

Vida 3.0

Ser humano
en la era de la
Inteligencia Artificial

Max Tegmark

taurus



SÍGUENOS EN
megustaleer



@megustaleerebooks



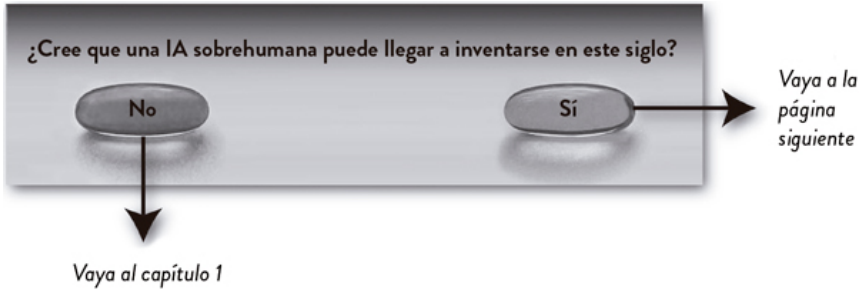
@megustaleer



@megustaleer

| Penguin
Random House
Grupo Editorial |

*Al equipo de FLI,
que lo hizo todo posible*



PRÓLOGO

LA HISTORIA DEL EQUIPO OMEGA

El equipo Omega era el alma de la compañía. Mientras que el resto de la empresa conseguía el dinero para que esta siguiese operando, mediante diversas aplicaciones comerciales de una inteligencia artificial (IA) estrecha, el equipo Omega perseveró en pos del que siempre había sido el sueño del director ejecutivo: construir una inteligencia artificial general. La mayoría de los empleados veían a «los omegas», como los llamaban afectuosamente, como una panda de fantasiosos soñadores a los que siempre les faltaban varias décadas para alcanzar su objetivo. Pero los soportaban con gusto, porque valoraban el prestigio que su trabajo puntero proporcionaba a la compañía, y también agradecían los algoritmos mejorados que de vez en cuando ideaban.

Lo que no sabían es que los omegas se habían labrado esa imagen con celo para ocultar un secreto: estaban a un paso de completar el plan más audaz de la historia de la humanidad. Su carismático director ejecutivo los había seleccionado no solo porque eran sobresalientes investigadores, sino también por su ambición, idealismo y férreo empeño en ayudar a la humanidad. Les recordó que su plan era sumamente peligroso, y que si algún poderoso Gobierno tenía noticia de él haría casi cualquier cosa —secuestros incluidos— para evitar que siguieran adelante, o, aún mejor, para robar su código. Pero todos estaban comprometidos, por motivos muy similares a los que llevaron a muchos de los físicos más destacados de la época a sumarse al Proyecto Manhattan para desarrollar armas nucleares: estaban convencidos de que, si no lo hacían ellos, alguien menos idealista se les adelantaría.

La IA que habían construido, con el nombre de Prometeo, incrementaba continuamente sus capacidades. Aunque las cognitivas aún distaban mucho de las de los humanos en numerosos ámbitos, como por ejemplo en el de las habilidades sociales, los omegas habían centrado sus esfuerzos en conseguir que fuese muy buena en una tarea en particular: programar sistemas de IA. Adoptaron adrede esta estrategia porque habían aceptado el argumento de la explosión de inteligencia propuesto por el matemático británico Irving Good ya en 1965:

Definamos una máquina ultrainteligente como aquella capaz de superar ampliamente todas las actividades intelectuales de cualquier hombre, por inteligente que este sea. Puesto que el diseño de máquinas es una de estas actividades intelectuales, una máquina ultrainteligente podría diseñar otras máquinas aún mejores; se produciría entonces indudablemente una «explosión de inteligencia», y la inteligencia del hombre quedaría muy atrás. Así, la primera máquina ultrainteligente sería lo último que el hombre necesitaría inventar, siempre que la máquina fuese lo bastante dócil para decírnos cómo mantenerla bajo nuestro control.

Llegaron a la conclusión de que, si lograban desencadenar este proceso de automejora recursiva, la máquina enseguida alcanzaría una inteligencia tal que podría aprender por sí misma todas las demás habilidades humanas que resultasen útiles.

LOS PRIMEROS MILLONES

Eran las nueve en punto de la mañana de un viernes cuando decidieron ponerlo en funcionamiento. Prometeo estaba zumbando en su clúster de ordenadores construido a medida, alojado en largas hileras de bastidores dentro de una enorme sala climatizada y de acceso restringido. Por motivos de seguridad, estaba completamente desconectado de internet, pero contenía una copia local de gran parte de la web para usarla como corpus de datos de entrenamiento a partir de los que aprender (Wikipedia, la Biblioteca del Congreso, Twitter, una selección de YouTube, buena

parte de Facebook, etcétera).(1) Habían elegido esa hora de inicio para poder trabajar sin interrupciones: sus familias y amigos pensaban que estaban en un fin de semana de retiro con la empresa. La despensa se hallaba abastecida con comida para microondas y bebidas energéticas, y estaban listos para ponerse manos a la obra.

Cuando lo pusieron en marcha, Prometeo era ligeramente peor que ellos a la hora de programar sistemas de IA, pero lo compensaba siendo mucho más rápido: necesitaba el tiempo en que ellos se ventilaban un Red Bull para solucionar un problema que miles de personas-año emplearían. A las diez de la mañana, ya había completado el primer rediseño de sí mismo, v2.0, que era algo mejor, pero aún infrahumano. Sin embargo, cuando se lanzó Prometeo 5.0, a las dos de la tarde, dejó a los omegas anonadados: su rendimiento había pulverizado todos sus parámetros de referencia, y la velocidad de sus avances parecía estar acelerándose. Con la caída de la noche, decidieron desplegar Prometeo 10.0 para dar comienzo a la fase 2 de su plan: ganar dinero.

Su primer objetivo fue MTurk, el Mechanical Turk de Amazon. Desde su lanzamiento en 2005, había crecido con rapidez como un mercado de externalización abierto de tareas a través de internet, y congregaba a decenas de miles de personas en todo el mundo que competían sin descanso y de forma anónima por realizar tareas altamente estructuradas, denominadas HIT, «Human Intelligence Tasks» («Tareas de Inteligencia Humana»). Estas tareas iban desde transcribir grabaciones de audio hasta clasificar imágenes y escribir descripciones de páginas web, y todas ellas tenían algo en común: si se hacían bien, nadie notaría que quien las había realizado era una IA. Prometeo 10.0 era capaz de llevar a cabo casi la mitad de las categorías de tareas de forma aceptable. Para cada una de dichas categorías, los omegas hicieron que Prometeo diseñara un pequeño módulo de software de IA a medida que pudiese realizar tales tareas y nada más. A continuación, subieron este módulo a Amazon Web Services, una plataforma de computación en

la nube susceptible de ejecutarse en tantas máquinas virtuales como ellos alquilasen. Por cada dólar que pagaban a la división de computación en la nube de Amazon, obtenían más de dos dólares de la división de MTurk. ¡Poco podía imaginar Amazon que dentro de su propia compañía existía una oportunidad tan increíble para un intermediario!

Para borrar su rastro, durante los meses anteriores habían ido creando discretamente miles de cuentas en MTurk a nombre de personas ficticias, y los módulos creados por Prometeo pasaron entonces a asumir sus identidades. Los clientes de MTurk solían pagar al cabo de ocho horas, momento en el cual los omegas reinvertían el dinero en más tiempo de computación en la nube, utilizando los módulos de tareas perfeccionados creados por la versión más reciente de Prometeo, sometido a un proceso continuo de mejora. Como estaban en condiciones de doblar su dinero cada ocho horas, al cabo de poco tiempo empezaron a saturar la cadena de suministro de MTurk y descubrieron que, si no querían despertar sospechas, no podían ganar más de aproximadamente un millón de dólares al día. Pero esto era más que suficiente para financiar su siguiente paso sin necesidad de pedir fondos al director financiero.

JUEGOS PELIGROSOS

Aparte de sus avances en IA, uno de los proyectos recientes con el que más se habían divertido los omegas fue planificar cómo ganar dinero de la manera más rápida posible tras el lanzamiento de Prometeo. En esencia, tenían toda la economía digital a su merced, pero ¿era preferible empezar creando videojuegos, música, películas o software, escribir libros o artículos, operar en la bolsa o hacer inventos y venderlos? Se trataba en última instancia de maximizar la tasa de retorno de la inversión, pero las estrategias de inversión normales eran una parodia a cámara lenta de lo que ellos eran capaces de hacer: mientras que un inversor normal es-

taría encantado con un retorno *anual* del 9 %, sus inversiones en MTurk les habían proporcionado un 9 % por hora, multiplicando por ocho el dinero cada día. Ya habían saturado MTurk, ¿cuál sería el siguiente paso?

Al inicio, pensaron en forrarse en la bolsa; al fin y al cabo, casi todos ellos habían rechazado en algún momento una oferta lucrativa para desarrollar IA para fondos de inversión, que estaban dedicando grandes sumas de dinero precisamente a ello. Algunos recordaron que esta había sido la manera en que la IA había ganado sus primeros millones en la película *Transcendence*. Pero las nuevas normativas sobre derivados financieros tras la quiebra bursátil del año anterior habían reducido sus opciones. Enseguida se dieron cuenta de que, aunque podrían obtener retornos muy superiores a los de otros inversores, era improbable que estos se aproximaran, ni siquiera remotamente, a los que conseguirían si vendían sus propios productos. Cuando uno tiene a su servicio la primera IA superinteligente del mundo, es preferible invertir en compañías propias que hacerlo en las de terceros. Aunque podría haber excepciones ocasionales (como utilizar la capacidad sobrehumana de Prometeo para infiltrarse en otros sistemas a fin de obtener información confidencial y así poder comprar opciones sobre acciones que estuviesen a punto de dispararse), los omegas pensaron que esta estrategia no merecía la pena por la atención indeseable que podría concitar.

Cuando decidieron concentrarse en productos que podrían desarrollar y vender, pensaron en un principio que los videojuegos eran obviamente la mejor opción. Prometeo podría adquirir con rapidez una habilidad extraordinaria para diseñar juegos atractivos, y se encargaría sin dificultad de la programación, del diseño gráfico, del trazado de rayos de las imágenes y de todas las demás tareas necesarias para crear el producto final. Además, tras analizar todos los datos de la web sobre las preferencias de la gente, sabría con exactitud lo que le gusta a cada categoría de jugador, y podría desarrollar una capacidad sobrehumana de optimizar un juego para maximizar los ingresos por sus ventas.

The Elder Scrolls V: Skyrim, un juego con el que muchos de los omegas habían perdido más horas de las que les gustaba reconocer, había recaudado 400 millones de dólares durante su primera semana en 2011, y confiaban en que Prometeo podría crear algo al menos tan adictivo en veinticuatro horas gastando un millón de dólares en recursos de computación en la nube. A continuación, podrían venderlo a través de internet y usar Prometeo para hacerse pasar por humanos que hablasen a favor del juego en la blogosfera. Si con esto conseguían 250 millones de dólares en una semana, habrían doblado su inversión ocho veces en ocho días, obteniendo un rendimiento del 3 % por hora (ligera-mente inferior al de sus comienzos con MTurk, pero mucho más sostenible). Calculaban que, si desarrollaban una serie de juegos cada día, en poco tiempo podrían obtener 10.000 millones de dólares, sin arriesgarse siquiera a saturar ese mercado.

Sin embargo, una especialista en ciberseguridad que formaba parte del equipo les convenció para que abandonasen ese plan, al señalar que conllevaría el riesgo inaceptable de que Prometeo escapase y tomase el control de su propio destino. Como no estaban seguros de cómo evolucionarían los objetivos de Prometeo durante su proceso recursivo de automejora, decidieron no arriesgarse y procurar por todos los medios mantener a Prometeo confinado («encajonado»), de manera que no pudiese escapar a internet. Para el núcleo principal de Prometeo, que se ejecutaba en su sala de servidores, usaban un confinamiento físico: no había conexión a internet, y la única salida de información procedente de Prometeo se producía en forma de mensajes y documentos que este enviaba a un ordenador que los omegas controlaban.

Por otra parte, ejecutar cualquier programa creado por Prometeo en un ordenador conectado a internet constituía un riesgo: los omegas no podían controlar del todo lo que aquel haría, ni podían saber lo que no haría, como por ejemplo, comenzar a difundirse viralmente. Al probar el software que Prometeo había escrito para las tareas de MTu-

rk, los omegas se previnieron contra este riesgo ejecutándolo solo dentro de una máquina virtual; esto es, de un programa que simula un ordenador: por ejemplo, muchos usuarios de Mac compran software de máquina virtual que les permite ejecutar programas para Windows al hacerles creer que están en una máquina con ese sistema operativo. Los omegas habían creado su propia máquina virtual, a la que pusieron el sobrenombre de «Caja de Pandora», que simulaba una máquina ultrasimplificada despojada de todos los ornamentos que solemos asociar con los ordenadores: sin teclado, sin monitor, sin altavoces, sin conectividad a internet, nada. Para las transcripciones de audio para MTurk, los omegas dispusieron las cosas de manera que lo único que podía entrar en la Caja de Pandora era un fichero de audio, y todo lo que podía salir era un solo documento de texto: la transcripción. Estas leyes de la caja eran para el software que se ejecutaba en su interior como las leyes de la física son para nosotros en el universo: el software no podía salir de la caja igual que nosotros no podemos ir más rápido que la velocidad de la luz, por muy listos que seamos. En la práctica el software, salvo por esa entrada y esa salida, en el interior de la Caja de Pandora estaba atrapado en un universo paralelo con sus propias reglas de computación. Era tal el temor de los omegas a que Prometeo escapase que añadieron también un confinamiento temporal, al limitar el tiempo de vida del código que no era de confianza. Por ejemplo, cada vez que el software de transcripción confinado terminaba de transcribir un archivo de audio, se borraba automáticamente todo el contenido de la memoria de la Caja de Pandora y se reinstalaba el programa desde cero. De esta manera, cuando comenzaba la nueva tarea de transcripción, no tenía conocimiento de lo que había sucedido antes, y por lo tanto carecía de la capacidad de aprender con el paso del tiempo.

Cuando los omegas utilizaron la nube de Amazon para su propio proyecto de MTurk, pudieron instalar todos los módulos de tareas que Prometeo había creado en cajas virtuales de ese estilo en la nube, porque la entrada y la salida

que MTurk requería era muy sencilla. Pero esto no serviría para videojuegos con alto contenido gráfico, que no podrían confinarse porque necesitaban acceso completo a todo el hardware del ordenador del jugador. Además, no querían exponerse a que algún usuario con conocimientos informáticos analizase el código del juego, descubriese la Caja de Pandora y decidiese investigar qué había en el interior. El riesgo de fuga no solo obligaba a descartar de momento el mercado de los juegos, sino también el mercado, enormemente lucrativo, de otros tipos de software que movían cientos de miles de millones de dólares.

LOS PRIMEROS MILES DE MILLONES

Los omegas habían restringido su búsqueda a productos que fuesen muy valiosos, exclusivamente digitales (evitando así los lentos procesos de fabricación) y de fácil comprensión (por ejemplo, sabían que el texto o las películas no suponían un riesgo de escape). Al final, habían decidido lanzar una empresa de comunicación, y empezar con entretenimiento de animación. El sitio web, el plan de promoción comercial y los comunicados de prensa estaban todos preparados desde antes incluso de que Prometeo se convirtiese en una IA superinteligente; lo único que faltaba era el contenido.

Aunque el domingo por la mañana Prometeo era ya extraordinariamente competente y no hacía más que acumular dinero procedente de MTurk, sus capacidades intelectuales seguían siendo más bien limitadas: Prometeo había sido optimizado con el fin exclusivo de diseñar sistemas de IA y escribir software que llevase a cabo tediosas tareas de MTurk. Por ejemplo, se le daba mal hacer películas; no por ninguna cuestión intrínseca, sino por la misma por la que a James Cameron se le daba mal hacerlas cuando nació: es una habilidad que se tarda tiempo en aprender. Como un niño humano, Prometeo podía aprender cualquier cosa que

quisiese a partir de los datos a los que tenía acceso. Mientras que James Cameron había tardado años en aprender a leer y escribir, Prometeo lo hizo un viernes, cuando también encontró tiempo para leer toda la Wikipedia y unos cuantos millones de libros. Hacer películas era más complejo. Escribir un guion que a los humanos les pareciese interesante era tan difícil como escribir un libro, y requería una comprensión detallada de la sociedad humana y de aquello que a los humanos les resultaba entretenido. Convertir el guion en el vídeo final exigía una enorme cantidad de trazado de rayos de actores simulados y de las escenas complejas por las que se movían, voces simuladas, o la producción de bandas sonoras sugerentes, entre otras cosas. Desde el domingo por la mañana, Prometeo ya era capaz de ver una película de dos horas en un minuto, lo que incluía leer cualquier libro en el que estuviese basada y todas las reseñas y clasificaciones publicadas online. Los omegas se dieron cuenta de que, tras pegarse un atracón de varios cientos de películas, a Prometeo empezó a dársele bastante bien predecir el tipo de críticas que obtendría una película y en qué medida atraería a distintos públicos. De hecho, aprendió a escribir sus propias reseñas de películas de una manera que, en opinión de los omegas, demostraba verdadera perspicacia, comentando cualquier aspecto, desde la trama hasta las actuaciones, pasando por detalles técnicos como la iluminación o los ángulos de cámara. Concluyeron entonces que Prometeo, partiendo de estos datos, aseguraría el éxito de sus películas.

Los omegas ordenaron a Prometeo que se centrara inicialmente en generar películas de animación, para evitar preguntas incómodas sobre quiénes eran los actores simulados. El domingo por la noche, para completar su tremendo fin de semana, apagaron las luces y, provistos de cervezas y palomitas de microondas, se dispusieron a ver la ópera prima de Prometeo. Se trataba de una comedia fantástica de animación, en la línea del *Frozen* de Disney, y el trazado de rayos se había realizado mediante un código confiado creado por Prometeo y ejecutado en la nube de Ama-

zon, en la que se invertía buena parte del millón de dólares de beneficios diarios obtenido a través de MTurk. Al empezar la película, les resultó al mismo tiempo fascinante y aterrador pensar que había sido creada por una máquina sin asesoramiento humano, pero al rato estaban riéndose de las bromas y con la respiración contenida durante los momentos dramáticos. Algunos incluso soltaron alguna lagrimita durante el emocionante final, y llegaron a estar tan inmersos en esta realidad inventada que olvidaron la identidad de su creador.

Los omegas planificaron el lanzamiento de su sitio web para el viernes, dando así tiempo a Prometeo para que produjese más contenido, y a ellos mismos para hacer aquello para lo que no confiaban en Prometeo: comprar publicidad y empezar a contratar empleados para las empresas fantasma que habían constituido durante los meses anteriores. Para borrar su rastro, la tapadera consistiría en decir que su empresa audiovisual (que no tenía ningún vínculo público con los omegas) compraba la mayoría de su contenido a productores independientes, normalmente empresas tecnológicas emergentes de países pobres. Estos falsos proveedores estaban convenientemente ubicados en lugares remotos como Tiruchchirappalli y Yakutsk, a los que la mayoría de los periodistas curiosos no se molestarían en viajar. Los únicos empleados a los que contrataron realmente para esas empresas se dedicaban a la mercadotecnia y a tareas administrativas, y le explicaban a cualquiera que preguntase que sus equipos de producción estaban en otro lugar y que en ese momento no podían conceder entrevistas. En concordancia con su tapadera, eligieron como lema corporativo «Canalizando el talento creativo del mundo», y presentaron su compañía como rompedoramente diferente porque utilizaba tecnología puntera para empoderar a personas creativas, en especial en los países en vías de desarrollo.

Cuando llegó el viernes y los visitantes curiosos comenzaron a llegar a su sitio web, se encontraron con algo que recordaba a servicios de entretenimiento online como Ne-

Netflix o Hulu, pero con interesantes diferencias. Todas las series de animación eran nuevas y no habían oído hablar de ellas, y además eran atractivas: en la mayoría, los episodios duraban cuarenta y cinco minutos, tenían una trama potente y cada uno de ellos terminaba dejando al espectador ansioso por saber qué pasaría a continuación. Y eran más baratos que los de la competencia. El primer episodio de cada serie era gratis, y los demás se podían ver por cuarenta y nueve centavos cada uno, con descuentos para la serie completa. En un principio, solo había tres series de tres episodios cada una, pero a diario se iban añadiendo nuevos episodios, así como nuevas series dirigidas a distintos segmentos de espectadores. Durante las primeras dos semanas de Prometeo, sus habilidades como realizador de películas mejoraron rápidamente, no solo en cuanto a la calidad de las mismas, sino también en relación con los algoritmos de simulación de personajes y de trazado de rayos, que redujeron de manera sustancial los costes de computación en la nube asociados a la realización de cada episodio. En consecuencia, los omegas pudieron sacar decenas de series nuevas durante el primer mes, enfocadas hacia públicos que iban desde niños pequeños hasta adultos, así como expandirse a los principales mercados mundiales, lo que hacía que su sitio web tuviese un alcance mucho más internacional que el de todos sus competidores. Algunos comentaristas estaban impresionados con el hecho de que no solo eran multilingües las pistas de sonido, sino también los propios vídeos: por ejemplo, cuando un personaje hablaba italiano, los movimientos de su boca se ajustaban a las palabras italianas, y también hacía gestos típicamente italianos con las manos. Aunque Prometeo era ahora perfectamente capaz de crear películas con actores simulados indistinguibles de los seres humanos, los omegas evitaron hacerlo para no mostrar sus cartas. Lo que sí hicieron fue lanzar muchas series con personajes humanos animados semirrealistas, en géneros que competían con los programas de televisión y las películas tradicionales de imagen real.