

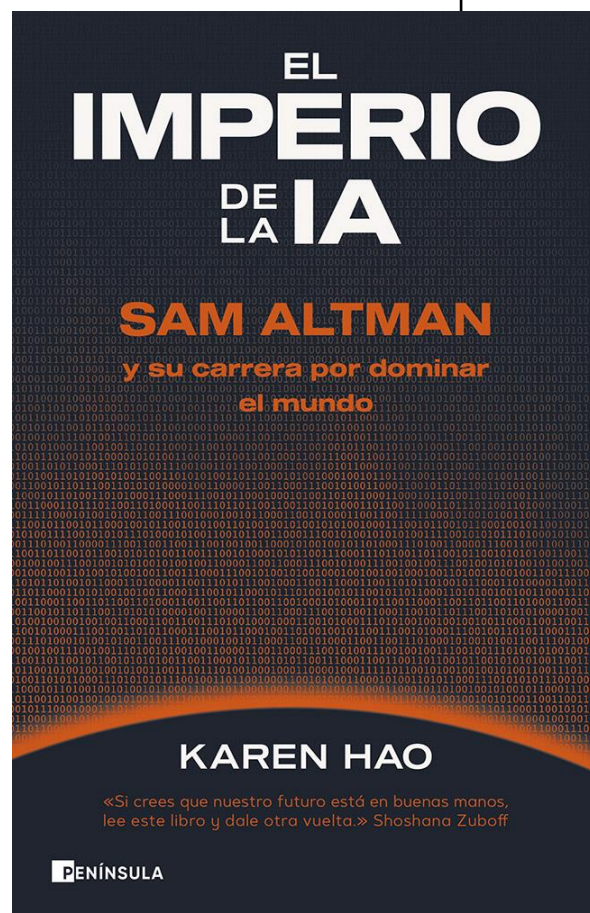
PENÍNSULA

EL IMPERIO DE LA IA

**Sam Altman y
su carrera por
dominar el
mundo**

KAREN HAO

La obra definitiva sobre la
carrera tecnológica que
está marcando el futuro
de nuestra historia.



A LA VENTA EL 26 DE NOVIEMBRE

**AUTORA DISPONIBLE PARA ENTREVISTAS
TELEMÁTICAS EL 26, 27 Y 28 DE NOVIEMBRE**

PARA AMPLIAR INFORMACIÓN, CONTACTAR CON:

Laura Fabregat (Responsable de Comunicación Área de Ensayo)
M: 682 69 63 61 / E: lfabregat@planeta.es

SINOPSIS

Cuando la periodista Karen Hao empezó a cubrir OpenAI en 2019, estaba convencida de que eran los «buenos» de la historia. Ideada como una organización sin ánimo de lucro focalizada en desarrollar una inteligencia artificial segura, su fundador Sam Altman aseguraba entonces que actuarían de contrapeso frente a otras multinacionales.

No ha tenido que transcurrir ni un lustro para que todas esas promesas se evaporaran y la compañía haya empezado a actuar como antaño lo hicieron los grandes imperios, conquistando nuevos territorios y explotando sus recursos y mano de obra, ya sea en forma de agua y energía o de trabajos precarios para recopilar y filtrar los datos que acumulan.

Con un acceso exclusivo al círculo cercano de Altman y basado en más de 250 entrevistas alrededor del mundo, *El imperio de la IA* narra el vertiginoso ascenso de OpenAI, da testimonio del lado oscuro de su creador, Sam Altman, y muestra cómo la carrera tecnológica por la inteligencia artificial pone en jaque nuestro futuro.

LA AUTORA



KAREN HAO es una periodista e ingeniera mecánica especializada en el impacto de la inteligencia artificial en la sociedad. Escribe para medios como *The Atlantic* y dirige la serie AI Spotlight del Centro Pulitzer. Anteriormente fue reportera tecnológica en *The Wall Street Journal* y editora sénior en IA en *MIT Technology Review*. Su trabajo ha sido reconocido por su profundidad, claridad y compromiso con la exposición de las complejidades del desarrollo y la implementación de la IA. Ha recibido numerosos galardones por su trabajo, entre ellos el American National Magazine Award for Journalists Under 30, y la revista *Time* la nombró una de las 100 personas más influyentes en el campo de la IA en 2025.

ALGUNOS FRAGMENTOS

«El anuncio se hizo público al poco tiempo. “La salida del señor Altman se produce tras un profundo proceso de deliberación por parte del consejo de administración, el cual ha llegado a la conclusión de que no siempre ha sido sincero en sus comunicaciones con el consejo, obstaculizando su capacidad para el ejercicio de sus responsabilidades. El consejo ya no confía en su capacidad para continuar al frente de OpenAI.” En apariencia, OpenAI se encontraba en su mejor momento. Desde el lanzamiento de ChatGPT en noviembre de 2022 se había convertido en la historia de éxito más espectacular de Silicon Valley. ChatGPT era la aplicación de consumo de más rápido crecimiento de la historia»

«Para el público, Altman era el hombre que había hecho que todo aquello sucediera. Había pasado la primavera y el verano de gira por el mundo, alcanzando un nivel de popularidad que estaba llevando a los medios de comunicación a compararlo con Taylor Swift. Había cautivado a prácticamente todo el mundo con su apariencia humilde, sus contundentes declaraciones y su aparente sinceridad. [...]. «¿Por qué dedicas tu vida a este trabajo?», le había preguntado Laurene Powell Jobs, fundadora y presidenta de Emerson Collective y viuda de Steve Jobs. “Creo que esta será la tecnología más transformadora y beneficiosa de la historia de la humanidad”, dijo. “En cuatro ocasiones en la historia de OpenAI — la última hace tan solo un par de semanas— he tenido ocasión de estar presente cuando, por así decirlo, hemos retirado el velo de ignorancia y hemos traspasado los límites del descubrimiento, y poder haberlo hecho constituye el máximo honor profesional de mi vida.” »

«¿Había hecho Altman algo ilegal? Quizá tenía que ver con su hermana, se preguntaban los empleados. Hacía un mes, ella había denunciado en unos tuits, que se habían vuelto virales, que su hermano había abusado de ella. Tal vez no se tratase de nada ilegal, sino éticamente reprochable, especulaban, acaso relacionado con otras de las inversiones de Altman o con su recaudación de fondos a través de inversores saudíes para financiar una nueva empresa de chips de inteligencia artificial. Sutskever publicó un anuncio en el Slack de OpenAI. Al cabo de dos horas celebraría una reunión corporativa virtual para responder a las preguntas de los empleados. “Fueron las dos horas más largas de mi vida», recuerda un empleado.” »

«OpenAI no era una empresa normal y su consejo de administración no era un consejo de administración normal. Tenía una estructura única diseñada por el propio Altman, que otorgaba al consejo amplios poderes para actuar en interés no de los accionistas de OpenAI, sino de su misión: garantizar que la AGI (inteligencia artificial general) beneficiara a la humanidad. Altman había pregonado durante mucho tiempo que el consejo tenía capacidad para despedirlo y que ese era su mecanismo de gobierno más importante. Toner subrayó ese dato: “si esta acción destruye la empresa, en realidad podría ser coherente con la misión”.»

«[...] La presión iba en aumento desde todas direcciones. Microsoft, otros inversores de OpenAI y pesos pesados de Silicon Valley se alinearon públicamente con Altman. Una fuente transmitió la estrategia a los medios de comunicación: no solo se marcharían los empleados en masa si la decisión no se revocaba, sino que Microsoft denegaría el acceso a su estructura informática y los inversores interpondrían demandas. Esa combinación haría que OpenAI fuera insostenible sin Altman. A pesar de todo, el consejo seguía resistiendo.»

«El martes, 21 de noviembre, la dirección se conectó con el consejo desde la casa de Altman. [...] El consejo, viendo que no había forma de mantener la empresa sin Altman, acabó aceptando su regreso. A última hora de la noche, llegaron a un acuerdo sobre la distribución de los tres puestos de los independientes en el consejo. D’Angelo continuaría; Toner y McCauley dejarían sus puestos; Bret Taylor, ex consejero delegado de Salesforce y exdirector de tecnología de Facebook, y Larry Summers, exsecretario del Tesoro y expresidente de Harvard, ocuparían los puestos vacantes. Como parte del acuerdo, el nuevo consejo añadiría más miembros con el

tiempo. Altman se sometería a una investigación.[...] En diciembre, Altman describió la experiencia a Trevor Noah en un pódcast como el segundo peor momento de su vida, superado únicamente por la muerte de su padre. El mes siguiente, en enero, se cerró la licitación, valorando OpenAI en 86.000 millones de dólares. Pero todo eso estaba por llegar. La noche del martes 21 de noviembre estaban de celebración.»

«La noticia de la destitución de Altman salió a la luz cuando me encontraba en medio de una entrevista para este libro. Había puesto mi móvil en silencio, felizmente inconsciente de la caótica semana que estaba a punto de acontecer. Veinte minutos más tarde, toqué la pantalla para ver la hora y vi un montón de notificaciones perdidas. Así se inició una serie de días confusos y llenos de adrenalina, mientras yo me apresuraba a entender qué estaba sucediendo. Durante las semanas siguientes, amigos, familiares y medios de comunicación me preguntaron un montón de veces: ¿qué significaba todo aquello, si es que significaba algo? ¿El tira y afloja era una simple distracción, o tendría consecuencias para el resto de nosotros? Para entonces, yo llevaba siguiendo a OpenAI cinco años. En 2019, fui la primera periodista en obtener amplio acceso a la empresa y en escribir su primera reseña. Para mí, aquellos acontecimientos no eran simples alardes frívolos de poder de Silicon Valley. El drama puso de relieve una de las preguntas más apremiantes de nuestra generación: ¿cómo gestionar la inteligencia artificial? La IA es una de las tecnologías más trascendentales de nuestra era. En poco más de una década, ha redefinido la columna vertebral de Internet, convirtiéndose en un mediador omnipresente de las actividades digitales. En menos tiempo todavía, ahora se dispone a modificar muchas otras funciones cruciales de la sociedad, desde la asistencia sanitaria a la economía, pasando por el periodismo y el Gobierno. El futuro de la IA — la forma que adopte esta tecnología— está inextricablemente ligada a nuestro futuro. La pregunta sobre cómo gestionar la IA es en realidad una pregunta sobre cómo asegurarnos de hacer que nuestro futuro sea mejor, no peor. Desde el principio, OpenAI se ha presentado como un experimento audaz para responder a esta pregunta. Fue fundada por un grupo entre cuyos miembros se encontraban Elon Musk y Sam Altman, junto con otros patrocinadores multimillonarios como Peter Thiel, para ser más que un simple laboratorio de investigación o una empresa. Los fundadores manifestaron un compromiso radical de desarrollar la llamada inteligencia artificial general, la cual describieron como la forma más poderosa de IA que nadie hubiera visto jamás, no en pos del lucro económico de los accionistas, sino en beneficio de la humanidad. Con ese fin, Musk y Altman la constituyeron como una organización sin ánimo de lucro y se comprometieron a donar 1.000 millones de dólares para ponerla en marcha. No se dedicaría a crear productos comerciales, sino que se consagraría exclusivamente a la investigación, impulsada únicamente por las más puras intenciones de dar lugar a una forma de inteligencia artificial general que desbloqueara la utopía global y no lo contrario. Musk y Altman se comprometieron también a compartir sus investigaciones tanto como fuera posible durante el proceso y a colaborar ampliamente con otras instituciones.»

«Cuando empecé a investigar a OpenAI, su compromiso con esos ideales se iba erosionando rápidamente. [...] En retrospectiva, la ruptura fue una señal inequívoca de que, en realidad, OpenAI no era un proyecto altruista, sino más bien de ego. La pérdida de su principal patrocinador sumió a OpenAI en la incertidumbre financiera. Para tapar el agujero, Altman reformuló la estructura legal de OpenAI. La encajó dentro de la organización sin ánimo de lucro y creó una rama con ánimo de lucro, OpenAI LP, para recaudar capital, comercializar productos y proporcionar beneficios a los inversores, en la línea de cualquier otra empresa. Cuatro meses más tarde, en julio de 2019, OpenAI anunció un nuevo inversor que aportaba 1.000 millones de dólares: el gigante del software y proveedor de servicios en la nube Microsoft.»

«[...]Atrás quedaban los ideales de transparencia y democracia, de sacrificio y colaboración. Los ejecutivos de OpenAI tenían una única obsesión: ser los primeros en lograr la inteligencia artificial general, hacerla a su imagen y semejanza.»

«[...] A lo largo de los cuatro años siguientes, OpenAI se convirtió en todo lo que había dicho que no sería. Se convirtió en una organización sin ánimo de lucro solo de nombre, comercializando agresivamente productos como ChatGPT y tratando de conseguir valoraciones sin precedentes. Se volvió aún más hermética, no solo denegando el acceso a sus propias investigaciones, sino modificando las normas de la industria para impedir que una parte significativa del desarrollo de la IA se sometiera a escrutinio público. Dio el pistoletazo de salida de la carrera hacia el abismo sobre la que había advertido, acelerando tremendamente la comercialización y el desarrollo de la tecnología, sin apuntalar sus defectos perjudiciales o las formas peligrosas en que podría amplificar y explotar las fallas de nuestra sociedad. Durante el proceso, los enfrentamientos entre directivos y empleados se volvían cada vez más virulentos, a medida que diferentes grupos dentro de la empresa trataban de hacerse con el control y reformar OpenAI según sus ideas. La destitución y reincorporación de Altman en noviembre de 2023 fue la prueba definitiva de que el experimento de gestión había fracasado. No simplemente porque el consejo sin ánimo de lucro de OpenAI cediera ante los intereses económicos, eliminando el último vestigio de la fachada altruista de la organización. Ilustraba a la perfección hasta qué punto una lucha de poder entre un minúsculo puñado de élites de Silicon Valley estaba determinando el futuro de la IA.»

«Nada de esta clase de IA que acapara la atención o incluso de cualquiera de las existentes era inevitable en absoluto; fue la culminación de miles de decisiones subjetivas, tomadas por las personas con poder para encontrarse en la sala de toma de decisiones. Del mismo modo, las futuras generaciones de las tecnologías de IA no están predeterminadas. Sin embargo, vuelve a plantearse la pregunta acerca de la gestión: ¿quién les dará forma? La segunda cosa que he aprendido es que esta manifestación actual de la IA y la trayectoria de su desarrollo ha tomado una dirección alarmante. A primera vista, la IA generativa es apasionante: una contribución a la propuesta de ideas y a la escritura creativa; una compañera con la que charlar hasta altas horas de la noche para alejar la soledad; una herramienta que tal vez un día podría ser tan eficaz a la hora de aumentar la productividad que incrementaría la actividad económica mundial. Sin embargo, igual que en su día pensamos que Facebook no era más que un lugar en el que publicar las fotos [...]»

«Bajo la superficie, los modelos de IA general son monstruosidades, construidas a partir del consumo de unas cantidades de datos, trabajo, capacidad de procesamiento y recursos naturales anteriormente inimaginables. GPT-4, el sucesor del primer ChatGPT, es, según un cálculo, supuestamente 15.000 veces más potente que la primera generación, GPT-1, lanzada cinco años antes. [...]»

«[...]A lo largo de los años, solamente he encontrado una metáfora capaz de resumir la naturaleza de los máximos exponentes del juego de poder de la IA: los imperios. Durante la larga época del colonialismo europeo, los imperios se adueñaron y extrajeron recursos que no eran de su propiedad y explotaron a los pueblos sometidos para extraer, cultivar y refinar dichos recursos para el enriquecimiento de los imperios. Proyectaron ideas racistas y deshumanizadoras acerca de su superioridad y modernidad para justificar — e incluso persuadir a los conquistados para que aceptaran— la invasión de la soberanía, el robo y el sometimiento. Justificaban su lucha por el poder por la necesidad de competir con otros imperios: en la carrera armamentística, todo vale.»

«OpenAI lidera actualmente nuestra carrera hacia ese moderno orden colonial mundial. En la búsqueda de una visión amorfa de progreso, su agresiva expansión de los límites de escala ha establecido las normas de una nueva era para el desarrollo de la IA. Ahora, todos los gigantes tecnológicos están compitiendo unos con otros, gastando sumas de dinero tan astronómicas que han llegado a rivalizar por redistribuir y consolidar sus recursos. Alrededor de la época en

que Microsoft invirtió 10.000 millones de dólares en OpenAI, despidió a 10.000 trabajadores para recortar gastos.»

«Mientras tanto, el resto del mundo empieza a desplomarse bajo el peso del creciente coste humano y material de esta nueva era. Los trabajadores de Kenia ganaban sueldos de miseria por filtrar los discursos de odio y violencia de las tecnologías de OpenAI, incluyendo a ChatGPT. Los artistas son sustituidos por los mismos modelos de IA contruidos a partir de sus trabajos sin su consentimiento y sin recibir compensación alguna. La industria periodística se va atrofiando a medida que las tecnologías de IA generan cada vez más desinformación. Vemos cómo la historia se repite ante nuestros ojos, y esto es solo el principio. OpenAI no aminora el paso. Sigue tratando de aumentar, si cabe, su volumen con unos recursos sin precedentes y el resto de la industria sigue su estela. Con el fin de reprimir las cada vez mayores preocupaciones acerca del actual funcionamiento de la IA generativa, Altman ha pregonado aún con más fuerza sus futuros beneficios. [...]»

«[...] «Creo que el futuro será tan brillante que nadie puede hacerle justicia intentando escribir ahora sobre él», escribió. «Aunque sucederá de manera gradual, triunfos asombrosos — frenar el cambio climático, establecer una colonia en el espacio y el descubrimiento de toda la física— serán con el tiempo algo habitual.» En este momento, la inteligencia artificial generativa es en gran medida retórica, una excusa fantástica polivalente para que OpenAI continúe tratando de conseguir cada vez más riqueza y poder. Muy pocos disponen de un capital comparable para invertir en opciones alternativas. OpenAI y su puñado de competidores constituirán un oligopolio en la tecnología que nos venden como clave para el futuro; todo aquel — ya sea una empresa o un gobierno— que quiera una porción de esa idea tendrá que recurrir a los imperios para que se lo suministren. Existe una salida diferente. La inteligencia artificial no tiene por qué ser lo que es en la actualidad. No tenemos por qué aceptar la lógica de una escala y un consumo sin precedentes para lograr avances y progreso. Gran parte de lo que nuestra sociedad necesita en realidad — mejor asistencia sanitaria y educación, aire puro y agua limpia, una transición más rápida para dejar atrás los combustibles fósiles— puede fomentarse y lograrse y, en ocasiones es incluso necesario que así sea, con modelos considerablemente más reducidos de IA y una variedad de enfoques diferentes. Además, la IA por sí sola no será suficiente: necesitaremos también más cohesión social y cooperación internacional, precisamente algunas de las cosas cuestionadas por la visión existente del desarrollo de la IA. Pero los imperios de la IA no renunciarán fácilmente a su poder.»

DERECHO DIVINO

«A Musk le pareció inteligente, motivado y, lo que es más importante, alguien que compartía sus opiniones acerca de la necesidad de desarrollar y gestionar cuidadosamente la inteligencia artificial. Como Musk describió años después en una demanda judicial, era como si Altman hubiera replicado todo lo que Musk había dicho alguna vez sobre el tema para ganarse su confianza [...].»

«Musk llevaba algún tiempo profundamente preocupado por la IA. En 2012, había conocido a Demis Hassabis, profesor y consejero delegado del laboratorio de IA DeepMind Technologies, con sede en Londres. Poco después, Hassabis visitó también a Musk en su fábrica de SpaceX. Cuando los dos hombres estaban sentados en la cantina, rodeados del ruido del transporte y el ensamblaje de enormes piezas de cohetes, Hassabis planteó la posibilidad de que una IA más avanzada, capaz de superar algún día a la inteligencia humana, pudiera representar una amenaza para la humanidad. Es más, el plan infalible de Musk de colonizar Marte para escapar no funcionaría en ese caso. La superinteligencia, dijo Hassabis divertido, simplemente seguiría a

los humanos por la galaxia. Musk, al que estaba claro que la cosa no le hacía tanta gracia, invirtió cinco millones de dólares en DeepMind para poder controlar la empresa. [...]»

«Cuando Altman y Musk dejaron de hablarse y Altman sustituyó a Musk como el nuevo “informático” de Silicon Valley, Altman modificó su discurso público acerca de sus creencias sobre los peligros de lo que estaba construyendo. “Ahora estoy bastante convencido de que la IA será una herramienta”, declaró a *Business Insider* en 2023, “aunque también pienso que los humanos del futuro y la sociedad humana del futuro serán extremadamente diferentes y tenemos una oportunidad de reflexionar detenidamente sobre cómo diseñar ese futuro.” Musk llegaría a tener la impresión de que Altman le había utilizado para que lo catapultase a la fama. Era el reflejo de una observación que había perseguido a Altman durante toda su vida. “Podrías lanzarlo en paracaídas en una isla llena de caníbales y volver al cabo de cinco años y sería el rey”, dijo en una ocasión su mentor Paul Graham. Años más tarde, Graham ahondó en ese punto: «Sam es extraordinario a la hora de conseguir poder».»

«Altman aprendió también la mejor manera de presentar las cosas a los medios de comunicación y el modo más seguro de lograr acuerdos extraordinarios. Aun siendo consejero delegado de una startup poco conocida, logró negociar con éxito alianzas empresariales con las principales operadoras de móviles estadounidenses. La clave de su fórmula, según se dice, es la combinación de su extraordinaria capacidad para escuchar, su disposición a ayudar y su habilidad para presentar lo que tenga que ofrecer exactamente como tú quieres. (Hoy en día, como pieza fundamental ultramillonaria de Silicon Valley, no está de más que tenga mucho que ofrecer.) Es el «Michael Jordan de la escucha», se ha dicho de él. Es el «Usain Bolt de la recaudación de fondos», dice Geoff Ralston, que ocupó la dirección de YC después de Altman.»

«De un modo que acabaría definiendo el resto de su carrera, Altman salió reforzado de la crisis. El consejo de Loopt se puso de su lado. A pesar de su mediocre trayectoria, Altman también saldría de Loopt mucho mejor que como había empezado. Utilizó la startup como trampolín para asentarse cada vez más arriba en las redes más poderosas de Silicon Valley y, posteriormente, utilizó esos contactos para orquestar una salida para su empresa que le hizo rico. A los veintiséis años, obtuvo un beneficio neto de cinco millones de dólares de la venta de Loopt. Altman lo consideró una decepción — Jobs tenía una fortuna de 256 millones de dólares a los veinticinco años—, pero pronto amasaría mucho más dinero. Esa riqueza cambiaría poco a poco su estilo de vida. Con el tiempo, dejaría de ir al supermercado. Viajaba en jet privado. Coleccionaba coches deportivos de lujo, incluyendo McLarens y un extraordinariamente exclusivo Koenigsegg valorado en cinco millones de dólares, y se convirtió en un gran aficionado a conducirlos. Durante un tiempo, asistió al festival artístico anual psicodélico y rebosante de sexo Burning Man de una semana de duración, que se celebra en el desierto. Como muchos peces gordos de Silicon Valley, se convirtió en un consumidor esporádico de ketamina, una droga de carácter recreativo que puede recetarse legalmente para mitigar la depresión. Una vez alcanzado el éxito, Altman trajo consigo a sus hermanos. En 2012, creó un fondo de inversiones personal llamado Hydrazine Capital con su hermano Jack, el cual había estudiado económicas en Princeton y estaba probando suerte en el sector de la banca de inversiones. [...]»

«“Sin crecimiento económico, la democracia no funciona, porque los votantes ocupan un sistema de suma cero”, dijo. Esta idea evolucionaría hasta transformarse en la tesis fundamental que impulsaría la carrera y las inversiones de Altman. “Lo mejor que puede hacer la gente del sector privado para ayudar a que el país vuelva a funcionar es conseguir que vuelva el crecimiento económico”, dijo en 2017. “En EE.UU. hemos tenido doscientos años de un crecimiento económico sin parangón. Hemos tenido cien años de expansión territorial y cien años de nuevas tecnologías que realmente funcionan”, añadió, pasando por alto la sangrienta historia colonial y las consecuencias medioambientales de una industrialización desbocada. “Y la mayoría de la gente estaba bastante contenta. Ahora ya no.” “El crecimiento económico

sostenible es casi siempre un bien moral”, añadiría en 2019. “Parte de lo que me motiva para trabajar en Y Combinator y en OpenAI es volver a eso, volver a tener un crecimiento económico sostenible, volver a un mundo en el que las vidas de la mayoría de la gente mejoren cada año y en el que sintamos el espíritu del éxito común.”»

UNA MISIÓN CIVILIZADORA

«Greg Brockman fue el primero que se comprometió a crear OpenAI. Como cofundador con Brockman, Altman seleccionó a Ilya Sutskever, que en aquel momento era un investigador de AI de Google al que Altman envió sin previo aviso un correo electrónico para invitarle a la cena de Rosewood en el verano de 2015; Sutskever aceptó entusiasmado en cuanto supo que Musk también asistiría. [...] A lo largo de cinco años, inventó los prototipos de muchos de los primeros productos de Stripe, contribuyendo a su transformación en una potente empresa de tecnología financiera que dota de infraestructura para pagos digitales a compañías como Amazon y Shopify. Su decisión le proporcionó una fortuna considerable y le hizo entrar a formar parte de la exclusiva categoría de quienes habían contribuido a crear una empresa multimillonaria en Silicon Valley. Sutskever era el científico. Delgado y enjuto, había nacido en la Unión Soviética y se había criado en Israel, donde destacó como un prodigio de las matemáticas. Tras tener problemas para encontrar profesores que pudieran mantenerse a su nivel, sus padres lo matricularon en el último año de la escuela secundaria en la Universidad Abierta de Israel. A los dieciséis años se trasladó a Toronto [...]. Allí fue donde Sutskever conoció a Geoffrey Hinton, un profesor británico-canadiense que había realizado influyentes trabajos de investigación sobre inteligencia artificial. [...] se convirtió en la única persona a la que consideraría su mentor [...]»

«En Silicon Valley hay un dicho popular: fundar una empresa con alguien es como contraer matrimonio. Como sucede en cualquier compromiso, Brockman y Sutskever se cortejaron mutuamente después de que Altman le dijera diplomáticamente a Brockman que necesitaba unirse a alguien que entendiese la investigación de la IA. [...] Investigó a fondo a cada candidato antes de intentar contratarlos de manera más convincente. Altman ensalzaría más adelante esos primeros pasos en una entrada de blog titulada «Greg». “Mucha gente me pregunta cómo sería el cofundador ideal”, escribió Altman. “Ahora tengo la respuesta: Greg Brockman.”»

«[...] Mientras Musk, Altman y Brockman debatían sobre la mejor forma de posicionar OpenAI en su lanzamiento, todos ellos eran plenamente conscientes de la importancia de su percepción pública. Estaban de acuerdo en la propuesta de Altman de convertirla en una organización sin ánimo de lucro y resaltar la apertura que indicaba su nombre. OpenAI, la anti Google, investigaría para todo el mundo, abogaba por una ciencia de código abierto y sería un modelo de transparencia. “Espero que nos introduzcamos en este campo como un grupo neutral, tratando de colaborar lo más ampliamente posible y haciendo que el objeto del diálogo pase de ser la victoria de un grupo o empresa en concreto a ser la victoria de la humanidad”, les escribió Brockman a Musk y a Altman en noviembre de 2015. (“Creo que esa es la mejor manera de obligarnos a ser una institución de investigación líder.”) «Para triunfar, es muy importante contar con el apoyo del público», respondió Musk ese mismo mes, proponiendo reescribir el anuncio de la creación de OpenAI para ganar más adeptos. “Necesitamos empezar con una cifra muy superior a los cien millones de dólares para no parecer insignificantes en comparación con lo que gastan Google o Facebook”, añadió. “Creo que deberíamos empezar con un compromiso de financiación de 1.000 millones de dólares. Esto va en serio.”»

«[...] este modelo existencial de seguridad de la IA, basado en experimentos de pensamiento filosófico, sería pronto objeto de críticas a medida que la comunidad investigadora de la IA fue consciente de los perjuicios menos apocalípticos y más inmediatos de la IA en el mundo real.»

«Entre los asistentes a la cena de Rosewood se encontraba Dario Amodei. Amodei, un neurocientífico computacional convertido en investigador de IA, trabajaba entonces en el laboratorio de inteligencia artificial de la empresa china Baidu con sede en Silicon Valley antes de pasar un breve periodo en Google. Su hermana, Daniela Amodei, había trabajado con Brockman en Stripe y, cuando este empezó a dedicarse en serio al desarrollo de la IA, recurrió a Dario en busca de recursos de aprendizaje. [...] Amodei se incorporó dos meses después para dirigir la investigación sobre seguridad de la IA. A continuación, Open Phil donaría 30 millones a OpenAI para garantizar que Holden Karnofsky tuviera un puesto en el consejo de administración durante tres años.[...] A finales de 2020, los hermanos Amodei quedaron tan afectados al ver cómo Altman y OpenAI se alejaban de su premisa original que abandonaron la empresa para crear otro laboratorio de IA, Anthropic, llevándose consigo a miembros esenciales de la plantilla y creando una rivalidad que tendría un papel fundamental en el frenético lanzamiento de Chat-GPT. Karnofsky abandonó el consejo de OpenAI una vez cumplido su mandato y a causa de un nuevo conflicto de intereses. En la lista de candidatos propuestos para sustituirle, incluyó a una de sus antiguas empleadas: Helen Toner. El problema era que OpenAI no tenía ni idea de lo que estaba haciendo. Al cabo de un año, había formado un equipo estelar arrebatando, suplicando y fichando personal en una agresiva caza de talentos en la industria, manteniendo el entusiasmo a nivel interno gracias a la enorme concentración de portentos. A pesar de todo, le costó encontrar una estrategia coherente. Y el impulso y el brillo empezaban a desaparecer.»

««Sus proyectos más grandes no parecían algo superinnovador», dice Nikhil Mishra, un investigador de IA que trabajó como becario en OpenAI en 2017. Las dotes de liderazgo de Brockman y Sutskever también se estaban llevando al límite. Mientras Brockman se pasaba la mayor parte de los días codificando, Sutskever merodeaba por la oficina preguntando repetidamente a los investigadores: «¿Cuál es tu próximo proyecto importante?». No parecía haber nadie al mando y el ambiente era muy estresante. [...] El laboratorio gastaba dinero sin cesar, en su mayoría para hacer frente a los salarios del equipo que había formado. En 2016, OpenAI gastó más de 7 millones de dólares de sus 11 millones de presupuesto en remuneraciones y beneficios. Musk se estaba impacientando. [...] Musk se pasaba periódicamente por la oficina para exigir más avances, a veces fijando plazos nada realistas, típicos de su filosofía de gestión. [...] Musk solamente hizo una pregunta: «¿Cuándo? ¿Cuándo vais a hacer esas cosas?». «No lo sé», dijo Zaremba. «O sea, que en realidad no tienes ningún plan», replicó Musk. [...]»

«Así que, en marzo de 2017, Brockman y Sutskever empezaron en serio a desarrollar una hoja de ruta de investigación más focalizada. La pregunta fundamental era: ¿qué se necesitaría para que OpenAI lograse la inteligencia artificial generativa y para ser la primera en lograrla? Sutskever intuía que, por encima de todo, tendría que ver con una dimensión clave: la cantidad de «computación», un término técnico para referirse a los recursos informáticos, que OpenAI necesitaría para lograr grandes avances en las capacidades de la IA. [...] La cantidad de computación se basa en tres cosas: el poder de procesamiento de un chip informático individual, o cuántos cálculos puede realizar por segundo; el número total de chips informáticos disponibles; y cuánto tiempo permanecen en funcionamiento para realizar sus cálculos. La primera viene determinada por la industria de la fabricación de chips que, durante décadas, ha duplicado la potencia de un chip cada dos años gracias a un proceso intensivo de investigación y desarrollo.[...] La constatación de este hecho llevó a la organización a perder su equilibrio financiero. Para Brockman y Sutskever representaba el cuestionamiento de la base misma de la estructura de OpenAI. ¿Cómo iba a poder una organización sin ánimo de lucro recaudar tanto dinero anualmente para mantener el ritmo necesario para seguir siendo la número uno? Se plantearon brevemente fusionarse con una startup de chips, pero, en el verano de 2017, empezaron a debatir seriamente con Altman y Musk si OpenAI tenía que transformarse en una organización con ánimo de lucro.»

«En caso de que OpenAI tuviera éxito, ¿no sería peligroso cederle a Musk pleno control de la AGI? Convencido, Brockman se lo comunicó a Sutskever, que seguía sin estar seguro. En septiembre de 2017, envió un correo electrónico a Musk y a Altman en su nombre y en el de Brockman, en un último intento por salvar la situación.[...] Al cabo de diez minutos, Musk ya había respondido. «Chicos, ya he tenido suficiente. Esta es la gota que colma el vaso», escribió. Si Sutskever y Brockman seguían queriendo optar por una organización con ánimo de lucro, tendrían que hacerlo por su cuenta. De lo contrario, OpenAI seguiría siendo una organización sin ánimo de lucro. “No pienso seguir financiando a OpenAI hasta que os comprometáis firmemente a continuar, porque si no seré un tonto que está financiando gratuitamente a una startup”, dijo Musk. Cincuenta minutos después, prosiguió. «Para que quede claro, esto no es un ultimátum para que aceptéis lo que ya se ha hablado anteriormente. Eso ya no está sobre la mesa.” Altman intervino en el hilo de correos a la mañana siguiente: “¡me sigue entusiasmando la estructura sin ánimo de lucro!”.»

«Sin embargo, lo cierto era que hacer que OpenAI siguiera siendo una organización sin ánimo de lucro no iba a solucionar sus problemas de dinero. Si bien Brockman y Sutskever seguían reuniéndose con inversores potenciales sin ánimo de lucro, les costaba siquiera acercarse al capital que creían que OpenAI iba a necesitar. Los caprichosos titubeos de Musk a la hora de cumplir con sus compromisos de financiación amenazaban también con sumir a OpenAI en un estado de crisis. Entre bastidores, Altman empezó a buscar alternativas de financiación y a desvincular a OpenAI de su dependencia de Musk. [...] Altman había abandonado sus proyectos políticos y había logrado convencer a Brockman y, a través de este a Sutskever, de que él sería el mejor líder. Ante la decisión del grupo, Musk ya no quería estar vinculado públicamente con la organización. «No estaré en una situación en la que la percepción de mi influencia y mi tiempo no se correspondan con la realidad», había escrito anteriormente. Algunas semanas después, Musk abandonó su cargo de copresidente de OpenAI y Altman se convirtió en presidente de la entidad sin ánimo de lucro.»

«[...]A pesar de todos sus esfuerzos, la gente de OpenAI que visionó la filmación coincidió en que era terrible. Contrataron a profesionales y Brockman empezó a financiarlos, en parte con su propio dinero. Mientras tanto, Altman desarrollaba su plan de recaudación de fondos. Tras plantearse una serie de estructuras con ánimo de lucro, acabó con una propuesta poco habitual para equilibrar la necesidad de capital con un compromiso continuado con la misión de OpenAI. Aunque las corporaciones con ánimo de lucro tenían incorporado un mecanismo para mantener ese equilibrio, también tenían muchas otras reglas. En su lugar, Altman quería crear una sociedad en comandita (Limited Partnership o LP) para que actuase como una sección con ánimo de lucro para recibir inversiones y comercializar las tecnologías de OpenAI. Esa sección fijaría un techo a las ganancias de los inversores y estaría regida por la entidad sin ánimo de lucro de OpenAI. La ventaja era que el acuerdo de funcionamiento de las LP podía redactarse en función de lo que quisiera su creador. OpenAI podría especificar que la misión tuviera prioridad sobre los inversores. Además, las LP limitaban el poder que podían ejercer los accionistas, de modo que nunca ostentaran el control mayoritario.»

«A principios de 2019, los altos directivos de Microsoft empezaron a pasarse por la oficina de OpenAI. [...] Más o menos en esa misma época, Altman empezó a tener problemas en YC. Después de cinco años como jefe de la organización, la frustración con Altman había alcanzado niveles críticos en relación con un asunto asombrosamente parecido a otro que había tenido que afrontar en Loopt: la aparente priorización de sus propios proyectos y aspiraciones por encima de los de la organización, en ocasiones incluso a expensas de esta. La cantidad de tiempo que dedicaba a las negociaciones de OpenAI y que dejaba de dedicar

a asesorar a las startups de YC no ayudaba. Para algunos, Altman estaba cosechando beneficios considerables a título personal, obteniendo ingresos enormes al invertir en empresas de YC con su fondo personal Hydrazine, mientras que su trabajo era limitado.»

«En una reunión corporativa posterior, Altman hizo pública la noticia, y proclamó que Microsoft era el inversor y socio adecuado. El gigante tecnológico tenía el dinero y la capacidad de computación que OpenAI necesitaba y los valores de su dirección estaban claramente alineados con la misión de garantizar una AGI beneficiosa. Además, OpenAI había asumido unos compromisos muy poco estrictos sobre qué ofrecer a Microsoft para su comercialización. El laboratorio no había tenido que hacer demasiadas concesiones, dijo Altman.[...]»

CENTRO NEURÁLGICO

«Este sería un tema recurrente durante mi visita: plantas que no podía ver, reuniones a las que no podía asistir, investigadores que miraban furtivamente al jefe de comunicaciones cada pocas frases para comprobar que no habían incumplido ninguna política de confidencialidad. Posteriormente, me enteré de que después de mi visita Jack Clark publicó una advertencia excepcionalmente severa a los empleados en Slack para que no hablasen conmigo más allá de las conversaciones autorizadas. El guardia de seguridad recibió una foto-grafía mía con instrucciones de que estuviera alerta en caso de que me presentase en las instalaciones sin autorización. Era una conducta extraña en general, más extraña si cabe teniendo en cuenta el compromiso de OpenAI con la transparencia. Empecé a preguntarme qué estaban ocultando, si se suponía que todo iba a ser una investigación beneficiosa que acabaría poniéndose a disposición del público. Durante el almuerzo y a lo largo de los días siguientes, indagué más en las razones por las que Brockman había cofundado OpenAI. Era un adolescente cuando empezó a obsesionarse con la idea de que se pudiera recrear la inteligencia humana.[...] Brockman se hizo famoso por su legendaria productividad a la hora de programar. Era un «ingeniero 10x», que en la jerga de Silicon Valley significaba que podía resolver problemas de programación diez veces más rápido que el programador medio. Las personas no se le daban tan bien. En su papel de director tecnológico, prefería mucho más programar que gestionar.»

«“Básicamente, quiero trabajar en la AGI el resto de mi vida”, me dijo Brockman. Lo abordaba todo con su intensidad característica. Era resolutivo y se fijaba en los detalles. Pasaba largas horas y noches en vela obsesionado por las tareas que tenía ante sí. Cuando no estaba trabajando, seguía trabajando para mejorar. Leía libros sobre oratoria y negociación, habilidades que creía que le ayudarían a cumplir la misión de OpenAI. Le avergonzaba admitir que también leía libros por diversión, novelas épicas de ciencia ficción como la trilogía de *El problema de los tres cuerpos* del escritor chino Liu Cixin y la serie *Fundación* de Isaac Asimov. Para los empleados, su incesante obsesión era una bendición a la vez que una maldición. [...] OpenAI no podría haber llegado a donde estaba sin él, dice un antiguo ingeniero que trabajó codo con codo con Brockman. Sin embargo, si se hubiera permitido que todo dependiera de él, las cosas habrían ido muy mal. “Greg no tiene una visión. No es un visionario como Sam Altman. Solo quiere resolver un buen problema difícil y demostrar que es diez veces más inteligente que los demás.” Le pregunté a Brockman qué lo motivaba. [...] En 2022, se convirtió en presidente de OpenAI.»

SUEÑOS DE MODERNIDAD

«Esas dos características de las revoluciones tecnológicas —su promesa de aportar progreso y, por otra parte, su tendencia a privar de él a las personas sin poder, especialmente a las más vulnerables— son quizá más reales que nunca en el momento en que nos encontramos con la inteligencia artificial. Desde su concepción, el desarrollo y el uso de la IA han sido impulsados por tentadores sueños de modernidad y modelados por una reducida élite con el dinero y la influencia necesarios para materializar su concepción de la tecnología. Esa concepción es la que ha provocado los costes de explotación social, laboral y medioambiental presentes en el mundo actual, especialmente, como veremos, en muchos países del sur, en los que las consecuencias de su expolio por los imperios históricos aún se aprecian en su retrasado desarrollo económico y sus débiles instituciones políticas. Y, a pesar de todo, [...]Silicon Valley presenta las experiencias de aquellos que son explotados y perjudicados por la tecnología como más felices gracias a ella. La promesa que impulsó el desarrollo de la IA está insertada en el propio nombre de la tecnología. En 1956, seis años después del artículo de Turing que empezaba con la frase «¿Pueden pensar las máquinas?», veinte científicos, todos ellos hombres blancos, se reunieron en el Dartmouth College para crear una nueva disciplina para el estudio de esa cuestión. [...] necesitaban un nombre nuevo que las unificase. [...] Al no despertar demasiado interés, se puso a buscar una expresión más evocadora. Se decidió por el término *inteligencia artificial*. El nombre de *inteligencia artificial* fue, por tanto, una herramienta de marketing desde el principio, que incluía la promesa de lo que la tecnología podría traer consigo. *Inteligencia* suena como algo intrínsecamente bueno y deseable, sofisticado e impresionante, algo de lo que, sin duda, la sociedad querrá más; algo que debería aportar un beneficio universal. El cambio de nombre cumplió su objetivo. Las dos palabras suscitaron inmediatamente más interés, no solo por parte de los inversores, sino también de los científicos, ansiosos por formar parte de un campo incipiente con unas ambiciones tan colosales.» [...] Esa tradición de antropomorfización continúa hasta hoy, respaldada por las historias de Hollywood que combinan la idea de la «IA» con descripciones ancestrales de creaciones hechas por el hombre que, de repente, cobran vida. Los desarrolladores de IA hablan a menudo de cómo su software «aprende», «lee» o «crea», exactamente igual que los humanos. Esto no solo ha alimentado la sensación de que las tecnologías de IA actuales tienen mucha más capacidad de la que tienen realmente, sino que se ha convertido en una herramienta retórica para que las empresas eludan su responsabilidad legal.»

«Ahora, la IA es el presagio de la cuarta revolución industrial. La piedra angular del superpoder moderno. [...] Lo que no se dice es que, ante la falta de un significado consensuado, la «inteligencia artificial» o la «inteligencia artificial general», puede ser lo que OpenAI quiera que sea. La historia de la IA nos enseña que el desarrollo de la inteligencia artificial siempre ha venido determinado por una élite poderosa.»

«A mediados de la primera década de la comercialización de la IA, la mayor parte de la investigación de alto nivel se desarrollaba en laboratorios conectados con empresas tecnológicas. [...]. A lo largo de la misma década, gigantes tecnológicos como Microsoft y Google triplicaron con creces sus artículos vinculados a empresas hasta el 66 por ciento. Irónicamente, esa fue precisamente la razón por la que Musk y Altman anunciaron que querían lanzar OpenAI. El ánimo de lucro de la industria tecnológica se había convertido en la arrolladora fuerza impulsora del desarrollo de la IA. El impacto de esta consolidación de financiación y talento en la primera era redujo de manera significativa la diversidad de ideas en el ámbito de la investigación de la IA. El aprendizaje profundo continuó reinando, no solo por sus méritos científicos, sino también porque se destinó muy poca inversión a explorar o promover otros paradigmas.[...]»

«“El cerebro humano tiene alrededor de un billón de parámetros o sinapsis”, me dijo Hinton en 2020. “Lo que ahora entendemos por un modelo realmente grande, como el GPT-3, tiene 175.000 millones. Es mil veces más pequeño que el cerebro.” “El aprendizaje profundo será capaz de hacer de todo”, dijo. Su bestia negra de la época moderna era Gary Marcus, un profesor emérito de psicología y neurociencia de la Universidad de Nueva York, que testificaría en el Congreso junto con Sam Altman en mayo de 2023. Cuatro años antes, Marcus coescribió un libro titulado *Rebooting IA*, en el que afirmaba que esos problemas eran inherentes al aprendizaje profundo. Atrapadas permanentemente en el mundo de las correlaciones, las redes neuronales no serían nunca capaces, independientemente de la cantidad de datos o computación, de entender las relaciones causales —por qué las cosas son como son— y así llevar a cabo un razonamiento causal.[...]»

«Los modelos de IA generativa son modelos de aprendizaje profundo entrenados para generar reproducciones de los datos introducidos. A partir de un texto antiguo, aprenden a sintetizar un texto nuevo; a partir de imágenes antiguas, aprenden a sintetizar imágenes nuevas. Sin embargo, para hacerlo con la fidelidad suficiente para parecer humanas, cosa que según OpenAI es la clave de la AGI, se entrenan con más datos y computación de los que se han utilizado nunca. La IA generativa es, por tanto, la forma maximalista del aprendizaje profundo»

«Esos problemas tienen el mismo origen que antes. Independientemente de su escala, las redes neuronales siguen siendo comparaciones de patrones estadísticos. Y dichos patrones siguen siendo también, en ocasiones, defectuosos o irrelevantes, solo que ahora más intrincados e inescrutables que nunca. A medida que las empresas han tratado de remodelar los modelos de IA generativa como motores de búsqueda, esas deficiencias han dado lugar a nuevos problemas. Los modelos no se basan en hechos, ni siquiera en informaciones individuales. Los generadores de texto se limitan a predecir la palabra más probable que puede venir a continuación en una frase y la frase más probable susceptible de venir a continuación en un párrafo. [...] La industria de la IA denomina a esas imprecisiones «alucinaciones».»

«En un ejemplo muy conocido del problema de las alucinaciones, un abogado utilizó el ChatGPT para llevar a cabo una investigación legal [...]. Posteriormente fue sancionado, multado y humillado públicamente tras descubrirse demasiado tarde que el chatbot se había inventado todo [...]. El patinazo no fue solo un caso de negligencia por parte del abogado, sino también un reflejo de cómo las empresas alimentan la confusión pública en relación con las capacidades de los modelos mediante un marketing ambiguo o exagerado. Altman ha tuiteado públicamente que “ChatGPT es increíblemente limitado”, especialmente en lo que respecta a la “veracidad”, pero la página web de OpenAI alardea de la capacidad de GPT-4 para aprobar el examen para poder ejercer la abogacía y las pruebas de ingreso en las facultades de Derecho de EE. UU. [...] Esta confianza equivocada en la IA generativa podría causar de nuevo verdaderos perjuicios, especialmente en contextos delicados. Las startups están presionando a las jefaturas de policía para que adopten software basado en los modelos de OpenAI para autogenerar informes de incidentes; actualmente, muchos pacientes tienden a hacer preguntas urgentes sobre temas de salud a los chatbots en lugar de acudir a su médico. En esos casos, las alucinaciones no identificadas podrían tener graves consecuencias *a posteriori*.»

[...] Los modelos de IA generativa amplifican los contenidos discriminatorios y de odio. *Bloomberg*, *Rest of the World*, *The Washington Post* y muchas otras publicaciones han mostrado cómo generadores de imágenes como Stable Diffusion y DALL-E cosifican y

regurgitan expresiones racistas y estereotipos culturales. Las «personas atractivas» son jóvenes y blancas. Las «empleadas domésticas» son negras y de color. Los «ingenieros» son hombres. Los «médicos de África» son blancos, a veces incluso cuando la pregunta especifica «médico africano negro».

ESCALA DE AMBICIÓN

«Si hay una persona a la que se le puede atribuir el mérito de ser la primera en imponer la filosofía del escalado en OpenAI, esa persona es su cofundador Ilya Sutskever. [...] No es que Sutskever fuera especialmente persuasivo. Si Altman era el político, Sutskever era todo lo contrario. Nunca medía sus palabras ni se andaba con rodeos para que su lenguaje calara mejor entre su público. Se limitaba a expresar sus opiniones con absoluta sinceridad y con una escandalosa confianza que hacía que la gente comulgara con ellas y las encontrara inspiradoras, o no. [...] Lo que llevaba a la gente a seguir a Sutskever era su fama y su veteranía. Muchos empleados de OpenAI eran plenamente conscientes de sus primeras contribuciones a la materia; algunos lo veían como una especie de profeta.»

«[...] Siguiendo la filosofía de Sutskever de aumentar la escala de las redes neuronales, la pregunta que se planteaba en los primeros tiempos de OpenAI era: ¿qué escala? Diversos investigadores plantearon y barajaron diferentes opciones, pero ninguna de las redes neuronales que habían ido ganando terreno en este campo parecían dar la talla. En agosto de 2017, la cosa cambió con la invención por parte de Google de un nuevo tipo de red neuronal conocida como Transformer. [...] Los Transformers pueden asimilar textos de gran volumen y analizar cada palabra, frase y párrafo en un contexto considerablemente más amplio. Google consideraba el Transformer como un medio de mejorar su motor de búsqueda y Google Translate, así como sus otros servicios basados en gran medida en el procesamiento del lenguaje. Sutskever les veía otra utilidad. Los Transformers son redes neuronales sencillas y escalables, un ejemplo de lo que estaba buscando.»

«[...] Durante la elaboración de mi reportaje, llegué a una conclusión aún más sorprendente. Ni siquiera en Silicon Valley otras empresas o inversores dieron ningún paso hasta después de que ChatGPT destinara sumas sin precedentes para llevar a cabo el proceso de escalado. Eso incluía a Google y a DeepMind, el rival original de OpenAI. Fue concretamente OpenAI, con sus orígenes multimillonarios, su ideología característica y el singular empuje, la red de contactos y el talento para recaudar fondos de Altman, lo que creó una combinación perfecta para que se revelara y se asentara su particular visión. «Tengo la sensación de que Sam es la persona más ambiciosa del planeta», dice un antiguo empleado. Dicho de otro modo, todo lo que hizo OpenAI era lo contrario de inevitable; los impresionantes costes globales de sus gigantescos modelos de aprendizaje profundo y la peligrosa carrera que desató en la industria para aumentar la escala de esos modelos hasta niveles estratosféricos solamente podían haber surgido en el lugar en que surgieron. »

«Para la demo de Gates en abril de 2019, OpenAI ya había aumentado la escala del GPT-2 hasta convertirlo en algo moderadamente más grande. Sin embargo, a Amodei no le interesaba una expansión moderada. Si el objetivo era ampliar el plazo de ejecución de OpenAI, GPT-3 tenía que ser tan grande como fuera posible. Microsoft estaba a punto de entregar un superordenador a OpenAI como parte de su inversión, con 10.000 Nvidia V100, que eran las unidades de procesamiento gráfico más potentes para entrenar modelos de aprendizaje profundo. (La V se debía al químico y físico italiano Alessandro

Volta.) Amodei quería utilizar todos los chips a la vez para crear el nuevo gran modelo lingüístico.[...]»

«Sutskever estaba ansioso por poner en práctica su hipótesis de escalado de los Transformers; Brockman quería seguir haciendo que la empresa creciera; Altman presionaba para lograr el máximo impulso posible. Poco después, Amodei fue ascendido a vicepresidente de investigación. Entre bastidores, Altman también era consciente de otro factor. La inversión de 1.000 millones de dólares de Microsoft tenía 1.000 millones de expectativas; OpenAI estaba trabajando para ofrecer algo que justificase el desembolso. [...] A lo largo de los meses siguientes, Amodei y Altman discreparon sobre cómo y cuándo presentar el GPT-3; Altman acabó imponiéndose, sacando el modelo a la luz en un plazo muy breve. Años antes de la aparición de ChatGPT, esas dos decisiones —la de ampliar el tamaño del GPT-3 y la de lanzarlo rápido— cambiarían el curso del desarrollo de la IA.»

ASCENSIÓN

«“De lo que estoy más orgulloso es de que realmente construimos un imperio”, dijo Altman después de dejar el cargo de presidente. El final de su era en YC marcó el inicio de su nueva era en OpenAI. En marzo de 2019, cuando pasó a dedicarse a OpenAI a tiempo completo, trajo consigo rápidamente la misma mentalidad agresiva que había empleado en YC. No quería que OpenAI estuviera entre las organizaciones de IA más importantes del mundo, quería que fuera la única. [...] “Todavía necesitamos multiplicar nuestros avances por mucho más de diez para llegar a la AGI”, añadía en la circular. “Tendríamos que aspirar siempre a resultados espectaculares, no a mejoras incrementales.” Esta fórmula de éxito requería otros aspectos esenciales. Para OpenAI era de capital importancia tener contenta a Microsoft para seguir a la cabeza en el ámbito de la computación. Si OpenAI tenía éxito, Microsoft había acordado que aportaría mucho más que 1.000 millones de dólares. “Nos gustaría que Microsoft fuera nuestro socio principal hasta el final”, dijo Altman.»

««Parafraseando la famosa cita de Disney», escribió Altman, «tenemos que ganar más dinero para poder realizar más investigaciones, no realizar más investigaciones para ganar más dinero.» Asimismo, la empresa tenía que empezar a reducir la transparencia. «El riesgo informativo de hablar de la AGI seguirá aumentando a medida que vayamos avanzando», sostenía Altman. Había llegado el momento de restringir las publicaciones sobre investigación y el despliegue de modelos, de adoptar una política de confidencialidad más estricta y revelar los avances únicamente cuando se refiriesen a habilidades específicas en lugar de a la AGI en general. [...] Por último, la empresa tenía que empezar a actuar con más seriedad y más unidad.»

«“Debemos poner esa misión por delante de cualquier preferencia individual. No deberíamos dramatizar las cosas de bajo riesgo y así guardarnos nuestra capacidad de dramatizar para las cosas de alto riesgo (que habrá muchas).”, [Altman] De hecho, el drama ya se estaba fraguando. Pequeñas desavenencias que habían surgido en la empresa estaban empezando a agravarse. Brockman y los hermanos Amodei, que en su día se habían considerado amigos, se enfrentaban ahora a causa de una lista de problemas cada vez más larga.[...] Mientras tanto, el frente de seguridad de la IA de Amodei se sentía cada vez más inquieto ante el comportamiento de Altman. Poco después de que se certificase el acuerdo de Microsoft con OpenAI, varios de ellos se quedaron atónitos al descubrir el alcance de las promesas de Altman a Microsoft sobre las tecnologías a las que tendría acceso a cambio de su inversión. Los términos del acuerdo no estaban en línea con lo que habían entendido de Altman. Les preocupaba que, si surgían problemas de seguridad en

la IA en los modelos de OpenAI, esos compromisos harían que fuera mucho más difícil, cuando no imposible, impedir el desarrollo de los modelos. Los partidarios de Amodeli empezaron a tener serias dudas sobre la honestidad de Altman.[...]

«El propio Altman estaba paranoico ante la posibilidad de que la gente filtrara información. A puerta cerrada, desconfiaba del personal de Neuralink, con el que OpenAI seguía compartiendo oficina, ahora con más inquietud tras la marcha de Musk. A Altman también le preocupaba Musk, que disponía de un amplio servicio de seguridad que incluía chóferes y guardaespaldas. Plenamente consciente de la diferencia de capacidad, en un momento dado Altman encargó en secreto una auditoría de contravigilancia electrónica en un intento de registrar la oficina en busca de cualquier micrófono que Musk hubiera podido dejar para espiar a OpenAI. Con los empleados, Altman recurrió al fantasma de que los enemigos de EE. UU. estaban avanzando en la investigación de la IA más rápido que OpenAI para así justificar por qué la empresa tenía que ser cada vez menos transparente y trabajar lo más rápidamente posible. [...]

«Paralelamente, Altman y Brockman desarrollaron un plan de comercialización. A finales de enero de 2020, Brockman empezó a escribir las primeras líneas de código de una interfaz para programación de aplicaciones, o API, para el GPT-3. La API proporcionaría a las empresas y a los usuarios acceso a las capacidades del modelo pero no a los puntos clave, permitiéndoles incorporar la tecnología en sus propios productos enfocados al consumidor. La empresa se dividió en dos. [...] Dicha separación agrandó la brecha que Altman había identificado. La formación de la división de Aplicación atrajo a un pequeño pero creciente grupo de personas procedentes de otras startups que reforzaron el clan Startup. Aunque el clan de Investigación Exploratoria veía con cierta ambivalencia la posibilidad de que OpenAI se convirtiera simplemente en una empresa más de productos de Silicon Valley, provocó la oposición cada vez más intensa de Amodeli y su clan de Seguridad, que también formaba parte de la división de Investigación. Para muchos en Seguridad, lanzar el GPT-3 rápidamente a través de una API o de cualquier otro medio socavaba el plazo de entrega —la clave del escalado acelerado— del que dispondría OpenAI para perfeccionar la seguridad del modelo.»

«En las reuniones corporativas, Altman jugaba a dos bandas: la API acabaría ayudando a cada grupo a lograr lo que quería; conseguir ingresos permitiría a OpenAI invertir aún más en investigación de seguridad de la IA. Cuando el GPT-3 finalizó su entrenamiento, los empleados empezaron a jugar con el modelo a nivel interno. Probaron los límites de sus capacidades y trastearon con la primera versión de la API. La empresa organizó un *hackathon* en que los empleados barajaban diferentes ideas de aplicación. Sin embargo, con cada nuevo prototipo, las tensiones se agudizaban. [...]

«Entre bastidores, más de uno, incluido Dario, discutía sobre sus preocupaciones acerca del comportamiento de Altman con miembros individuales del consejo: Altman había tomado todas las decisiones de OpenAI sobre el acuerdo con Microsoft y el despliegue del GPT-3 como si se tratara de algo inevitable, pero había maniobrado y manipulado a los disidentes para hacerles creer que su voz sería escuchada hasta que fue demasiado tarde para cambiar de rumbo. No solo creían que ese enfoque sería catastrófico y existencialmente peligroso, sino que había demostrado ser doloroso a nivel personal para algunos y había erosionado la cohesión del equipo directivo. Los hermanos Amodeli les describían las tácticas de Altman a las personas de su entorno como «luz de gas» y «maltrato psicológico». [...]

CIENCIA EN CAUTIVIDAD

«A medida que el equipo de generación de código obtenía cada vez mejores resultados, Altman fue apareciendo regularmente en las reuniones, animando a los investigadores a que siguieran adelante y dieran lo mejor de sí mismos a Microsoft. En primavera, después de impresionar a los ejecutivos del gigante tecnológico con varias demos, estaba claro que el modelo iba a ser el segundo proyecto comercial de OpenAI por detrás del GPT-3. [...] Según la circular, Microsoft debería plantearse cancelar el producto o, como mínimo, destinar un porcentaje de las ganancias generadas por el producto a la comunidad de código abierto.»

«Mientras OpenAI concentraba sus apuestas, Altman aplicaba la misma estrategia a sus otros proyectos e inversiones. A lo largo de los años, a medida que se había ido desplazando hacia la innovación tecnológica, había desarrollado la convicción de apostar fuerte y durante mucho tiempo por los proyectos más importantes. [...] En 2021, Altman llevó a cabo las dos mayores inversiones de su vida: 180 millones de dólares en una empresa antiedad llamada Retro Biosciences, que trabajaba para aumentar la esperanza de vida mediante el rejuvenecimiento celular, y 375 millones en Helion Energy, que se dedicaba a la comercialización de la fusión nuclear. «Básicamente, cogí todo mi efectivo y lo metí en esas dos empresas», le dijo Altman a Antonio Regalado, de *MIT Technology Review*. Altman describió ambas tecnologías en un lenguaje que reflejaba la hoja de ruta de OpenAI: parecían imposibles en ese momento, pero, si se aumentaba su escala drásticamente, podrían estar a la vuelta de la esquina. La apuesta por Retro Biosciences reflejaba la obsesión de Altman por la longevidad. [...] 2021 fue también el año en que Altman trasladó a OpenAI su predilección por invertir. Aquel mayo lanzó el OpenAI Start-up Fund, un fondo de inversiones para apoyar a empresas que se encontraban en su primera fase con lo que describió como «grandes ideas sobre cómo usar la IA para transformar el mundo». Una vez más, Microsoft se convirtió en uno de los inversores. Según algunos observadores, la creación del fondo fue una decisión extraña. OpenAI apenas generaba ingresos, así que, ¿por qué recaudar aún más dinero para que lo utilizaran empresas separadas? A otros les parecía que era la forma de Altman de recrear los poderosos efectos de red de YC en torno a OpenAI.»

CAPITALISMO DE DESASTRE

«A medida que OpenAI avanzaba a toda velocidad, guiada por las convicciones de Altman, los límites de las arrolladoras consecuencias de la visión de la empresa se iban ampliando. Con la incorporación de conjuntos de datos cada vez más grandes y contaminados en sus modelos, había creado el “cambio de paradigma” que me explicó Ryan Kolln, de Appen: pasar de filtrar las entradas de datos a controlar los resultados de los modelos. El lenguaje abstracto disfraczaba de nuevo una triste realidad: lo que significaba realmente ese cambio para las personas que ahora se llevaban la peor parte del control de esos resultados.[...]»

«Para crear el filtro automático, OpenAI necesitaba en primer lugar trabajadores humanos que pudiesen revisar y catalogar cuidadosamente cientos de miles de ejemplos del contenido (sexo, violencia y abusos) que la empresa quería evitar que generasen sus modelos. Tras seis meses de búsqueda, encontró un proveedor que parecía adecuado para hacerse cargo del proyecto: una empresa externa que había estado realizando tareas de moderación de contenidos para Meta desde 2019 y que, casualmente, compartía el mismo apodo que Altman: Sama. OpenAI envió un correo electrónico a

Sama preguntándole si aceptaba proyectos que implicasen contenidos delicados o explícitos y cuál era su estrategia habitual para abordarlos. Sama respondió de manera pormenorizada. OpenAI firmó cuatro contratos con la empresa por valor de 230.000 dólares, dejando el proyecto en manos de unos cuantos trabajadores de Kenia. No es casualidad que Kenia se convirtiera en el centro de la que acabaría siendo una de las formas de explotación laboral más evidentes que intervino en la creación del ChatGPT. Era uno de los principales destinos en los que Silicon Valley había estado subcontratando sus trabajos más sucios durante años. Kenia comparte un denominador común con los muchos otros países que la industria tecnológica relega a ese papel: es pobre, se encuentra en el Sur Global y tiene un Gobierno ávido de inversiones extranjeras de países ricos.[...]

«De manera que Kenia se convirtió en un centro neurálgico de la labor de control de Internet. Diversas empresas como Sama —intermediarios en la cadena de suministro de la mano de obra encargada del control de datos— establecieron sus operaciones en Nairobi, creando fondos de asalariados para trabajar en empresas tecnológicas extranjeras, principalmente de la zona de la Bahía de San Francisco. A OpenAI le parecía que Sama cumplía todos los requisitos.[...] »

«[...] Casi doscientos trabajadores interpusieron múltiples demandas contra Sama y Meta, alegando condiciones laborales traumáticas y rescisiones ilegales por tratar de organizarse para exigir mejoras salariales y mejores condiciones laborales. El portavoz de Sama rechazó las acusaciones. En este contexto, OpenAI inició la primera fase de su proyecto a finales de 2021. Bajo los nombres en clave PBJ1, PBJ2, PBJ3 y PBJ4, empujó a equipos de trabajadores de Sama a realizar un trabajo de moderación de contenidos más traumático por un salario medio de entre 1,46 y 3,74 dólares la hora. Los trabajadores no tenían ni idea de para quién ni por qué se encargaban del proyecto.»

«Del mismo modo que la primera era de la comercialización de la IA sentó las bases de la recopilación de datos y la capitalización de la computación de la era de la IA generativa, también estableció los principios de su explotación laboral a gran escala. En este sentido, es importante conocer estas bases para entender la experiencia de los trabajadores kenianos con- tratados por OpenAI. Solo entonces será posible reconocer que sus experiencias distaban mucho de ser anómalas, sino que eran más bien una consecuencia directa del permanente trato dispensado por la industria de la IA a sus trabajadores ocultos y sus opiniones sobre qué trabajo se valora y cuál no, con el objetivo imperialista de OpenAI de alcanzar una escala sin precedentes. [...]»

«Esa «extraña coincidencia» puso de relieve una fórmula inquietante. Al enfrentarse al colapso económico, en Venezuela se dio la combinación perfecta de condiciones para ofrecer una fuente inagotable de mano de obra barata: su población tenía un alto nivel educativo, buena conexión a Internet y, ahora, un ferviente deseo de trabajar por cualquier salario. No era el único país que se ajustaba a esa descripción. Otras poblaciones disponían de una conexión a Internet de más calidad. Con el agravamiento del cambio climático y la creciente inestabilidad geopolítica, era difícil apostar en contra de que más poblaciones se sumieran en la crisis. «Es bastante probable que haya otra Venezuela», dijo Schmidt. En ese momento, la predicción de Schmidt me hizo preguntarme si la segunda vez seguiría siendo una coincidencia o si las empresas de anotación de datos elaborarían un manual de estrategia a partir de lo que había funcionado allí. [...] Varios años más tarde, si echamos la vista atrás, eso es exactamente lo que sucedió y lo que se ha convertido en uno de los paralelismos más asombrosos entre los imperios de la antigüedad y los imperios de la IA.

Una de las características definitorias de la rápida acumulación de riqueza de un imperio es su capacidad de pagar muy poco o nada para recoger los beneficios económicos de una amplia base de mano de obra humana.»

«Lo cierto es que Scale estaba incorporando a nuevos usuarios. A mediados de 2021, conforme los venezolanos se quemaban y abandonaban la plataforma, Scale localizaba e incorporaba a decenas de miles de usuarios procedentes de otras economías que se habían desplomado durante la pandemia. Para cubrir las necesidades de sus clientes, cada vez más amplias y diversas, recurrió a países con gran densidad de población que se enfrentaban a graves dificultades económicas y que, además, hablaban las lenguas más valiosas desde un punto de vista económico: inglés, francés, italiano, alemán, chino, japonés y español. Buscó hablantes de francés en las antiguas colonias francesas de África, recuerda un expleado que trabajaba en expansión internacional; buscó hablantes de chino mandarín en lugares con grandes poblaciones de la diáspora china como el Sudeste Asiático. Scale siguió repitiendo la estrategia desarrollada en Venezuela una y otra vez. [...] Ofrecía elevados ingresos en cada mercado nuevo para atraer trabajadores y, una vez se asentaba, reducía dichos ingresos. Variaba el importe de los pagos a los trabajadores mediante una serie de experimentos que los empleados fijos desde su sede central de San Francisco, ahora de 17.000 metros cuadrados, calificaban como optimización e innovación. Los trabajadores, mientras tanto, veían cómo sus medios de vida disminuían drásticamente con los cambios impredecibles.»

«En ese sentido, el trabajo era claramente distinto de la moderación de contenidos tradicional. Mientras que los moderadores de contenidos de Meta revisaban entradas reales generadas por usuarios para determinar si debían permanecer o no en Facebook, Okinyi y su equipo se dedicaban a anotar contenidos para entrenar al filtro de moderación de contenidos de OpenAI, con el objetivo de evitar de entrada que los modelos de la empresa generasen dichos resultados. Para que los ejemplos fueran suficientemente amplios, algunos de ellos eran concebidos, al menos en parte, por el propio software de la empresa para imaginar lo peor de lo peor.[...] Cuando escribí el artículo sobre Okinyi y los otros tres trabajadores kenianos para *The Wall Street Journal*, OpenAI intentó eludir la responsabilidad del peaje que tuvieron que pagar los trabajadores por su proyecto.[...] Sin embargo, la coincidencia de las experiencias de los trabajadores en el tiempo y espacio demuestra que la explotación laboral que sustenta la industria de la IA es sistémica. [...]»

«El acuerdo de OpenAI con Sama fue solo una parte de la extensa red de mano de obra humana reunida durante dos años para producir lo que sería ChatGPT. La empresa dijo que también utilizaba a más de mil trabajadores autónomos en EE. UU. y el resto del mundo para ajustar sus modelos con un aprendizaje por refuerzo a partir de la retroalimentación humana, la técnica que había desarrollado para incrementar la seguridad de la IA. Para conseguir a esos trabajadores, re- currió en gran medida a la misma plataforma que se convirtió en el elemento principal de la primera era de la comercialización de la IA mediante la aplicación de su manual para tiempos de crisis: Scale AI.»

«[...] La asociación entre OpenAI y Scale se fraguó en parte gracias a una relación personal: Alexandr Wang, el actual consejero delegado de Scale que se convirtió en el multimillonario hecho a sí mismo más joven del mundo en 2021, es buen amigo de Altman. En 2016, Wang había entrado a formar parte de la última promoción de emprendedores de YC con la idea de fundar una startup diferente y la puso en práctica con Scale, otorgándole a Altman una participación indirecta en la empresa a través de YC. En un momento dado, durante la pandemia, ambos compartieron piso durante va- rios

meses. Según *The Information*, en otoño de 2023, llegaron incluso a discutir sobre la posibilidad de que OpenAI adquiriera Scale.»

DIOSES Y DEMONIOS

«Con la identificación de una teórica IA descontrolada como un riesgo para la existencia, el altruismo eficaz promulgaba el mismo tipo de seguridad de la IA entrelazado en el ADN de OpenAI desde el principio y había desempeñado un papel decisivo en El Divorcio. Amodei y sus compañeros fundadores de Anthropic discrepaban radicalmente con Altman y el resto de los ejecutivos de OpenAI sobre hasta qué punto tomarse en serio la posibilidad de que la IA devastara la civilización. A Amodei, que sí se la tomaba muy en serio, el comportamiento de Altman —su falta de transparencia en cuanto al acuerdo con Microsoft; su aparente tendencia a decirle siempre a la gente lo que quería oír para contar con su aprobación, solo para que el enredo se descubriera cuando ya era demasiado tarde— no solo le parecía la típica maquinación de un ejecutivo de Silicon Valley, sino un comportamiento alarmante e inhumano que podía poner en peligro el destino de la humanidad. Cuando Anthropic se asentó, se apoyó en ese rasgo diferencial: mientras que la OpenAI de Altman jugueteaba de manera negligente con el futuro de la humanidad, Anthropic era la empresa regida por principios y que situaba la seguridad de la IA como su primera prioridad»

«Muchos no estaban seguros de la fidelidad de Altman y citaban diferentes momentos en los que parecía simpatizar con ambas. Los más partidarios de la beneficencia consideraban que estaba a medio camino, lidiando con el duro trabajo de representar los diferentes puntos de vista de la empresa. Sin embargo, a partir del Divorcio y del distanciamiento entre Altman y los cofundadores de Anthropic, cada vez más Doomers empezaron a tener la peor opinión posible de Altman. Muchas de las cosas que situaron a OpenAI en el mapa y que le aportaron cada vez más éxito comercial habían empezado como proyectos de seguridad de la IA: leyes de escudo, generación de códigos, aprendizaje por refuerzo a partir de la retroalimentación humana, la combinación de las tres cosas en un gran lenguaje increíblemente persuasivo y luego los modelos multimodales. Muchos Doomers tenían la impresión de que su trabajo estaba siendo utilizado y retorcido para lograr algo radicalmente opuesto a sus valores fundamentales. En su opinión, Altman era quien lo estaba utilizando y retorciendo. Y eso lo convertía en un mentiroso patológico, un abusador y manipulador y una amenaza en sí mismo para la humanidad.[...]»

«Al poco tiempo, el choque entre esas ideologías polarizadas dentro de OpenAI y su entorno amenazaría con hacer trizas la empresa que había hecho más que ninguna otra por marcar la pauta de la nueva era del desarrollo de la IA. Sin embargo, por mucho que cada ideología profesara lo contrario de la otra, en realidad las dos predicaban lo mismo. Las dos debatían sobre la IA como una conclusión inevitable y con una vehemencia religiosa; ambas se centraban en el largo plazo y reivindicaban la autoridad moral de mantener el desarrollo de la IA bajo el control de sus adeptos.[...]»

«El GPT-4 disponía ahora de capacidades multimodales integradas y, en contra de los cálculos internos de OpenAI, estaba generando código más pulido que nunca y era más ágil a la hora de reconocer la intención del usuario y de proporcionar respuestas útiles. En una impresionante muestra de dichas habilidades, Brockman retransmitiría más adelante en directo una demo en la que aparecía él mostrándole al GPT-4 una foto de un sencillo esquema de una página web garabateado en su cuaderno de notas. [...] Sin embargo, cuando OpenAI empezó a presentar el modelo en círculos de confianza, incluyendo

inversores y clientes selectos, hubo al menos una persona a la que no le impresionó ni lo más mínimo. Una vez más, se trataba del siempre difícil de complacer Bill Gates. [...] Brockman se tomó el comentario de Gates como un reto.»

«OpenAI lanzaría el GPT-4 en la API junto con el producto orientado al consumidor basado en el GPT-4. A los empleados, Altman les vendió la decisión de retrasar el lanzamiento del modelo como una prueba de la actitud cautelosa y centrada en la seguridad de OpenAI. La empresa se estaba tomando su tiempo para llevar a cabo el despliegue del sistema, dijo. De hecho, la directiva ni siquiera estaba segura de desplegarlo. En Aplicación, la mayoría de la gente interpretó que lo que Altman quería decir era que detener el lanzamiento supondría un problema extraordinario para la directiva, pero que, si llegaba a darse la situación, estaban dispuestos a hacerlo. Para muchas personas dentro de Seguridad, las palabras de Altman tuvieron un sentido diferente: la idea era que únicamente cuando el GPT-4 superara todas las pruebas, la dirección daría luz verde a su lanzamiento. En reuniones privadas, Altman reafirmaba la percepción de cada bando. A los dos les planteaba dudas de que el GPT-4 pudiera «despertar a Google», evocando un escenario en el que, el día que cayese el modelo y captase la atención de Larry Page y Sergey Brin, los dos acabarían eliminando la burocracia política de Google.[...]»

«Desde hacía ya algún tiempo, los directivos tenían la impresión de que la lista de espera se había descontrolado y el proceso de revisión no aumentaba de escala. Los desarrolladores se quejaban de lo mucho que se tardaba en tener acceso a las tecnologías de OpenAI, y enviaban constantemente correos electrónicos a Altman, Brockman y Peter Welinder, o los etiquetaban en Twitter mencionando grandes retrasos en las aplicaciones que los ejecutivos consideraban totalmente inocuas. En Aplicación, muchos pensaban que OpenAI estaba impidiendo el acceso a los beneficios de su tecnología, lo cual era contrario a su misión.[...]»

«El GPT-4 no fue solo un punto de inflexión para Gates y Microsoft. Posteriormente, aquel mismo verano, tras cautivar al filántropo multimillonario, Altman y Brockman lo presentaron ante el consejo de OpenAI, donde Brockman hizo una demostración en directo del modelo que incluía chistes sobre Gary Marcus. Los chistes les encantaron a los miembros independientes del consejo; la desenfadada exposición les mostró también un avance considerable de las capacidades de la tecnología y les transmitió la sensación de que lo que se jugaban en sus futuras decisiones era cada vez más importante. Hasta ese momento, Altman había reunido al consejo aproximadamente una vez cada trimestre y prefería exponer sus actualizaciones oralmente. Trataba sin dificultad temas de investigación complejos, convocando a veces a investigadores de la empresa para que presentaran sus avances, y resumía rápidamente los últimos acuerdos que estaban en proceso de negociación con Microsoft u otros socios. Algunos de los miembros independientes del consejo presionaban para que las reuniones fueran más frecuentes y la información más estructurada, e insistían en que Altman presentara informes escritos y les permitiera el acceso a más documentos. A Altman le irritaba el aumento de la supervisión.»

CIMA

«El desenfrenado éxito instantáneo del ChatGPT fue más allá de lo que nadie había soñado en OpenAI. Los ingenieros de la empresa se quedaron completamente mosqueados, incluso años después. El GPT-3.5 no había representado una mejora tan importante respecto al GPT-3, que ya llevaba en el mercado dos años. Y el GPT-3.5 ya había sido puesto a disposición de los desarrolladores. La interfaz y el formato habían hecho que el

modelo fuera más accesible, sin duda, pero no representó el cambio decisivo que los empleados habían notado con el GPT-4. Altman dijo más adelante que él creía que el ChatGPT sería popular, pero algo así como «un orden de magnitud menos». «Fue impactante que a la gente le gustara», recuerda un ex empleado. «Para todos nosotros, lo que habían hecho era rebajar lo que utilizábamos internamente y lanzarlo.»

Al cabo de cinco días, Brockman tuiteó que el ChatGPT había superado el millón de usuarios. Al cabo de dos meses, había alcanzado los cien millones, convirtiéndose en la aplicación para consumidores de más rápido crecimiento de la historia hasta la fecha. (Threads, el rival de X de Meta, reclamó posteriormente ese título al alcanzar el mismo número de usuarios en menos de cinco días; los expertos sostuvieron que no era así, ya que Meta estaba recurriendo básicamente a una base de usuarios ya existente.) ChatGPT catapultó a la fama a OpenAI, que pasó de ser una startup de moda, conocida en el ámbito de la industria tecnológica [...] Inmediatamente después, toda la empresa estaba apagando fuegos. Los servidores de OpenAI se colapsaron repetidamente mientras el equipo de infraestructura luchaba por aumentar su capacidad lo más rápidamente posible, en el plazo más ajustado de la historia de Silicon Valley. El equipo reutilizó parte de la computación de la división de Investigación para contribuir al crecimiento del ChatGPT y aun así no tenía suficiente para hacer que la aplicación siguiera funcionando.»

«Altman no se durmió en los laureles. Les recordó a los empleados que, en definitiva, la empresa tenía una misión que cumplir; algo mucho más importante que crear «el producto más importante de la historia de Silicon Valley». [...] Los equipos no daban abasto, así que los directivos le suplicaron a Altman que les proporcionara más personal. No faltaron candidatos. Tras la aparición del ChatGPT, el número de solicitantes de empleo que clamaban por subirse al cohete se había multiplicado rápidamente. Sin embargo, a Altman le preocupaba qué pasaría con la cultura de la empresa y el alineamiento de la misión si la empresa aumentaba su plantilla demasiado rápido. Creía firmemente en mantener una plantilla reducida con una gran densidad de talento.»

««Si lo que quieres es crear Meta, estás haciendo un gran trabajo», le dijo lanzándole una pulla a Altman, aludiendo a los propios temores de los que el consejero delegado había advertido, con relación a que la empresa diluyese rápidamente su densidad de talento y su orientación, incrementando la burocracia. La velocidad de expansión estaba provocando también un repunte de los despidos. Durante su incorporación, a un directivo se le pidió que documentara e informara rápidamente de cualquier miembro de su equipo que presentara un bajo rendimiento, siendo él mismo despedido al poco tiempo. Las rescisiones casi nunca se comunicaban al resto de la empresa [...]»

«A los nuevos empleados, plenamente convencidos de que entraban a formar parte de una startup dinámica y rentable, tanto movimiento les parecía una manifestación especialmente caótica y, en ocasiones, brutal, de los habituales problemas empresariales: mala gestión, prioridades poco claras y la crueldad despiadada de una empresa capitalista dispuesta a tratar a sus empleados como material de usar y tirar. [...]»

«ChatGPT codificó firmemente el paso de OpenAI de organización sin ánimo de lucro hacia la comercialización. Altman y otros directivos presionaron en favor de aprovechar el éxito del chatbot para lanzar un montón de productos de pago. En febrero de 2023, lanzó una versión de pago del ChatGPT; en marzo, lanzó sucesivamente una versión API, la API Whisper y, por último, el GPT-4. «Después del ChatGPT, había una vía que conducía

claramente a ingresos y beneficios», dice un antiguo empleado. «Ya no se podía argumentar que era un laboratorio de investigación idealista. Había clientes que esperaban un servicio aquí y ahora.» La explosión de nuevos productos superó de nuevo las previsiones del equipo de confianza y seguridad [...]

TIERRA SAQUEADA

«A partir del momento en que se comprometió con la idea de escalado, OpenAI trató de lograr una infraestructura informática sin precedentes. «En la IA, quien tiene el ordenador más grande es el que tiene más beneficios», me dijo Brockman en 2019. Así que, con Microsoft, Altman desarrolló un plan sobre cómo el gigante tecnológico podría cubrir las necesidades informáticas de OpenAI que iban aumentando de manera exponencial. [...]

«[...] Altman y otros directivos nunca sacaban a colación el tema del coste medioambiental de los centros de datos en las reuniones corporativas. Mientras OpenAI entrenaba al GPT-4 en Iowa, el estado llevaba dos años de sequía. [...] Lo que Altman sí puso sobre la mesa fue su impaciencia. En marzo de 2024, tras dormirse en los laureles durante los primeros años de la carrera por la IA generativa, Meta irrumpió con nuevos planes agresivos para tener 350.000 H100 funcionando como parte de una infraestructura aún mayor, construida para respaldar su repentina oleada de inversiones en IA generativa. Eran más GPU de las que OpenAI tenía a su disposición. Altman no estaba contento. Le parecía que Microsoft iba demasiado despacio y le estaba costando a OpenAI su ventaja competitiva. [...]

LOS DOS PROFETAS

«En mayo de 2023, Altman llegó a Washington D.C. para testificar ante el Congreso. Fue una intervención extraordinaria. Reiteró la promesa de que la AGI solucionaría el cambio climático y curaría el cáncer, argumentó de manera convincente por qué las tecnologías de OpenAI mejorarían y crearían nuevos empleos «fantásticos», esquivó las preguntas sobre problemas de derechos de autor, falta de transparencia y garantías de privacidad de sus datos de entrenamiento, e hizo un llamamiento sincero a la regulación, es decir, a una regulación auspiciada por OpenAI, invocando el espectro de China para instar a los legisladores a que no ralentizaran la innovación. A los senadores les encantó. En un revelador diálogo que reflejó la ágil retórica de Altman y la confianza y el entusiasmo que despertaba, planteó tres recomendaciones políticas que desviaron la conversación de problemas existentes como la mano de obra, el medioambiente y la propiedad intelectual para centrarse en la regulación de los sistemas de IA y los riesgos extremos futuros. [...]

«Según *The New York Times*, a principios de junio, Altman se había reunido personalmente con al menos cien responsables políticos de EE. UU., algunos de los cuales hicieron referencia con orgullo a esas conversaciones durante la vista [...]

«El mismo discurso esgrimido durante mucho tiempo por Altman en OpenAI para justificar la ocultación de sus investigaciones y el hecho de avanzar lo más rápidamente posible, se utilizaba ahora hábilmente para enfocar el debate regulador de la IA de EE. UU. hacia propuestas que eludiesen la responsabilidad de OpenAI y, en algunos casos, reafirmar su monopolio. La carta ganadora de Silicon Valley de «¿Qué pasa con China?» había obrado maravillas sistemáticamente para eludir la regulación. Ahora tenía más fuerza que nunca, [...]

«Un punto esencial de las recomendaciones de Altman y la idea del modelo fronterizo era la asociación del escalado de un modelo con capacidades emergentes y, por tanto, posiblemente peligrosas. Ese argumento se basaba en la filosofía de las leyes de escala, según la cual más entrenamiento conllevaría previsiblemente modelos más potentes, así como en la creencia de los círculos Doomer de que una IA muy avanzada podría descontrolarse.[...]»

«[...]La secuencia completa de los acontecimientos —la declaración de Altman, el informe, la campaña sin cuartel de influencia política, la desmedida reacción de Washington por temor a China y la apresurada consagración de los umbrales de cómputo en documentos políticos relevantes en EE. UU. y en el extranjero— fue una escueta ilustración de, entre otras cosas, hasta qué punto se había deteriorado gran parte del conocimiento independiente sobre IA.»

«Se dio cuenta de que las personas que estaban sentadas a su lado y sus enormes equipos políticos habían monopolizado el mensaje en Washington durante tanto tiempo que ahora muchos legisladores pensaban que sus palabras iban a misa. Un portavoz de Schumer declaró posteriormente a la prensa que el senador consultaba personalmente con Altman y otros directivos de OpenAI a medida que se aproximaba a la regulación. [...] Washington fue únicamente el punto culminante del tramo estadounidense de la ofensiva política de Altman. En marzo de ese año, después de tuitear que tenía previsto viajar al extranjero para reunirse con usuarios, su viaje se había convertido en una odisea por múltiples ciudades y continentes para intentar fotografiarse aparentemente con todos los presidentes del G20. Ahora se llamaba la «gira mundial» de Sam Altman»

«[...] Era una manifestación de una dinámica que siempre había estado presente: Altman a su manera.[...] Mientras Altman daba vueltas por Europa, Latinoamérica, Oriente Próximo, Asia y África, deslumbrando —y solo en contadas ocasiones ofendiendo— a audiencias cuidadosamente seleccionadas, compuestas por estudiantes, inversores tecnológicos y fans, la falta de claridad estratégica exacerbaba las antiguas líneas divisorias de OpenAI e impulsaba a la empresa a posiciones extremas más enfrentadas que nunca.»

«[...]Durante casi ocho años, la analogía que hizo Altman en sus primeros correos electrónicos a Musk, entre OpenAI y el Proyecto Manhattan, había sido un tema recurrente en la empresa, utilizado incluso durante la fase de orientación a los nuevos contratados. A Altman le gustaba. Su fecha de nacimiento era la misma que la de Oppenheimer, cosa que mencionaba a los periodistas. También le gustaba parafrasear al creador de la bomba que creía que «la tecnología sucede por lo que es posible». Al parecer, nunca añadía que Oppenheimer se pasó la segunda mitad de su vida sumido en el arrepentimiento y haciendo campaña en contra de la difusión de su propia creación.»

«[...]También confiaban en el propio Altman como medio para mantenerse al día de las informaciones importantes. Lo que les preocupaba cada vez más era lo mucho que el discurso de Altman difería a menudo de otras informaciones que les llegaban. Mientras que, habitualmente, Altman planteaba una situación de color de rosa, los consejeros recibían cada vez más informes de sus propias fuentes sobre problemas, incluyendo la falta de preparación antes del ChatGPT y el considerable revuelo que tuvo lugar posteriormente, las constantes dudas acerca de la seguridad de la IA que rodeaban el lanzamiento del GPT-4 y el ritmo sin precedentes al que avanzaba OpenAI antes de haber solucionado muchos de sus problemas. [...] Al mismo tiempo, había otros ejemplos

preocupantes del comportamiento de Altman. En marzo de 2023, envió un correo electrónico al consejo sin incluir a D'Angelo, anunciando que creía que había llegado el momento de que D'Angelo dejase su cargo. El hecho de que Quora estuviese diseñando su propio chatbot, Poe, argumentaba Altman, planteaba un conflicto de intereses. La afirmación parecía improvisada y escasamente motivada. Tanto Toner como McCauley y D'Angelo le habían formulado en diversas ocasiones preguntas incómodas a Altman sobre las prácticas de seguridad de OpenAI, la solidez de la organización sin ánimo de lucro y otros asuntos. Cada vez con menos aliados en el consejo, a los tres les parecía que Altman estaba buscando una excusa para echar a alguno de ellos. En respuesta al correo electrónico de Altman, un consejero independiente contraatacó. Poe representaba un conflicto de intereses moderado en comparación con lo que Altman les había permitido durante tiempo a Hoffman y a Zilis. La moción de Altman fracasó; D'Angelo se quedó.»

«Cuando lo oyó por segunda vez, los consejeros independientes presionaron a Altman para que les mostrara documentación sobre la estructura del fondo. Cuando, finalmente, Altman se la facilitó, los consejeros descubrieron que la estructura del fondo no solo era extraña; Altman era su *propietario* legal, cuando debería haber sido propiedad de OpenAI [...] Para los consejeros independientes, cada ejemplo se sumaba a un panorama preocupante: poco a poco, Altman estaba intentando reducir su visibilidad y maniobrar de tal manera que el consejo no pudiera controlarlo. Durante años, Altman había proclamado la capacidad del consejo para actuar como contrapeso e incluso despedirlo como el mecanismo de gobierno más importante de OpenAI. Y ahora estaba pregonando ese hecho por todo el mundo para asegurarse la confianza del público y de los Gobiernos.[...]»

LIBERACIÓN

«[...]El brusco cambio de imagen de Altman respecto a su magnífica acogida desde el lanzamiento de ChatGP y el hecho de ver cómo dolorosos asuntos familiares salían a la luz, hicieron mella en él. Los efectos colaterales del artículo empeoraron las cosas. Tras su publicación, antiguos tuits de Annie en los que acusaba a Sam de haber abusado de ella de diversas formas reaparecieron y empezaron a hacerse virales Sin embargo, por muy dura y difícil que fuera la situación para Altman, para Annie fue una liberación.»

«[...]La historia de Annie ahonda en los retratos contradictorios que pinta la gente de Sam. Es a la vez generoso y egoísta, agradable y amenazador, un benefactor para muchas personas y fuente de un gran dolor personal para otras. Alguien que transmite sinceridad y altruismo en público, pero que muestra una vena más complicada y calculadora a puerta cerrada. Alguien que puede dar y quitar, dejando a muchos con la impresión de que son peones de una gran partida de ajedrez en la que solo Sam ve el tablero completo y que consiste en que él conserve su poder como rey. La historia de Annie también complica el gran relato de Sam y otros ejecutivos de OpenAI, según el cual la IA abre la puerta a un mundo de abundancia. Altman ha dicho que espera que la IA acabe con la pobreza: Brockman ha repetido, en sus historias sobre su amigo y su mujer, Anna, que la AGI mejorará drásticamente la asistencia sanitaria. Sutskever ha dicho que dará lugar a una psicoterapia extraordinariamente eficaz y tirada de precio. Y, pese a todo, frente a la realidad de las vidas de los trabajadores de Kenia, los activistas de Chi- le y la experiencia de la propia hermana de Altman, que tienen que soportar todos esos problemas, esos sueños parecen vacíos. A pesar de los avances a pasos agigantados de las capacidades de la IA, ninguno de ellos contribuyó a aliviar ni siquiera un poco la desesperación de Annie. En todo caso, puede que la IA haya contribuido a atraparla aún más.»



Para ampliar información, contactar con:

Laura Fabregat (Responsable de Comunicación Área de Ensayo)

M: 682 69 63 61 / E: lfabregat@planeta.es