabajo crítico verdaderamente democrático e independiente. Sin embargo, la organización dentro de la academia y de los lugares de trabajo tecnológicos también puede ayudarnos a protegernos y a proteger el interés público a corto plazo. Nos prepara para el apoyo mutuo ante la presión institucional y nos hace entrenar el músculo del cuidado y la responsabilidad colectiva, lo que nos permite señalar las dinámicas de coacción y captura más fácilmente. No será fácil; exigirá enfrentarse a la cultura de la competición y al

²⁰. Quiero dar las gracias especialmente a Lilly Irani y a Nantina Vgontzas por su atención y generosidad intelectual, que me ha ayudado a situar mis ideas.

reclamo de territorios que atraviesan los lugares de trabajo tecnológicos y académicos. Sin embargo, es mucho lo que está en juego y quienes intentan definir lo que sabemos (y lo que no) sobre la IA y su industria están bien organizados, además de extraordinariamente bien provistos de recursos. En resumen, se trata de una batalla de poder, no es solo un combate de ideas; y tener la razón no nos protegerá si no contamos con la estrategia y la solidaridad para defender nuestra postura.

Plantaciones, ordenadores y control industrial

Historias de origen¹

El prototipo de la computación digital moderna fue codiseñado por Charles Babbage, un gran defensor de las preocupaciones de la clase capitalista industrial emergente que condenaba la organización obrera y consideraba que la democracia y el capitalismo eran incompatibles. Las historias sobre Babbage divergen mucho en sus aspectos más destacados. Sus influyentes teorías sobre cómo el «capitalismo emprendedor» podría subyugar mejor a los trabajadores están bien documentadas en estudios convencionales sobre el trabajo. Sin embargo, resulta extraño que estas mismas historias no se mencionen en muchos de los principales

¹ Publicado originalmente en *Logic(s)*, n.º 19, 17 de mayo de 2023.

relatos sobre su imprescindible contribución a la computación digital, que llevó a cabo junto con la matemática Ada Lovelace en el siglo XIX.² Al leer en paralelo ambas crónicas sobre su figura, descubrimos que las ideas prototayloristas de Babbage sobre cómo disciplinar a los trabajadores están inextricablemente relacionadas con las máquinas de cálculo que intentó construir durante toda su vida.

Desde su concepción, estas máquinas (o motores) —«los principios en los que se basan todas las máquinas de computación modernas»—³ se idearon como herramientas para automatizar y disciplinar el trabajo. Su arquitectura codifica de manera muy directa las

² Las obras siguientes representan interesantes excepciones a este patrón, pues destacan esta conexión y teorizan al respecto: Dan Mcquillan, *Resisting AI. An anti-fascist approach to artificial intelligence*, Bristol University Press, Bristol, 2022, p. 25; Nathan Rosenberg, *Exploring the black box. Technology, economics, and history*, Cambridge University Press, Cambridge, 1994, pp. 24-46; Sun-ha Hong, «Prediction as extraction of discretion», *Big Data & Society*, vol. 10, n.º 1, 2023; Matteo Pasquinelli, «On the origins of Marx's general intellect», *Radical Philosophy*, n.º 206, 2019, pp. 43-56; Dorothy K. Stein, «Lady Lovelace's notes: Technical text and cultural context», *Victorian Studies*, vol. 28, n.º 1, 1984, pp. 33-67; Wolfgang Schivelbusch, «World machines. The steam engine, the railway, and the computer», *Log*, n.º 33, 2015, pp. 54-61.

³ Philip Morrison y Emily Morrison (coords.), *Charles Babbage. On the principles and development of the calculator and other seminal writings*, Dover Publications, Nueva York, 1989, p. xi.

teorías del economista Adam Smith sobre la división del trabajo y toma prestadas funciones básicas de las tecnologías de control laboral que ya se utilizaban tiempo atrás. Estas máquinas eran en sí mismas herramientas para el control de la mano de obra, pues automatizaban y disciplinaban no ya el trabajo manual, sino el mental.⁴ Babbage no inventó las teorías que dieron forma a estos motores, pero tampoco lo hizo Smith. Se encontraban prefiguradas en el sistema de plantaciones, desarrolladas en primer lugar como tecnologías para controlar a la población esclavizada. Cuestiones vigentes hoy en día —como son la vigilancia de la mano de obra, la automatización de los lugares de trabajo y la reestructuración tecnológica del trabajo tradicional en forma de «microtrabajos por encargo»— recuerdan a la forma en que el pensamiento computacional surge históricamente como modelo de control durante la «era del abolicionismo de la esclavitud», a principios del siglo XIX. Gran Bretaña abolió oficialmente la esclavitud en las Indias Occidentales en 1833 y Babbage era muy consciente del debate que generaba el fin de la esclavitud. También estaba familiarizado con las cuestiones que inquietaban a las élites británicas, que buscaban alternativas al trabajo de la población negra esclavizada. Les

⁴ Por supuesto, esta división es problemática en sí misma y a menudo se utiliza más para connotar el estatus del trabajo que las capacidades necesarias para llevarlo a cabo.

preocupaba especialmente la cuestión de cómo controlar la mano de obra industrial blanca, que no dejaba de rebelarse contra la industrialización, pues tenían el objetivo de mantener la producción al ritmo necesario para sostener el Imperio británico. Tanto las influyentes teorías de Babbage sobre el trabajo como sus máquinas pueden considerarse intentos de dar respuesta a estas preguntas; cuestiones que, a sabiendas o no, rearticularon tecnologías de control que se habían desarrollado en tiempos de las plantaciones.

Buscar la disciplina del trabajo industrial en las plantaciones

En legislación y política laboral el concepto de libertad está ampliamente enraizado en el contrato: en él se estipula por escrito la (in)capacidad de las personas de entrar y salir, de aceptar unas condiciones o rechazarlas; constituye un instrumento amparado por la ley y, en última instancia, por la violencia estatal.⁵ Según este planteamiento, se entiende que trabajadores y empresarios acceden al contrato como iguales; cada parte ejerce su libertad de aceptar o rechazar las condiciones que en él se recogen. Sin embargo, en la práctica, las asimetrías de poder estructurales, impuestas a través

⁵ Gracias, Veena Dubal, por este punto de vista y este enfoque.

de tecnologías de control de los y las trabajadoras y diseñadas para lograr el objetivo del empresario de extraer toda la energía productiva posible, restringen esa presunta libertad. Estas tecnologías disciplinan los cuerpos, los movimientos y los hábitos mentales de los trabajadores. Su aplicación permite regímenes de trabajo que se pueden considerar «libres», pero que al mismo tiempo limitan la agencia y la expresión del empleado en su lugar de trabajo, amparándose en el contrato como prueba de que los trabajadores han elegido libremente esas condiciones.

Los métodos industriales de control de la mano de obra estaban prefigurados en el sistema de plantaciones, cuyo objetivo era exprimir al máximo la fuerza de trabajo de la población negra esclavizada, que de otro modo no tendría motivación alguna para producir valor en beneficio de quienes la mantenían cautiva. A pesar de que la relación entre el control de la mano de obra industrial y la de las plantaciones resulta ineludible, también es necesario reconocer que no se puede equiparar fácilmente el régimen de trabajo racializado de las plantaciones, basado en el terror de la esclavitud, con los procesos de trabajo industriales que recurrieron a las tecnologías desarrolladas en aquel entonces.⁶ La gestión de las plantaciones —y las relaciones

⁶ Agradezco a Stefan Ouma y a Saumya Premchander por insistir en la importancia de conservar las particularidades de estos distintos regímenes. Véase Stefan Ouma y Saumya

de dominación que las estructuraban— se basaba en la concepción de la población negra como mercancía, como algo no del todo humano. Así, las condiciones de cautiverio de las plantaciones definieron la categoría de «no libre» frente a la que se podía clasificar a los trabajadores blancos como «libres».

La similitud estructural entre la esclavitud de las plantaciones y las fábricas industriales favoreció que métodos desarrollados en tiempos esclavistas fueran transferidos al trabajo blanco «libre» y regido por un contrato. Las plantaciones británicas de las Indias Occidentales de principios del siglo XIX no eran lugares de trabajo «rudimentario» en contraste con las fábricas modernas y eficientes. Muy al contrario, las plantaciones eran «innovadoras iniciativas industriales»⁷ cuyos propietarios y capataces desarrollaron muchas formas «modernas» de gestión y disciplina laboral tiempo antes de que se implantaran con la misma escala y rigor en las fábricas.⁸ El historiador Thomas C. Holt apunta que

Premchander, «Labour, efficiency, critique. Writing the plantation into the technological present-future», *Environment and Planning A: Economy and Space*, vol. 54, n.º 2, 2022, pp. 413-421.

⁷ *Ibid.*

⁸ Caitlin Rosenthal, *Accounting for slavery. Masters and management*, Harvard University Press, Cambridge, 2018, p. 14. En sus investigaciones, Simone Browne también apunta con claridad que estas formas totalizadoras de gestión punitiva de los trabajadores tenían su origen en la vigilancia y la clasificación de la población negra, y que se aplicaban a lo largo de toda

... gran parte de la parafernalia de la nueva disciplina industrial guardaba [...] una sorprendente semejanza con la de las plantaciones esclavistas. La vigilancia centralizada, la reglamentación, la división de las tareas, el lugar de trabajo estrictamente controlado, las reglas y normativas por escrito eran los estándares a los que aspiraba cualquier dueño de una plantación.⁹

Los magnates industriales británicos se inspiraron a sabiendas en la esclavitud. Las directrices sobre gestión de las plantaciones circulaban habitualmente entre capitalistas británicos y, durante los siglos XVIII y XIX, en Gran Bretaña existía un abanico de libros de contabilidad, plantillas y manuales prediseñados ampliamente disponible, una oferta que a menudo convivía con literatura sobre gestión industrial.¹⁰

A pesar de estos claros vínculos, la historia del trabajo y de los trabajadores, así como la investigación sobre gestión y negocios, han borrado casi por completo la centralidad de la esclavitud a la hora de definir las

una vida de formas que destacaban la insuficiencia de la categoría «trabajador» para dar cuenta de todas las implicaciones de estos regímenes.

⁹. Thomas C. Holt, *The problem of freedom. Race, labor, and politics in Jamaica and Britain, 1832-1938*, Johns Hopkins University Press, Baltimore, 1992, p. 38.

¹⁰. Rosenthal, *Accounting for slavery, op. cit.*, pp. 100, 112, 119.

formas modernas de gestión y los procesos de trabajo. A su vez, negar la relación de la fábrica industrial con las plantaciones genera una visión de la historia de la industrialización desvinculada de las prácticas de dominación racial. Estas brechas estratégicas del relato sirven para clasificar ciertas formas de trabajo asalariado como «libre» mientras se recurre a tecnologías propias de las plantaciones para controlarlo. Separar las historias y las tecnologías de control de trabajadores de sus orígenes en las plantaciones limita nuestra capacidad de identificar las complejidades contemporáneas en torno a la naturaleza del trabajo y su relación con la raza. Este revisionismo provoca una distorsión habitual que presenta el desarrollo de la maquinaria industrial que tuvo lugar durante los siglos XVIII y XIX como *el germen* de la industrialización. A pesar de que la industrialización conoció un gran impulso *gracias* al desarrollo tecnológico (al que Babbage contribuyó significativamente), esa maquinaria estaba inevitablemente ligada a las lógicas de disciplina laboral de las plantaciones previas a las fábricas. Si ignoramos raza y mano de obra, estas historias sesgadas dibujan la fábrica y sus regímenes de trabajo «libre» como el resultado de un avance científico que no guarda ninguna relación con los modelos de disciplina de las plantaciones que la fábrica replicó y amplió.¹¹

¹¹. Se parece mucho al planteamiento que hacen las empresas de «microtrabajos por encargo» como Uber o Lyft. El marketing

Babbage y la gestión de las plantaciones

Las teorías de Babbage sobre el control de los trabajadores, enunciadas a principios del siglo XIX, ayudaron a definir la gestión de las fábricas industriales y se anticiparon a los métodos que más tarde Frederick Winslow Taylor englobaría bajo el término «gestión científica». Babbage documentó sus ideas sobre disciplina laboral en su famosa obra *On the economy of machinery and manufactures*, publicada un año antes de que Gran Bretaña aboliera la esclavitud en las Indias Occidentales. Su trabajo se apoyaba en el de Adam Smith y ensalzaba los métodos de división, vigilancia y racionalización del trabajo surgidos en las plantaciones.

En *Dark matters. On the surveillance of blackness*, Simone Browne demuestra que el poder sobre la población esclavizada se ejercía a través de tecnologías burocráticas que dividían a los trabajadores esclavizados, les imponían rutinas y desplazamientos y calibraban sus movimientos; todo ello con el objetivo de gestionar y controlar

de estas compañías dibuja modelos de negocio que dependen en gran medida de una tecnología innovadora que, según prometen, generará una nueva economía. Como apunta Veena Dubal, esta narrativa tecnófila esconde el hecho de que estas empresas se basan principalmente en el incumplimiento del arbitraje reglamentario y la legislación laboral para promover su negocio. Véase Veena Dubal, «A brief history of the gig», *Logic(s)*, n.º 10, 4 de mayo de 2020.

«cada instante de la vida esclavizada».¹² Su investigación nos ayuda a comprender cómo la acción conjunta de la estricta división y cuantificación de la vida y el trabajo en las plantaciones permitió que supervisores y capataces examinaran a la población esclavizada.¹³ La fragmentación de la producción, tanto en el campo como en la fábrica, aparta el poder de quienes realizan el trabajo para otorgárselo a los propietarios, que se benefician de definir y supervisar una visión coherente de los trabajadores y del proceso de trabajo en su conjunto. Esta visión no surge por sí sola. Más bien, se reproduce a través de registros, métricas y evaluaciones estandarizadas (y el término «registro» ha de entenderse como un sinónimo de «vigilancia»). La supervisión y la cuantificación del trabajo y de los trabajadores fue el primer paso, y presumiblemente el más importante, para elaborar los registros de las plantaciones. Estos registros exigían datos e información, lo que a su vez definió cómo se dividía y se organizaba la mano de obra, todo ello con el objetivo de que el trabajo y los trabajadores fueran lo más observables y cuantificables posible.

La supervisión y el control de los trabajadores también era un elemento central de las teorías de Babbage. Un capítulo de uno de sus tratados recomienda al

¹². Simone Browne, *Dark matters. On the surveillance of blackness*, Duke University Press, Durham, 2015, p. 51.

¹³. Rosenthal, *Accounting for slavery*, *op. cit.*, p. 111.

lector qué «datos» tendrá que recabar quien desee comprender y gestionar el funcionamiento de una fábrica. Existen paralelismos claros entre los datos que Babbage anhelaba controlar y los valores y métricas que obtenían propietarios, capataces y supervisores de las plantaciones, que atendían al diseño de las prácticas laborales. Ambos planteamientos aconsejaban crear registros con el número de trabajadores necesarios para completar determinada tarea y supervisar su velocidad, sus resultados individuales por día y por tarea, las herramientas e instrumentos requeridos para llevar a cabo el trabajo y las capacidades necesarias para realizar determinado esfuerzo.¹⁴ Para Babbage y para los capataces y supervisores de las plantaciones, esta vigilancia inspiró el diseño y el rediseño de la administración del trabajo, además de los distintos regímenes de violencia y disciplina calculados para aumentar los beneficios y la productividad.

Aparte de la apelación directa que hacía Babbage a la recopilación de datos (también conocida como vigilancia), su obra reconoce que dividir a los trabajadores y segmentar los procesos laborales permite un mayor control. La división del trabajo exige que, en primer lugar, empleadores y supervisores definan y especifiquen cada elemento de una determinada tarea, creando

¹⁴ Charles Babbage, *On the economy of machinery and manufactures*, Augustus M. Kelley, Nueva York, 1963 [1832], p. 65; Rosenthal, *Accounting for slavery*, *op. cit.*, p. 88.

así un sistema estandarizado y cuantificable que ante todo impone a los trabajadores el cumplimiento de las normas.¹⁵ Esto hace que los procesos laborales (y las personas que los llevan a cabo) puedan ser estudiados, cuantificados y controlados más fácilmente. Cada tarea, y las expectativas que conlleva, puede especificarse de antemano y, así, compararse con unos parámetros o una serie de expectativas previamente definidas. La división y la estipulación de estos procesos también sirve para realizar «ajustes» basados en el conjunto de datos obtenidos a través de esa vigilancia. Esto genera un ciclo de supervisión y disciplina gerenciales que se retroalimenta, una dinámica que en esencia sigue las lógicas de las plantaciones, pues entiende que los trabajadores son cuantificables, fungibles y susceptibles de ser organizados y reorganizados desde arriba. Por supuesto, cuando esta «organización del trabajo» se ponía en práctica, se veía atravesada por la categoría de la raza, lo que justificaba y naturalizaba recurrir a la violencia y al cautiverio para la población negra esclavizada, mientras que con los trabajadores blancos «libres» esa brutalidad se atenuaba.

Estas prácticas de registro y vigilancia también permitían la «gestión desde la distancia» (o el «control desde arriba»), ya se tratara de un espacio de unos cientos de metros o de miles de hectáreas. Al representar a las

¹⁵ Ursula M. Franklin, *The real world of technology*, House of Anansi Press, Toronto, 1999, p. 8.

personas y sus actividades como mercancías cuantificables, quienes deseaban ejercer el control podían hacerlo sin perder la posibilidad de negar de manera creíble las consecuencias colaterales de sus decisiones. En el caso de las plantaciones esclavistas, la gestión desde la distancia otorgaba a los propietarios cierto margen para negar la violencia y el horror, lo que ofrecía una imagen abstracta de las plantaciones como un dispositivo mecánico controlado por su operador. Nicholas Fiori cita un manual contemporáneo para la explotación de plantaciones de azúcar que hace referencia a estas ideas fantasiosas de control mecánico abstracto: «Como en una máquina bien construida, el éxito del conjunto [de la plantación] depende ante todo de la energía y de la correcta disposición de los resortes principales o partes fundamentales».¹⁶

Tanto en las fábricas como en las plantaciones, las medidas de vigilancia propiciaban mecanismos de evaluación que determinaban la vida de las personas y estructuraban el proceso de trabajo. En el contexto esclavista, se catalogaba, se valoraba y se ponía un precio de venta o de alquiler a la población negra, esclavizada o víctima de trata, en virtud de una clasificación cuyo criterio fundamental era el valor de la capacidad productiva percibida de la persona. Traficantes de personas, gerentes

¹⁶. Nicholas Fiori, «Plantation energy. From slave labor to machine discipline», *American Quarterly*, vol. 72, n.º 3, septiembre de 2020, p. 563.

de las plantaciones y otros beneficiarios del sistema esclavista clasificaban a la población esclavizada según su productividad estimada, su aspecto físico, su fertilidad y otros atributos: «el criterio fundamental de evaluación era la cantidad de trabajo que podían desempeñar».¹⁷

Babbage también propuso mecanismos para la evaluación de los trabajadores. En lo que Harry Braverman, experto en materia laboral, denominó el «principio de Babbage», este último detallaba cómo el hecho de dividir una tarea compleja en elementos más sencillos y designar cada una de esas partes más asequibles como «de baja cualificación» podía justificar pagar menos a quienes se ocuparan de cada una de dichas partes.¹⁸ La premisa fundamental de este principio era el derecho implícito del empleador a definir el valor del trabajo y del trabajador, y a hacerlo controlando el alcance y el método del proceso laboral. El valor, para Babbage, se correspondía con la clasificación de las «cualificaciones», que se evaluaban según la repercusión de la tarea desempeñada; una repercusión que definía el empleador mediante la división del trabajo.¹⁹ En un

17. Daina Ramey Berry, *The price for their pound of flesh. The value of the enslaved, from womb to grave, in the building of a nation*, Beacon Press, Boston, 2017, p. 41.

18. Babbage, *On the economy of machinery and manufactures*, op. cit., p. 101.

19. «Al dividir el trabajo que ha de llevarse a cabo en diferentes procesos, cada uno requerirá distintos grados de destreza o de

intento de definir la cualificación, Babbage recurre a una tautología que resulta reveladora: la «cualificación» de un trabajador vendrá determinada por el importe al que ascienda el salario que el trabajador pueda exigir por su trabajo; la cualificación «se medirá de acuerdo con esa suma»,²⁰ es decir, en función de la cantidad de dinero que un capitalista esté dispuesto a pagar por ella. Con esta definición circular, Babbage reconoce que la cualificación, igual que el cálculo del valor en las plantaciones, en última instancia es un índice de cuánto beneficio se presupone que es capaz de producir una persona. Es algo que refleja los mandatos y los juicios del capital, no a la persona que desempeña el trabajo, ni la naturaleza del trabajo que desempeña.

Para comprender del todo el planteamiento que hace Babbage de la «cualificación», hemos de regresar

fuerza; así, el patrón puede adquirir exactamente la cantidad precisa de ambas que necesita para cada proceso. Mientras que, si todo el trabajo lo ejecuta un solo trabajador, este ha de contar con la destreza suficiente para ocuparse de la labor más difícil y, a la vez, con la fuerza suficiente para ejecutar la tarea más laboriosa» (Babbage, *ibid.*, p. 100; Harry Braverman, *Labor and monopoly capital. The degradation of work in the twentieth century*, Monthly Review Press, Nueva York, 1998, p. 55. [Hay trad. cast.: *Trabajo y capital monopolista. La degradación del trabajo en el siglo XX*, Nuestro Tiempo, Ciudad de México, 1981, pp. 76-77]).

²⁰. Babbage, *On the economy of machinery and manufactures*, op. cit., p. 107.

a la idea de «libertad» que vertebró el régimen laboral industrial emergente que el propio Babbage ayudó a conformar; además, hemos de interpretar el concepto en el contexto de una libertad supuestamente garantizada por contrato. Natasha Iskander, en su análisis de las políticas de cualificación y trabajo migrante en Catar, demuestra que el concepto de cualificación está estrechamente vinculado al de libertad, pues, con él, se apuntalan regímenes de trabajo en los que «la falta de libertad se transforma en un reflejo o incluso en un rasgo del carácter fundamental de la persona no cualificada».²¹ Iskander ilustra cómo las categorías de cualificación —y el poder que el capital se otorga para definir lo que es y lo que no es «cualificado»— producen y naturalizan condiciones de esclavitud, dando lugar a una jerarquía de «merecimiento» que justifica situaciones de precariedad y dominación de la población «no cualificada».²² El concepto de cualificación también está racializado. En un contexto de trabajo «libre», la «cualificación» se interpreta como algo que los trabajadores (blancos) poseen y que funciona como un índice del salario que un trabajador puede deducir de los beneficios que desea el capitalista; un importe que, en teoría, aquel puede negociar o rechazar. En las

²¹. Natasha Iskander, *Does skill make us human? Migrant workers in 21st century Qatar and beyond*, Princeton University Press, Princeton, 2021, pp. 13, 46.

²². *Ibid.*, p. 14.

plantaciones, a la población negra esclavizada no se le confería capacidad para la cualificación. Se la consideraba incapaz de estar cualificada y cualquier proeza que mostrara se atribuía a diferencias biológicas que, sin embargo, la señalaban como inferior; era capacidad animal, no ingenio humano. Las categorías raciales estructuran, en primer lugar, a quién se considera capaz de poseer cualificación y, al mismo tiempo, asocian su ausencia con la condición de falta de libertad; por tanto, condición de la negritud.²³

El planteamiento que hace Babbage de la cualificación, sumado al contrato, sirve para crear una categoría de trabajo «libre» pero disciplinada por la tecnología de las plantaciones. Igual que el contrato y sus nociones de libertad, el concepto de cualificación también apuntala la idea del trabajo «libre». Según el marco de Babbage, se presupone que los trabajadores industriales (blancos) cuentan con (alguna) cualificación y, por lo tanto, quedan fuera de la categoría de no libres adscrita a la población negra esclavizada. La aplicación que hace de la división del trabajo para «descualificar» a los trabajadores puede interpretarse como un esfuerzo por *reducir*, no eliminar, la libertad (cualificación) de los trabajadores. Babbage triunfa en su propósito al reclamar que los empleadores tengan derecho a definir la «cualificación» y a estructurar los procesos

²³ *Ibid*, p. 38.

laborales para despojar a los trabajadores «libres» de la máxima «cualificación» posible. Estas prácticas de «descualificación» teorizadas por Babbage sirven para lograr y naturalizar importantes conquistas de control y degradación sin perder la presunción de libertad. Al definir el salario y las condiciones de trabajo en función de la «cualificación», se crea un marco (muy parecido a la meritocracia) en el que el salario, las condiciones de trabajo y la (falta de) agencia de un trabajador pueden presentarse como merecidas (es lo que le corresponde); al mismo tiempo, el contrato garantiza la naturaleza voluntaria del trabajo de un empleado y su agencia para desempeñarlo (es lo que ha elegido). Entre ambos supuestos media el espectro de la falta de libertad negra, que actúa como el objeto frente al que comparar el trabajo industrial blanco que, indefectiblemente, siempre se considera «libre».

Esclavitud, imperio y la cuestión del trabajo

Para Babbage, la función disciplinaria de la división y la supervisión del trabajo que plantea en *On the economy of machinery and manufactures* no resulta en absoluto desafortunada. Los objetivos que persigue su obra parten del supuesto de que los beneficios son bienes en estado puro para el capitalismo emergente y que cualquier cosa que pueda entorpecerlos es un obstáculo que hay que eliminar. En este punto de vista,

coincide con muchos de sus compañeros de las élites británicas: a principios del siglo XIX, la cuestión de cómo disciplinar la mano de obra en el contexto de la conflictividad laboral nacional, sumada a la inminente abolición de la esclavitud británica en las Indias Occidentales, era un asunto acuciante. El futuro del Imperio británico dependía de una respuesta que conservara la capacidad productiva necesaria para mantener su posición económica.

Aquí es importante recordar que defender la industria y el capitalismo no era una postura que se considerara de sentido común a principios y mediados del siglo XIX, cuando Babbage desarrollaba su obra. En 1830, precisamente cuando el autor terminaba de escribir *On the economy of machinery and manufactures*, surgió una nueva y potente ola de protestas bajo la bandera de la revuelta agraria que arrasó el país. Trabajadores británicos blancos quemaron equipamiento agrícola, destruyeron maquinaria industrial y difundieron críticas radicales contra la industrialización; se oponían a que la industria usurpara las formas de trabajo tradicionales.²⁴ Tampoco se consideraba una idea descabellada que esa rebelión pudiera triunfar: el derrocamiento de la aristocracia que había logrado la Revolución francesa estaba reciente y aquello agitaba las revueltas al tiempo que alarmaba a las élites británicas, temerosas de una repetición.

²⁴ Carl J. Griffin, *The rural war. Captain Swing and the politics of protest*, Oxford University Press, Oxford, 2012, pp. 5, 192-193.

La industria británica no fue el único escenario de rebeliones. El comercio esclavista británico se prohibió en 1807 y, desde entonces hasta 1833 —cuando la esclavitud en las Indias Occidentales se abolió oficialmente—, se vivió un periodo crítico en Gran Bretaña, acompañado de frecuentes sublevaciones en distintas plantaciones británicas. También fueron años en los cuales Babbage desarrolló activamente sus máquinas y sus teorías de control de los trabajadores. Las sucesivas «oleadas de rebeliones y resistencia» habían incrementado el coste de la esclavitud en las plantaciones, lo que ayudó a que «los cálculos del capital pasaran a estar en el lado del “trabajo libre”». El tablero cambió definitivamente en 1831.²⁵ En lo que se ha denominado la Rebelión de Navidad, la población negra esclavizada en Jamaica, ayudada por las misiones baptistas, se alzó en una huelga general armada. Su objetivo era acabar con los gerentes y capataces blancos y tomar las plantaciones. Thomas C. Holt destaca el paralelismo entre los levantamientos agrario y jamaicano: «Los pajares en llamas en los condados al sur de Inglaterra replicaban las plantaciones de azúcar arrasadas por el fuego en las parroquias occidentales de Jamaica»; una resonancia que la clase política e industrial británica no pasó por alto.²⁶ Además, si los objetivos rebeldes en

^{25.} Holt, *The problem of freedom*, *op. cit.*, p. 14.

^{26.} *Ibid.*, p. 51.

Gran Bretaña tenían su precedente en la Revolución francesa, las aspiraciones de los trabajadores esclavizados de Jamaica no iban desencaminadas: la Revolución haitiana, que había logrado la eliminación y el derrocamiento de los esclavistas franceses, era un recuerdo reciente para muchos.

Aunque las rebeliones obreras en Gran Bretaña y las de los pueblos esclavizados de las Indias Occidentales no eran idénticas, ni tampoco estaban conectadas por nada parecido a una causa común, juntas ejercían una presión colectiva sobre las élites británicas. Aquello despertó una serie de preguntas y preocupaciones concretas. Para la clase política británica, la cuestión de la esclavitud y su abolición en el extranjero, así como la creciente industrialización nacional, eran «fundamentalmente un problema de mano de obra».²⁷ La viabilidad del trabajo industrial «libre» no estaba resuelta por aquel entonces, aun cuando era un asunto de creciente importancia para los intereses británicos. Por su parte, la política de Gran Bretaña buscaba modelos efectivos de disciplina laboral con los que pacificar y controlar a la población trabajadora, con la esperanza de lograr una mano de obra capaz de alcanzar el nivel y los estándares necesarios para que el país conservara su posición.

Babbage no era ajeno a estas cuestiones y debates. Y aunque sus ideas sobre las rebeliones obreras británicas

²⁷ *Ibid.*, p. 33.

son claras y están bien documentadas (se oponía a ellas con vehemencia y señalaba a los luditas con especial oprobio), su compromiso con la esclavitud y la raza no es tan firme. Prestar atención a esto nos ayuda a relacionar su obra con las cuestiones laborales contemporáneas que entonces agitaban Gran Bretaña. Babbage menciona explícitamente la raza y la esclavitud en sus escritos y, al menos en una ocasión, afirma su postura abolicionista.²⁸ Esto no resulta incongruente: el abolicionismo era una posición mayoritaria en la época y connota, en su caso, no tanto un compromiso con la liberación como una visión optimista del capitalismo industrial que confiaba en que los regímenes de trabajo «libre» lograrían abastecer al Imperio británico.²⁹

28. En un extracto de *The Times*, que cubría un acto de campaña de su desafortunada candidatura de 1832, vemos cómo expresa un punto de vista abolicionista [de la esclavitud] de manera casual. Al ser preguntado acerca de si la población católica irlandesa debería estar obligada a pagar impuestos a la Iglesia de Inglaterra, el «Sr. Babbage respondió que aborrecía la esclavitud negra, pero ¿acaso estaba dispuesto a emancipar al jornalero inglés del tiránico control de la justicia parroquial?» (citado en Anthony Hyman, *Charles Babbage. Pioneer of the computer*, Princeton University Press, Princeton, 1985, p. 84).

29. Muchos capitalistas en Gran Bretaña también estaban a favor de la abolición debido a que querían acabar con los aranceles aplicados a las mercancías producidas por las colonias no británicas. Se consideraba que estos aranceles inflaban los precios de las mercancías que propietarios de industrias y fábricas necesitaban. Véase Archana Tewari, «The Reform Bill (1832)

Tomemos otro ejemplo a raíz de una nota que Babbage incluye en la segunda edición de su *Ninth Bridgewater treatise. A fragment* (en el que se propone demostrar matemáticamente la existencia de los milagros). Esta nota hace referencia a un fragmento en el que se debate la crueldad de la esclavitud, extraído de un testimonio que hace el capitán Hayes al almirantazgo británico en 1832 sobre la cuestión de la abolición.³⁰ El testimonio pertenece a un artículo más extenso titulado «The foreign slave-trade», publicado en la edición de diciembre de 1835 de *London Quarterly Review*.

Este artículo no solo recoge el testimonio de Hayes. En particular, incluye también un sólido razonamiento a favor de la obligación moral —y económica— de Gran Bretaña de abolir la esclavitud en todo el mundo, pues el país había acabado con el comercio esclavista en 1807. La verdadera preocupación que se escondía tras esta postura era la posibilidad de que el país se viera perjudicado en su empresa de dominio internacional, si era el único que no se beneficiaba del trabajo esclavizado; lo que provocaría «una reducción de la producción británica».³¹ Que Babbage

and the abolition of slavery (1833). A Caribbean link», *Proceedings of the Indian History Congress*, vol. 73, 2012, pp. 1140-1147.

^{30.} John Barrow, «The foreign slave-trade», *London Quarterly Review*, vol. 55, n.º 109, diciembre de 1835, p. 136.

^{31.} *Ibid.*, p. 145.

cite este artículo deja claro que entendía la relevancia de la «cuestión de la mano de obra» en el debate sobre la abolición; concretamente la preocupación por cómo disciplinar a los trabajadores, sin recurrir a la brutal esclavitud, para que siguieran produciendo para el capital y que los beneficios no se vieran amenazados. Nicholas Fiori plantea esta fricción con claridad:

En la fábrica blanca, hacía falta un mecanismo diferente para afinar la producción y disciplinar el trabajo, uno que atenuara la violencia de las plantaciones, pero que mantuviera la relación que estas habían establecido entre el humano y la máquina; una relación que consideraba que el cuerpo era (solo) una pieza más de la maquinaria de producción.³²

Las plantaciones en el ordenador

El trabajo de Babbage en el desarrollo de teorías de control laboral industrial y su ambición vital con las *máquinas de cálculo* pueden ser interpretadas conjuntamente como dos planteamientos que responden a una misma cuestión: cómo estandarizar y disciplinar el trabajo al servicio del capitalismo y del Imperio

³² Fiori, «Plantation energy...», op. cit., p. 567.

británico. En su empeño por construir estos complejos motores, modelos de la computación moderna,³³ Babbage pasó las primeras décadas del siglo XIX visitando fábricas y talleres en Reino Unido y la Europa continental. Las observaciones que realizó durante estas visitas sentaron las bases para *On the economy of machinery and manufactures*.³⁴ Babbage concebía sus máquinas como herramientas del Imperio y así justificaba sus demandas de generosa financiación gubernamental, alegando el uso que se podría hacer de ellas en el diseño de cartas náuticas para la Armada británica. Por aquel entonces, las deficientes cartas de navegación provocaban muchos naufragios, lo que ponía en peligro el comercio y al Imperio que defendía esa proeza naval.

Para comprender el vínculo entre las máquinas de Babbage y sus teorías de control laboral, podemos atender primero a sus ideas sobre la automatización. En tiempos de Babbage, el término «motor» era un sinónimo de «máquina» y hacía referencia a la avalancha

33. Joseph Henry, del Instituto Smithsonian, subraya la importancia de esta obra para el Imperio británico cuando afirma que «todo aquello que representa un elemento importante para el comercio internacional en los tiempos modernos depende de la exhaustividad y de la precisión de las cartas [de navegación]» (citado en Hyman, *Charles Babbage, op. cit.*, p. 49).

34. Babbage, *On the economy of machinery and manufactures, op. cit.*, p. 5.

de maquinaria industrial que se utilizaba para transformar las prácticas de trabajo tradicionales. Sus motores, o máquinas, conviven con otras herramientas mecánicas para la automatización del trabajo y se caracterizan por su intención de automatizar el trabajo mental (más que el manual). Babbage entendió que la automatización en general —incluidas sus máquinas— dependía de la división laboral.³⁵ Se dio cuenta de que «la división del trabajo exige la invención de herramientas y maquinaria para llevar a cabo este proceso», razonando que «cuando cada procedimiento se haya reducido al uso de una simple herramienta, la suma de todas esas herramientas, accionadas por una potencia motriz, constituirá una máquina».³⁶ La división y la racionalización del trabajo, es decir, la especificación de cada elemento de una tarea determinada para hacer que el proceso de trabajo (y la

^{35.} *Ibid.*, pp. 164, 173-174; Adam Smith, *The wealth of nations*, Neeland Media, 2004 [1776], p. 22. [Hay trad. cast.: *La riqueza de las naciones*, trad. Carlos Rodríguez Braun, Alianza, Madrid, 1994, p. 33]. Smith también se interesó por la descualificación y la atomización que provocaban estas divisiones en los trabajadores. Dividir el trabajo convertía cada parte del proceso laboral en su «oficio particular», lo que provocaba que los trabajadores tuvieran que estar cualificados en un aspecto concreto del proceso de trabajo, en lugar de en el proceso en su conjunto.

^{36.} Babbage, *On the economy of machinery and manufactures*, *op. cit.*, p. 173.

persona que lo lleva a cabo) sea observable, cuantificable y controlable «desde arriba» era, para Babbage, la condición de posibilidad para la automatización. Así, para diseñar máquinas que automatizaran el trabajo mental, Babbage primero tenía que tomar prestados (o desarrollar) sistemas de división y control del trabajo.

Babbage creó dos «versiones» fundamentales de sus motores, con múltiples iteraciones de cada una de ellas: la *máquina diferencial* y, más tarde, la *máquina analítica*, que desarrolló junto con Ada Lovelace.³⁷ Para diseñar la máquina diferencial, Babbage se basó en las plantillas de división del trabajo creadas por Gaspard de Prony, el ingeniero civil francés por excelencia. Babbage

³⁷. La máquina diferencial se ideó como una compleja calculadora automatizada que recurría al método de las diferencias divididas para computar funciones polinómicas y que servía para producir tablas matemáticas complejas muy usadas en navegación. La máquina analítica se diseñó para ofrecer la flexibilidad que la máquina diferencial no tenía, pues permitía que quienes la operaban pudieran «programar» la máquina mediante tarjetas perforadas. La computación moderna no abandonó las tarjetas perforadas hasta la década de 1970 y la máquina analítica sentó las bases arquitectónicas de la computación digital, las mismas que se conservan en la actualidad: tenía componentes independientes de memoria («molino») y de procesamiento («almacenamiento»), así como sistemas de entrada y de salida, todo ello coreografiado a través de la programación codificada en las tarjetas. Ninguno de los dos diseños se consumó durante la vida de Babbage.

entendía sus máquinas como «el análogo mecánico del sistema de De Prony».³⁸ En un esfuerzo posrevolucionario del Gobierno francés por estandarizar las mediciones en toda la nación, De Prony se comprometió a desarrollar un enorme conjunto de complejas tablas logarítmicas y trigonométricas para el catastro galo. Cuando asumió la tarea, De Prony no estaba seguro de cómo la llevaría a cabo, pues no había matemáticos suficientes para ejecutarla. Sin embargo, la historia cuenta que De Prony encontró una copia de *La riqueza de las naciones* de Adam Smith e, inspirado por el texto, aplicó la división del trabajo a su tarea de cálculo. Segmentó a los «calculadores» humanos en tres niveles jerárquicos, en los cuales el más amplio y «menos cualificado» comprendía entre sesenta y ochenta funcionarios (muchos de ellos, antiguos peluqueros de la aristocracia desempleados tras la Revolución francesa).³⁹ Un selecto grupo de los matemáticos «más cualificados» trabajaría para desarrollar las instrucciones que quienes ocuparan los niveles medio e inferior deberían cumplir. Codificando plantillas de división y control del trabajo, la máquina diferencial de Babbage se diseñó explícitamente para automatizar el trabajo de

³⁸. Hyman, *Charles Babbage*, *op. cit.*, p. 50; James Essinger, *Jacquard's web. How a hand-loom led to the birth of the information age*, Oxford University Press, Oxford, 2007, pp. 62 y 73.

³⁹. Essinger, *ibid.*, pp. 60-62; Babbage, *On the economy of machinery and manufactures*, *op. cit.*, pp. 109-111.

los matemáticos «menos cualificados»: los sesenta u ochenta expeluqueros.⁴⁰

El segundo motor de Babbage, la máquina analítica, fue diseñado junto con Ada Lovelace. Aunque supuso una prefiguración de la computación moderna, estaba en deuda con otro dispositivo de la era industrial para la automatización del trabajo, una herramienta que también se basaba en rígidas estructuras de división laboral: el telar de Jacquard. Este aparato, una evolución de otros diseños de telares mecánicos más antiguos, utilizaba tarjetas perforadas para estandarizar y aplicar patrones de tejido complejos y, al mismo tiempo, reglamentaba y disciplinaba a los trabajadores encargados de ejecutarlos (en una ocasión, los empleados de Jacquard, epónimo inventor, furiosos al sentirse desplazados por el nuevo telar, arrojaron a su creador al río).⁴¹ Estas tarjetas perforadas, inspiradas en el telar, constituían el «programa» que la máquina analítica programable permitía ejecutar.⁴² Las ventajas de la máquina analítica frente a la máquina diferencial anterior también se pueden describir en términos de su capacidad de hacer (o automatizar) más trabajo que su predecesora, a la vez que reglamentaba el esfuerzo de sus operadores humanos. En una reflexión sobre estas ventajas, Luigi Federico Menabrea, un científico

⁴⁰. Babbage, *ibid.*, p. 111.

⁴¹. Essinger, *Jacquard's web*, *op. cit.*, p. 40.

⁴². *Ibid.*, pp. 86-88.

italiano cuya descripción de la máquina analítica nos ofrece una de las reseñas contemporáneas más completas, sugiere que sería capaz de sustituir no solo el tercero, sino también el segundo nivel de los trabajadores matemáticos de De Prony.⁴³ Asimismo aclara que «una vez construida la máquina, [...] será fácil, mediante algunas anotaciones sencillas, encomendar su ejecución a un trabajador». Esto señala el papel de la máquina como facilitadora de la «gestión desde la distancia». El motor codifica instrucciones de los supervisores («el primer nivel») y los trabajadores no solo obedecen esas instrucciones, sino que se ocupan de una máquina programada para imponer obediencia. Aquí también está implícito el papel de la máquina en la «descualificación»: si el trabajo de especificación y programación lo

⁴³. Atribuyo este punto de vista a Ron Eglash, «Broken metaphor. The master-slave analogy in technical literature», *Technology and Culture*, vol. 48, n.º 2, 2007, pp. 360-369. Véase también Luigi Federico Menabrea, *Sketch of the analytical engine invented by Charles Babbage*, trad. al inglés Ada Lovelace, Quaternion, s. l., 2020 [1843], p. 3. En las detalladas notas que acompañan a su traducción del artículo de Menabrea, Lovelace asegura al lector que, a pesar de contar con excelentes capacidades (automatizar y desplazar el trabajo), la máquina permanecerá bajo «nuestro» control: «Puede hacer todo lo que sepamos pedirle que haga». Aquí vemos un resumen perfecto de la tensión entre la voluntad de automatizar y controlar a los demás, y el miedo de que esa automatización pueda potencialmente ampliarse y controlarnos a «nosotros» además de «al resto».

realizan los trabajadores «cualificados» «desde arriba», la máquina en sí misma se puede «encomendar» a un trabajador, que será capaz de supervisar fácilmente el dispositivo sin contar con conocimientos ni «cualificación» especializados.

Las dos máquinas de Babbage también se diseñaron para supervisar a los trabajadores encargados de ocuparse de ellas. Ante la insistencia de Babbage en que se imprimieran los resultados de sus cálculos a lo largo del proceso, la complejidad de su diseño aumentó significativamente, y su viabilidad se redujo.⁴⁴ Esta documentación reflexiva sirve para depurar errores. Pero, indudablemente, también es un mecanismo de supervisión de trabajadores, pues registra el progreso y los posibles pasos en falso de quienquiera que se haya ocupado de operar las máquinas. El empeño de Babbage no solo por vigilar, sino, como ya hemos visto, por automatizar la vigilancia, se hace patente de más formas. Entre sus muchas contribuciones mecánicas se encuentra un precoz reloj de registro, el «soplón», que servía para registrar la presencia o la ausencia de un trabajador e «informaba al propietario si el empleado había incumplido su turno».⁴⁵

Mientras el grueso de los textos de Babbage alaban las grandezas de la automatización y la mecanización,

⁴⁴. Hyman, *Charles Babbage*, *op. cit.*, p. 51.

⁴⁵. Babbage, *On the economy of machinery and manufactures*, *op. cit.*, p. 36.

los beneficios que menciona el autor siempre se corresponden con los intereses capitalistas; incluido el control disciplinario que permite la automatización «frente a la falta de atención, la inactividad o la inmoralidad del agente humano».⁴⁶ En un capítulo de *On the economy of machinery and manufactures* aborda brevemente las desventajas de la automatización para los trabajadores, a la vez que critica la organización obrera. Argumenta que cuando los trabajadores hacen huelga, el capitalismo se desespera y que esta desesperación ofrece mayores incentivos para sustituir el trabajo de los huelguistas. En esas circunstancias, surgen nuevas innovaciones mecánicas que sirven para sustituir a la mano de obra, lo que permite a los empleadores acabar con una huelga a través de la automatización. Para Babbage, esto demuestra que la organización y el desplazamiento de la mano de obra no beneficia a los trabajadores, a pesar de que pueda promover la innovación.⁴⁷ Así, se hace evidente que Babbage era consciente de la capacidad de observancia que imponía la automatización en sí misma: una herramienta de control que construye y amplía el carácter disciplinante de la división del trabajo.

La arquitectura de las máquinas de Babbage está ligada a sus teorías de control del trabajo y sus motores

^{46.} *Id.*

^{47.} *Ibid.*, pp. 162-163.

representan uno de los muchos mecanismos con los que intentó disciplinar a los trabajadores. En la raíz de su proyecto más ambicioso de obediencia laboral industrial, encontramos las lógicas y las tecnologías de las plantaciones.

Redefinir la libertad

Lo que conecta las teorías de Babbage sobre el control de los trabajadores y su arquitectura de la computación es un objetivo común: disciplinar el trabajo supuestamente «libre» con el objetivo de que siguiese produciendo para el Imperio británico. Los modelos de control laboral que definen la aportación de Babbage a la computación se basan en tecnologías desarrolladas en las plantaciones, que ya se aplicaban activamente para controlar a empleados rebeldes en las fábricas industriales. Los vínculos entre la computación, la tecnología de las plantaciones y el control laboral industrial plantean preguntas que van mucho más allá de quién controla los sistemas de automatización y computación hoy en día, dando por sentado que los sistemas controlados por quienes tienen buenas intenciones producirán resultados positivos. Este escenario exige que emprendamos indagaciones más profundas, que examinemos las tecnologías de control que estructuran las lógicas centrales de la computación y que prestemos atención a las condiciones de posibilidad con

las que se diseñan las tecnologías computacionales: el paisaje imaginario al que ajustamos nuestras relaciones y hábitos. Como se aprecia con las máquinas de Babbage, este paisaje asume la presencia de tecnologías de división, vigilancia y control del trabajo «desde arriba» propias de las plantaciones: los motores de Babbage solo «funcionan» en estos contextos. El espectro de las plantaciones que sobrevuela la computación y los regímenes de trabajo industrial también nos habla de la necesidad de revisar las condiciones del trabajo industrial «libre» y de reconocer el controvertido proceso por el que se creó y se consolidó esta categoría concreta de «libertad». Para ello, debemos confrontar directamente la disimulada presencia de la falta de libertad negra que acecha al trabajo «libre» y volver a tejer los lazos que se han roto estratégicamente entre la raza, el trabajo y las tecnologías computacionales. Tengo la esperanza de que este análisis nos ayude a identificar puntos de apoyo para el cambio y que dejemos de centrarnos en hacer retoques en los márgenes de las tecnologías de control, para ocuparnos de articular futuros que exijan el derecho a redefinir las categorías de libertad.

Abierto (para las empresas)

*Big tech, concentración de poder
y economía política de la IA abierta¹*

Introducción

Este artículo analiza la IA «abierta» en el contexto del reciente interés por los sistemas de IA abierta y de código abierto. Observamos que los términos «abierto» y «código abierto» se utilizan de forma confusa y diversa, a menudo más como descriptores aspiracionales o de marketing que técnicos, y muchas veces mezclando conceptos del software de código abierto y de la ciencia abierta. Esto complica aún más un panorama ya de por sí complejo en el que actualmente no existe una definición de consenso para «abierto» en el contexto de la IA, de modo que el término se

aplica a una gama de servicios profundamente divergentes que apenas hacen referencia a un descriptor estable.¹

Por norma general, «abierto» se refiere a sistemas que ofrecen transparencia, son reutilizables y ampliables; es decir, que se pueden analizar, volver a usar y desarrollar. Sin embargo, lo que esto significa en la práctica varía enormemente.² Los sistemas de IA están formados por múltiples componentes interconectados. Como apunta Irene Solaiman, la IA catalogada como «abierto» incluye un amplio abanico de casos. La realidad es que este término se utiliza para describir sistemas que ofrecen una transparencia y una capacidad de reutilización mínimas —que solo proporcionan accesibilidad y posibilidad de intervención sobre uno de sus muchos componentes—, así como para referirse a otros que asumen los mayores estándares de transparencia, capacidad de reutilización y de desarrollo.³

1. Recientemente, la Open Source Initiative ha propuesto una definición que aún no ha sido ampliamente adoptada: Stefano Maffulli, «Towards a definition of “open artificial intelligence”. First meeting recap», *Open Source Initiative*, 13 de julio de 2023, [lc.cx/ob65gh](https://l.c.x/ob65gh).

2. Andreas Liesenfeld y Mark Dingemans, «Rethinking open source generative AI: open-washing and the EU AI Act», en *FACCT'24: Proceedings of the 2024 ACM Conference on Fairness, accountability and transparency*, Association for Computing Machinery, Nueva York, 2024, [lc.cx/zp04JO](https://l.c.x/zp04JO).

3. Aleksandrina Goeva, Sara Stoudt y Ana Trisovic, «Toward

Los sistemas de IA no funcionan como el software tradicional; exigen procesos de desarrollo concretos y dependen de recursos especializados y onerosos que actualmente se acumulan en manos de unos pocos gigantes tecnológicos. Aun así, muchas de las promesas originales que hizo el movimiento por el código abierto, formuladas en relación con ese tipo de software, ahora se proyectan en la IA «abierta»: la idea de que el código abierto democratizaría el desarrollo de software; la premisa de que, si todas las miradas se ponían en el código abierto, se podría garantizar su uso para lo que se suponía que debía servir y no tendría agujeros de seguridad; o la hipótesis de que el código abierto equilibra el terreno de juego, pues permite que triunfe lo más innovador. El software de código abierto sirvió, en distinta medida, para muchas de estas cosas. Pero la IA «abierta» es otra historia.

Algunas IA «abiertas» sí ensanchan el reducido círculo de quién tiene acceso a los grandes sistemas de IA ya desarrollados y bien documentados: extienden su alcance más allá del puñado de grandes empresas y acercan estos proyectos a ecosistemas más amplios de desarrolladores y académicos. Son entornos que tienen la posibilidad de analizar, reutilizar y desarrollar esos mismos sistemas. En algunos casos hasta permiten auditorías y supervisiones provechosas que no se podrían

reproducible and extensible research. From values to action», *HDSR*, 16 de diciembre de 2020, lc.cx/eYzoST.

llevar a cabo de otro modo. Sin embargo, ni siquiera las opciones con parámetros más maximalistas de IA «abierta» democratizan ni amplían el acceso a los recursos necesarios para desarrollarla desde cero; un proceso en el que se toman decisiones «editoriales» muy relevantes. Se trata de una diferencia fundamental entre la «apertura» en IA y en otros tipos de software. Esta condición de «abierto» tampoco reduce el coste de implementar grandes sistemas de IA a escala, lo que limita la definición práctica de la idea de «reutilización». Por último, incluso la máxima transparencia en el contexto de los sistemas de IA no equivale al nivel de escrutinio que posibilitan un código abierto y una documentación clara en el software tradicional. En el caso de la IA, ni el código ni la documentación pueden dar cuenta de forma precisa de cómo se comportará un modelo en un contexto determinado y tampoco permiten predecir las propiedades emergentes del sistema.⁴

⁴ Zachary Lipton, «The mythos of model interpretability», *arXiv*, 6 de marzo de 2017, [lc.cx/weeOoo](https://arxiv.org/abs/1703.00974); Forough Poursabzi-Sangdeh, Daniel G. Goldstein, Jake M. Hofman, Jennifer Wortman Vaughan, Hanna Wallach, «Manipulating and measuring model interpretability», en *CHI'21: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery, Nueva York, 2021, [lc.cx/mR4VaM](https://arxiv.org/abs/2104.02522); Cynthia Rudin, «Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead», *Nature Machine Intelligence*, vol. 1, 2019, pp. 206-215.

Entonces, ¿qué es exactamente lo que está «abierto» en la IA «abierta» y qué es lo que la IA «abierta» nos permite? Para responder mejor a estas preguntas, este capítulo comienza con un análisis de los múltiples recursos necesarios para crear e implementar sistemas de IA y de los componentes que conforman estos sistemas. Trabajamos con la mirada puesta en cuáles de ellos pueden, o no, abrirse para el escrutinio, la reutilización y la ampliación. ¿Qué significa «abierto» en la práctica y cuáles son sus límites en el contexto de la IA?

Existen algunos sistemas de IA con amplios parámetros de apertura y que cuentan con criterios de transparencia, de capacidad de reutilización y de desarrollo profundos e intencionados, pero observamos que los recursos necesarios para desarrollar IA desde cero y para implementar grandes sistemas a escala permanecen «cerrados», a la exclusiva disposición de quienes cuentan con unos recursos considerables (casi siempre corporativos). A partir de aquí, abrimos el foco y repasamos la historia del código abierto, su escisión del software libre a mediados de la década de 1990 y el controvertido proceso por el cual el código abierto se ha incorporado a las grandes tecnológicas, y estudiamos además cómo estas lo han instrumentalizado. Con un ejemplo actual del uso vago e impreciso que las empresas tecnológicas hacen del término, analizamos el vocablo «abierto» (*open*) en el contexto de la empresa OpenAI. Repasamos cómo esta se ha transformado desde su constitución sin ánimo de lucro y con vocación

humanista hasta su carácter lucrativo y su vinculación con Microsoft, y revisamos también la evolución de su postura sobre la IA «abierta». Por último, estudiamos el actual discurso en torno a la IA «abierta», con la vista puesta en cómo el término y las interpretaciones (erróneas) acerca de lo que permite la «apertura» se utilizan para definir la percepción pública y política sobre la IA, sus capacidades y el poder de la industria que hay tras ella. En particular, examinamos algunos argumentos a favor y en contra de la IA «abierta» y de código abierto, quién los plantea y cómo son utilizados en el debate sobre su regulación.

En conjunto, descubrimos que la IA «abierta» puede, en sus instancias más maximalistas, ofrecer un nivel de transparencia, capacidad de reutilización y desarrollo suficiente para que terceros implementen y desarrollen potentes modelos de IA listos para su uso, modelos que, además, permiten ciertas formas de auditoría y supervisión. Sin embargo, incluso los sistemas más abiertos de IA «abierta» no garantizan por sí mismos el acceso democrático a la IA ni la competencia significativa del sector; igual que la apertura por sí sola tampoco resuelve el problema de la supervisión y del escrutinio. A pesar de que encontramos una pujante comunidad de honestos colaboradores que desarrollan y participan en iniciativas de IA «abierta» con el único propósito de ampliar el acceso y los puntos de vista, también descubrimos que muchas empresas poderosas aprovechan el marketing en torno a la

apertura y la inversión en sistemas de IA (relativamente) abiertos para afianzar sus posiciones ante el creciente interés por la regulación. Asimismo, vemos como algunas empresas han optado por abrazar la IA «abierta» como mecanismo para consolidar su posición dominante. Estas organizaciones se sirven de la retórica de la IA «abierta» para ampliar su poder de mercado e invierten en este tipo de iniciativas para poder definir los estándares de desarrollo; todo ello sin dejar de beneficiarse del trabajo gratuito de los colaboradores de código abierto.

Definir la IA «abierta»

Empecemos por las definiciones. La apertura en IA es un concepto engorroso. En parte porque la definición de la propia IA y de lo que significa «abierto» en el contexto de estos complejos sistemas emergentes son cuestiones controvertidas y que no están claramente definidas.

Podemos comenzar con la inestabilidad del término «IA», que a lo largo de sus setenta años de historia ha sido aplicado a toda una suerte de planteamientos, no tanto como vocablo técnico, sino de marketing y, como un significante aspiracional.⁵ Hoy por hoy, el término se suele utilizar más para describir grandes sistemas

⁵ «What is AI? Part 2, with Lucy Suchman», *AI Now Institute*, entrevista, [lc.cx/lgdKB6](https://www.lc.cx/lgdKB6).

de aprendizaje automático que hacen un uso intensivo de recursos, entre los cuales la inteligencia artificial generativa se lleva la palma como la variante más de moda. Sin embargo, dada la flexibilidad del concepto de IA también se usa para comercializar una avalancha de productos informáticos que, por lo general y dado su carácter reservado, no pueden ser analizados ni validados.

Puesto que los sistemas de IA generativa son los que más claramente alteran las fronteras y cuestionan la ideología tradicional del código abierto y de la ciencia abierta, en este artículo nos centramos en ellos fundamentalmente. Cuando utilizamos el término «IA», lo hacemos para referirnos a estos grandes sistemas.

La falta de una definición de consenso de la IA y la consiguiente dificultad de precisar qué se considera «parte de» un sistema de IA multipartito hacen que dilucidar qué es la IA «abierta» resulte muy laborioso. De hecho, actualmente no contamos con una definición de consenso para la IA «abierta» o de «código abierto», a pesar de que el interés sobre el tema alcanza altísimas cotas. Sí existe un proceso emergente, liderado por la Open Source Initiative (OSI, en castellano Iniciativa para el Código Abierto), que empieza a debatir cómo podría ser esa definición; para ello, aborda cuestiones polémicas, como determinar si los datos de entrenamiento potencialmente sensibles deberían abrirse o no, si es seguro ofrecer modelos potentes para que se reutilicen o si restringir el uso de modelos «abiertos» por cuestiones de

«seguridad» o de «ética» se considera aceptable o no bajo la bandera del «código abierto».⁶ Lo que la OSI expone con claridad, como hacen otros observadores y estos autores, es que la definición tradicional de «código abierto», redactada para referirse al software en un momento muy diferente de la historia de la industria tecnológica, no abarca la IA.

Aun así, la ideología y los supuestos sobre el código abierto forjados hace décadas en el contexto del software tradicional se están trasladando a la IA «abierta», lo cual da pie a muchos errores de concepto y a afirmaciones hiperbólicas sobre lo que esta permite y lo que no.

Este es el escenario en el que escribimos y el que define los términos que hemos elegido.

Salvo que nos refiramos a una licencia concreta o citemos declaraciones literales, evitamos el término «código abierto». En su lugar, usamos el término «IA “abierta”» para connotar los sistemas de IA en los que uno o más componentes se ofrecen con criterios de transparencia o capacidad de reutilización, o que permiten la ampliación o el desarrollo por parte de terceros. En la práctica, esto combina definiciones de código abierto y ciencia abierta tradicionales y, de esta manera, queda reflejada la confusa utilización de la apertura en la IA. Con esta decisión, no pretendemos aplicar el espíritu de la definición tradicional de código abierto

⁶ Maffulli, «Towards a definition of “open artificial intelligence”...», *op. cit.*

a iniciativas de IA, sino que incluimos explícitamente reivindicaciones de «apertura» que resultan polémicas, poco profundas y casi deshonestas.⁷ Por ejemplo, las declaraciones de Meta sobre su recién estrenado modelo LLaMA-2, que la empresa catalogó como de código abierto, a pesar de que la OSI indicara que no lo es.⁸ Al optar por incluir en nuestra terminología denominaciones engañosas y poco profundas del concepto «abierto», en modo alguno respaldamos el *openwashing*,⁹ ni pretendemos tomar las declaraciones de Meta (ni de otros) al pie de la letra. Al contrario, dado que nos interesa el papel que la retórica de la apertura representa para la interpretación que hacen la opinión pública y la clase política sobre la IA y sobre los debates normativos

7. Liesenfeld y Dingemans, «Rethinking open source generative AI...», *op. cit.*

8. Stefano Maffulli, «Meta's LLaMa 2 license is not open source», *Open Source Initiative*, 20 de julio de 2023, [lc.cx/37hubF](https://oc.opensourceinitiative.org/llama-2-is-not-open-source/); Michael Nolan, «Llama and ChatGPT are not open-source», *IEEE Spectrum*, 27 de julio de 2023, [lc.cx/EnCYMc](https://spectrum.ieee.org/llama-chatgpt-are-not-open-source/).

9. Este término utiliza la estructura *open* (abierto) + *washing* (lavar) para referirse a una estrategia propia del marketing con la que se promueve una imagen que no se corresponde con la realidad. Es decir, las empresas o iniciativas «lavan» su imagen para presentarse como «abiertas» y lograr una mejor acogida entre el público, a pesar de que sus prácticas no son realmente abiertas. Es la misma estructura que encontramos en otros términos populares como *greenwashing* o *pinkwashing*, los cuales hacen referencia a este mismo mecanismo en relación con el ecologismo o los derechos LGTBQ+, respectivamente. (*N. de la T.*)

que la rodean, nos damos cuenta de que tanto las reivindicaciones más honestas como las más engañosas ayudan a conformar esta idea. Asimismo, reconocemos que discernir los matices de una reivindicación «legítima» sobre IA abierta, aunque en algunos aspectos pueda resultar intuitivo, en última instancia es difícil dada la falta de definiciones asentadas.

En general, en el contexto de la IA los términos «abierto» y «código abierto» se utilizan de muy distintas formas para referirse a un abanico de capacidades que, a grandes rasgos, pueden englobarse como aquellas que ofrecen cualidades de *transparencia* (acceder y analizar el código fuente, la documentación y los datos), *capacidad de reutilización* (las licencias necesarias para que terceros reutilicen el código fuente o los datos) y *capacidad de desarrollo* (ampliar modelos existentes y listos para ser usados, modificándolos para diversos propósitos concretos). Aunque los términos «abierto» y «código abierto» se emplean de distintas formas para referirse a estas cualidades, en la práctica existen gradientes de apertura que ofrecen niveles de acceso sumamente diversos.¹⁰ Algunos sistemas que se describen como «abiertos» ofrecen poco más que una API (*application programming interface*)¹¹ o la posibilidad

^{10.} Irene Solaiman, «The gradient of generative AI release. Methods and considerations», *arXiv*, 5 de febrero de 2023, [lc.cx/XrHixD](https://arxiv.org/abs/2302.01411).

^{11.} Herramientas y protocolos que permiten a diversas aplicaciones

de descargar un modelo hospedado.¹² En estos casos, hay muchas voces que cuestionan si se debería utilizar el término «abierto» o si nos encontramos ante casos de *openwashing* de sistemas que deberían considerarse «cerrados». Otras propuestas con planteamientos más maximalistas de IA «abierta» van más allá y permiten acceder al código fuente, a los datos de entrenamiento subyacentes y a toda la documentación, y, además, ofrecen sus sistemas de IA en virtud de licencias que permiten una amplia reutilización, con unas condiciones que se corresponden con la definición de «código abierto» que hace la OSI.¹³

Estas diferencias tan precisas son importantes, en parte porque hoy por hoy la ideología de la «apertura» y del «código abierto» se proyecta en sistemas de IA «abiertos» incluso aunque no se corresponda con ellos; una ideología fraguada hace décadas en el contexto

comunicarse para hacer efectivo su uso en diferentes softwares y dispositivos, como, por ejemplo, para la descarga y ejecución de aplicaciones en teléfonos y ordenadores. (N. de la E.)

12. Liesenfeld y Dingemans, «Rethinking open source generative AI...», *op. cit.*

13. La Open Source Initiative dispone de una definición de «código abierto» que exige que se cumplan los siguientes criterios: libre redistribución, código fuente, disposiciones que permitan modificaciones y obra derivada, varias disposiciones de no discriminación y no restricción, además de neutralidad tecnológica, entre otros. Véase: opensource.org/osd.

del software de código abierto, cuando la actual industria tecnológica era un sector emergente, y no extraordinariamente poderoso. En un reflejo de lo que fue esa ideología ancestral (más que de lo que es la IA «abierta» o su industria), vemos como aumentan las referencias a la apertura en la IA como algo que puede equilibrar el terreno de juego y resolver la concentración de poder en el sector.¹⁴ De igual manera, en los debates sobre normativa y gobernanza, se suele apelar a la existencia de sistemas de IA «abiertos» o de «código abierto» para sugerir que los beneficios de la IA y los recursos necesarios para crear esos sistemas están (o podrían ponerse) a disposición de quienes no pertenecen a las grandes tecnológicas.¹⁵ Hace poco, Bruce Schneier y Jim Waldo hacían esta reivindicación de forma explícita:

Nos hemos adentrado en la era de la democratización de los LLM. Se ha demostrado que los modelos más pequeños pueden resultar muy efectivos, permiten una experimentación sencilla, diversifican

¹⁴. Mark Surman, «Introducing Mozilla.ai. Investing in trustworthy AI», *The Mozilla Blog*, 22 de marzo de 2023, [lc.cx/dVG-Oz](#); Jane Zhang y Saritha Rai, «AI presents nearly level playing field for startups and big tech», *Bloomberg*, 21 de marzo de 2023, [lc.cx/jnXNYg](#).

¹⁵. Nick Clegg, «Openness on AI is the way forward for tech», *Financial Times*, 11 de julio de 2023, [lc.cx/S8ILdt](#).

*el control y ofrecen incentivos que no tienen carácter lucrativo; las iniciativas de código abierto nos conducen hacia un escenario más dinámico e inclusivo de la IA.*¹⁶

Sin embargo, la realidad no es tan reconfortante. Aunque una IA «abierta» en su máximo exponente es condición necesaria para cualquier terreno de juego hipotéticamente democratizado y equilibrado, no es condición suficiente.¹⁷ Cuando analizamos los recursos que hacen falta para desarrollar el proceso íntegro de crearla e implementarla, y vemos en manos de quién están, se desvanecen las ideas de acceso democrático y de un mercado de IA competitivo. Descubrimos que, a pesar de que existan unos pocos sistemas de IA profundamente transparentes y con capacidad de reutilización y de desarrollo, tanto estos como todas las demás IA «abiertas» se inscriben en un escenario de profunda concentración de empresas tecnológicas. Con contadas salvedades (las excepciones que confirman la regla), apenas unos pocos gigantes tecnológicos pueden afrontar el proceso íntegro de crear e implementar grandes sistemas de IA a escala; algo que queda muy lejos de la infraestructura descentralizada

¹⁶. Bruce Schneier y Jim Waldo, «Big tech isn't prepared for A.I.'s next chapter», *Slate*, 30 de mayo de 2023, [lc.cx/-0l2h9](https://www.slate.com/technology/article/big-tech-ai-prepared-2023-05-30/).

¹⁷. Algunas de las personas que han contribuido a este capítulo cuestionan la posibilidad de que la IA se pueda democratizar.

y modificable que en su día promovió el sueño del movimiento por el software de código abierto/libre.¹⁸ Dada la inmensa importancia de la escala en la actual trayectoria de la inteligencia artificial, esto significa que las iniciativas «abiertas» no pueden, por sí mismas, «democratizar» de manera significativa la IA; y supone además que tampoco plantean ningún desafío importante para la concentración de poder de la industria tecnológica.

Comprender la multiplicidad de definiciones que pueden hacerse de la IA «abierta» o de «código abierto», y entender lo que esos sistemas permiten o no, también es importante porque considerar un sistema «abierto» o de «código abierto» podría afectar *a posteriori* a su regulación. En los debates en torno a la Ley de Inteligencia Artificial de la Unión Europea y otras medidas reguladoras potenciales, existen grupos de presión activos que defienden grandes exenciones para la IA de código abierto.¹⁹ En cierta medida, estas demandas se basan en alegaciones de que la IA «abierta»

^{18.} Gracias a Chris Kelty por este punto de vista.

^{19.} Véanse Jude Karabus, «GitHub CEO says EU AI Act shouldn't apply to open source devs», *The Register*, 7 de febrero de 2023, [lc.cx/AvJvvB](https://www.theregister.com/2023/02/07/github_eu_ai_act/); BSA Foundation, «BSA leads joint industry statement on the EU Artificial Intelligence Act and high-risk obligations for general purpose AI», *Business Software Alliance*, 27 de septiembre de 2022, [lc.cx/bH9vC7](https://www.bsa.org.uk/news/bsa-leads-joint-industry-statement-on-the-eu-artificial-intelligence-act-and-high-risk-obligations-for-general-purpose-ai); y la propuesta de Google sobre la Ley de IA de la UE, [lc.cx/t1GIfe](https://www.google.com/policy/terms/ai-act/).

permite un acceso más democrático, innovador y competitivo a los grandes y potentes sistemas de IA que presumiblemente la regulación impedirá. En concreto, estas alegaciones suelen ser planteadas, con cierta agresividad, por las grandes tecnológicas.

Nos decidimos a escribir este texto al darnos cuenta de que las voces más estridentes que abogan por excluir el código abierto del escrutinio regulador proceden de muchas de las mismas empresas que, según la lógica de estos argumentos, se verían amenazadas por esa IA «abierta». Planteamos que las grandes reivindicaciones de que la IA «abierta» permite un acceso más democrático a aquellos grandes y potentes sistemas son muy limitadas. De hecho, la evidencia sugiere que, sin una revisión en profundidad del poder de las grandes tecnológicas, la IA «abierta» y un entorno regulador definido por las ideas de esas empresas, defensoras de la exclusión del código abierto, lejos de dispersar la concentración de poder, generaría una mayor aglomeración en las grandes tecnológicas.

No nos posicionamos aquí respecto de si la apertura de la IA es «buena» o «mala» en términos generales. Reconocemos que una IA «abierta» que sea transparente y permita la reutilización y el desarrollo de manera auténtica puede ofrecer puntos de vista valiosos y propiciar distintas prácticas de rendición de cuentas, a la vez que favorecer que quienes cuentan con los recursos puedan implementar y desarrollar componentes de IA previamente diseñados.

No obstante, lo que aquí nos interesa no es analizar pros y contras de la IA «abierta» en sí misma, sino comprender mejor todo lo que se engloba bajo esa definición, y estudiar cómo se utilizan los actuales argumentos (favorables o contrarios a su apertura) en el actual discurso de la opinión pública y de las autoridades reguladoras. Quién plantea qué argumentos y cómo se corresponde cada uno de ellos con las posibilidades que ofrecen las distintas definiciones de IA abierta. Esto nos condujo a las siguientes conclusiones:

1. En un extremo de un gradiente amplio, existen una serie de iniciativas de IA abierta con parámetros maximalistas; se trata de proyectos no corporativos que hacen todo lo posible por ofrecer transparencia, capacidad de reutilización y de desarrollo. Sin embargo, desplegar estos modelos exige, en cualquier caso, el acceso a una onerosa infraestructura computacional que se suele alquilar a las grandes empresas tecnológicas.²⁰ Además, estas propuestas requieren una importante financiación para

²⁰. Por lo general, esto sucede incluso en las infraestructuras de financiación pública que suelen tener acuerdos contractuales con empresas de hiperescala. Véase AI Now Institute and Data & Society Research Institute, «Democratize AI? How the proposed national AI research resource falls short», *AINow*, 5 de octubre de 2021, [lc.cx/et-aAg](https://l.c.cx/et-aAg). La única

promover el desarrollo y el mantenimiento de muchos componentes complejos, como conjuntos de datos abiertos minuciosamente elaborados o una amplia documentación.

2. Existe un dilatado historial y una estrategia clara en la captura e instrumentalización industrial de los proyectos de código abierto; las grandes empresas de IA reconocen el valor de este tipo de iniciativas, pues disfrutan de las ventajas de ser propietarias del ecosistema, recogen los frutos del trabajo comunitario y definen las condiciones de participación.
3. La ideología del movimiento de software de código abierto suele vincularse al concepto de IA «abierto» de formas que no dan cuenta de las importantes diferencias entre los grandes sistemas de IA y el software tradicional. Esto ofrece una narrativa que asume que la IA «abierto» puede por sí misma equilibrar el terreno de juego, promover la innovación y democratizar el desarrollo y el uso. Aunque es probable que

excepción que confirma la regla es la supercomputadora Jean Zay, diseñada para el Gobierno francés por Hewlett Packard Enterprise y utilizada por BigScience para entrenar al modelo BLOOM. Véase «Jean Zay, France's most powerful supercomputer for research», [lc.cx/fcl9m_](https://lsc.cx/fcl9m_).

parte de estas proyecciones tengan su origen en la confusa definición popular de la IA, y muchos colaboradores de proyectos de IA abierta no comulgan con estas ideas, los agentes con más peso en la industria recurren estratégicamente a esta retórica para afianzar su poder y a menudo lo hacen bajo la bandera de la democratización del acceso.

Qué está (y qué no está) abierto en la IA de código abierto

En la práctica, el significado de apertura varía considerablemente. Puede referirse a todas estas cosas: que se permita el acceso público a los conjuntos de datos de entrenamiento y evaluación utilizados para definir modelos de IA; que se publique el código que define la arquitectura de un modelo o la configuración de sus hiperparámetros, en virtud de una u otra licencia de código abierto; que se pongan a disposición los parámetros aprendidos de un modelo (los pesos del modelo);²¹

²¹. En el original, *model weights*; este término se refiere a los parámetros numéricos con los cuales la tecnología mide los datos con los que trabaja. Durante el proceso, la evolución y el aprendizaje, esta atribuye de manera variable diferente peso o importancia a los distintos conjuntos de datos que contiene, estableciendo jerarquías entre ellos. (*N. de la E.*)

o que se ofrezca documentación como las tarjetas del modelo²² o las fichas técnicas.²³ Sin embargo, algunos modelos que se describen como «abiertos» ofrecen poco más que una API) y una licencia que autoriza su reutilización, incluida la comercialización de la tecnología.

Por ejemplo, tanto GPTNeo como LLaMA-2 han sido descritos como sistemas de IA «abiertos», pero en la práctica son muy distintos. Si bien LLaMA-2 está disponible para descarga gratuita y permite acceder a los pesos del modelo, no cumple determinados criterios clave para que convencionalmente pueda considerarse de código abierto.²⁴ Meta redactó la licencia del modelo desde cero y la OSI no la ha reconocido; se puede afirmar que no se ajusta a los principios de la *Open Source Definition* (el documento de la OSI que funciona como estándar definitorio consolidado en el contexto del software de código abierto).²⁵ Tampoco ofrece

22. Margaret Mitchell, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, Timnit Gebru, «Model cards for model reporting», *arXiv*, 14 de enero de 2019, [lc.cx/j67gbA](https://arxiv.org/abs/1901.02628).

23. Timnit Gebru, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé III, Kate Crawford, «Datasheets for datasets», *arXiv*, 1 de diciembre de 2021, [lc.cx/NGIaTu](https://arxiv.org/abs/2106.04601).

24. Maffulli, «Meta's LLaMa 2 license is not open source», *op. cit.*

25. *Ibid.* Téngase en cuenta que la OSI está en proceso de revisar esta definición para que se ajuste a las particularidades del desarrollo de IA de código abierto.

una transparencia significativa, especialmente si atendemos a los datos que utiliza para entrenar al sistema. Ante un alcance tan limitado, hay quien argumenta que LLaMA-2 no debería considerarse de código abierto.²⁶ En cambio, GPTNeo, desarrollado y conservado por EleutherAI, un proyecto sin ánimo de lucro, se ofrece en virtud de la licencia de código abierto del MIT (Massachusetts Institute of Technology), reconocida por la OSI desde hace mucho tiempo; además, permite consultar los pesos del modelo, los parámetros y la información detallada sobre el entrenamiento y la configuración. Asimismo, fue entrenado con un conjunto de datos abiertamente disponible: The Pile,²⁷ también desarrollado por EleutherAI.

El artículo de Irene Solaiman en el que propone el concepto de «gradientes de apertura» ofrece una referencia muy útil para reflexionar en torno a los muchos niveles posibles de transparencia, capacidad de reutilización y de desarrollo de los distintos elementos que conforman la IA (véase cuadro de la página 103).

A continuación, revisamos los recursos necesarios —marcos, computación, datos, mano de obra y modelos— para crear y utilizar grandes sistemas de IA. Hacerlo nos permite dibujar una perspectiva común de qué elementos conforman un sistema de IA, con la que podremos

²⁶. Nolan, «Llama and ChatGPT are not open-source», *op. cit.*

²⁷. «GPT-Neo», *EleutherAI*, [lc.cx/UI9aOL](https://lcs.cmu.edu/~jferret/ai/).

evaluar qué partes de ese sistema son (o pueden ser) transparentes y con capacidad de reutilización y desarrollo, cuáles no lo son (o no lo pueden ser) y en qué medida.²⁸

Marcos de desarrollo

Los marcos de desarrollo de software permiten que los desarrolladores puedan compilarlo e implementarlo más fácilmente de formas reglamentadas, predecibles y apropiadas. Actualmente forman parte de las prácticas de desarrollo estándar, crean el entorno que define cómo tiene lugar el desarrollo y no son exclusivos de la IA. Ofrecen elementos de código preescritos, plantillas de flujos de trabajo, herramientas de evaluación y otros métodos y componentes estandarizados para las tareas de desarrollo habituales. Esto ayuda a crear sistemas computacionales más fungibles, interoperables y testeables, a la vez que reduce el tiempo invertido en «volver a inventar la rueda» y evita los errores que se pueden introducir fácilmente al implementar sistemas desde cero. Estos componentes casi siempre son de código abierto; eso es lo que permite que se usen como elementos comunes para el desarrollo.

²⁸. Dieuwertje Luitse y Wiebke Denkena, «The great Transformer. Examining the role of large language models in the political economy of AI», *Big Data & Society*, vol. 8, n.º 2, julio de 2021, [lc.cx/B909F3](https://doi.org/10.1093/bds/bgab03).

<p>Consideraciones</p> <ul style="list-style-type: none"> ➤ Solo investigación interna ➤ Control de alto riesgo ➤ Poca capacidad de auditoría ➤ Perspectivas limitadas 	<p style="text-align: center;">Gradientes de apertura de la IA abierta</p> <p style="text-align: center;">Fuente: Irene Solaiman, «The gradient of generative AI release», arXiv, 2023</p>				<ul style="list-style-type: none"> ➤ Investigación comunitaria ➤ Control de bajo riesgo ➤ Gran capacidad de auditoría ➤ Perspectivas más amplias 	
<p>Nivel de acceso</p>	<p>Completamente cerrado</p>	<p>Publicación gradual/escalonada</p>	<p style="text-align: center;">Cerrado al público</p>			<p>Completamente abierto</p>
<p>Sistema (desarrollador)</p>	<ul style="list-style-type: none"> ➤ PaLM (Google) ➤ Gopher (DeepMind) ➤ Imagen (Google) ➤ Make-A-Video (Meta) 	<ul style="list-style-type: none"> ➤ GPT-2 (OpenAI) ➤ Stable Diffusion (Stability AI) 	<p>Acceso hospedado</p>	<p>Acceso basado en la nube/a través de API</p>	<p>Descargable</p>	<ul style="list-style-type: none"> ➤ BLOOM (Big-Science) ➤ GPT-J (EleutherAI)

Suelen reutilizarse en virtud de una u otra licencia de código abierto que define las condiciones en las que se permite la reutilización.

Como sucede con el software en general, la evolución de la IA depende de un puñado de marcos de desarrollo de código abierto muy populares. Se trata de repositorios de conjuntos de datos cada vez más enormes, herramientas de validación de datos, de evaluación, de construcción de modelos, herramientas para el entrenamiento y la exportación de modelos, librerías de preentrenamiento y mucho más, que en su conjunto definen cómo se crea y se implementa la IA.

Los dos marcos de desarrollo de IA predominantes son PyTorch y TensorFlow. Ambos fueron elaborados por grandes empresas tecnológicas comerciales, Meta y Google respectivamente, que siguen ocupándose de abastecerlos y mantenerlos. Desde 2022, existen muchos más modelos de IA preentrenados que funcionan exclusivamente en el marco de PyTorch (incluidos los modelos GPT de OpenAI) que aquellos que se ejecutan en TensorFlow. PyTorch también es el marco más popular en la investigación académica sobre IA y se utiliza en la mayoría de artículos de investigación.²⁹

PyTorch fue desarrollado inicialmente para uso interno de Meta, pero se hizo público en el año 2017.

²⁹. Kelsey Foster, «PyTorch vs TensorFlow. Who has more pre-trained deep learning models?», *Hackernoon*, 21 de marzo de 2022, [lc.cx/kQeCMZ](https://l.ccx/kQeCMZ).

Aunque este modelo funciona como una plataforma de investigación bajo el paraguas de la Fundación Linux, sigue financiándose a través de Meta,³⁰ tres de sus cuatro colaboradores principales son empleados de esta empresa y el cuarto lo fue en el pasado (una práctica bastante habitual dentro de la industria tecnológica cuando se trata de proyectos de código abierto de importancia estratégica).³¹ TensorFlow fue desarrollado por Google Brain y publicado por primera vez en 2015;³² la empresa sigue ocupándose de su dirección y financiación, y también emplea a muchos de sus principales colaboradores.³³

Los marcos de desarrollo de código abierto ofrecen herramientas que hacen que el proceso de desarrollo

30. Andrew Tarantola, «Meta is spinning off the Pytorch framework into its own AI research foundation», *Engadget*, 12 de septiembre de 2022, [lc.cx/4LGjII](https://www.engadget.com/meta-is-spinning-off-the-pytorch-framework-into-its-own-ai-research-foundation-1222092222.html).

31. «Contributors to PyTorch», *PyTorch*, [lc.cx/gQh-yN](https://pytorch.org/docs/master/contributors.html).

32. Cade Metz, «Google just open sourced TensorFlow, its artificial intelligence engine», *Wired*, 9 de noviembre de 2015, [lc.cx/BtVyN2](https://www.wired.com/2015/11/google-just-open-sourced-tensorflow-its-artificial-intelligence-engine/).

33. Véase «Contributors to TensorFlow», *Google*, [lc.cx/sU4v1y](https://www.tensorflow.org/contributors). La decisión de Google de ofrecer TensorFlow en código abierto provocó un cambio significativo en el escenario de la IA de código abierto. Aunque existían precedentes de marcos en código abierto, como Chainer, Theano o Caffe, la calidad del paquete TensorFlow de Google era considerablemente mayor e hizo que Amazon, Microsoft y Meta lanzaran sus propios marcos. Véase Maximilliam Langenkamp, *How open source machine learning software shapes AI*, tesis, MIT, Massachusetts, 2022, [lc.cx/NP2jPA](https://www.mit.edu/~mlangenk/papers/2022-open-source-ml-ai/).

e implementación de IA sea más rápido, más predecible y más robusto. También aportan beneficios importantes a las empresas que los desarrollan. Lo más relevante es que permiten que Meta, Google o quien sea que oriente el desarrollo de los marcos estandaricen la construcción de IA para que sea compatible con sus propias plataformas corporativas; así, garantizan que sus marcos condicionen a los desarrolladores para que creen sistemas de IA que, como piezas de Lego, encajen en los entornos de su propia empresa.³⁴ En el caso de Meta, esto le permite integrar y comercializar más fácilmente modelos de IA académicos y de otro tipo que se desarrollan, se ajustan o se implementan con PyTorch. Sucede lo mismo con Google y TensorFlow; en su caso, este marco ha sido creado para que funcione fácil e intuitivamente con el hardware de las TPU³⁵ de

^{34.} Alex Engler, «How open-source software shapes AI policy», *Brookings*, 10 de agosto de 2021, [lc.cx/VACv3Y](https://www.brookings.edu/blog/tech-policy/2021/08/10/how-open-source-software-shapes-ai-policy/).

^{35.} Una unidad de procesamiento tensorial (TPU, por sus siglas en inglés) es una tipología de microprocesador diferente a la clásica CPU (unidad de procesamiento central) o GPU (unidad de procesamiento gráfico), más popularmente conocida. Desarrollada por Google específicamente para las tareas de aprendizaje automático, está pensada para realizar cálculos de baja precisión a gran escala y realizar inferencias (recoger datos y convertirlos en conclusiones). Se emplean, por ejemplo, en el reconocimiento y procesamiento de imágenes (pero no de gráficos, videojuegos o similares) o de lenguaje natural (pero no de texto), en simulaciones científicas o datos financieros. (*N. de la E.*)

Google, la potente infraestructura informática que vertebra su negocio de computación en la nube. Así, Google puede optimizar su oferta comercial en la nube para el desarrollo de IA y colocar estas rentables ofertas como el motor del desarrollo de la IA.

Los marcos de desarrollo de la IA de código abierto permiten que quienes los financian y controlan creen vías de acceso a ofertas de computación rentables. También favorecen que Google y Meta, en este caso, definan las prácticas de trabajo de investigadores y desarrolladores, para que los nuevos modelos de IA se puedan integrar y comercializar fácilmente. Esto hace que la empresa que ofrece el marco ejerza un importante poder indirecto dentro del ecosistema: forma a desarrolladores, investigadores y estudiantes que interactúan con estas herramientas según las normas del marco que prefiera la compañía; así, interviene en la definición (y, de algún modo, en la captura) del campo de la IA.³⁶ En apartados posteriores ofrecemos ejemplos más concretos de este fenómeno.³⁷

^{36.} Langenkamp, *How open source machine learning software shapes AI, op. cit.*; véase también el primer capítulo de este libro, «El elevado coste de la captura», p. 17.

^{37.} En particular, las empresas también suelen patentar los nuevos métodos de IA en el momento en que liberan el código correspondiente en abierto; esto las protege ante demandas de propiedad intelectual, a la vez que favorece la adopción masiva y «el desarrollo derivado». Véanse, por ejemplo, Noam M. Shazeer, Aidan Nicholas Gomez, Lukasz Mieczyslaw Kaiser,

Desarrollar poderosos sistemas de IA exige conjuntos de datos gigantescos que a su vez requieren una enorme capacidad computacional para poder ser procesados.³⁸ El desarrollo contemporáneo de IA se traduce en una carrera por la escala; los conjuntos de datos aumentan cada año en un orden de magnitud y las necesidades computacionales siguen ese mismo ritmo.³⁹ El acceso a la computación representa una barrera importante para la reutilización, incluso en los sistemas de IA con criterios de «apertura» más maximalistas, puesto que el entrenamiento y la ejecución de inferencias a escala en grandes modelos

Jakob D. Uszkoreit, Llion Owen Jones, Niki J. Parmar, Illia Polosukhin y Ashish Teku Vaswani, «Attention-based sequence transduction neural networks», Google Patents, [lc.cx/6vNiZY](https://patents.google.com/patent/lc.cx/6vNiZY); Geoffrey E. Hinton, Alexander Krizhevsky, Ilya Sutskever y Nitish Srivastva, «System and method for addressing overfitting in a neural network», Google Patents, [lc.cx/djooCz](https://patents.google.com/patent/lc.cx/djooCz).

³⁸. Jared Kaplan, Sam McCandlish, Tom Henighan, Tom B. Brown, Benjamin Chess, Rewon Child, Scott Gray, Alec Radford, Jeffrey Wu y Dario Amodei, «Scaling laws for neural language models», *arXiv*, 23 de enero de 2020, [lc.cx/J5n4HL](https://arxiv.org/abs/2001.01687); Jai Vipra y Sarah Myers West, «Computational power and AI», *AINow*, 27 de septiembre de 2023, [lc.cx/3D2g6H](https://ainow.com/2023/09/27/computational-power-and-ai/).

³⁹. Dylan Patel, «The AI brick wall. A practical limit for scaling dense transformer models, and how GPT 4 will break past it», *Semianalysis*, 24 de enero de 2023, [lc.cx/A4ifmi](https://semianalysis.com/2023/01/24/the-ai-brick-wall/).

de IA (es decir, incorporarlos en un producto o API de uso público generalizado) tienen un coste muy elevado.

Estos notables requisitos computacionales no tienen por qué reducirse una vez superada la fase preliminar, cuando el modelo de IA se entrena y se calibra inicialmente. De hecho, los requisitos de computación iniciales pueden verse eclipsados por la computación necesaria para que grandes modelos de IA ofrezcan respuestas o generen imágenes en el mundo real.⁴⁰ Es difícil obtener cifras exactas, ya que este sigue siendo un asunto muy custodiado (por lo general, escondido en los contratos que se celebran entre proveedores de infraestructura y empresas), pero una estimación aproximada sugiere que la computación necesaria para ejecutar la interfaz de ChatGPT supera semanalmente el coste de formación de su modelo subyacente GPT-4.⁴¹ Otra estimación de una empresa de análisis financieros calculó que Microsoft tendría que invertir cuatro mil millones de dólares estadounidenses en infraestructura

⁴⁰. Mientras que los modelos en dispositivos o «periféricos» se utilizan con aplicaciones limitadas para inferencia a pequeña escala, en la mayoría de los casos la computación centralizada ofrece ventajas importantes que resultan especialmente relevantes para la IA a gran escala.

⁴¹. Dylan Patel y Afzal Ahmad, «The inference cost of search disruption. Large language model cost analysis», *Semianalysis*, 9 de febrero de 2023, lc.cx/w6L8bu.

para implementar ChatGPT dentro de la herramienta de búsqueda de Bing.⁴² El elevado coste de inferencia provocó que el director general de OpenAI, Sam Altman, afirmara ante el Congreso: «Intentamos desarrollar sistemas que no promuevan el interés del público. De hecho, estamos tan escasos de GPU, que cuanto menos gente utilice nuestros productos, mejor».⁴³

Esta necesidad de más y más computación no tiene visos de reducirse. En un perfil de OpenAI publicado recientemente (y después eliminado), la empresa expone su esfuerzo por obtener cada vez más recursos de computación y reconoce que la limitación de GPU (procesadores especializados que se utilizan para entrenar IA) es su principal obstáculo en la carrera hacia modelos más grandes y potentes.⁴⁴

Además, conseguir la máxima capacidad computacional del hardware especializado exige sistemas de software específicos y, en algunos casos, patentados. La mayor parte del entrenamiento de los modelos de IA se implementa en CUDA, un marco desarrollado y patentado por Nvidia que solo permite el entrenamiento

^{42.} *Ibid.*

^{43.} Will Oremus, «AI chatbots lose money every time you use them. That is a problem», *The Washington Post*, 5 de junio de 2023, [lc.cx/vsVFEx](https://www.washingtonpost.com/technology/ai-chatbots-lose-money-every-time-you-use-them-that-is-a-problem/2023/06/05/).

^{44.} Raza Habib, «OpenAI's plans according to Sam Altman», *Humanloop*, 29 de mayo de 2023, [lc.cx/jXc8yc](https://www.humanloop.com/blog/openai-plans-according-to-sam-altman/).

en sus propias GPU.⁴⁵ OpenAI creó Triton, presumiblemente como alternativa de código abierto a CUDA, que podría permitir una futura portabilidad entre distintos tipos (plataformas) de GPU; sin embargo, actualmente Triton solo es compatible con las GPU de Nvidia.⁴⁶ Por su parte, OpenAI invita a sus colaboradores comunitarios a intentar que Triton funcione en otras plataformas de computación; ahora bien, la empresa no invierte los recursos necesarios para hacerlo posible. Los programas TensorFlow de Google están escritos para que el código se pueda ejecutar directamente en las TPU de la misma marca; además, TensorFlow cuenta con varios compiladores adicionales específicos del dominio, como XLA, capaz de acelerar la ejecución de modelos concretos en el hardware de Google. Esto quiere decir que, aunque podamos acceder en abierto a cierto software de nivel inferior que optimiza capacidad computacional para el desarrollo de IA y podamos también inspeccionar su código, en la práctica este software suele estar

^{45.} GPU (siglas de *graphics processing unit*) significa «unidad de procesamiento gráfico». Los procesadores especializados de computación se desarrollaron inicialmente para los videojuegos y ahora se utilizan para entrenamiento de IA, al permitir realizar muchos cálculos en paralelo rápidamente.

^{46.} «Introducing Triton. Open-source GPU programming for neural networks», *OpenAI*, 28 de julio de 2021, [lc.cx/VuWoH4](https://lcccx.com/VuWoH4).

diseñado para que resulte más eficiente en entornos de hardware patentado, desarrollado y controlado por las empresas que venden recursos de computación o que licencian modelos de IA.

En resumen, los recursos de computación necesarios para construir nuevos modelos de IA y utilizar los existentes a escala (fuera del contexto de la empresa privada y de los experimentos individuales) son escasos, extremadamente caros y se concentran en manos de un puñado de corporaciones, la cuales, a su vez, se benefician de la economía de escala, de controlar el software que optimiza la computación y de vender el costoso acceso a los recursos de computación. Esta importante asimetría de recursos socava cualquier pretensión de democratización que pueda ser defendida con la disponibilidad de modelos de IA «abiertos».

Datos

Para construir sistemas de IA a gran escala, es necesario contar con datos que se correspondan con precisas (y exigentes) especificaciones. Algunos investigadores incluso afirman que, a la hora de desarrollar IA a gran escala, el acceso a los datos podría ser más importante que el acceso a la computación.⁴⁷ Se trata de dos

⁴⁷. Micah Musser, Rebecca Gelles, Ronnie Kinoshita, Catherine Aiken y Andrew Lohn, «The main resource is the human», *CSET*, abril de 2023, [lc.cx/cP8UzK](https://www.cset.org/2023/04/04/the-main-resource-is-the-human/).

elementos esenciales y, en la actual carrera por escalar el paradigma, cuanto mayor sea el volumen de cada uno de ellos de que dispongan los modelos, «mejor» funcionarán.⁴⁸

Sabemos muy poco acerca de los datos a los que se recurrió para entrenar la actual generación de sistemas de IA generativa disponibles (GPT-4, Bard, PaLM-2). En concreto, OpenAI/Microsoft y Google se niegan a revelar información importante sobre los datos de entrenamiento que se utilizaron para definir GPT-4⁴⁹ y PaLM-2⁵⁰ respectivamente; tampoco han facilitado el número de *tokens*⁵¹ (fragmentos de datos concretos) con los que se entrenaron estos modelos y menos aún los detalles sobre el origen de los datos u otra información necesaria para comprender (hasta cierto punto) las tendencias de dichos modelos.

48. Kaplan *et al.*, «Scaling laws for neural language models», *op. cit.*

49. OpenAI, «GPT-4 technical report», *arXiv*, 27 de marzo de 2023, [lc.cx/7_P3v0](#).

50. Jennifer Elias, «Google's newest A.I. model uses nearly five times more text data for training than its predecessor», *CNBC*, 16 de mayo de 2023, [lc.cx/WxQowR](#).

51. Los denominados «tókenes» son una suerte de máscara que permite transmitir datos sensibles entre diferentes puntos en cuyo recorrido se les da apariencia numérica o gráfica que no transmite su contenido real, por ejemplo, en la transmisión de datos bancarios durante un pago o una transferencia. (*N. de la T.*)

Sí sabemos algo más de los datos que se usaron para entrenar a las generaciones anteriores de sistemas de IA disponibles en el mercado y de modelos de IA «abiertos» más habituales. Los actuales modelos de IA «abiertos» suelen depender de dos repositorios de datos extraídos de internet y puestos a disposición para su uso generalizado: el conjunto de datos Common Crawl y The Pile. Common Crawl es un repositorio de datos extraídos de internet que una organización homónima y sin ánimo de lucro, con sede en Estados Unidos, actualiza cada mes con el objetivo de «democratizar [...] los datos para que todo el mundo, no solo las grandes empresas, pueda desarrollar investigación y análisis de gran calidad».⁵² Common Crawl está disponible a través de Amazon, cuya plataforma AWS hospeda los datos en el marco del Programa de Patrocinio de Datos Abiertos que desarrolla la empresa.⁵³ The Pile es un conjunto de datos de amplia disponibilidad diseñado para producir sistemas de lenguaje de IA. Se creó a partir de la compilación de veintidós conjuntos de datos más pequeños, entre ellos Wikipedia, los subtítulos de YouTube y el sitio web *HackerNews*. La comunidad The Eye, con sede en Europa, hospeda este conjunto de datos «en apoyo al constante esfuerzo por democratizar el acceso y

^{52.} «Frequently asked questions», *Common Crawl*, [lc.cx/3Jw6PB](https://commoncrawl.org/faq/).

^{53.} «Common Crawl, index server», *Common Crawl* [lc.cx/mxMZEi](https://commoncrawl.org/index-server/).

la investigación»;⁵⁴ se trata de una entidad creada por EleutherAI,⁵⁵ la organización sin ánimo de lucro centrada en desarrollar y conservar IA «abierta». EleutherAI cuenta con el apoyo de empresas con intereses comerciales, entre otras, Stability AI y Hugging Face, que le ofrecen financiación y acceso a computación en la plataforma AWS de Amazon (en virtud de acuerdos que mantienen estas últimas con AWS).⁵⁶

Extraer datos para crear conjuntos de datos con la finalidad de desarrollar IA también genera conflictos de extracción y propiedad intelectual. Estos conjuntos, ya sean abiertos o cerrados, se suelen elaborar a partir de imágenes, texto y código de internet protegidos por derechos de autor, o bien copiando y reutilizando conjuntos de datos compilados por grupos lingüísticos del mundo mayoritario, como Ghana NLP o Lesan AI.⁵⁷

⁵⁴. Leo Gao, Stella Biderman, Sid Black, Laurence Golding, Travis Hoppe, Charles Foster, Jason Phang, Horace He, Anish Thite, Noa Nabeshima, Shawn Presser y Connor Leahy, «The Pile. An 800GB dataset of diverse text for language modeling», *arXiv*, 31 de diciembre de 2020, [lc.cx/d98-3S](https://arxiv.org/abs/2001.01629).

⁵⁵. *Ibid.*

⁵⁶. Kyle Wiggers, «Stability AI, Hugging Face and Canva back new AI research nonprofit», *TechCrunch*, 2 de marzo de 2023, [lc.cx/DoOZNp](https://techcrunch.com/2023/03/02/stability-ai-hugging-face-canva-back-new-ai-research-nonprofit/).

⁵⁷. «Ghana natural language processing (NLP)», *Ghana NLP*, [lc.cx/_v_71O](https://ghana-nlp.com/); Andrew Deck, «The AI startup outperforming Google Translate in Ethiopian languages», *Rest of World*, 11 de julio de 2023, [lc.cx/yV9qbW](https://restofworld.com/news/ai-startup-outperforming-google-translate-in-ethiopian-languages/).

Esto significa que, en la práctica, quienes utilizan estos conjuntos de datos para entrenar y evaluar modelos de IA a menudo se sirven de la propiedad intelectual y del trabajo ajenos para hacerlo, aduciendo un uso legítimo incluso cuando esos alegatos se basan en fundamentos jurídicos poco firmes.⁵⁸ Sea legal o no, la práctica de arrasarse indiscriminadamente con los datos de internet para crear sistemas que hoy por hoy logran debilitar el sustento de escritores, artistas y programadores —con cuyo esfuerzo se crearon esos datos «de internet» en primera instancia— ha generado alarma y enfado;⁵⁹ asimismo, las partes afectadas han interpuesto múltiples demandas que avanzan en su tramitación.⁶⁰

Estas preocupaciones resultan especialmente acuciantes si tenemos en cuenta el contexto colonial en el

58. Mehtab Khan y Alex Hanna, «The subjects and stages of AI dataset development. A framework for dataset accountability», *SSRN*, 13 de septiembre de 2022, [lc.cx/pyhlcL](https://www.ssrn.com/lc/cx/pyhlcL).

59. Harry H. Jiang, Lauren Brown, Jessica Cheng, Mehtab Khan, Abhishek Gupta, Deja Workman, Alex Hanna, Johnathan Flowers y Timnit Gebru, «AI art and its impact on artists», en *AIES'23 Conference*, Association for Computing Machinery, 2023, [lc.cx/eRc5Kw](https://www.acm.org/lc/cx/eRc5Kw).

60. Deck, «The AI startup ...», *op. cit.*; Chloe Xiang, «A photographer tried to get his photos removed from an AI dataset. He got an invoice instead», *Vice*, 28 de abril de 2023, [lc.cx/3MBp6h](https://www.vice.com/en/article/3MBp6h); Riddhi Setty, «Sarah Silverman, authors hit OpenAI, Meta with copyright suits», *Bloomberg Law*, 10 de julio de 2023, [lc.cx/eUH1aR](https://www.bloomberglaw.com/lc/cx/eUH1aR).

que se produce gran parte de esta explotación. Los sistemas de IA suelen ser construidos por una minoría de empresas del mundo, con recursos de datos y de mano de obra obtenidos de la mayoría de la población mundial, a la que luego se imponen sin consulta previa.⁶¹ Como planteaba en septiembre de 2022 Paul Azunre, fundador del proyecto de código abierto Ghana NLP:

*Si la comunidad africana que investiga la inteligencia artificial y el aprendizaje automático no está atenta, este nuevo movimiento de «código abierto», promovido por las empresas tecnológicas internacionales más poderosas, se convertirá en un mecanismo de explotación continua de nuestro capital humano y nuestro continente.»*⁶²

Las consecuencias de estas dinámicas coloniales agravan los patrones históricos de desigualdad, pues refuerzan estereotipos y sesgos preocupantes.⁶³

^{61.} Grace Browne, «AI is steeped in big tech’s “digital colonialism”», *Wired*, 25 de mayo de 2023, [lc.cx/Hzoh1X](https://www.wired.com/story/ai-digital-colonialism/).

^{62.} Paul Azunre, «If African AI/ML researchers are not careful, this new “Open Source” movement championed by the richest global tech companies will become a mechanism for continued exploitation of our human capital and continent... [1/3]», *X*, 13 de septiembre de 2022, [lc.cx/GA5yFM](https://twitter.com/pazunre/status/1570000000000000000).

^{63.} Abeba Birhane, Vinay Uday Prabhu y Emmanuel Kahembwe, «Multimodal datasets. Misogyny, pornography and malignant stereotypes», *arXiv*, 5 de octubre de 2021, [lc.cx/ES-JSH](https://arxiv.org/abs/2110.00000).

Esta explotación también avanza diametralmente en contra de las nociones de soberanía de datos surgidas en el mundo mayoritario, que se reflejan en proyectos como Te Hiku Media:

Aunque reconocemos la importancia de la tecnología de código abierto, somos conscientes de que puede que la mayoría del pueblo tangata whe-nua, así como otros pueblos indígenas, no tengan acceso a los recursos que les permitirían beneficiarse de las tecnologías de código abierto [...]. Si nos limitamos a poner nuestros datos y nuestro conocimiento a disposición del código abierto, permitiremos una colonización digital de nuestro pueblo aún más profunda en el mundo moderno.⁶⁴

De acuerdo con los principios maoríes, Te Hiku desarrolló, y hoy mantiene, la licencia Kaitiakitanga, un ejemplo de custodia de datos por parte del pueblo del que proceden.⁶⁵

Con esto no pretendemos hacer una defensa de los conjuntos de datos cerrados, los cuales agravan profundamente el problema, pero hemos de ser claros

⁶⁴. Ngāti Kuri, Te Aupōuri, Ngai Takoto, Te Rarawa y Ngāti Kahu, «He reo tuku iho, he reo ora. Living language transmitted intergenerationally», *Mai Journal*, vol. 11, n.º 1, 2022, ic.cx/_lxonP.

⁶⁵. «Kaitiakitanga license», *Whare kōrero*, ic.cx/ux4q6Y.

acerca de qué permiten y qué no. Cuando los conjuntos de datos no están disponibles para escrutinio o cuando resultan inescrutables debido a su gran tamaño, es muy difícil comprobar si dichos conjuntos de datos blanquean la propiedad intelectual ajena o si utilizan con fines comerciales datos con licencias específicas de uso no comercial u otorgadas en virtud de ciertas normas de soberanía. Por ejemplo, se ha demostrado que el asistente de programación de Microsoft, GitHub Copilot —un sistema de IA generativa que produce código—, ha sido entrenado con código sujeto a una licencia GPL⁶⁶ (licencia de código abierto que exige que el código derivado se publique en virtud de las mismas condiciones), código que el sistema después regurgita. Sin embargo, incluso usar código con una licencia permisiva (por ejemplo, la licencia MIT) para entrenar IA generativa podría violar de forma similar las disposiciones que exigen atribución; algo que los actuales sistemas de IA generativa podrían proporcionar, pero que no proporcionan.

Los conjuntos de datos como The Pile y Common Crawl se ofrecen sujetos a licencias de uso generalizado. Sin embargo, no están listos para ser utilizados. Hace falta un mayor esfuerzo de mano de obra para que esos conjuntos de datos sirvan para desarrollar grandes modelos de IA. Para crear una IA eficiente, es

⁶⁶. Karabus, «GitHub CEO says EU AI Act shouldn't apply to open source devs», *op. cit.*

necesario construir y combinar los conjuntos de datos de forma minuciosa: el modelo LLaMA de Meta obtiene muestras de siete conjuntos de datos distintos en proporciones variables durante el proceso de calibración del rendimiento.⁶⁷ El modelo BLOOM de BigScience fue entrenado por una combinación de 498 conjuntos de datos, lo que supuso un complejo proceso de gobernanza y un filtrado de calidad, realizado manualmente, para eliminar código, *spam* y otros contenidos innecesarios.⁶⁸ El modelo Falcon 40B (que actualmente se considera el mejor de su clase entre los modelos de IA de código abierto para una serie de tareas de evaluación limitadas), entrenado por el Technology Innovation Institute, con sede en Abu Dabi, prioriza la calidad de los datos en su documentación. En este caso, partiendo de Common Crawl, se filtró y se deduplicó el conjunto de datos, antes de ampliarlo con fuentes «como artículos de investigación y conversaciones de redes sociales», subrayando que «se prestó especial atención a la calidad de

^{67.} Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, Aurelien Rodriguez, Armand Joulin, Edouard Grave y Guillaume Lample, «LLaMA: open and efficient foundation language models», *arXiv*, 27 de febrero de 2023, [lc.cx/u_BSCp](https://arxiv.org/abs/2302.13971).

^{68.} Teven Le Scao, Angela Fan, Christopher Akiki *et al.*, «BLOOM: a 176b-parameter open-access multilingual language model», *arXiv*, 27 de junio de 2023, [lc.cx/5r21pQ](https://arxiv.org/abs/2306.01533).

los datos a escala». ⁶⁹ Aunque todavía no se ha publicado documentación que detalle el modelo Falcon, sí existe un artículo sobre su conjunto de datos que repasa sus canales de datos y los procesos con los que se llevó a cabo la cuidadosa selección de datos de internet. ⁷⁰

Las prácticas de etiquetado de datos (que atribuyen «significado» a los datos de entrenamiento y evaluación de la IA, y definen la ontología que a su vez reflejarán los sistemas de IA durante sus inferencias) no suelen ser transparentes y normalmente dependen de plataformas patentadas. Como veremos a continuación, también exigen una ingente mano de obra. Las empresas más conocidas que trabajan en el ámbito de las plataformas de etiquetado de datos son Amazon Mechanical Turk, Sama y ScaleAI, todas ellas con fines de lucro (a diferencia de aquellas de gestión comunitaria o de los proyectos académicos); así, se establece un vínculo entre los clientes que pagan por el uso y los trabajadores a los que se paga (poco) por etiquetar datos a demanda.

La preparación y gestión de los datos que se utilizan para entrenar y calibrar los modelos líderes de IA a gran escala implica ciertos procesos que hacen un uso

⁶⁹. «Falcon LLM», *Technology Innovation Institute*, [lc.cx/-1gYep](https://lcccx-1gYep).

⁷⁰. Guilherme Penedo, Quentin Malartic, Daniel Hesslow, Ruxandra Cojocar, Alessandro Cappelli, Hamza Alobeidli, Baptiste Pannier, Ebtesam Almazrouei, Julien Launay, «The RefinedWeb dataset for Falcon LLM. Outperforming curated with web data, and web data only», *arXiv*, 1 de junio de 2023, [lc.cx/8YMWt8](https://lcccx/8YMWt8).

intensivo de los recursos, algo mucho más complicado que descargar un conjunto de datos de acceso libre. La transparencia y la capacidad de reutilización de conjuntos de datos como The Pile y Common Crawl permiten una mejor evaluación del entrenamiento del modelo y de sus limitaciones. Pero, además del coste y del tiempo necesario para desarrollarlos, hace falta mucha mano de obra para preparar esos datos antes de que sean utilizados con el fin de que el entrenamiento permita un mejor desempeño del modelo. Además, como ya hemos visto, el coste y la escasez de la computación necesaria para procesar estos datos limitan en la práctica quién puede utilizarlos con el propósito de crear sistemas de IA a gran escala.

Mano de obra

La insaciable necesidad de datos seleccionados, etiquetados y cuidadosamente organizados de los sistemas de IA a gran escala supone que su desarrollo a escala exige una importante aportación de mano de obra humana. Ese esfuerzo es el que crea la «inteligencia» que los sistemas de inteligencia artificial comercializan como automatización y computación.⁷¹ A grandes

⁷¹. Adrienne Williams, Milagros Miceli y Timnit Gebru, «The exploited labor behind artificial intelligence», *Noēma*, 13 de octubre de 2022, [lc.cx/vZicv2](https://l.ccx/vZicv2); Astra Taylor, «The automation

rasgos, este trabajo se puede clasificar en las siguientes tareas:

- Etiquetado y clasificación de datos.
- Calibración de modelos (aprendizaje reforzado con retroalimentación humana y procesos similares).
- Moderación de contenido, confianza y seguridad y otros soportes posteriores a la implementación.
- Ingeniería y desarrollo y mantenimiento del producto.

Los sistemas de IA generativa, que son los grandes sistemas a gran escala que actualmente despiertan mayor interés, se entrenan y se evalúan con una amplia gama de texto, discurso e imágenes generados por humanos. El proceso de definir un modelo de tal forma que pueda imitar un resultado similar al de una persona, y hacerlo sin replicar material ofensivo o peligroso, exige una gran implicación humana para garantizar que lo que produzca el modelo esté dentro de los límites de lo que se considera «acceptable»;⁷² de este modo,

charade. The rise of the robots has been greatly exaggerated. Whose interests does that serve?», *Logic(s)*, n.º 5, 1 de agosto de 2018, lc.cx/gKlIfE.

⁷². Nanna Thylstrup y Zeerak Talat, «Detecting “dirt” and “toxicity”. Rethinking content moderation as pollution behaviour», *SSRN*, enero de 2020, lc.cx/WL9vuo.

empresas y otras instituciones podrán comercializarlo, venderlo y aplicarlo en el mundo real conservando su clientela y su reputación. Este proceso se suele denominar *aprendizaje de refuerzo con retroalimentación humana* (o RLHF, por sus siglas en inglés), que suena mucho más técnico aunque, en la práctica, se refiera a miles de horas de trabajo humano por lo general traumático. En un proceso habitual de RLHF, la tarea de la mano de obra puede ser seleccionar cuál de entre varios fragmentos de texto producidos por un sistema de IA generativa se parece más a un texto de autoría humana. Las elecciones del trabajador son devueltas al sistema para «enseñarle» qué es y qué no es «humano». ⁷³ Estos procesos también se usan para que el modelo «aprenda» a reconocer y filtrar resultados inapropiados, ofensivos o peligrosos, un trabajo que en repetidas ocasiones se ha demostrado perjudicial para la salud mental de quienes lo llevan a cabo. ⁷⁴ Estas tareas suelen externalizarse, lo que ofrece una distancia entre

⁷³. «Learning from human preferences», *OpenAI*, 13 de junio de 2017, [lc.cx/Mey81t](https://openai.com/research/learning-from-human-preferences).

⁷⁴. Los trabajadores de Meta y Microsoft han presentado demandas por las pésimas consecuencias de unas prácticas laborales similares en la moderación de contenidos. Véanse Casey Newton, «Facebook will pay \$52 million in settlement with moderators who developed PTSD on the job», *The Verge*, 12 de mayo de 2020, [lc.cx/hK591j](https://www.theverge.com/2020/5/12/21251111/facebook-settlement-ptsd-moderators); y Sam Levin, «Moderators who had to view child abuse content sue Microsoft, claiming PTSD», *The Guardian*, 12 de enero de 2017, [lc.cx/joHB4a](https://www.theguardian.com/technology/2017/jan/12/microsoft-ptsd-moderators).

la empresa que desarrolla y comercializa el modelo y las pésimas condiciones de trabajo que se dan en el proceso de entrenamiento.⁷⁵ Bard, el modelo de Google, depende de trabajadores contratados por empresas de externalización como Accenture y Appen, cuya tarea es tomar decisiones oportunas sobre los límites de lo que se considera una expresión «aceptable», a pesar de que, al parecer, reciben una formación mínima y trabajan con unos plazos frenéticos.⁷⁶ Para sus modelos GPT, OpenAI contrató trabajadores en Kenia a través de la empresa de externalización Sama. Este proceso tuvo unas nefastas consecuencias, puesto que los trabajadores se vieron obligados a leer y ver ideas e imágenes horribles de forma reiterada, a cambio de bajos salarios y sin contar con ningún apoyo significativo.⁷⁷ Desde entonces, los trabajadores se han sindicado⁷⁸ y han

75. Matteo Wong, «America already has an AI underclass», *The Atlantic*, 26 de julio de 2023, [lc.cx/kMwZf8](https://www.theatlantic.com/technology/archive/2023/07/ai-underclass/).

76. Davey Alba, «Google's AI chatbot is trained by humans who say they're overworked, underpaid and frustrated», *Bloomberg*, 12 de julio de 2023, [lc.cx/u8-goP](https://www.bloomberg.com/news/articles/2023-07-12/google-ai-chatbot-trained-by-overworked-humans).

77. Billy Perrigo, «OpenAI used Kenyan workers on less than \$2 per hour to make ChatGPT less toxic», *Time*, 18 de enero de 2023, [lc.cx/tlmHl_](https://www.time.com/time/technology/article/6411111/openai-used-kenyan-workers-on-less-than-2-per-hour-to-make-chatgpt-less-toxic/); Kate Linebaugh y Ryan Knutson, «The hidden workforce that helped filter violence and abuse out of ChatGPT», *The Journal*, 7 de noviembre de 2023, [lc.cx/dyxJ8a](https://www.thejournal.com.au/news/ai-chatbot-workers-1000000).

78. Billy Perrigo, «150 African workers for ChatGPT, TikTok and Facebook vote to unionize at Landmark Nairobi Meeting», *Time*, 1 de mayo de 2023, [lc.cx/YPVgU](https://www.time.com/time/technology/article/6411111/150-african-workers-for-chatgpt-tiktok-and-facebook-vote-to-unionize-at-landmark-nairobi-meeting/).

presentado una demanda ante la Asamblea Nacional de Kenia para investigar el bienestar y las condiciones laborales de quienes prestan estos servicios; estas pesquisas abordarán también si se trata de una situación de explotación, si se cumple el derecho a una remuneración justa y si el trabajo se desarrolla en unas condiciones aceptables.⁷⁹

La precariedad, el perjuicio y las dinámicas coloniales de estas prácticas laborales plantean una serie de preguntas importantes sobre el coste y las consecuencias del desarrollo de IA a gran escala en su conjunto. A pesar de que la preparación de los datos y la calibración de los modelos exigen este gran esfuerzo de mano de obra (que rara vez se publicita, pero que es imprescindible para crear el significado de los datos que definen los sistemas de IA), las empresas suelen proporcionar poca —cuando no nula— información sobre las prácticas laborales que sostienen este trabajo de datos. Asimismo, el hecho de que no se publique esta información tampoco se suele criticar como una forma de «cierre» del sistema. En gran medida, lo que

⁷⁹. Véase el *post* de Mercy Mutemi al respecto: «En nombre de los jóvenes kenianos cuya vida se ha visto destrozada por haber hecho el trabajo sucio de entrenar al algoritmo de #ChatGPT, hemos presentado una demanda ante la @NAssemblyKE para investigar cómo @OpenAI y @samasource han podido eludir las consecuencias de una explotación como esta y para regular con urgencia el trabajo tecnológico», X, T.Co/9seeYGKqFM.

sabemos es fruto del periodismo de investigación⁸⁰ o de la organización de investigadores y trabajadores.⁸¹

La mano de obra necesaria para seleccionar, preparar los datos y calibrar los sistemas se paga mal y, aun así, dado el número de trabajadores y el tiempo que hace falta para perfilar los datos con los que construir los sistemas de IA contemporáneos, representa un importe total nada desdeñable. Esto supone otra barrera para el acceso democrático y «abierto» a los recursos necesarios con los que se crean e implementan los grandes modelos de IA (aunque no se puede aceptar el término «democrático» para una estructura que depende de un trabajo precario y mal pagado, cuyo personal obtiene beneficios mínimos mientras sufre los perjuicios y queda excluido de esa supuesta democracia).

Modelos de IA

Un modelo se refiere a un sistema algorítmico «entrenado» y evaluado con grandes cantidades de datos

^{80.} Karen Hao y Andrea Paola Hernández, «How the AI industry profits from catastrophe», *MIT Technology Review*, 20 de abril de 2022, [lc.cx/7VL4lY](#); Perrigo, «OpenAI used Kenyan workers...», *op. cit.*

^{81.} Perrigo, «150 African workers for ChatGPT, TikTok and Facebook...», *op. cit.*; Lilly Irani y M. Six Silberman, «Turkopticon. Interrupting worker invisibility in Amazon Mechanical Turk», Universidad de California en San Diego, 27 de abril de 2013.

estructurados que produce resultados estadísticamente probables en respuesta a una premisa aportada. Por ejemplo, ChatGPT funciona aplicando los modelos GPT, entrenados con grandes cantidades de datos de texto, en su mayoría extraídos de internet, para ejecutar la tarea de predecir una respuesta estadísticamente plausible en forma de texto como contestación a una pregunta o indicación. Una vez entrenado, un modelo de IA puede publicarse del mismo modo que se publicaría otro código de software: en virtud de una licencia de código abierto para su reutilización, mediante una filtración o puesto a disposición en línea de otra manera. Reutilizar un modelo de IA ya entrenado no exige tener acceso al entrenamiento ni a la evaluación de datos subyacentes. En este sentido, muchos sistemas de IA que se clasifican como «abiertos» hacen un uso muy ambiguo del término. Lejos de proporcionar documentación y acceso significativos, esconden modelos cerrados, heredan datos indocumentados, no ofrecen datos de formación de RLHF con anotaciones y rara vez publican sus hallazgos; aún más excepcional es que documenten esos hallazgos en publicaciones con revisión independiente.⁸²

Hay muchos ejemplos de modelos de IA «abiertos» a gran escala disponibles para reutilización pública; algunos de ellos son el LLaMA-2 de Meta; el Falcon 40B,

⁸². Liesenfeld y Dingemans, «Rethinking open source generative AI...», *op. cit.*

desarrollado por el Technology Innovation Institute de Emiratos Árabes Unidos y entrenado en AWS; los modelos MPT de MosaicML, ahora relacionados con el modelo BLOOM de BigScience de Microsoft Azure, entrenado en la supercomputadora francesa Jean Zay y ejecutado en AWS; y Stable Diffusion de Stability AI, entrenado en AWS, además de varios modelos proporcionados por la organización sin ánimo de lucro LAION, entrenados con el hardware de Stable Diffusion que se ejecuta en AWS. Englobarlos todos bajo la única etiqueta de «abierto» hace un flaco favor a las importantes diferencias que existen entre ellos.

Varias empresas como Hugging Face y Stability AI ofrecen modelos de IA de código abierto a sus clientes y al público. Sus modelos de negocio no se basan en la concesión de licencias de modelos propios, sino en cobrar por funciones adicionales y servicios complementarios de los modelos abiertos; funciones como el acceso a la API, entrenar al modelo con datos personalizados o el soporte técnico y de seguridad se ofrecen como servicios de pago para los clientes.⁸³ También ofertan el servicio de perfeccionamiento de modelos privados para su clientela, el cual consiste en pulir y calibrar el rendimiento de modelos ya entrenados para una tarea o ámbito concretos.

⁸³. «Hugging Face», *Contrary Research*, 14 de septiembre de 2022, lc.cx/oZiu0s; Tim Smith, «Leaked deck raises questions over Stability AI's Series A pitch to investors», *Sifted*, 21 de abril de 2023, lc.cx/EFJ-wt.

La organización sin ánimo de lucro EleutherAI también ofrece modelos de IA de código abierto a gran escala, además de la documentación y las bases de código utilizadas para entrenarlos. La entidad se centra exclusivamente en promover la investigación sobre IA a gran escala y admite el uso de sus modelos en virtud de la (muy permisiva) licencia de código abierto Apache 2.0, a fin de que sean utilizados por investigadores de IA. Entre quienes conforman la escena de la IA «abierta», EleutherAI ofrece presumiblemente el sistema de IA «abierto» con parámetros más maximalistas. Funciona gracias a donaciones y patrocinios de CoreWeave, Hugging Face, Stability AI, Google TRC, Nat Friedman (antiguo director general de GitHub) y Lambda Labs.

Existen distintos proyectos académicos que también han producido grandes modelos de IA «abierto» a escalas menores. Entre ellos, el modelo de IA «abierto» Vicuna, desarrollado por un equipo de investigación de la Universidad de California en San Diego, la Universidad Carnegie Mellon y la Universidad de Berkeley; el modelo de IA abierto Koala, desarrollado en Berkeley; y el modelo Alpaca de Stanford. Este último es famoso por haber sido desarrollado en un único ordenador portátil, algo que resulta sorprendente dada la intensidad de la computación necesaria para implementar estos modelos.⁸⁴ Sin embargo, incluso un chatbot basado en este

⁸⁴ Benj Edwards, «You can now run a GPT-3-level AI model on

modelo extremadamente eficiente a nivel computacional resultaba demasiado caro —y peligroso, dadas las alucinaciones del modelo— para mantenerlo en funcionamiento, motivo por el cual fue retirado.⁸⁵

La escala cada vez mayor (de los datos y de la computación) de los modelos es el paradigma actual en el desarrollo de IA. Los modelos más grandes son más potentes, por lo que hacen un uso más intensivo de los recursos y, por tanto, resultan más difíciles de producir fuera de las grandes empresas tecnológicas. La escala se puede medir de distintas formas, pero suele evaluarse en función del número de fragmentos de datos concretos (o tókenes) del conjunto de datos de entrenamiento. Los modelos de IA de mayor escala abiertamente disponibles (LLaMA-2, MosaicML, OpenLLaMA) se entrenan con conjuntos de datos de cerca de un billón de tókenes. Otros famosos modelos de IA «abiertos» (como BLOOM o el paquete de modelos Pythia) se entrenaron con cerca de doscientos o trescientos mil millones de tókenes, comparables a los de GPT-3 de OpenAI. Ninguno de ellos puede equipararse con los modelos cerrados más grandes disponibles en el mercado: OpenAI no ha revelado

your laptop, phone, and Raspberry Pi», *Ars Technica*, 14 de marzo de 2023, [lc.cx/5xfmZ3](#).

⁸⁵. Thomas Germain, «Stanford researchers take down Alpaca AI due to “hallucinations” and rising costs», *Gizmodo*, 21 de marzo de 2023, [lc.cx/2oEV9I](#).

el tamaño de GPT-4, pero PaLM-2 de Google supuestamente cuenta con 3,6 billones de tókenes.⁸⁶

Gran parte de la actividad actual en torno a la IA «abierta» se centra en la tarea de reducir el coste computacional y de perfeccionar los modelos públicamente disponibles, una labor que llevan a cabo los pocos agentes poderosos que se pueden permitir entrenarlos desde cero. Este perfeccionamiento permite que terceros tomen modelos previamente entrenados y los corrijan para usarlos en un contexto o ámbito concretos.⁸⁷ Esto no equivale a reunir la capacidad de construir un modelo de IA desde cero y supone una diferencia fundamental, dado que el proceso de creación desde el inicio implica muchas decisiones editoriales sumamente importantes que definen el modelo de manera sustancial.

Breve historia del código abierto

La estrecha relación entre la apertura y el poder corporativo no empezó con la IA. Hacer un breve repaso de la historia del código abierto y de su relación con la

^{86.} Elias, «Google's newest A.I. model...», *op. cit.*

^{87.} Edward J. Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang y Weizhu Chen, «LoRA: low-rank adaptation of large language models», *arXiv*, 16 de octubre de 2021, [lc.cx/UzpKG_](https://arxiv.org/abs/2104.09800).

industria tecnológica dibuja un tablero aproximado que nos ayuda a entender más fácilmente los complejos incentivos actuales y los relatos a menudo contradictorios en torno a la apertura.

La Free Software Foundation fue fundada en 1985 como parte de un movimiento que aspiraba a permitir que la gente utilizara la nueva generación de ordenadores personales sin tener que depender del software patentado bajo control de las corporaciones.⁸⁸ Indudablemente, esta tendencia supuso una amenaza para las empresas tecnológicas dominantes como Microsoft, cuyo beneficio y crecimiento dependían de la venta de software. En memorandos internos filtrados de 1998, la ejecutiva de Microsoft era clara al respecto cuando describía el software de código abierto como una «amenaza para los ingresos a corto plazo y para la plataforma», y expresaba su preocupación por que «el libre intercambio de ideas» del movimiento «alterara la mentalidad de los desarrolladores a largo plazo».⁸⁹ Para hacer frente a esta amenaza percibida a sus beneficios y a su posición dominante, Microsoft desplegó «todo su arsenal de marketing para desacreditar la fiabilidad [del software de código abierto]», adoptando una retórica propagandística que nos resulta familiar y que

^{88.} Sam Williams, *Free as in freedom. Richard Stallman's crusade for free software*, O'Reilly, Sebastopol (California), 2002.

^{89.} «Open source software: a (new?) development methodology», *Microsoft*, lc.cx/F7g4QO.

definía las licencias de código abierto con «tres de las palabras más temidas en Estados Unidos: cáncer, comunismo y antiestadounidense».⁹⁰

Sin embargo, la animadversión de la industria hacia el código abierto no era total. Algunas empresas supieron beneficiarse del software que se desarrollaba en abierto y estaba disponible gratuitamente, y así fue como los esfuerzos por alinear los incentivos empresariales con las prácticas de código abierto adquirieron mayor fuerza. La Open Source Initiative se fundó en 1998 y adoptó la etiqueta de «código abierto» para distanciarse de lo que consideraba una «actitud moralizante y agresiva que se había asociado al software libre» y, en su lugar, defender los principios del software en abierto con «argumentos pragmáticos y de relevancia empresarial».⁹¹

La premisa de la relevancia empresarial resulta evidente con un ejemplo contemporáneo. En 1999, IBM invirtió mil millones de dólares estadounidenses en el sistema operativo de código abierto Linux. No se trataba de un gesto de buena voluntad ilustrada, sino que en gran medida formaba parte de una maniobra

^{90.} Enid Gabriella Coleman, *Coding freedom. The ethics and aesthetics of hacking*, Princeton University Press, Princeton, 2012.

^{91.} «History of the OSI», *Open Source Initiative*, 19 de septiembre de 2006, lc.cx/w5nxJ4; véase también «Open source case for business: advocacy», *Open Source Initiative*, 31 de marzo de 2007, lc.cx/a3MrL1.

para interferir en el dominio del mercado que ostentaba Microsoft. MarketWatch describió la inversión de IBM como un intento de «robar cuota de mercado a versiones patentadas de Unix o de Windows, propiedad de Microsoft».⁹² En este caso, la inversión corporativa en código abierto funcionó, entre otras cosas, como una forma de debilitar a la competencia.

Si nos fijamos en Google y en el sistema operativo Android, encontramos dinámicas similares. Google adquirió Android en 2005 y en 2007 puso a disposición su sistema operativo como software de código abierto. Este movimiento le otorgó a Google una serie de ventajas en el mercado móvil y le permitió posicionarse como alternativa a Apple. Al brindar el sistema operativo Android en abierto y ofrecer amplios marcos de desarrollo, incentivos y soporte, Google logró atraer el interés de numerosos desarrolladores que dedicaron su tiempo a crear y mantener aplicaciones para Android. Gracias a sus contribuciones, el sistema operativo resultaba más atractivo para el consumidor que quería utilizar estas aplicaciones, lo que promovió su adopción. A este respecto, algunos organismos de control han dejado claro que la condición de código abierto no exime a una entidad de incurrir en conductas contrarias a la competencia: el estudio de mercado del ecosistema móvil que realizó en 2022 la Autoridad de

⁹². MarketWatch, «Market dynamics. The battle for enterprise Linux», *The Free Library*, [lc.cx/QLbnwF](https://www.thefreelibrary.com/cx/QLbnwF).

Competencia y Mercados del Reino Unido (CMA, por sus siglas en inglés) concluyó que, en la práctica, existe un duopolio en el mercado móvil. A pesar de que la decisión de Google de ofrecer el sistema operativo Android en código abierto promovió la adopción y la competencia en lo que respecta a las aplicaciones, no sirvió para reducir las barreras de acceso al desarrollo del sistema operativo; asimismo, el cambio del sistema Android a iOS sigue siendo complejo.⁹³ La CMA tiene previsto desarrollar nuevas indagaciones⁹⁴ y, por su parte, la Comisión Europea ha llevado a cabo una investigación de mercado sobre Android por prácticas anticompetencia,⁹⁵ que ha concluido recientemente con una multa a la empresa de 4.300 millones de dólares estadounidenses por «restringir la competencia móvil y la elección del consumidor».⁹⁶

^{93.} Competition and Markets Authority, «Mobile ecosystems market study final report», *Gov.uk*, 10 de junio de 2022, [lc.cx/ADl0v1](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/107411/lc.cx/ADl0v1).

^{94.} Competition and Markets Authority, «Mobile ecosystems market study» y «Mobile browsers and cloud gaming», *Gov.uk*, 15 de junio de 2021, [lc.cx/f704Ji](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/107411/lc.cx/f704Ji), y 10 de junio de 2022, [lc.cx/DsWUjx](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/107411/lc.cx/DsWUjx), respectivamente.

^{95.} Comisión Europea, «Antitrust: Commission opens formal investigation against Google in relation to Android mobile operating system», 15 de abril de 2015, [lc.cx/Ui4DfQ](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/107411/lc.cx/Ui4DfQ).

^{96.} Kelvin Chan, «Google appeals huge Android antitrust fine to EU's top court», *AP*, 1 de diciembre de 2022, [lc.cx/biOi4C](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/107411/lc.cx/biOi4C).

de ganancias de 2023, el director general de Meta, Mark Zuckerberg, expuso abiertamente la lógica corporativa por la que ofrecía su popular marco de IA, PyTorch, en abierto:⁹⁹

[PyTorch] se ha estandarizado en el sector, es la herramienta que está utilizando mucha gente que desarrolla modelos de IA y otras cuestiones en este ámbito; en general, ha resultado muy útil para nosotros [...] pues, al estar integrado con nuestro stack tecnológico [o conjunto de soluciones], cuando se presenta la ocasión de hacer alguna integración con un producto, es mucho más fácil asegurarse de que los desarrolladores y las demás partes implicadas son compatibles con lo que necesitamos para el funcionamiento de nuestros sistemas.¹⁰⁰

La librería de PyTorch, que incluye herramientas de fácil acceso, funciona como un estándar *de facto* y, a su vez, permite que Meta integre más fácilmente otras innovaciones en sus productos con fines lucrativos, algo que en muchos casos hace sin pagar u ocupándose directamente de gestionar su desarrollo.

Analizados en conjunto, estos ejemplos demuestran que, aunque algunas empresas inicialmente se oponían al código abierto por considerarlo una amenaza

⁹⁹. Foster, «PyTorch vs TensorFlow...», *op. cit.*

¹⁰⁰. «Meta Q1 2023 Earnings Call Transcript», *Meta*, pp. 17-18, lc.cx/lfXKv9.

para su propia oferta, ahora optan por abrazarlo. Se trata de un mecanismo que puede permitirles afianzar su dominio, pues son ellas las que definen los estándares de desarrollo, a la vez que se benefician del trabajo gratuito de los colaboradores de código abierto. Con este repaso de la historia, comprobamos que, aunque un «régimen abierto puede generar una mayor innovación y actividad económica [...], al final los mayores beneficios los recoge un pequeño grupo de empresas».¹⁰¹

Este no es un argumento en contra del código abierto, ni pretende negar sus muchos beneficios, los cuales suelen enredarse de formas complejas con dinámicas de captura e influencia corporativa. Pero, para analizar la relación entre la IA «abierta» y la concentración corporativa a través de una lente clara, tenemos que reconocer que la «apertura» no es inherente a la democratización, especialmente en un ecosistema que cada vez se define más por la concentración del capital y de la capacidad de infraestructura.

El caso de OpenAI LP: la apertura como marketing

En 2015, OpenAI se ponía en marcha como laboratorio de investigación sin ánimo de lucro con el propósito

¹⁰¹. Duncan McCann, «Commoning intellectual property. Public funding and the creation of a knowledge commons», *Common Wealth*, 29 de julio de 2020, lc.cx/MIehly.

de «generar valor para todo el mundo y no solo para los accionistas»,¹⁰² contaba además con una política que animaba a todos sus empleados a «publicar su trabajo, ya fuera en forma de artículos, entradas de blog o como código». ¹⁰³ La empresa también anunció que compartiría con cualquier persona en cualquier lugar del mundo todas las patentes que presentara. En otras palabras, en sus comienzos, OpenAI promovía un modelo de «apertura» que aspiraba a compartir la propiedad intelectual, los hallazgos de sus investigaciones y determinada información financiera mediante las declaraciones fiscales obligatorias para entidades sin ánimo de lucro.

Incluso en esta etapa inicial, se trataba de una versión limitada de «apertura». No incluía una gobernanza abierta ni contribuciones democráticas o colaborativas, tampoco datos abiertos, ni existía un acceso abiertamente disponible a los onerosos recursos necesarios para crear modelos de IA a gran escala. Además, los fundadores de OpenAI —a excepción de Elon Musk, uno de sus primeros creadores—¹⁰⁴ han hecho declaraciones

¹⁰². «Introducing OpenAI», *OpenAI*, 11 de diciembre de 2015, [lc.cx/XQf9Qd](https://openai.com/blog).

¹⁰³. Entre los proyectos iniciales se encontraban OpenAI Gym, un kit de herramientas para desarrollar y comparar algoritmos de aprendizaje de refuerzo, y Universe, que permitía convertir cualquier programa en un entorno de Gym.

¹⁰⁴. Elon Musk: «@GRDecter OpenAI se creó como empresa de código abierto (por eso la llamé “OpenAI”), sin ánimo de lucro y para hacer de contrapeso a Google, pero ahora se ha

sistemáticamente ambivalentes sobre los riesgos asociados a la divulgación de su tecnología:

*Desarrollar toda la investigación en abierto no es necesariamente la mejor opción [...]. Produciremos un montón de código en abierto, pero también habrá muchas cosas que aún no estamos listos para publicar.*¹⁰⁵

Un ejemplo de esta ambivalencia fue el lanzamiento de GPT-2 en 2019. En febrero de ese año, la empresa anunció que había creado GPT-2, un gran modelo de lenguaje, pero lo describió como «demasiado peligroso»¹⁰⁶ y optó por no publicarlo, preocupada por que se pudiera hacer un uso indebido «para generar lenguaje

convertido en una empresa de código cerrado que busca el máximo beneficio y que está controlada por Microsoft. Nada que ver con lo que yo pretendía», [lc.cx/tqXGII](https://lccx.tqXGII).

^{105.} Greg Brockman, presidente de OpenAI, citado en Cade Metz, «Inside OpenAI, Elon Musk's wild plan to set artificial intelligence free», *Wired*, 27 de abril de 2016, [lc.cx/WUviLC](https://lccx/WUviLC).

^{106.} En aquel momento la empresa sí publicó una versión más pequeña del modelo, demostrando que asociaba la escala con la seguridad: «Hemos entrenado un modelo de lenguaje no supervisado que puede generar párrafos coherentes y desarrollar tareas básicas de comprensión lectora, traducción automática, respuesta a preguntas y resumen; todo ello sin un entrenamiento específico: blog.openai.com/better-language-models». Véase OpenAI, *X*, [T.Co/360bGgoea3](https://t.co/360bGgoea3).

engañoso, sesgado u ofensivo a gran escala».¹⁰⁷ Esta retórica también cumple una función de marketing, pues retrata unos modelos capaces y potentes (aunque subraye el peligro de ese poder). Si atendemos al verdadero modelo de negocio de IA que persiguen Microsoft y OpenAI, también encontramos contradicciones: por un lado, Greg Brockman describe estos modelos como demasiado potentes para divulgarlos con una licencia abierta pero, por otro, según pone de manifiesto su oferta comercial, ni Microsoft ni OpenAI consideran que eso sea un problema para comercializarlos mediante un contrato con la API de Microsoft Azure (en virtud del cual ahora mismo se encuentran disponibles para uso de terceros).

¹⁰⁷. «Better language models and their implications», *OpenAI*, 14 de febrero de 2019, [lc.cx/tYrpw3](https://l.c.cx/tYrpw3). Téngase en cuenta que OpenAI no era la única entidad que trabajaba con este tipo de sistemas. En respuesta a esas declaraciones, Margaret Mitchell, la entonces responsable del Equipo de Inteligencia Artificial Ética de Google, afirmó: «No podemos ignorar que durante muchos años ha habido gente increíble que ha lidiado con los problemas potenciales que el progreso inevitable traía consigo y que ha sabido anticiparse a ellos. Los expertos importantes no son solo las personas que hicieron algo, sino también quienes previeron el desencadenante evidente y optaron por no hacerlo». Según declararon Mitchell y sus colaboradores en su estudio sobre loros estocásticos (término que describe sistemas de IA que pueden generar respuestas aprendidas sin comprenderlas realmente), Google también contaba con sistemas de desarrollo similares.

El plan de OpenAI de hacer una publicación controlada fracasó rápido. En junio de 2019, Connor Leahy, estudiante de posgrado, había logrado replicar GPT-2; sin embargo, decidió no publicarlo tras una serie de debates con OpenAI e investigadores del Machine Intelligence Research Institute, institución dedicada a estudiar el «riesgo existencial» de la IA. «No se trata de GPT-2. Lo importante es que, en algún momento en el futuro, alguien *creará* algo realmente peligroso y habrá que contar con normas de seguridad consensuadas *antes* de que eso ocurra», afirmó Leahy.¹⁰⁸ En agosto, un grupo de investigadores de la Universidad Brown lo replicó también y, en esta ocasión, optaron por publicarlo; insistían en que

*...dada la relativa facilidad para replicar GPT-2, un número abrumador de partes interesadas podría hacerlo. [...] Dado que nuestro trabajo de réplica no es excepcional [...], consideramos que publicarlo es un primer paso razonable para combatir el posible abuso que se haga de este tipo de modelos en el futuro.*¹⁰⁹

^{108.} En particular, el coste de computación en la nube para entrenar un modelo del tamaño de GPT 1.5 era de cuarenta mil dólares estadounidenses. Véase OpenAI, «Better language models and their implications», *op. cit.* En una de sus publicaciones, Leahy afirma: «Me siento como si le debiera a Google mi primer hijo o algo así por la cantidad de apoyo gratuito que me han prestado».

^{109.} Vanya Cohen, «OpenGPT-2: we replicated GPT-2 because you can too», *Medium*, 22 de agosto de 2019, [lc.cx/jqiK2F](https://www.lc.cx/jqiK2F).

Más tarde, en noviembre de 2019, OpenAI publicó su propia versión del modelo, afirmando que no consideraba que hubiera «pruebas significativas» de uso indebido.¹¹⁰ Desde entonces, Ilya Sutskever, cofundador de OpenAI, ha confesado arrepentirse de esa decisión: «Nos equivocamos [...]; creo firmemente que en unos años resultará evidente para todo el mundo que ofrecer IA con código abierto no es una decisión inteligente».¹¹¹ Es posible que los acontecimientos de los años intermedios influyeran significativamente en su reticencia a la apertura: en marzo de 2019 OpenAI anuncia que abandona su carácter sin ánimo de lucro y pasa a ser OpenAI LP, una empresa de «beneficio limitado»; lo que significa que sus inversores de la primera ronda no pueden percibir rendimientos que superen cien veces su aportación inicial. En el contexto de los miles de millones de dólares de inversión en OpenAI LP, esto no representa una limitación importante (motivo por el cual a lo largo de este artículo nos referimos a OpenAI como empresa con ánimo de lucro). Para justificar esta transformación, la compañía afirmó que el cambio le permitiría «aumentar rápidamente la inversión en computación y talento, y, al mismo

¹¹⁰. «GPT-2: 1.5B release», *OpenAI*, 5 de noviembre de 2019, [lc.cx/J1TODO](#).

¹¹¹. James Vincent, «OpenAI co-founder on company's past approach to openly sharing research: "We were wrong"», *The Verge*, 15 de marzo de 2023, [lc.cx/qveUhz](#).

tiempo, llevar a cabo comprobaciones y ajustes para revisar su misión». ¹¹² En julio, OpenAI anunció que recibiría una inversión de mil millones de dólares estadounidenses de parte de Microsoft ¹¹³ a cambio de concederle una licencia exclusiva de su modelo GPT-3; así, OpenAI se incorporaba definitivamente como una división de la corporación Microsoft. ¹¹⁴

Junto con este cambio en la estructura organizativa de OpenAI, se produjo una transición de la apertura limitada y aspiracional a un cierre total. Cuando la empresa publicó su modelo GPT-4, a raíz de la integración de OpenAI en Microsoft, hubo detalles fundamentales del sistema que se ocultaron al público. En el informe técnico de GPT-4, se negaba explícitamente a divulgar detalles sobre «la arquitectura (incluido el tamaño del modelo), el hardware, la computación de entrenamiento, la construcción de los conjuntos de datos o el método de entrenamiento» del modelo. ¹¹⁵ Algunos de estos detalles se mantuvieron en secreto no solo por la preocupación que despertaban sus consecuencias

¹¹². «OpenAI LP», *OpenAI*, 11 de marzo de 2019, [lc.cx/_bNYdR](#).

¹¹³. «OpenAI forms exclusive computing partnership with Microsoft to build new Azure AI supercomputing technologies», *Microsoft News Center*, 22 de julio de 2019, [lc.cx/Dx-vPg](#).

¹¹⁴. Karen Hao, «OpenAI is giving Microsoft exclusive access to its GPT-3 language model», *MIT Technology Review*, 23 de septiembre de 2020, [lc.cx/-SqpWZ](#).

¹¹⁵. «GPT-4 technical report», *OpenAI*, *op. cit.*

para la seguridad, sino por el deseo de OpenAI de mantenerse por delante de sus competidores; algo que ilustra muy bien su evolución desde aquel mandato inicial de «generar valor para todo el mundo y no solo para los accionistas».

Los argumentos a favor y en contra de la IA «abierta» y quién los defiende

Después de haber repasado qué está (y qué no está) «abierto» en los actuales sistemas de IA «abiertos» y de haber revisado una breve historia del código abierto y el software libre, ahora ampliamos nuestra mirada para analizar cómo se utiliza la retórica de la apertura en el actual debate sobre las políticas de IA.

En la discusión en torno a la normativa, existen poderosos agentes que, en un intento por definir la trayectoria de las políticas que la regularán, recurren a la apertura. Las conversaciones y representaciones sobre IA «abierta» tienen consecuencias de primer orden que van más allá de las limitadas cláusulas que conceden las licencias de código abierto, de los mecanismos de transparencia disponibles o de su capacidad de ampliación y, hoy por hoy, se están empleando para definir el debate en torno a las políticas de IA en su conjunto. Como sucede al analizar cualquier grupo de presión e influencia, la retórica que defiende las políticas de IA «abierta» debe interpretarse a la luz de los

intereses particulares de la entidad que la plantea. Y, como en el pasado, las empresas hoy también utilizan la apertura como una varita retórica con la que ejercer presión para consolidar y reforzar sus postulados.

El debate sobre si excluir o no la «IA de código abierto» de la normativa impuesta por la Ley de IA de la Unión Europea nos ofrece un ejemplo ilustrativo. En los días previos a la votación del borrador del texto en el Parlamento Europeo, una serie de organizaciones expresaron su preocupación por que la normativa propuesta resultara demasiado onerosa para los desarrolladores de IA de código abierto. Este argumento fue liderado por la organización sin ánimo de lucro LAION, que emitió una carta abierta en la que aseguraba que una intervención regulatoria ahogaría la innovación.¹¹⁶ Entre las exigencias a las que se oponía estaba la obligación de realizar una evaluación de riesgos de los sistemas que se implementaran en contextos considerados «de alto riesgo», así como la de garantizar una adecuada documentación y trazabilidad de la toma de decisiones automatizada de los modelos de IA (algo que, para empezar, se corresponde

¹¹⁶. Kyle Wiggers, «This startup is setting a DALL-E 2-like AI free, consequences be damned», *TechCrunch*, 12 de agosto de 2022, [lc.cx/yAHdL1](https://l.cx/yAHdL1). La propia LAION cuenta con financiación de Hugging Face y Stability AI, dos empresas con ánimo de lucro cuyo modelo de negocio gira en torno a la IA «abierta».

con la definición de «transparencia y capacidad de reutilización» de la «apertura».¹¹⁷

Aunque LAION desempeñó un papel fundamental, las presiones más importantes procedían de la industria. En su alegato ante la Unión Europea sobre la Ley de IA, Google garantizó que la norma tendría «consecuencias espeluznantes para la colaboración abierta en el ecosistema de la IA».¹¹⁸ El director general de GitHub, propiedad de Microsoft, afirmó que la IA de código abierto estaba creando un «torrente de innovación por todo el mundo y aquí en Europa»;¹¹⁹ por su parte, la Business Software Alliance —un grupo sectorial que representa, entre otras, a Cisco, IBM y Oracle— argumentó que ajustarse a la ley «tendría una grave repercusión y desincentivaría el desarrollo de software de código abierto e IA en Europa».¹²⁰

¿Por qué los representantes industriales se mostraron tan activos en su intento por excluir la IA de código abierto del alcance de la Ley de IA de la Unión Europea?

Como hemos visto, muchas empresas grandes y afianzadas como Google, Microsoft, Meta y varias más tienen

117. «Regulatory framework proposal on artificial intelligence. Shaping Europe's digital future», *Comisión Europea*, [lc.cx/NtEcAJ](#).

118. «Submission to the European Commission on the EU AI Act», *Google*, [lc.cx/wGOnAQ](#).

119. Karabus, «GitHub CEO says...», *op. cit.*

120. BSA Foundation, «BSA leads joint industry statement...», *op. cit.*

intereses adquiridos en el desarrollo de la IA «abierta». El autor del memorando filtrado sobre la ventaja competitiva de Google lo puso en evidencia: «El valor que aporta [a la empresa] que el ecosistema sea suyo no se puede sobrestimar»; controlar ese ecosistema ha contribuido directamente a que Google domine los distintos ámbitos de búsqueda, móvil y publicidad.¹²¹ De igual manera, como plataforma en la que se hospeda gran parte del código abierto (un código que se ha extraído y utilizado para entrenar modelos como Copilot de Microsoft), GitHub, también de Microsoft, tiene sus propios intereses en el desarrollo de código abierto sin restricciones; por su parte, OpenAI, a pesar de estar financiada por Microsoft, se mantiene mucho más cautelosa a la hora de impulsar la concesión de licencias (lo que permite que Microsoft tácitamente respalde opiniones contradictorias que, en última instancia, le benefician igualmente). Con el objetivo de proteger sus intereses, muchas empresas presionan para conseguir excepciones en la documentación básica y en los mecanismos de rendición de cuentas.

Con esta visión de conjunto, pasamos ahora a analizar los postulados que subyacen a los argumentos planteados por los grupos de presión sobre la IA de código abierto.

¹²¹. Dylan Patel y Afzal Ahmad, «Google “We have no moat, and neither does OpenAI”. Leaked internal Google document claims open source AI will outcompete Google and OpenAI», *Semi-analysis*, 4 de mayo de 2023, lc.cx/yC9ur1.

«La IA abierta genera seguridad gracias a la transparencia»

La carta abierta de LAION defiende que la IA de código abierto promueve la seguridad permitiendo que investigadores y autoridades auditen el rendimiento del modelo, identifiquen los riesgos y apliquen soluciones o medidas correctivas. En general, estamos de acuerdo con esta afirmación. Sin embargo, los recursos y el acceso necesarios para llevar a cabo esas auditorías y evaluaciones son prácticamente los mismos que los proyectos de IA «abierta» se verían obligados a producir y aplicar en virtud de la Ley de IA; esto hace que la negativa a incluir estos modelos en dicha ley resulte confusa, si asumimos que la seguridad es el objetivo principal.

La eficacia de las auditorías como medida de seguridad depende en gran medida de la disponibilidad de recursos relevantes y de que existan incentivos para que esas valiosas auditorías efectivamente se lleven a cabo y tengan la firmeza necesaria para explicar los verdaderos riesgos que plantea el desarrollo de modelos de IA. El mero hecho de que el código sea abierto no garantiza que expertos voluntarios vayan a dedicar su tiempo y sus recursos a revisarlo a conciencia. El caso del agujero de seguridad Heartbleed¹²² nos sirve para demostrar que el hecho de que un código sea

¹²² «A call to protect open-source AI in Europe», *Laion*, 28 de abril de 2023, lc.cx/_tyBdE.

abiertamente analizable no significa que se audite ni que esté libre de errores (incluso cuando se trata de código de una importancia crítica como OpenSSL, el protocolo de encriptado que hizo posible el comercio en internet).¹²³ En caso de explotarse, el agujero de seguridad Heartbleed¹²⁴ permitiría extraer información transmitida a través de la red, como contraseñas personales y otro tipo de información privada. Este agujero de seguridad no se detectó antes debido, en parte, a la insuficiente dotación de recursos para auditar el código abierto: OpenSSL funcionaba con un presupuesto de dos mil dólares estadounidenses al año en donaciones, a pesar de que se utilizaba para asegurar

¹²³. «Heartbleed bug», *Synopsys*, heartbleed.com.

¹²⁴. Según la entrada de Wikipedia dedicada al tema: «En 2011, uno de los autores del RFC, Robin Seggelmann, en ese tiempo estudiante de doctorado en la Universidad de Duisburgo-Essen, implementó la extensión Heartbeat para OpenSSL. Siguiendo la petición de introducir su trabajo en OpenSSL, su cambio fue revisado por Stephen N. Henson, uno de los cuatro desarrolladores del núcleo de OpenSSL. Henson aparentemente no se dio cuenta del error en la implementación de Seggelmann e introdujo el código con fallo en el repositorio de OpenSSL el 31 de diciembre de 2011. El código vulnerable fue adoptado y usado ampliamente con el lanzamiento de la versión 1.0.1 de OpenSSL el 14 de marzo de 2012. El soporte de Heartbeat estaba habilitado de forma predeterminada, provocando que las versiones afectadas fueran vulnerables por defecto», lc.cx/RLoDw6.

dos tercios de las páginas web del mundo.¹²⁵ A raíz de la revelación del error, Amazon, Google y Microsoft, entre otras, desarrollaron la Core Infrastructure Initiative, a la que aportaron cada una trescientos mil dólares estadounidenses para conservar el código de OpenSSL,¹²⁶ una cifra que sigue siendo una miseria comparada con los recursos que se dedican al desarrollo corporativo en estas empresas.

Lo que Heartbleed pone en evidencia es que solo porque un código *pueda* ser auditado no significa que *lo vaya a ser*. Técnicamente no basta con un código y una documentación «abiertas» para garantizar una revisión y una reparación meticolosas; y esta dinámica es aún más grave en el caso de la IA, en la que grandes modelos de probabilidad funcionan produciendo resultados y decisiones que no se pueden predecir a partir del código, la documentación y los datos únicamente. Así, una auditoría que estudiara el impacto y las implicaciones de los sistemas de IA implementados en contextos complejos necesitará muchos más recursos y permisos que ofrecer un código abierto y unos archivos README. No disponer de incentivos ni de procesos claros y bien definidos hace menos probable que

¹²⁵ Nichole Perlroth, «Heartbleed highlights a contradiction in the web», *The New York Times*, 18 de abril de 2014, [lc.cx/9fTA3W](#).

¹²⁶ Reuters, «Big tech companies offer millions to help with Heartbleed crisis», *Reuters*, 24 de abril de 2014, [lc.cx/mOgqR2](#).

esas auditorías se lleven a cabo. De igual manera, el argumento de que la IA de código abierto lograría una mayor seguridad también debe abordar cómo se efectuarían la auditoría, el mantenimiento y la inspección, y no ceñirse al hecho de que los distintos atributos de transparencia de la IA de código abierto *podrían* permitir dicho análisis; también tendría que ocuparse de definir en quién recae la responsabilidad cuando no se solucionan los errores o los agujeros de seguridad.

«*La IA abierta aumenta la inseguridad*»

Los argumentos en la dirección opuesta sitúan a la IA de código abierto como una fuente de profunda *inseguridad*, pues permite que una potente tecnología esté ampliamente disponible para ser reutilizada y potencialmente quede en manos de usuarios con malas intenciones. Como plantean los investigadores Bruce Schneier y Jim Waldo (que sin embargo defienden la IA de código abierto): «Que una tecnología sea de código abierto significa que quienes quieran utilizarla para objetivos no contemplados, ilegales o infames puedan acceder a ella igual que cualquier otra persona».¹²⁷

Supuestamente este fue el temor que motivó la decisión inicial de OpenAI de no publicar GPT-2 como código abierto, un miedo que sigue reflejándose en

¹²⁷. Schneier y Waldo, «Big tech isn't prepared for A.I.'s next chapter», *op. cit.*

declaraciones de responsables de OpenAI como Ilya Sutskever:

*Estos modelos son muy potentes y lo van a ser cada vez más. En algún momento será muy fácil, si alguien así lo quisiera, causar grandes perjuicios con ellos. A medida que las capacidades aumentan, tiene sentido que se opte por no ponerlos a disposición.*¹²⁸

A pesar de esto, en mayo de 2023 vimos como OpenAI se posicionaba en contra de «engorrosos mecanismos como licencias y auditorías» para los proyectos actuales de IA corporativa y de código abierto que, según afirmaba la empresa, «crean un inmenso valor para el mundo» (como su GPT-4); no obstante, la compañía es partidaria de que este tipo de normativa se aplique a proyectos de IA con capacidades muy superiores a las actuales. Este planteamiento utiliza la regulación como un foso: se opone a aplicarla a su modelo predominante, pero la defiende para cualquier otra iniciativa significativamente más potente que pudiera desafiarlo.

Merece la pena señalar que hemos descrito dos casos de grupos de presión vinculados a dos entidades estrechamente ligadas a Microsoft, pero que parecen avanzar en sentidos opuestos. El argumento de GitHub,

¹²⁸ Vincent, «OpenAI co-founder on company's past...», *op. cit.*

presentado en la primera sección, redundante en su propio interés, pues la empresa (y Microsoft) dependen del desarrollo de código abierto: es el modelo de negocio de la plataforma GitHub y la fuente de entrenamiento de datos para sistemas rentables como Copilot. Les interesa este punto de vista. Por su parte, el argumento principal de OpenAI es que los modelos «por encima de determinado umbral» no deberían estar abiertos; un umbral que, en la práctica, definen ellos mismos dado el monopolio de recursos del cual se beneficia Microsoft. Así, eximir de regulación al código abierto resulta beneficioso para ellos. Por otro lado, la premisa de que ofrecer sus potentes modelos en código abierto resultaría peligroso también beneficia a OpenAI, pues una afirmación como esta reafirma la potencia de sus modelos y les permite equiparar la concentración de recursos con el desarrollo científico de vanguardia.

La preocupación por cuestiones de seguridad ocupaba un papel destacado en la carta que los senadores Blumenthal y Hawley enviaron a Meta. En ella cuestionaban a la empresa por su toma de decisiones previa a la publicación del primer modelo de LLaMA, poniendo de manifiesto un escepticismo evidente ante la decisión de divulgarlo:

La determinación de Meta de distribuir LLaMA con tan pocas restricciones y de manera tan permisiva despierta importantes y complejas preguntas

*sobre cuándo y cómo resulta apropiado publicar abiertamente sofisticados modelos de IA. Ante las aparentemente mínimas protecciones incorporadas a la publicación de LLaMA, Meta debería haber tenido en cuenta que LLaMA se divulgaría ampliamente y tendría que haber anticipado el potencial de uso indebido. Sin embargo, Meta se ha referido a la publicación como una filtración [véase recuadro de las páginas 158-159]; el científico responsable de IA de la compañía ha indicado que los modelos abiertos son fundamentales para su éxito comercial.*¹²⁹

La preocupación en torno a la inseguridad, surgida a partir de la proliferación de modelos de IA de código abierto, se justifica, entre otras cosas, porque estos modelos *permiten que la IA se ajuste a pequeña escala* sin una curva de aprendizaje demasiado pronunciada. Lo que sigue pasando inadvertido y resulta poco claro de este argumento, sobre todo si tenemos en cuenta quiénes lo defienden, es por qué el acceso a esos mismos modelos (o a otros igual de potentes) a través de un contrato en la nube con Microsoft o con Google —que es el patrón actual— resulta menos peligroso que reutilizar un modelo de IA publicado abiertamente. Además, suponiendo que los modelos de IA de

¹²⁹ Richard Blumenthal y Josh Hawley, «Letter to Meta», 6 de junio de 2023, [lc.cx/4lb8Sf](https://www.congress.gov/records/117/16444).

acceso abierto sean efectivamente más peligrosos, eso no significa que las ventajas compensatorias de un mercado más concentrado ofrezcan el mejor camino a seguir: la concentración de la industria genera una competición tóxica entre empresas de IA, lo que las lleva a comercializar modelos antes de que estén listos y sin haberlos sometido al escrutinio o la mitigación de riesgos necesarios.¹³⁰ Un pequeño número de grandes jugadores no garantiza necesariamente una IA más segura.

«La IA abierta reducirá la concentración del sector»

Otra serie de reivindicaciones defienden que la regulación incorporará requisitos demasiado onerosos para que los desarrolladores de código abierto puedan cumplirlos, lo que contribuirá a la concentración de IA. Por ejemplo, una de las propuestas de Sam Altman en su declaración ante el Congreso fue la de crear una suerte de régimen de licencias para los sistemas de IA generativa. Serían distintas de las licencias que dan derecho a uso y que suelen acompañar al desarrollo de software de código abierto. Un enfoque como este probablemente instauraría una entidad de control que garantizaría cierto nivel de escrutinio de los sistemas o de las empresas, antes de que los modelos de IA se pusieran a disposición para uso comercial.

¹³⁰. *Ibid.*

La filtración de LLaMA-2 y el perfeccionamiento de los modelos de IA de código abierto

LLaMA, el modelo de Meta, se filtró al público después de que la empresa lo compartiera con un número limitado de investigadores y hoy constituye la base de muchos proyectos de perfeccionamiento que se desarrollan a partir de la arquitectura de Meta, según se detalla a continuación.

Muchos proyectos de IA «abierta» se desarrollan a partir del modelo de lenguaje de Meta: LLaMA. Se trata de un modelo de lenguaje que la empresa facilitó como un paquete de código abierto, inicialmente destinado en exclusiva a aquellos investigadores que se habían puesto en contacto con la organización para solicitar acceso a él. De manera previsible, una semana después de su lanzamiento se publicó un *torrent* descargable de LLaMA en 4chan y rápidamente se puso a disposición para su uso generalizado. LLaMA no es precisamente de uso fácil para terceros, pero al tratarse de un modelo preentrenado —en el que ya se han invertido un trabajo y un

gasto considerables para definir los datos y llevar a cabo el entrenamiento—, permite que desarrolladores de IA reduzcan las barreras de computación y lo puedan utilizar como punto de partida para sus desarrollos.

La filtración de LLaMA también beneficia a Meta, la empresa creadora. Un memorando filtrado de un ingeniero de Google explica el porqué: «Paradójicamente, quien sale ganando con todo esto es Meta. Dado que el modelo filtrado era suyo, en realidad han obtenido la mano de obra gratuita de todo un planeta. *Puesto que la mayoría de la innovación en código abierto se desarrolla a partir de una arquitectura que es suya, nada les impide incorporarla directamente a sus productos*» [énfasis añadido]. El propio Mark Zuckerberg ha hecho declaraciones similares en la última presentación de resultados de Meta, que analizamos más adelante. Por todo ello, cuando Meta publicó posteriormente LLaMA-2, lo hizo en virtud de una licencia que permitía retocar el modelo para uso comercial, algo que tiene sentido desde un punto de vista empresarial.

Implementar LLaMA-2 (o cualquier otro modelo de IA de gran tamaño) para uso público —esté ajustado o no— también comporta una utilización intensiva de recursos, pues exige un gran despliegue de computación que casi siempre supone alquilar infraestructura en la nube de una o varias de las grandes tecnológicas.

En un artículo reciente, Sayash Kapoor y Arvind Narayanan afirmaban que exigir que los proveedores de IA obtengan licencias de este tipo agravaría los riesgos de la IA en su conjunto, pues promovería un monocultivo y una homogeneización de los resultados. Es decir, que un régimen de licencias como ese haría que solo se desarrollaran cierto tipo de modelos que encajaran en dicha normativa. En contraposición, proponen la IA de código abierto como alternativa viable para garantizar una mayor diversidad en el desarrollo de IA.¹³¹ Sin perder de vista que la dependencia de escasos modelos de propiedad corporativa no deja de ser un problema acuciante, la homogeneización es algo que ya caracteriza el actual entorno de IA. Este fenómeno es fruto de los elevados niveles de concentración y captura que hoy se dan en el ámbito académico de la inteligencia artificial (y de la infraestructura necesaria para desarrollarla);¹³² no se trata de un futuro desenlace provocado por la intervención reguladora. Como se destacaba anteriormente, la IA abierta existe dentro de ese entorno, no fuera de él: según ilustra un reciente estudio sobre los agujeros de seguridad de la IA de código abierto, la proliferación de unos pocos modelos de IA

^{131.} Sayash Kapoor y Arvind Narayanan, «Licensing is neither feasible nor effective for addressing AI risks», *AI Snake Oil*, 10 de junio de 2023, [lc.cx/bk7Sb9](https://lcamtuf.com/post/bk7Sb9).

^{132.} Véase el primer capítulo de este libro: «El elevado coste de la captura», p. 17.

de código abierto generaría los mismos problemas de punto único de fallo que una fuerte dependencia de unos pocos modelos cerrados fabricados por grandes tecnológicas.¹³³

La carta de LAION representa una versión modificada del argumento de Kapoor y Narayanan: reivindica que la IA de código abierto permitirá la entrada de pequeños negocios, que podrán desarrollar los modelos existentes perfeccionándolos en lugar de tener que adquirir el acceso a modelos creados por grandes firmas.¹³⁴ Lo que este argumento pasa por alto son los muchos elementos críticos presentes en los modelos originales a gran escala y en las bibliotecas de código abierto: los productos perfeccionados resultantes suelen convertirse en un lastre para las grandes tecnológicas, no en alternativas viables. Por lo general, estas iniciativas siguen dependiendo de la infraestructura de las grandes tecnológicas, a las que ceden la capacidad de definir y crear las lógicas centrales del modelo, pues son las grandes empresas quienes cuentan con los recursos para crearlos desde cero.

El rechazo del informático Zeerak Talat a la carta de LAION señala, además, que lejos de reducir las

¹³³. Tianyu Gu, Brendan Dolan-Gavitt y Siddharth Garg, «Bad-Nets. Identifying vulnerabilities in the machine learning model supply chain», *arXiv*, 11 de marzo de 2019, [lc.cx/RXX6MT](https://arxiv.org/abs/1903.01333).

¹³⁴. Laion.ai, «A call to protect open-source AI in Europe», *op. cit.*

Los numerosos caminos para el control y el beneficio corporativo

Breves estudios de caso

Las empresas han controlado o utilizado los proyectos de código abierto en beneficio propio de distintas formas:

1. *Invertir en código abierto para atacar a tus competidores: IBM y Linux.* En 1999, IBM invirtió mil millones de dólares en el sistema operativo de código abierto Linux para frenar el dominio de mercado de Microsoft.
2. *Liberar código abierto para controlar una plataforma: Google y Android.* En 2007, Google realizó una importante inversión en el sistema operativo Android y lo puso a disposición en código abierto, lo que le permitió alcanzar gran trascendencia como sistema operativo móvil frente a su competidor Apple; aquello atrajo también el escrutinio de organismos reguladores por prácticas anticompetencia.
3. *Reimplementarlo y venderlo en forma de software como servicio: Amazon y MongoDB.* En 2019, Amazon implementó su propia versión de la famosa base de datos de código abierto MongoDB y después la vendió como servicio en su plataforma AWS.
4. *Desarrollar un marco de código abierto que permita que la empresa integre productos de código abierto en su sistema: Meta y PyTorch.* El director general de Meta, Mark Zuckerberg, ha explicado que ofrecer el marco PyTorch en código abierto les permitió capitalizar de forma más sencilla y gratuita nuevas ideas desarrolladas externamente.

obligaciones de las pequeñas empresas, las exenciones del código abierto podrían en realidad *aumentarlas*. Eximir a los modelos de código abierto del escrutinio normativo podría crear una presión competitiva excesiva a las empresas más pequeñas, en caso de que los requisitos de cumplimiento recayeran sobre la entidad que perfecciona un producto final, en lugar de hacerlo sobre el agente de mayor tamaño que desarrolló el modelo o la biblioteca original a gran escala (véase recuadro de la página 162).¹³⁵

«La IA abierta es clave para la innovación»

La última serie de argumentos defiende las excepciones de la regulación basándose en que la IA de código abierto es fundamental para la innovación. Por ejemplo, Derek Slater y Betsy Masiello insisten en que «la IA de código abierto impulsa la innovación en potentes modelos de IA», aunque matizan esta reivindicación apuntando que la innovación vendrá en forma de productos de consumo basados en grandes modelos de lenguaje (la mayoría de los cuales han sido desarrollados por gigantes tecnológicos o mediante proyectos que utilizaron la infraestructura en la nube de estas grandes empresas).¹³⁶ En declaraciones más prudentes,

^{135.} Zeerak Talat, «Response to the LAION letter», *Medium*, 17 de mayo de 2023, [lc.cx/GzRn1f](https://www.zeerak.com/cx/GzRn1f).

^{136.} Derek Slater y Betsy Masiello, «Will open source AI shift

Alex Engler ha apuntado que la IA de código abierto puede acelerar su adopción al reducir el nivel de conocimientos técnicos necesarios para implementarla.¹³⁷ En una línea similar, Schneier y Waldo señalan que «en lugar de necesitar decenas de miles de máquinas y millones de dólares para entrenar un modelo nuevo, ahora se puede *personalizar* uno ya existente con un portátil de precio medio en unas horas».¹³⁸

Estas afirmaciones son ciertas. Pero también realizan una interpretación muy limitada de la innovación; es una lectura que no cuestiona en profundidad las estructuras de incentivos, los modelos de negocio ni las dinámicas que dictan «quién usa» los sistemas de IA ni «a costa de quién se usan». Tampoco altera de modo significativo quién se beneficia del desarrollo de los sistemas de IA y quién se arriesga a sufrir los perjuicios.

En lugar de ofrecer una verdadera alternativa a una IA controlada por un puñado de grandes empresas, el locus fundamental del desarrollo y la «innovación» de la IA de código abierto está en el perfeccionamiento de modelos desarrollados por grandes empresas tecnológicas, lo que a su vez exige contratos de computación continuados para ejecutar inferencias (es decir, para usarlos en la

power from “big tech”? It depends», *Tech Policy*, 16 de junio de 2023, [lc.cx/8pzkKD](https://www.lc.cx/8pzkKD).

137. Engler, «How open-source software shapes AI policy», *op. cit.*

138. Schneier y Waldo, «Big tech isn't prepared for A.I.'s next chapter», *op. cit.*

vida real) a una escala que resulte relevante. Mientras tanto, la obsesión por perfeccionar modelos entrenados por empresas y ampliarlos a distintos ámbitos también funciona como desarrollo gratuito de un producto para esas mismas empresas, quienes, como hemos visto, tienen un largo historial de explotación y captura de los proyectos de código abierto más exitosos.¹³⁹ Esta dinámica se ve agravada por el actual control corporativo y su capacidad para definir la dirección de los principales marcos de desarrollo de IA que, como hemos analizado, se dedican a divulgar elementos imprescindibles para su construcción, de tal forma que garantizan que los modelos e implementaciones resultantes sean compatibles con los sistemas propiedad de las grandes empresas.

Otras reivindicaciones en torno a la innovación impulsada por el código abierto tienen aún menos matices. Las demandas de Thibault Schrepel y Alex Pentland en un artículo reciente ofrecen un claro ejemplo, a saber, que la IA «de código y acceso abiertos» quede exenta de escrutinio antimonopolio: «Si los agentes de código abierto y acceso abierto pudieran compartir los costes iniciales, las redes de comercialización y los conocimientos técnicos —exponen—, se encontrarían en una posición más fuerte para competir contra los sistemas patentados que capturan más

¹³⁹. Véase, por ejemplo, Amazon bifurcando MongoDB y colocándolo bajo su marca blanca DocumentDB.

valor del que crean».¹⁴⁰ La combinación de «código y acceso abiertos» que aquí se plantea podría referirse a cualquier modelo de IA que se ofreciera a través de una API; un marco que engloba casi la totalidad del mercado de los grandes modelos de lenguaje.

Por último, como se desprende del análisis anterior, la «apertura» suele permitir la explotación sistemática del trabajo de desarrolladores y creadores, mientras conserva el dominio de la infraestructura y el ecosistema en manos de las empresas más grandes.¹⁴¹ En un contexto en el que los ingredientes necesarios para desarrollar sistemas de IA conocen altos niveles de concentración y control corporativos, la IA «abierta», tal y como se concibe en este momento, está predestinada a ser capturada por las empresas.

Conclusión

Incluso en sus instancias más maximalistas, en las que los sistemas de IA «abiertos» ofrecen robusta

^{140.} Thibault Schrepel y Alex Pentland, «Competition between AI foundation models. Dynamics and policy recommendations», *MIT Connection Science Working Paper*, 10 de julio de 2023, [lc.cx/Oms7iS](https://ic.cx/Oms7iS).

^{141.} Véanse, por ejemplo, el post de Paul Azunre citado anteriormente ([lc.cx/GA5yFM](https://ic.cx/GA5yFM)), o Xiang, «A photographer tried to get his photos removed...», *op. cit.*

transparencia, capacidad de reutilización y de desarrollo, estas prestaciones no garantizan por sí mismas un acceso democrático ni una competencia significativa en el ámbito de la IA. Tampoco la apertura por sí sola resuelve el problema de la supervisión y el escrutinio. Aun así, las empresas poderosas se aprovechan de la retórica y la promesa de apertura de estos sistemas para afianzar su posición ante el creciente interés por la regulación del sector. Algunas empresas abrazan la IA «abierta» como mecanismo para apuntalar su posición dominante: invierten en ella para poder definir los estándares de desarrollo mientras se benefician del trabajo gratuito de los colaboradores de código abierto.

La clase política debe afrontar la tarea de regular la IA habiendo comprendido con claridad las muchas cosas que la IA es, y las que no es, con una conciencia firme de lo que la IA «abierta» puede y no puede ofrecer. Esto generará una imagen profundamente distinta de la retórica actual sobre su alcance. También exigirá prestar atención a las importantes diferencias entre el software de código abierto y la IA «abierta»; y reconocer que los procesos de desarrollo, la necesidad de recursos y la centralización inherente a la IA implican que esta no puede ser descrita fácilmente en los términos que se acuñaron en su día para promover y definir el software de código abierto. E igual que la IA es distinta del software tradicional, también los primeros años de la década de 2020 son distintos de los idílicos días de finales de la década de 1990, cuando

el software de código abierto cimentaba sus raíces ideológicas en medio de un frenesí por la informática comercializada en red.

Aunque habría que abordar las preocupaciones sobre seguridad y concentración de mercado, ya sea a través o a pesar de la IA de código abierto, deberían entenderse como parte de un escenario mucho mayor. En todo caso, el propio aumento de los ya elevados niveles de concentración de mercado crea perjuicios de seguridad, como la presencia de puntos únicos de fallo mediante los cuales puede diseminarse un riesgo sistémico.¹⁴² Y existen muchos otros perjuicios que van más allá de consideraciones de seguridad y que deben tenerse en cuenta al pensar en el desarrollo de grandes sistemas predictivos en entornos sociales y políticos delicados.

De igual manera, la «apertura» en IA no conduce *directamente* a la innovación, aunque haya quien afirme lo contrario. No cabe duda de que no es así cuando el telón de fondo de la investigación en IA «abierta» se define por la importante presencia de la industria, la onerosa infraestructura de propiedad corporativa que resulta necesaria para crear nuevos sistemas de IA y el dominio empresarial del ámbito de la investigación

¹⁴² FTC Office of Technology, «An inquiry into cloud computing business practices. The Federal Trade Commission is seeking public comments», Federal Trade Commission, 22 de marzo de 2023, [lc.cx/xfqxll](https://www.ftc.gov/ftc/cx/xfqxll).

académica.¹⁴³ Las pruebas señalan que estos factores ya contribuyen al desarrollo de un monocultivo de IA,¹⁴⁴ pues promueven exclusivamente la creación de los sistemas de IA que más le interesan a la industria y de los que tiene más posibilidades de beneficiarse.¹⁴⁵

Desarrollar alternativas significativas a la tecnología definida y dominada por las grandes empresas monopolizadoras es y sigue siendo un objetivo loable; un propósito que exige la creación de alternativas a la infraestructura dominada por las empresas. Es una necesidad especialmente acuciante, ya que los sistemas de IA están integrados en muchos entornos muy delicados con especial impacto público: en la sanidad, las finanzas, la educación o los lugares de trabajo, la IA provoca efectos difusos y profundos que no deberían estar definidos por un puñado de empresas con ánimo de lucro; efectos que no se pueden comprender con un mero análisis del sistema de código y de la documentación en el vacío. Para construir alternativas verdaderamente relevantes, no bastará con perseguir el desarrollo de IA «abierta» según los parámetros más

¹⁴³. Véase el primer capítulo de este libro: «El elevado coste de la captura», p. 17.

¹⁴⁴. Nur Ahmed, Muntasir Wahed y Neil C. Thompson, «The growing influence of industry in AI Research», *Science*, 2 de marzo de 2023, [lc.cx/ocIh9P](https://doi.org/10.1126/science.1258888).

¹⁴⁵. Sara Hooker, «The hardware lottery», *arXiv*, 21 de septiembre de 2020, [lc.cx/aAP8m2](https://arxiv.org/abs/2009.08811).

maximalistas. Necesitamos una mirada más amplia de su desarrollo y una mayor diversidad de métodos, además de la construcción de una IA y de otros sistemas que resuelvan las necesidades del público de forma más significativa, no atendiendo a intereses comerciales. Crear las condiciones para hacer que estas alternativas sean posibles es un proyecto que puede convivir con la regulación, e incluso verse respaldado por ella.¹⁴⁶ Pero depositar nuestras esperanzas en la IA de código abierto de forma aislada no nos llevará hasta ese universo y, en muchos aspectos, podría hacer que las cosas empeoren aún más.

¹⁴⁶ Véanse, por ejemplo, las inversiones de la Agencia de Protección Ambiental de Estados Unidos en proyectos científicos de financiación participativa que incluyen el desarrollo de becas para organizaciones científicas comunitarias y ciudadanas: US EPA, «Funding opportunities for participatory science projects», *EPA.gov*, [lc.cx/chmiwF](https://www.epa.gov/lc/cx/chmiwF).

Bibliografía

- AHMED, Nur, Muntasir WAHED y Neil C. THOMPSON, «The Growing Influence of Industry in AI Research», *Science*, 2 de marzo de 2023, [lc.cx/ocIh9P](#).
- AI NOW INSTITUTE y DATA & SOCIETY RESEARCH INSTITUTE, «Democratize AI? How the national AI research resource falls short», *AINow*, 5 de octubre de 2021, [lc.cx/et-aAg](#).
- AI NOW INSTITUTE, «What is AI? Part 2, with Lucy Suchman», entrevista, [lc.cx/lgdKB6](#).
- ALBA, Davey, «Google's AI chatbot is trained by humans who say they're overworked, underpaid and frustrated», *Bloomberg*, 12 de julio de 2023, [lc.cx/u8-goP](#).
- AZUNRE, Paul, «If African AI/ML researchers are not careful, this new "Open Source" movement championed by the richest global tech companies will become a mechanism for continued exploitation of our human capital and continent... [1/3]», *X*, 13 de septiembre de 2022, [lc.cx/GA5yFM](#).
- BABBAGE, Charles, *On the economy of machinery and manufactures*, Augustus M. Kelley, Nueva York, 1963 [1832].
- BARROW, John, «The foreign slave-trade», *London Quarterly Review*, vol. 55, n.º 109, diciembre de 1835.

- BENDER, Emily M., Timnit GEBRU, Angelina McMILLAN-MAJOR y Shmargaret SHMITCHELL, «On the dangers of stochastic parrots. Can language models be too big?», actas en *Proceedings of the 2021 ACM Conference on Fairness, Accountability and Transparency*, Association for Computing Machinery, Nueva York, 2021, pp. 610-623, lc.cx/N16AyA.
- BIRHANE, Abeba, Vinay Uday Prabhu y Emmanuel Kahembwe, «Multimodal datasets. Misogyny, pornography and malignant stereotypes», *arXiv*, 5 de octubre de 2021, lc.cx/ES-JSH.
- BLUMENTHAL, Richard y Josh HAWLEY, «Letter to Meta», 6 de junio de 2023, lc.cx/4lb8Sf.
- BOMMASANI, Rishi, Drew A. HUDSON, Ehsan ADELI *et al. et al.*, «On the opportunities and risks of foundation models», *arXiv*, 2021, lc.cx/XaRcUZ.
- BRAVERMAN, Harry, *Labor and monopoly capital. The degradation of work in the twentieth century*, Monthly Review Press, Nueva York, 1998 [Hay trad. cast.: *Trabajo y capital monopolista. La degradación del trabajo en el siglo XX*, Ciudad de México, Editorial Nuestro Tiempo, 1981].
- BROWNE, Grace, «AI is steeped in big tech’s “digital colonialism”», *Wired*, 25 de mayo de 2023, lc.cx/Hzoh1X.
- BROWNE, Simone, *Dark matters. On the surveillance of blackness*, Duke University Press, Durham, 2015.
- BSA FOUNDATION, «BSA leads joint industry statement on the EU Artificial Intelligence Act and high-risk obligations for general purpose AI», *Business Software Alliance*, 27 de septiembre de 2022, lc.cx/bH9vC7.
- CHAN, Kelvin, «Google appeals huge Android antitrust fine to EU’s top court», *AP*, 1 de diciembre de 2022, lc.cx/biOi4C.
- CLABURN, Thomas, «Fed up with cloud giants ripping off its database, MongoDB forks new “open-source license”» *The Register*, 16 de octubre de 2018, lc.cx/kx5bw2.

- CNRS, «Jean Zay, France's most powerful supercomputer for research», [lc.cx/fcl9m_](#).
- CLEGG, Nick, «Openness on AI is the way forward for tech», *Financial Times*, 11 de julio de 2023, [lc.cx/S8ILdt](#).
- COHEN, Vanya, «OpenGPT-2: we replicated GPT-2 because you can too», *Medium*, 22 de agosto de 2019, [lc.cx/jqiK2F](#).
- COLEMAN, E. Gabriella, *Coding freedom. The ethics and aesthetics of hacking*, Princeton University Press, Princeton, 2012.
- COMISIÓN EUROPEA, «Antitrust: Commission opens formal investigation against Google in relation to Android mobile operating system», 15 de abril de 2015, [lc.cx/Ui4DfQ](#).
—«Regulatory framework proposal on artificial intelligence. Shaping Europe's digital future», [lc.cx/5kbGBm](#).
- COMMON CRAWL, «Common Crawl, index server», [lc.cx/mxMZEi](#).
—«Frequently asked questions», [lc.cx/3Jw6PB](#).
- COMPETITION AND MARKETS AUTHORITY, «Mobile ecosystems market study», Gobierno del Reino Unido, 15 de junio de 2021, [lc.cx/f7o4Ji](#).
—«Mobile browsers and cloud gaming», 10 de junio de 2022, [lc.cx/DsWUjx](#).
—«Mobile ecosystems market study final report», Gobierno del Reino Unido, 10 de junio de 2022, [lc.cx/ADlov1](#).
- CONTRARY RESEARCH, «Hugging Face», 14 de septiembre de 2022, [lc.cx/oZiu0s](#).
- DAVE, Paresh y Jeffrey DASTIN, «Google told its scientists to “strike a positive tone” in AI research – documents», *Reuters*, 23 de diciembre de 2020, [lc.cx/OGCCYa](#).
- DECK, Andrew, «The AI startup outperforming Google Translate in Ethiopian languages», *Rest of World*, 3 de julio de 2023, [lc.cx/yV9qbW](#).

- DUBAL, Veena, «A brief history of the gig», *Logic(s)*, n.º 10, 4 de mayo de 2020.
- EDWARDS, Benj, «You can now run a GPT-3-level AI model on your laptop, phone, and Raspberry Pi», *Ars Technica*, 14 de marzo de 2023, [lc.cx/5xfmZ3](#).
- EGLASH, Ron, «Broken metaphor. The master-slave analogy in technical literature», *Technology and Culture*, vol. 48, n.º 2, 2007.
- ELEUTHERAI, «GPT-Neo», [lc.cx/UI9aOL](#).
- ELIAS, Jennifer, «Google's newest A.I. model uses nearly five times more text data for training than its predecessor», *CNBC*, 16 de mayo de 2023, [lc.cx/WxQowR](#).
- ESSINGER, James, *Jacquard's web. How a hand-loom led to the birth of the information age*, Oxford University Press, Oxford, 2007.
- FIORI, Nicholas, «Plantation energy: From slave labor to machine discipline», *American Quarterly*, vol. 72, n.º 3, septiembre de 2020.
- FOSTER, Kelsey, «PyTorch vs TensorFlow. Who has more pre-trained deep learning models?», *Hackernoon*, 21 de marzo de 2022, [lc.cx/kQeCMZ](#).
- FRANKLIN, Ursula M. *The real world of technology*, House of Anansi Press, Toronto, 1999.
- FTC OFFICE OF TECHNOLOGY, «An inquiry into cloud computing business practices. The Federal Trade Commission is seeking public comments», 22 de marzo de 2023, [lc.cx/xfqxLL](#).
- GAO, Leo, Stella BIDERMAN, Sid BLACK, Laurence GOLDING, Travis HOPPE, Charles FOSTER, Jason PHANG, Horace HE, Anish THITE, Noa NABESHIMA, Shawn PRESSER y Connor LEAHY, «The Pile. An 800GB dataset of diverse text for language modeling», *arXiv*, 31 de diciembre de 2020, [lc.cx/d98-3S](#).
- GEBRU, Timnit, Jamie MORGENSTERN, Briana VECCHIONE, Jennifer WORTMAN VAUGHAN, Hanna WALLACH, Hal DAUMÉ III,

- Kate CRAWFORD, «Datasheets for datasets», *arXiv*, 1 de diciembre de 2021, [lc.cx/NgIaTu](#).
- GERMAIN, Thomas, «Stanford researchers take down Alpaca AI due to hallucinations and rising costs», *Gizmodo*, 21 de marzo de 2023, [lc.cx/2oEV9I](#).
- GHANA NLP, «Ghana Natural Language Processing (NLP)», [lc.cx/_v_71O](#).
- GOEVA, Aleksandrina, Sara STOUTD y Ana TRISOVIC, «Toward reproducible and extensible research», *HDSR*, 16 de diciembre de 2020, [lc.cx/eYz0ST](#).
- GOOGLE, «Contributors to TensorFlow», [lc.cx/sU4v1y](#).
—«Submission to the European Commission on the EU AI Act».
- GRIFFIN, Carl J., *The rural war. Captain Swing and the politics of protest*, Oxford University Press, Oxford, 2012.
- GU, Tianyu, Brendan DOLAN-GAVITT y Siddharth GARG, «Bad-Nets. Identifying vulnerabilities in the machine learning model supply chain», *arXiv*, 11 de marzo de 2019, [lc.cx/RXX6MT](#).
- HABIB, Raza, «OpenAI's plans according to Sam Altman», *Humanloop*, 29 de mayo de 2023, [lc.cx/jXc8yc](#).
- HAO, Karen, «OpenAI is giving Microsoft exclusive access to its GPT-3 language model», *MIT Technology Review*, 23 de septiembre de 2020, [lc.cx/-SqPWZ](#).
- HAO, Karen y Andrea Paola HERNÁNDEZ, «How the AI industry profits from catastrophe», *MIT Technology Review*, 20 de abril de 2022, [lc.cx/7VL4IY](#).
- HIATT, Fred «Official seeks like minds in “Star Wars”», *The Washington Post*, 13 de mayo de 1986, [lc.cx/KX4mJm](#).
- HINTON, Geoffrey E., Alexander KRIZHEVSKY, Ilya SUTSKEVER y Nitish SRIVASTVA, «System and method for addressing overfitting in a neural network», Google Patents, [lc.cx/djooCz](#).

- KAPOOR, Sayash y Arvind NARAYANAN, «Licensing is neither feasible nor effective for addressing AI risks», *AI Snake Oil*, 10 de junio de 2023.
- KARABUS, Jude, «GitHub CEO says EU AI Act shouldn't apply to open source devs», *The Register*, 7 de febrero de 2023, [lc.cx/AvJvvB](#).
- KHAN, Mehtab y Alex Hanna, «The subjects and stages of AI dataset development. A framework for dataset accountability», *SSRN*, 13 de septiembre de 2022, [lc.cx/pyhlcL](#).
- KRINSKY, Robert, «Swords and sheepskins. Militarization of higher education in the United States and prospects of its conversion», *Bulletin of Peace Proposals*, vol. 19, n.º 1, 1988, pp. 33-51, [lc.cx/4_Mq4A](#).
- KRIZHEVSKY, Alex, Ilya SUTSKEVER y Geoffrey E. HINTON, «ImageNet classification with deep convolutional neural networks», en Peter L. Bartlett, Fernando C. N. Pereira, Christopher J. C. Burges, Léon Bottou y Kilian Q. Weinberger (eds.), *Advances in neural information processing systems 25*, Curran Associates, 2013, [lc.cx/x1s-Pg](#).
- KURI, Ngāti, Te AUPŌURI, Ngai TAKOTO, Te RARAWA y Ngāti KAHU, «He Reo Tuku Iho, He Reo Ora. Living language transmitted intergenerationally», *Mai Journal*, vol. 11, n.º 1, 2022, [lc.cx/_lxonP](#).
- LAION, «A call to protect open-source AI in Europe», 28 de abril de 2023, [lc.cx/_tyBdE](#).
- LANGENKAMP, Maximilliam, *How open source machine learning software shapes AI*, tesis, MIT, Massachusetts, 2022, [lc.cx/NP2jPA](#).
- LARDNER JR., George, «Army accuses SDI critic of falsifying credentials», *The Washington Post*, 14 de abril de 1992, [lc.cx/1sRjvV](#).

- LE SCAO, Teven, Angela FAN, Christopher AKIKI *et al.*, «BLOOM: a 176b-parameter open-access multilingual language model», *arXiv*, 27 de junio de 2023, [lc.cx/5r21pQ](#).
- LEVIN, Sam «Moderators who had to view child abuse content sue Microsoft, claiming PTSD», *The Guardian*, 12 de enero de 2017, [lc.cx/joHB4a](#).
- LIESENFELD, Andreas, Alianda LOPEZ y Mark DINGEMANSE, «Rethinking open source generative AI: open-washing and the EU AI Act», *FAccT '24*, Río de Janeiro, 3-6 de junio de 2024, [lc.cx/zp04JO](#).
- LINEBAUGH, Kate y Ryan KNUTSON, «The hidden workforce that helped filter violence and abuse out of ChatGPT», *The Journal*, 7 de noviembre de 2023, [lc.cx/dyxJ8a](#).
- LIPTON, Zachary «The mythos of model interpretability», *arXiv*, 6 de marzo de 2017, [lc.cx/weeOoo](#).
- LUITSE, Dieuwertje, y Wiebke DENKENA, «The great Transformer. Examining the role of large language models in the political economy of AI», *Big Data & Society*, vol. 8, n.º 2, julio de 2021, [lc.cx/B909F3](#).
- MAFFULLI, Stefano, «Towards a definition of “Open Artificial Intelligence”. First meeting recap», *Open Source Initiative*, 13 de julio de 2023, [lc.cx/ob65gh](#).
—«Meta’s LLaMa 2 license is not open source», *Open Source Initiative*, 20 de julio de 2023, [lc.cx/37hubF](#).
- MARKETWATCH, «Market dynamics. The battle for enterprise Linux», *The Free Library*, s. f., [lc.cx/QLbnwF](#).
- MARTIN, Brian «Science: contemporary censorship», en Derek Jones (coord.), *Censorship. A World Encyclopedia*, vol. 4, Fitzroy Dearborn, Londres, 2001, pp. 2167-2170, [lc.cx/MOy1-7](#).
- MATIAS, J. Nathan, «Why we need industry-independent research on tech & society», *Citizens and Tech Lab*, enero de 2020, [lc.cx/CrsIDq](#).

- McCANN, Duncan, «Commoning intellectual property. Public funding and the creation of a knowledge commons», *Common Wealth*, 29 de julio de 2020, lc.cx/MIehIy.
- MCQUILLAN, Dan, *Resisting AI. An anti-fascist approach to artificial intelligence*, Bristol University Press, Bristol, 2022
- MELMAN, Seymour *Pentagon capitalism. The political economy of war*, McGraw-Hill, Nueva York, 1970. [Hay trad. cast.: *El capitalismo del Pentágono. La economía política de la guerra*, trad. Alejandro Licon, Siglo XXI, Ciudad de México, 1972].
- MENABREA, Luigi Federico, *Sketch of the analytical engine invented by Charles Babbage*, trad. al inglés Ada Lovelace, Quaternion, s. l., 2020 [1843].
- METZ, Cade, «Google just open sourced TensorFlow, its artificial intelligence engine», 9 de noviembre de 2015, lc.cx/BtVyN2.
—«Inside OpenAI, Elon Musk's wild plan to set artificial intelligence free», *Wired*, 27 de abril de 2016, lc.cx/WUviLC.
- MICROSOFT, «Open Source Software: a (new?) development methodology», lc.cx/F7g4QO.
—«OpenAI forms exclusive computing partnership with Microsoft to build new Azure AI supercomputing technologies», 22 de julio de 2019, lc.cx/Dx-vPg.
- MITCHELL, Margaret, Simone WU, Andrew ZALDIVAR, Parker BARNES, Lucy VASSERMAN, Ben HUTCHINSON, Elena SPITZER, Inioluwa DEBORAH RAJI, Timnit GEBRU, «Model cards for model reporting», *arXiv*, 14 de enero de 2019, lc.cx/j67gbA.
- MORRISON, Philip y Emily MORRISON (coords.), *Charles Babbage. On the principles and development of the calculator and other seminal writings*, Dover Publications, Nueva York, 1989.

- MURGIA, Madhumita, «AI academics under pressure to do commercial research», *Financial Times*, 13 de marzo de 2019, [lc.cx/J_Q5br](#).
- MUSSER, Micah, Rebecca GELLES, Ronnie KINOSHITA, Catherine AIKEN y Andrew LOHN, «The main resource is the human», CSET, abril de 2023, [lc.cx/cP8UzK](#).
- NEWTON, Casey, «Facebook will pay \$52 million in settlement with moderators who developed PTSD on the job», *The Verge*, 12 de mayo de 2020, [lc.cx/hK591j](#).
- NOLAN, Michael «Llama and ChatGPT are not open-source», *IEEE Spectrum*, 27 de julio de 2023, [lc.cx/EnCYMc](#).
- OPENAI, «Learning from human preferences», *OpenAI*, 13 de junio de 2017, [lc.cx/Mey81t](#).
- «Better language models and their implications», 14 de febrero de 2019, [lc.cx/tYrpw3](#).
 - «OpenAI LP», 11 de marzo de 2019, [lc.cx/_bNYdR](#).
 - «GPT-2», 5 de diciembre de 2019, [lc.cx/J1TODo](#).
 - «Introducing Triton», 28 de julio de 2021, [lc.cx/VuWoH4](#).
 - «GPT-4 technical report», 27 de marzo de 2023, [lc.cx/7_P3vo](#).
- OPEN SOURCE INITIATIVE, «History of the OSI», 19 de septiembre de 2006, [lc.cx/w5nxJ4](#);
- «Open Source case for business», 31 de marzo de 2007, [lc.cx/a3MrL1](#).
 - «Introducing OpenAI», 11 de diciembre de 2015, [lc.cx/XQf9Qd](#).
- REMUS, Will, «AI chatbots lose money every time you use them. That is a problem», *The Washington Post*, 5 de junio de 2023.
- UMA, Stefan y Saumya PREMCHANDER, «Labour, efficiency, critique. Writing the plantation into the technological

- present-future», *Environment and Planning A: Economy and Space*, vol. 54, n.º 2, 2021.
- PASQUINELLI, Matteo, «On the origins of Marx's general intellect», *Radical Philosophy*, n.º 206, 2019.
- PATEL, Dylan, «The AI brick wall. A practical limit», *Semianalysis*, 24 de enero de 2023, lc.cx/A4ifmi.
- PATEL, Dylan y Afzal AHMAD, «The inference cost of search disruption. Large language model cost analysis», *Semianalysis*, 9 de febrero de 2023, lc.cx/w6L8bu.
- «Google “We have no moat, and neither does OpenAI”», *Semianalysis*, 4 de mayo de 2023, lc.cx/yC9ur1.
- PENEDO, Guilherme, Quentin MALARTIC, Daniel HESSLOW, Roxandra COJOCARU, Alessandro CAPPELLI, Hamza ALOBEIDLI, Baptiste PANNIER, Ebtessam ALMAZROUEI y Julien LAUNAY, «The RefinedWeb dataset for Falcon LLM. Outperforming curated with web data, and web data only», *arXiv*, 1 de junio de 2023, lc.cx/8YMWt8.
- PERLROTH, Nichole, «Heartbleed highlights a contradiction in the web», *The New York Times*, 18 de abril de 2014, lc.cx/9fTA3W.
- PERRIGO, Billy, «OpenAI used Kenyan workers on less than \$2 per hour to make ChatGPT less toxic», *Time*, 18 de enero de 2023, lc.cx/tlmHL_.
- «150 African workers for ChatGPT, TikTok and Facebook vote to unionize at Landmark Nairobi Meeting», *Time*, 1 de mayo de 2023, lc.cx/YPVGgU.
- POURSABZI-SANGDEH, Forough, Daniel G. GOLDSTEIN, Jake M. HOFMAN, Jennifer WORTMAN VAUGHAN, Hanna WALLACH, «Manipulating and measuring model interpretability», en *CHI'21: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery, Nueva York, 2021, lc.cx/mR4VaM.

- PYTORCH, «PyTorch contributors», [lc.cx/NlyiVB](https://lcamtuf.coredump.cx/NlyiVB).
- RAMEY BERRY, Daina, *The price for their pound of flesh. The value of the enslaved, from womb to grave, in the building of a nation*, Beacon Press, Boston, 2017.
- REUTERS, «Big tech companies offer millions to help with Heartbleed crisis», 24 de abril de 2014, [lc.cx/mOgqR2](https://lcamtuf.coredump.cx/mOgqR2).
- ROSENBERG, Nathan, *Exploring the black box. Technology, economics, and history*, Cambridge University Press, Cambridge, 1994.
- ROSENTHAL, Caitlin, *Accounting for slavery. Masters and management*, Harvard University Press, Cambridge, 2018.
- RUDIN, Cynthia, «Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead», *Nature Machine Intelligence*, vol. 1, 2019.
- SATARIANO, Adam y Matina STEVIS-GRIDNEFF, «Big tech turns its lobbyists loose on Europe, alarming regulators», *The New York Times*, 14 de diciembre de 2020, [lc.cx/uf3gnh](https://lcamtuf.coredump.cx/uf3gnh).
- SETTY, Riddhi, «Sarah Silverman, authors hit OpenAI, Meta with copyright suits», *Bloomberg Law*, 10 de julio de 2023, [lc.cx/eUH1aR](https://lcamtuf.coredump.cx/eUH1aR).
- SCHIVELBUSCH, Wolfgang, «World machines. The steam engine, the railway, and the computer», *Log*, n.º 33, 2015.
- SCHNEIER, Bruce y Jim WALDO, «Big tech isn't prepared for A.I.'s next chapter», *Slate*, 30 de mayo de 2023, [lc.cx/ol2h9](https://lcamtuf.coredump.cx/ol2h9).
- SCHREPEL, Thibault y Alex PENTLAND, «Competition between AI foundation models», *MIT Connection Science Working Paper*, 10 de julio de 2023, [lc.cx/Oms7iS](https://lcamtuf.coredump.cx/Oms7iS).
- SHAZBEER, Noam M. Aidan NICHOLAS GOMEZ, Lukasz MIECZYSLAW KAISER, Jakob D. USZKOREIT, Llion OWEN JONES, Niki J. PARMAR, Illia POLOSUKHIN y Ashish Teku VASWANI JA, «Attention-based sequence transduction neural networks», Google Patents, [lc.cx/6vNiZY](https://lcamtuf.coredump.cx/6vNiZY).

- SLATER, Derek y Betsy MASIELLO, «Will open source AI shift power from “big tech”? It depends», *Tech Policy*, 16 de junio de 2023, [lc.cx/8pzkKD](#).
- SMITH, Adam, *The wealth of nations*, Neeland Media, 2004 [1776] [Hay trad. cast.: *La riqueza de las naciones*, trad. Carlos Rodríguez Braun, Alianza, Madrid, 1994].
- SMITH, Tim, «Leaked deck raises questions over Stability AI’s Series A pitch to investors», *Sifted*, 21 de abril de 2023, [lc.cx/EFJ\]-wt](#).
- SOLAIMAN, Irene, «The gradient of generative AI release. Methods and considerations», *arXiv*, 5 de febrero de 2023, [lc.cx/XrHixD](#).
- STEIN, Dorothy K., «Lady Lovelace’s notes: Technical text and cultural context», *Victorian Studies*, vol. 28, n.º 1, 1984.
- SURMAN, Mark, «Introducing Mozilla.ai. Investing in trustworthy AI», *The Mozilla Blog*, 22 de marzo de 2023, [lc.cx/dVG-Oz](#).
- SYNOPSIS, «Heartbleed bug», [heartbleed.com](#).
- TALAT, Zeerak, «Response to the LAION letter», *Medium*, 17 de mayo de 2023, [lc.cx/GzRn1f](#).
- TARANTOLA, Andrew, «Meta is spinning off the Pytorch framework into its own AI research foundation», *Engadget*, 12 de septiembre de 2022, [lc.cx/4LGjIl](#).
- TAYLOR, Astra, «The automation charade. The rise of the robots has been greatly exaggerated. Whose interests does that serve?», *Logic(s)*, n.º 5, 1 de agosto de 2018, [lc.cx/gKLLfE](#).
- TECHNOLOGY INNOVATION INSTITUTE, «Falcon LLM», [lc.cx/1gYep](#).
- TEWARI, Archana, «The Reform Bill (1832) and the abolition of slavery (1833). A Caribbean link», *Proceedings of the Indian History Congress*, vol. 73, 2012.

- THYLSTRUP, Nanna y Zeerak TALAT, «Detecting “dirt” and “toxicity”. Rethinking content moderation as pollution behaviour», *SSRN*, enero de 2020, [lc.cx/WL9vuo](#).
- TOUVRON, Hugo, Thibaut LAVRIL, Gautier IZACARD, Xavier MARTINET, Marie-Anne LACHAUX, Timothée LACROIX, Baptiste ROZIERE, Naman GOYAL, Eric HAMBRO, Faisal AZHAR, Aurélien RODRIGUEZ, Armand JOULIN, Edouard GRAVE y Guillaume LAMPLE, «LLaMA: open and efficient foundation language models», *arXiv*, 27 de febrero de 2023, [lc.cx/u_BSCp](#).
- US EPA, «Funding opportunities for participatory science projects», [lc.cx/chmiwF](#).
- VANIAN, Jonathan, «Why Microsoft and Twitter are turning to bug bounties to fix their A.I.», *Fortune*, 10 de agosto de 2021, [lc.cx/QWPFQ](#).
- VINCENT, James, «OpenAI co-founder on company’s past approach to openly sharing research: “We were wrong”», *The Verge*, 15 de marzo de 2023, [lc.cx/qveUhz](#).
- VIPRA, Jay y Sarah MYERS WEST, «Computational power and AI», *AINow*, 27 de septiembre de 2023, [lc.cx/3D2g6H](#).
- WHARE KŌRERO, , «Kaitiakitanga license», [lc.cx/ux4q6Y](#).
- WIGGERS, Kyle, «This startup is setting a DALL-E 2-like AI free, consequences be damned», *TechCrunch*, 12 de agosto de 2022, [lc.cx/yAHdL1](#).
- «Stability AI, Hugging Face and Canva back new AI research nonprofit», 2 de marzo de 2023, *TechCrunch*, [lc.cx/DoOZNp](#).
- WILLIAMS, Adrienne, Milagros MICELI y Timnit GEBRU, «The exploited labor behind artificial intelligence», *No ma*, 13 de octubre de 2022, [lc.cx/vZicv2](#);
- WILLIAMS, Sam, *Free as in freedom. Richard Stallman’s crusade for free software*, O’Reilly, Sebastopol (California), 2002.

WONG, Matteo «America already has an AI underclass», *The Atlantic*, 26 de julio de 2023, [lc.cx/kMwZf8](https://www.theatlantic.com/technology/archive/2023/07/ai-underclass/).

XIANG, Chloe, «A photographer tried to get his photos removed from an AI dataset. He got an invoice instead», *Vice*, 28 de abril de 2023;

ZHANG, Jane y Saritha RAI, «AI presents nearly level playing field for startups and big tech», *Bloomberg*, 21 de marzo de 2023, [lc.cx/jnXNYg](https://www.bloomberg.com/news/articles/2023-03-21/ai-startups-big-tech).





Impreso en febrero de 2025
en Tauro Gráfica,
Madrid

