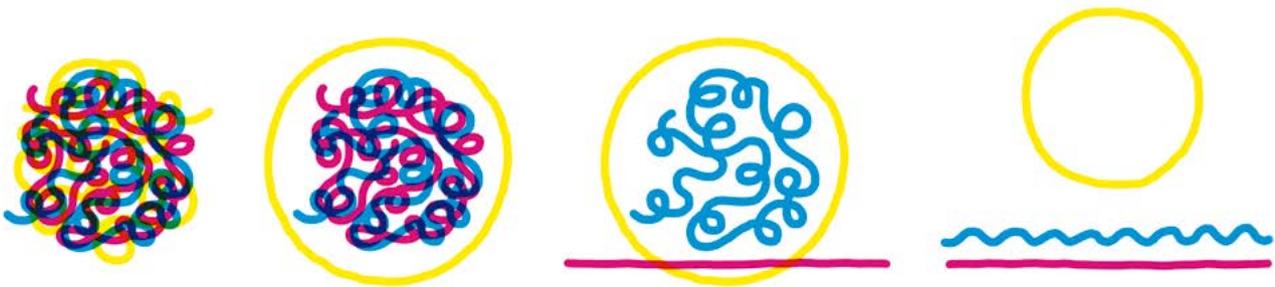


Reportaje

Del invierno de la IA al verano de la lucha por nuestros derechos de autoría

Luis Demano

Ilustrador y profesor



*Me dijeron: – O te subes al carro
o tendrás que empujarlo.*

*Ni me subí ni lo empujé.
Me senté en la cuneta
y alrededor de mí,
a su debido tiempo,
brotaron las amapolas.*

Gloria Fuertes

Demis Hassabis tenía 34 años cuando fundó Deepmind junto con el neurocientífico Shane Legg y Mustafa Suleyman, un amigo de la infancia. El principal objetivo de la empresa era crear un programa informático que implementara el llamado «aprendizaje por refuerzo», una técnica de *machine learning* para calcular la probabilidad del éxito de una acción imitando el proceso de aprendizaje humano de ensayo y error. La principal novedad con la que cautivaron a los recelosos inversores, que por aquel entonces percibían la inteligencia artificial como el *chiquipark* de los informáticos más excéntricos, fue combinar el aprendizaje por refuerzo con la simulación de un sistema de redes neuronales, la especialidad de Legg.

En la tentadora telaraña de aquella idea cayeron, no sin dudarlo, dos de los principales inversores de la historia de la tecnología, Elon Musk y Peter Thiel. Una vez formalizado el contrato de financiación, Thiel le propuso a Hassabis trasladar su domicilio de Londres, donde vivía por aquel entonces, a Silicon Valley, esa zona de confort tributario en la que los inversores se mueven como pez en el agua. Ante la negativa de Hassabis al cambio, el equipo de abogados de Thiel se trasladó a Londres para formalizar el contrato de financiación. La primera pregunta que le hicieron dejó a Hassabis un tanto desconcertado: «¿Londres tiene una ley sobre la propiedad intelectual?». «¡Creo que pensaron que vivíamos en Tombuctú!» pensó un sorprendido Hassabis.

Este episodio, en apariencia inocente, lo cuenta Marcus du Sautoy en su libro *Programados para crear. Cómo está aprendiendo a escribir, pintar y pensar la inteligencia artificial*. Y así, añadiría, es cómo empezó una de las aventuras que marcarían un

antes y un después de la historia de la Inteligencia Artificial, la creación de AlphaGo: con unos señores multimillonarios y su flamante equipo de abogados protegiendo su inversión, gracias a un sólido marco jurídico y regulatorio bajo el que poder proteger sus futuros descubrimientos. De momento, el «moverse rápido y romper cosas», ese estilo de vida que Mark Zuckerberg convirtió en el slogan motivacional de la nueva oleada emprendedora, quedaba relegado en favor del más práctico «nadar y guardar la ropa» de nuestro querido refranero popular. Una industria, la de la Inteligencia Artificial, cuyo valor aproximado es de 48.000 millones de dólares¹ no se levanta regalando nada a nadie.

En 2014, Google compró Deepmind por 500 millones de dólares.

Moveos rápido y romped cosas, pero que no sean las vuestras

El verano de 2022 fue especialmente caluroso. Así que la noticia de la salida de DALL-E 2, el programa de Inteligencia Artificial Generativa (IAG) propiedad de Open AI, me cogió debajo de las aspas de un ventilador tratando de aplacar el bochorno. De repente, mis tranquilas vacaciones se volvieron un cruce vertiginoso de imágenes a través de los grupos de wasap compartidos con mi comunidad profesional. Entre los muchos comentarios de fascinación, se sucedía el intercambio de unas imágenes que, aunque de un

1 Kentaro Takeda, Akira Oikawa, and Yukiko Une. 2023. Generative AI mania brings billions of dollars to developers. Nikkei Asia (March 4, 2023). <https://asia.nikkei.com/Spotlight/Datawatch/Generative-AI-mania-brings-billions-of-dollars-to-developers>

realismo torpe, estaban a años luz de las generadas por la primera versión de DALL-E lanzada en 2021. «A este paso la IA nos acabará quitando el trabajo XD» comentó alguien. Muchas de las ideas preconcebidas que teníamos sobre la inteligencia artificial comenzaron a tambalearse. ¿Cómo una tecnología que había mostrado claros síntomas de agotamiento había podido evolucionar tanto en apenas un año? Muy pronto entenderíamos el porqué.

Uno de los primeros artistas en hacerse preguntas aquel mismo 2022 fue el dibujante Dave McKean². Su libro *Prompt: Conversations with AI*, desarrollado en tan sólo 12 días, exploraba las posibilidades técnicas de Midjourney, otra de las varias IAGs que salieron a competir por su cuota de mercado ese mismo año. McKean también reflexionaba sobre las implicaciones éticas y el impacto que estos competidores artificiales tendrían en la industria. Pero no fue mucho más allá en cómo eran capaces de generar aquel contenido. Quién sí lo hizo, pocos meses más tarde, fue el artista Steven Zapata en su viralizado video *The End of Art: An Argument Against Image AIs*³. Un punto de inflexión desde el que la comunidad artística en concreto, y la industria cultural en general, comenzaría su toma de conciencia del verdadero problema al que se enfrentaba y las dimensiones abismales del mismo.

En su video de casi una hora de duración Zapata analizaba, además del impacto ético de la IAG más ampliamente de lo que lo había hecho McKean, algunas de las claves técnicas de la tecnología. Entre las más importantes, se encontraba la creación de las bases de datos a partir de las cuales los modelos de IAGs se habían entrenado. Era muy sencillo de entender: para que pueda generar una imagen necesitan disponer de una base de datos previa. En el caso de Stable Diffusion y Google's Imagen, su base de datos era LAION 5B, una colección de 5.000 millones de URLs de imágenes y texto descriptivo protegido, cómo veremos más detalladamente, en una gran cantidad por derechos de autoría. Efectivamente, ese mismo marco jurídico por el que preguntaban los abogados de Peter Thiel a Hassabis. ¿Qué hicieron entonces estas compañías tecnológicas para dinamitar el marco?

LAION 5B es difundida por LAION, una organización alemana sin ánimo de lucro, mientras que las propias imágenes son rastreadas mediante la práctica del *web scraping* por otra organización sin ánimo de lucro llamada CommonCrawl. Dependiendo de dónde se encuentren estas empresas, al estar registradas como organizaciones sin ánimo de lucro con

finés de investigación, gozan de exenciones legales. No serán declaradas culpables de recopilar y utilizar datos que, de otro modo, sí que estarían protegidos por derechos de autoría. ¿Y quién financió LAION para crear su base de datos? Exacto: Stability AI, la empresa dueña de Stable Diffusion.

De este modo, a través de una empresa sin ánimo de lucro, con fines de investigación y creada en un país donde disfruta de exenciones legales, se canaliza posteriormente todo el trabajo de los y las artistas protegidos por derechos de autoría a empresas con ánimo de lucro como Stability AI: una empresa que llegó a alcanzar un valor de mercado de mil millones de dólares gracias a su producto estrella Stable Diffusion.

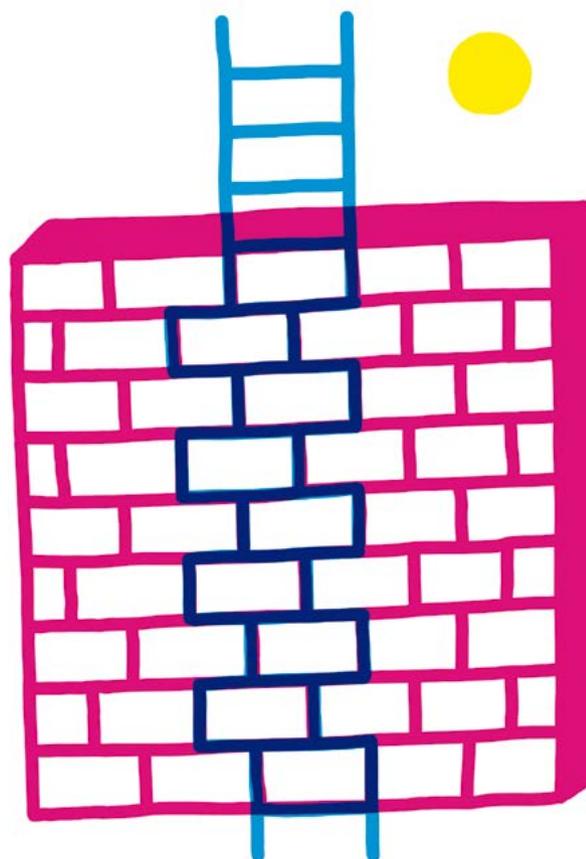
«En el caso de DALL-E», continuaba Zapata en su video, «su modelo de evasión de responsabilidades es el llamado OpenAI, formado por la corporación con ánimo de lucro OpenAI LP y su empresa matriz, la organización sin ánimo de lucro OpenAI Inc». Quienes meses más tarde lanzarían otro famoso competidor al mercado: ChatGPT (Chat Generative Pre-Trained Transformer). En el caso de Midjourney, otra popular IAG, ni tan siquiera han publicado la información sobre el origen de su base de datos.

Otra base de datos revelada gracias a una investigación del diario estadounidense The Washington Post, ha sido la C₄ (Colossal Clean Crawled Corpus) de Google. Recopilada también a través de la organización sin ánimo de lucro CommonCrawl, está formada a partir del contenido de «sitios web de industrias como el periodismo, el entretenimiento, el desarrollo de software, la medicina y la creación de contenidos, lo que ayuda a explicar por qué estos campos pueden verse amenazados por la nueva ola de inteligencia artificial. Los tres sitios más importantes fueron patents.google.com, que contiene texto de patentes emitidas en todo el mundo; wikipedia.org, la enciclopedia gratuita en línea; y scribd.com, una biblioteca digital de acceso restringido por suscripción. También ocupa un lugar destacado en la lista b-ok.org, un famoso mercado de libros electrónicos pirateados confiscado posteriormente por el Departamento de Justicia de Estados Unidos. Al menos otros 27 sitios identificados por el gobierno de EEUU como mercados de piratería y falsificaciones están presentes en el conjunto de datos C₄»⁴. Curiosamente, el símbolo de copyright señalando que una obra está registrada como propiedad intelectual, aparece más de 200 millones de veces en el conjunto de datos C₄.

Pero lo que realmente hizo saltar las alarmas sobre el contenido de las bases de datos de

² <https://www.youtube.com/watch?v=oqOnohx211A&t=29s>
³ <https://www.youtube.com/watch?v=tjSxPAGP9S&t=7s>

⁴ <https://www.washingtonpost.com/technology/interactive/2023/ai-chatbot-learning/>



entrenamiento, fue el informe publicado por el *Stanford Internet Observatory* y difundido por varios medios⁵. Según la investigación, LAION 5B contenía ingente material ilegal de abuso sexual infantil (CSAM), con el que Stable Diffusion habría sido entrenado. Advirtiendo, además, que la cantidad de material pedófilo podría ser mucho mayor, «dado que nuestra evaluación se centró en una pequeña porción del conjunto de miles de millones de datos». Otros modelos como Midjourney, cuya transparencia en el entrenamiento de los modelos es nula, podrían contener también este tipo de material.

Como es evidente, la lucha de los y las artistas en favor de la regulación y la estricta supervisión ética de los modelos de IAG, va más allá de su interés, que algunos tachan de egoísmo, por defender su parcela de mercado profesional. Constata que, sin

control democrático y transparencia, las empresas privadas de IA seguirán utilizando cualquier recurso que les dé ventaja sobre el resto de sus competidores. Moviéndose rápido y rompiendo muchas cosas que en el futuro serán muy difíciles de reconstruir.

Nadar y esconder la ropa robada

Una vez sale a la luz la información sobre las bases de datos necesarias para el entrenamiento de las IAGs, sin las cuales sería imposible la escalada de precisión de las imágenes y textos generados, las reacciones tanto de las empresas tecnológicas como de los defensores acérrimos de la IAG fueron variopintas.

En un primer momento, se negaron en rotundo a asumir la evidencia. Pero hasta las propias empresas de IAG admitían de forma velada la vulneración de los derechos de autoría de los y las artistas, y la falta total de consentimiento en el uso de las mismas. Tal era el caso de Stability AI, haciendo la siguiente

⁵ <https://www.forbes.com/sites/alexandralevine/2023/12/20/stable-diffusion-child-sexual-abuse-material-stanford-internet-observatory/>

declaración a propósito del lanzamiento de su IAG de audio Dance Diffusion dos meses después de la salida de Stable Diffusion:

Dance Diffusion también se basa en conjuntos de datos compuestos en su totalidad por muestras de música y audio libres de derechos de autor y proporcionadas voluntariamente. Dado que los modelos de difusión son propensos a la memorización y al sobreajuste, la publicación de un modelo entrenado con datos protegidos por derechos de autor podría dar lugar a problemas legales. Para honrar la propiedad intelectual de los artistas y, al mismo tiempo, cumplir en la medida de lo posible con las normas de derechos de autor, a menudo estrictas, de la industria musical, era imprescindible mantener cualquier tipo de material protegido por derechos de autor fuera de los datos de entrenamiento.

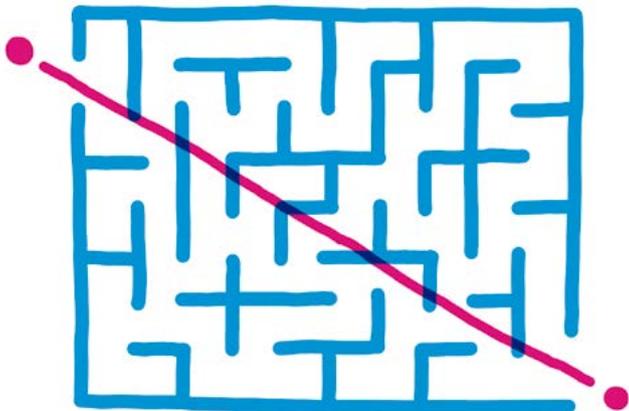
La respuesta legal del colectivo de artistas y las empresas afectadas por el rastreo masivo de datos sin su consentimiento no se hizo esperar. Las primeras en organizarse y reaccionar fueron las ilustradoras Kelly McKernan, Karla Ortiz, and Sarah Andersen, con la ayuda de la Concept Art Association de EEUU. Su demanda colectiva implicaba a Stability AI (Stable Diffusion), Midjourney y la web DeviantArt por entrenar a su IAG, Dreamup, con las obras alojadas en su web propiedad de las artistas. La denuncia incide en la posibilidad de incluir el propio nombre de las artistas en la descripción

de texto (prompt) para poder replicar el estilo de las artistas y generar contenido que perjudique su imagen profesional. Tanto es así que, en el caso de McKernan, un artículo del New York Times desveló cómo los miembros de la extrema derecha habían imitado su estilo artístico para crear historietas pronazis difundidas en el foro online *4chan*.

Algunos meses después, a esta misma demanda se sumaron siete nuevos artistas: Hawke Southworth, Grzegorz Rutkowski, Gregory Manchess, Gerald Brom, Jingna Zhang, Julia Kaye y Adam Ellis. Incluyendo además a una nueva empresa en la denuncia, Runway AI, la empresa que entrenó al modelo Stable Diffusion 1.5. Uno de los elementos más destacable en la nueva ampliación de la demanda es el artículo académico del investigador de Google DeepMind, Nicholas Carlini⁶, donde él y su equipo demuestra que las IAGs se entrenan intencionadamente para reconstruir las obras que forman su base de datos:

Los modelos de difusión de imágenes como DALL-E 2, Imagen y Stable Diffusion han atraído una gran atención debido a su capacidad para generar imágenes sintéticas de alta calidad. En este trabajo, demostramos que los modelos de difusión memorizan imágenes individuales a partir de sus datos de entrenamiento y las emiten en el momento de la generación. Mediante un proceso de generación y filtrado, extraemos más de mil ejemplos de entrenamiento de los modelos más avanzados, desde fotografías de

6 <https://arxiv.org/abs/2301.13188>



personas hasta logotipos de marcas comerciales. También entrenamos cientos de modelos de difusión en diversos entornos para analizar cómo afectan a la privacidad las distintas decisiones sobre modelos y datos. En general, nuestros resultados muestran que los modelos de difusión son mucho menos privados que los modelos generativos anteriores, como los GAN, y que para mitigar estas vulnerabilidades pueden ser necesarios nuevos avances en el entrenamiento para preservar la privacidad.

Recientes declaraciones de OPEN AI recogidas por el Telegraph, han reforzado este argumento: «Dado que actualmente los derechos de autoría cubren prácticamente todo tipo de expresión humana, incluidas publicaciones de blogs, fotografías, publicaciones en foros, fragmentos de código de software y documentos gubernamentales, sería imposible entrenar los principales modelos de IA sin utilizar materiales protegidos por derechos de autoría». A lo que añaden, por si quedaba alguna duda de la gran importancia que tiene la calidad de los datos de entrenamiento en el resultado final del contenido generado por los modelos de IAG: «Limitar los datos de formación a libros y dibujos de dominio público creados hace más de un siglo podría producir un experimento interesante, pero no proporcionaría sistemas de IA que satisfagan las necesidades de los ciudadanos de hoy»⁷. Alentadas por la demanda actual del New York Times a OPEN AI, estas declaraciones demuestran como la industria de IAG ha recorrido un ciclo muy similar al de las fases del luto: etapa de negación, ira, negociación, depresión y, por último, aceptación.

Podría resultar gracioso, pero es la misma fase por la que están pasando, o ya han completado, muchos de los y las artistas afectadas por la absoluta falta de ética de la industria tecnológica de IAG. Empresas dirigidas en su gran mayoría por hombres que se niegan a aceptar que, si utilizas propiedad ajena dentro de la cadena de desarrollo de tu producto con fines de lucro, debes compensar a sus creadoras por ello. Recordemos que la ilustración es, en proporción, el sector creativo más feminizado que existe en la actualidad. Así, la lógica vertical del mercado en un contexto capitalista, nos vuelve a confirmar el doble rasero que aplica a sus sectores más vulnerables: privatizar las ganancias y socializar las pérdidas.

Referencias

[Provisional]

Luis Demano. Nacido en Alicante (1976). En Valencia ganó el primer premio del Cartel de la Fira del Llibre, iniciando su carrera. Sus primeros encargos fueron con estudios de diseño, agencias de publicidad y periódicos como *El País*, *eldiario.es*, *El Mundo* o *ABC Cultural*. Ha trabajado en la mayoría de ámbitos de su disciplina: elaborando cartelería institucional, campañas publicitarias, animación corporativa y álbum ilustrado. Sus libros han sido editados en EEUU, México, Corea, Inglaterra, Argentina, Francia, Alemania, Italia, Rusia, Polonia y Rumanía. Su trabajo forma parte del archivo permanente del Museu del Disseny de Barcelona.

luis.demano@gmail.com

7 <https://www.theguardian.com/technology/2024/jan/08/ai-tools-chatgpt-copyrighted-material-openai>