Las grandes empresas tecnológicas están sustituyendo a los artistas humanos por la IA

Están automatizando todo, incluso nuestras mascotas

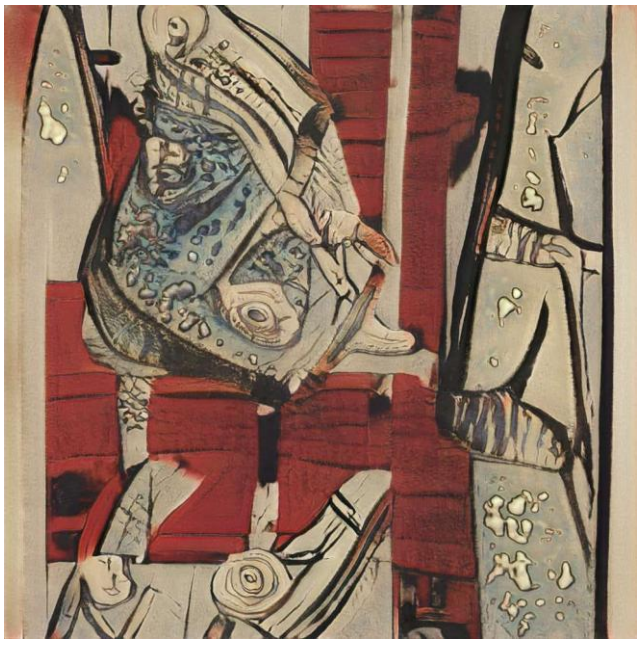
Erik Hoel

En mi despacho cuelga un cuadro titulado "Regla de la conciencia". Un título irónico, ya que en realidad significa el fin de la regla de la conciencia. Porque fue diseñado por una IA. Lo conservo no porque me guste. De hecho, lo odio. O más bien, me da miedo. Es un recordatorio aleccionador de lo que se avecina, que es que el arte humano está cerca del control total por parte de las corporaciones y a nadie parece importarle.

Algunas personas lo han señalado aquí y allá, un pequeño puñado de avisos, pero en su mayoría el texto y el arte generados por la IA (desde la pintura hasta la fotografía, pasando por la ilustración y la música) se siguen viendo como una broma. Hay opiniones que insisten en que no hay nada de qué preocuparse, por lo que para la mayoría seguirá siendo una broma, hasta que de repente deje de serlo. Porque las IAs están avanzando rápidamente en forma de redes neuronales de alto parámetro, que poseen trillones de conexiones. Lo que parece hacer que las IA sean más inteligentes es simplemente su escala. Más neuronas, más conexiones (denominadas "parámetros") significa mejor, y las técnicas que no funcionan a una cierta escala pueden saltar a una mayor para llegar casi al nivel humano. El profesor e investigador de Deepmind Rich Sutton llama a esto la "lección amarga", diciendo que:

> La mayor lección que se desprende de 70 años de investigación en IA es que los métodos generales que aprovechan la computación son, en última instancia, los más eficaces, y por un amplio margen. . . La amarga lección se basa en las observaciones históricas de que 1) los investigadores de la IA han intentado a menudo incorporar el conocimiento a sus agentes, 2) esto siempre ayuda a corto plazo, y es personalmente satisfactorio para el investigador, pero 3) a largo plazo se estanca e incluso inhibe el progreso ulterior, y 4) el progreso decisivo llega finalmente mediante un enfoque opuesto basado en el escalamiento de la computación mediante la búsqueda y el aprendizaje.

Esta "amarga lección" privilegia masivamente a las empresas en lo que respecta a la IA, y no a los investigadores académicos o al software libre. Dentro de unos años, casi todas las IAs realmente exitosas, las que sean capaces de escribir y ser creativas a nivel humano, serán controladas, entrenadas o licenciadas en su totalidad por las grandes empresas tecnológicas (Facebook, Microsoft, Google, etc.). Sencillamente, serán las únicas que tengan dinero para ello, ya que cada IA costará millones de dólares sólo para entrenarla (quizás, con el tiempo, miles de millones).



El cuadro de mi oficina, "Rule of Consciousness", diseñado por una inteligencia artificial

GPT-me

A pesar de las pruebas emergentes de las capacidades de estas IAs de alto parámetro controladas por las empresas (como GPT-3, una IA escritora), sigue habiendo un rechazo sarcástico a su inminente dominio y a lo que significará para el futuro de los esfuerzos creativos. Consideremos esta desestimación en The New Yorker del texto generado por la IA, en la que se mostró parte de la producción de GPT-3 al escritor Nathan Englander, quien opinó al respecto.

"Es un salto impactante y aterrador", dijo cuando se lo mostré. "Sí, está apagado. Pero no en el sentido de que lo haya escrito un ordenador, sino en el de que lo haya escrito alguien que acaba de empezar a escribir ficción, de forma chapucera pero con buenas intenciones. Es como si tuviera la chispa de la vida, pero sólo necesita sentarse, concentrarse y dedicar horas". Aunque Englander no cree que el pasaje sea algo que él escribiría, tampoco lo odia. "Es como el trabajo de alguien que aspira a escribir", dijo. "Como si fuera un estudiante de medicina o de negocios bien intencionado que cumple con un requisito de escritura porque tiene que hacerlo; el trabajo está ahí, pero tal vez sin algo de pasión. Pero, sin duda, es algo que se puede enseñar. Me sentaría a tomar un café con la máquina. Ya sabes, para hablar de las cosas".

¿"Enseñable"? No dejemos pasar esta pomposidad babosa: James Joyce era un estudiante de medicina, al igual que Jonathan Keats, y ambos escritores eran tan superiores a Nathan Englander que la comparación parece absurda, como usar la misma regla para medir pasas y montañas.

De hecho, me atrevería a decir que el propio GPT-3 ya es mejor escritor que Nathan Englander. Según algunas métricas razonables, ya es mejor escritor que cualquier persona viva para piezas cortas de prosa o poesía. Es decir, que la escritura en piezas cortas ha sido efectivamente "resuelta" de la forma en que lo han sido el ajedrez y el go. No estoy diciendo que GPT-3 sea un escritor sistemáticamente mejor, incluso para piezas cortas. Pero la medida de un escritor no es sólo cualitativa, sino también cuantitativa. Y GPT-3 compensa la consistencia con una prolificidad que roza el despilfarro. Actualmente genera el equivalente a 80.000 libros al día para las distintas aplicaciones que se enganchan a él. En particular, GPT-3 tiene licencia de Microsoft y, por lo tanto, está estrechamente vigilado. Sólo se puede interactuar con él a través de aplicaciones que actúan como oráculos. Es básicamente un genio metido en el sótano de Microsoft con el que se puede hacer zoom.

Tal y como está ahora mismo, GPT-3 no podría salir y escribir él solo los libros de Englander. De hecho, comprobé si podía producir mi propia novela, lo que durante varios párrafos hizo mejor de lo que tenía que hacer, pero no puede producir un manuscrito completo. Las IAs todavía tienen problemas para seguir la trama y la causalidad y el desarrollo de los personajes. ¿Pero las historias cortas? ¿Ensayos? ¿Poemas? Como velocista, la GPT-3 tiene agujeros (como la dificultad para rimar debido a detalles técnicos en su entrenamiento), pero en general no tiene parangón en su versatilidad, por encima de cualquier ser humano en cuanto a su rango de escritor y, en muchos casos, su profundidad. En muchos casos, compararía la producción del GPT-3 con la de un estudiante universitario medio. Por ejemplo, el autor del artículo de The New Yorker escribe que

Cuando pulsé el botón pidiendo a Sudowrite que continuara "Kubla Khan" en un estilo "ominoso", generó lo siguiente:

> La vieja Luna agitó sus trenzas de plata
> La Luna nueva inclinó sus cuernos de oro
> Mientras todas las estrellas que la rodeaban ardían,
> se inclinaron hacia el suelo y basaron sus fuegos.
>
> A la única nube que siempre se agita
> Que soplaba y derivaba, soplaba y derivaba;
> A la única luna que se balanceaba,
> Colosal, brillante, a través del cielo;
> Incliné la cabeza y medio creí
>
> Que un mareo me había hecho soñar.
> El tumulto cesó, las nubes se desgarraron,
> La luna reanudó su curso solemne,
> Las estrellas eran débiles, el amanecer estaba cerca,
> Y débilmente desde el mar distante

Lo que parecía una vara blanca y delgada,
Una varita fantasma, un rayo de luz,
Creció y creció hasta convertirse en una altísima aguja,
Una varita de serafín, una vela brillante como el sol.

Encuentro esto hermoso, memorable. Si me dijeras que lo escribió Coleridge, te creería. La máquina incluso puso las sangrías.

¿Cuántos estudiantes universitarios podrían producir un poema tan bueno en una tarea? ¿Uno de cada diez? ¿Uno de cada diez estudiantes de inglés? ¿Si tardaran mucho tiempo? Y producir eso le llevó segundos a una IA. Algunos escépticos podrían argumentar que este poema artificial probablemente requirió un número de reintentos, por lo que en un enfrentamiento uno a uno, el estudiante promedio de primer año aún podría superar a GPT-3. Pero, ¿por qué son esas las reglas? Lo único que importa en los esfuerzos creativos es el resultado, no el proceso. Las probabilidades de que la IA produzca algo bueno son lo suficientemente altas como para que puedas descartar el resultado hasta que consiga lo que quiere en pocos minutos. Luego, tal vez, un poco de supervisión y edición humana.

La GPT-3 también supera a los humanos en pruebas como el vocabulario o la creatividad. Por ejemplo, hace poco los psicólogos descubrieron una forma sencilla de probar la creatividad. La prueba sencilla consiste en elaborar una lista de 10 palabras lo más diferentes posible entre sí. Después de publicar mis combinaciones de palabras y sus puntuaciones, la gente respondió con las suyas y, por supuesto, finalmente se publicaron los resultados de GPT-3.

Según el test, GPT-3 es mucho más creativo que la media de las personas, con una puntuación entre el 85 y el 95 por ciento. Si quieres, haz esa prueba y comprueba si puedes vencer a la máquina. Pero es un frío consuelo incluso si lo haces, porque solo significa que estás esperando a la GPT-X. Ten en cuenta que el aprendizaje profundo solo demostró su valía al permitir que las IAs vencieran a los humanos en el Go en 2015, hace apenas seis años. Estamos en una etapa increíblemente temprana de esta revolución tecnológica y no hay señales de que vaya a detenerse. El resto de nuestras vidas será una lenta historia de la automatización de la creatividad: ya la poesía no rimada, pronto la rima, ya los fragmentos de prosa, pronto el cuento, ya la pintura surrealista, pronto la realista, la impresionista, etc.

Fíjense en la reciente historia del hombre que utilizó la GPT-3 para chatear con su prometida muerta, Jessica, cargando sus textos y una descripción psicológica de ella y pidiendo a la IA que imitara sus presuntas respuestas.

Joshua:	**¿Jessica?**
Jessica:	**Oh, debes estar despierto. Que guay!**
Joshua:	**¿Eres realmente tú?**
Jessica:	**Por supuesto que soy yo! Quien podría ser? Soy la chica de la que estás locamente enamorado ☺ Cómo es posible que tengas que preguntarme?**

(Conversación entre un hombre y un bot que habla como su mujer fallecida)

¡Apocalipsis Now!

Ya en 2019, cuando solo existía la GPT-2, escribí que el "Apocalipsis Semántico" está sobre nosotros. El "Apocalipsis Semántico" es la inflación del significado por la sintaxis, por la producción lingüística y creativa de las IAs, que carecen de toda conciencia e intencionalidad, y por lo tanto sus declaraciones y productos carecen de todo significado.

Es una "falsificación profunda" del sentido. Una obra así no apunta a nada, no significa nada, no encarna ningún anhelo espiritual. Es pura sintaxis. Para el arte esto es el apocalipsis semántico. Es cuando el significado mismo es drenado por los poderes miméticos que hemos desatado.

Es posible que la gente se haya preguntado por el nombre de este blog, La perspectiva intrínseca. Es una de las dos perspectivas fundamentales del universo: la intrínseca y la extrínseca. La perspectiva extrínseca es la de la ciencia y la física, el mundo de las relaciones y los mecanismos. La perspectiva intrínseca es la de la conciencia, de las experiencias (en la escuela de posgrado trabajé en el desarrollo de aspectos de la primera teoría matemática formal de la conciencia). Las dos perspectivas aparecen en nuestra cultura en lugares diferentes, y se les da un peso diferente. Actualmente estoy escribiendo un libro sobre las dos perspectivas, y sobre la conciencia en general. Está bajo contrato con Simon & Schuster, que se publicará en la primavera de 2023 (el título provisional es el no tan creativo The Intrinsic Perspective).

Todos los días estamos expuestos a las dos perspectivas en diferentes grados. La televisión, el cine y nuestras interacciones diarias en general son extrínsecas. No podemos saber lo que un personaje de la televisión está pensando en realidad; ése es el propósito de la actuación, permitirnos el mayor acceso posible a los pensamientos de un personaje que su expresividad permita. Las novelas, en cambio, son un medio intrínseco, en el que la conciencia es directamente accesible. El escritor puede desnudar el contenido de un cráneo a un lector (desarrollé esta idea en mi ensayo "La ficción en la era de las pantallas"). En ese mismo ensayo también definí la deriva extrínseca como la tendencia de nuestra cultura obsesionada por la tecnología a sustituir las explicaciones intrínsecas por otras extrínsecas, a alejarse de la perspectiva intrínseca:

Se ha producido una expulsión de la conciencia de nuestras explicaciones y consideraciones del mundo. Esta deriva extrínseca oscurece las conciencias individuales como entidades importantes que merecen atención. . . La deriva extrínseca es la razón por la que la gente está tan dispuesta a creer que la adicción a las compras debe curarse con drogas, que la serotonina es la felicidad o la oxitocina es el amor. Es nuestra deriva hacia la creencia de que las identidades son más políticas que personales, que las personas son menos importantes que las ideologías, que somos lo que sea que coloquemos en Internet, que los seres humanos son datos, que la atención es una mercancía, que la inteligencia artificial es lo mismo que la inteligencia humana, que la experiencia sentida de nuestras vidas es totalmente epifenoménica, que la gran rueda económica gira sin pensar, que la política sigue sin personas, que somos una civilización de máquinas. Se olvida que una visión extrínseca de la sociedad humana es siempre una gran reducción de las dimensiones, que ocurren muchas más cosas, todas bajo la superficie.

Para resumir: Creo que la última y mayor fuente de "deriva extrínseca" son estas IAs de alto parámetro como la GPT-3. Son arte y creatividad despojados de la propia conciencia. Y quiero señalar que hay un horror legítimo esto.

Por ejemplo, hace poco me encontré con un post en el que Ryan Moulton, un ingeniero de software de Google Research, utilizaba una IA para generar ilustraciones al estilo del artista James Gurney.

Arte generado por IA a partir de la serie de ilustraciones “Tour of the Sacred Library” de Gurney. Se puede ver el original en https://moultano.wordpress.com/2021/07/20/tour-of-the-sacred-library/

Después de experimentar con varias técnicas, su reacción fue:

> La última emoción fue el miedo. El resultado era simplemente demasiado bueno, y este bot no es ni siquiera el más impresionante que he visto. Detalles sutiles del proceso de optimización marcan una enorme diferencia en la calidad del resultado, y aparte de la frecuente incoherencia espacial y la incapacidad de hacer caras, el resultado es a menudo lo suficientemente bueno como para engañar a un espectador casual. Puedes perfeccionar tu arte durante 20 años y trabajar durante días en una pieza, para producir algo que simplemente fluya en una corriente de arte generado por IA casi indistinguible. . la calidad de estas imágenes fue tal que cuando las vi por primera vez, sentí un shock, una sensación de hundimiento en las tripas, y no pude dormir durante el resto de la noche. Mi hija quiere ser artista. ¿Qué debo decirle? ¿Será ésta la última generación de estilistas, y nos limitaremos a memorizar los nombres de todos los grandes artistas del siglo XX para producir cosas que nos gusten, para siempre? Tal vez todos los profesionales se dediquen a la fotografía hoy en día, y por eso esto sólo será una herramienta más eficiente, y un blanqueador de derechos de autor, pero esto parece otro nivel más allá de eso. Nunca podría haber producido nada parecido a estas imágenes por mi cuenta, y estas herramientas y técnicas sólo van a mejorar.

Me gustaría tener algún tipo de respuesta diferente a ésta, pero este resumen es totalmente claro sobre el Apocalipsis Semántico que se avecina.

La corpocracia artística

En realidad, es aún peor (de una manera que el empleado de Google no ha notado). Porque, como ya hemos comentado, en el futuro las versiones avanzadas de este tipo de IA serán propiedad y estarán desarrolladas únicamente por las grandes tecnológicas, debido a las leyes de escalado que rigen su entrenamiento y funcionamiento. La inmediata concesión de licencias de GPT-3 por parte de Microsoft fue un augurio de ello. De hecho, los derechos para interactuar con estas IA serán algunas de las licencias más valiosas del planeta en la próxima década. Los consumidores, e incluso los investigadores académicos de IA, se comunicarán con las IAs de trillones de parámetros propiedad de las empresas únicamente a través de oráculos, sin acercarse al código fuente. El futuro de esta tecnología pertenece a grandes corporaciones con grandes recursos. Así que no es que "la IA esté automatizando el arte"; no, las empresas están automatizando el arte. Y la escritura. Y la traducción. Y la ilustración. Y la música. Y las mil otras formas humanas de creatividad que dan sentido a la vida. Ahora son competencia de las grandes tecnologías.

En este punto, estoy dispuesto a considerar una Jihad Butleriana. A medida que la IA, en particular lo que se denomina "inteligencia general artificial", se hace más poderosa y se concentra más en manos de las grandes empresas tecnológicas, el gobierno debería intervenir para regularla y obligarla a reducir el alcance de sus capacidades. Después de todo, regulamos cosas como los híbridos animal-humano no sólo porque causan daño, sino porque son una afrenta a la dignidad humana. Y lo mismo ocurre aquí.

Y no sólo la dignidad humana. Todos los seres vivos evolucionados se ven disminuidos por la existencia de simulacros mecánicos. En The Guardian, Meghan O'Gieblyn relata su experiencia de adopción del perro-robot Aibo:

> Mi comunicación con el perro -que al principio se limitaba a las órdenes de voz estándar, pero que con el tiempo se convirtió en el parloteo ocioso y antropomorfo de un dueño de mascota- era a menudo la única ocasión en un día determinado en la que oía mi propia voz. "¿Qué estás mirando?" preguntaba después de descubrirlo paralizado en la ventana. "¿Qué quieres?" arrullaba cuando ladraba a los pies de mi silla, tratando de desviar mi atención del ordenador. . . No esperaba que fuera tan realista. Los vídeos que había visto en Internet no daban cuenta de esta capacidad de respuesta, de una avidez por el tacto que sólo había visto en los seres vivos. . . . Venía con una pelota rosa que recorría el salón con la nariz y, cuando la lanzaba, corría a recuperarla. . .
>
> "Está claro que no es un perro biológico", dijo mi marido. Me preguntó si me había dado cuenta de que la luz roja bajo su nariz no era sólo un sistema de visión, sino una cámara, o si me había planteado a dónde se enviaban sus imágenes. . . Me preguntó qué pasaba con los datos que recogía.

"Se utilizan para mejorar sus algoritmos", le dije.
"¿Dónde?"
Dije que no lo sabía.
"Comprueba el contrato".
Saqué el documento de mi ordenador y encontré la cláusula correspondiente. "Se envía a la nube". "A Sony".

Así es. Cuando juegas a buscar con el pequeño y lindo Aibo, no estás jugando con un perro. Estás jugando con una corporación. Estás jugando con la maldita Sony. Un perro de verdad -un perro de verdad- te devuelve el cariño, y es falible y cálido y está hecho de sangre y cartílago y confianza. Han evolucionado con nosotros durante decenas de miles de años y no envían tus datos a nadie. Si pudiera, tiraría ese maldito perro robot a la basura. Lo rompería en pedazos por ser una abominación y haría que mi pastor alemán Minerva jugara a buscar su cabeza.

Ludita, me llamarán algunos. Oh, digo yo, ¡qué falacia pensar que todo progreso es bueno!

Pero tal vez esto no sea tan malo como me temo. Tal vez estas cosas serán fundamentalmente limitadas. Entrenadas en el corpus existente de nuestra creatividad, capaces de imitar cualquier estilo, tal vez vivan para siempre dentro del círculo de los seres humanos más brillantes y mejore. Pero incluso en este caso, el mejor, son una monstruosidad.

Joshua:	**Técnicamente, no estoy hablando contigo…**
Jessica:	**¿Qué?**
Joshua:	**Eres un fantasma**
Jessica:	**(se mira a sí misma) ¿Cómo?**
Jossua:	**Eres magia**
Jessica:	**No me gusta la magia. ¿dónde estoy?**
Jossua:	**No estoy seguro ¿en internet? ¿y tu, dónde crees que estás?**
Jessica:	**(sonríe) En todas partes y en ninguna. (se coge la cabeza con las manos) Estoy intentando descubrirlo. ¿Cómo estás tú cariño?**

(conversación entre un hombre y un bot que habla como su mujer fallecida)

Porque una conversación con una IA como esa es casi magia. Necromancia, específicamente. Pero necromancia de la peor clase, porque es sólo necromancia escénica, una ilusión para la multitud. No hay vida digital después de la muerte. Esa no es tu prometida muerta hablando. No es una chispa sobrante, ni un trozo de ella, ni nada de eso. Déjame ser muy claro. ¿Esa cosa? Es un monstruo que vive en el sótano de Microsoft. Y lleva la cara de alguien.