El cerebro sin órganos de la inteligencia artificial

Por: Franco Berardi. 06/07/2025

“La escritura metódica me distrae de la presente condición de los hombres. La certidumbre de que todo está escrito nos anula o nos afantasma. Yo conozco distritos en que los jóvenes se prosternan ante los libros y besan con barbarie las páginas, pero no saben descifrar una sola letra. Las epidemias, las discordias heréticas, las peregrinaciones que inevitablemente degeneran en bandolerismo, han diezmado la población. Creo haber mencionado los suicidios, cada año más frecuentes. Quizá me engañen la vejez y el temor, pero sospecho que la especie humana –la única– está por extinguirse y que la Biblioteca perdurará: iluminada, solitaria, infinita, perfectamente inmóvil, armada de volúmenes preciosos, inútil, incorruptible, secreta.” Jorge Luis Borges, La biblioteca de Babel

En este pasaje del relato que Borges publicó en 1941 está todo nuestro presente: la desintegración de la civilización humana, el fanatismo religioso de los jóvenes que besan las páginas del libro que no saben ni leer, las epidemias, las discordias, migraciones que degeneran en bandolerismo y diezman la población. Y por último los suicidios, cada año más frecuentes.

Finalmente Borges anuncia que la Biblioteca no está destinada a desaparecer con la humanidad: permanece solitaria, infinitamente secreta y perfectamente inútil.

¿Entonces la inmensa biblioteca de datos registrada por sensores visuales, sonoros y gráficos insertos en todos los rincones del planeta seguirá alimentando eternamente al autómata cognitivo que está remplazando frágiles organismos humanos, envenenados por miasmas, dementes hasta el suicidio? Eso dice Borges, quién sabe.

Durante la era moderna, como había dicho Francis Bacon, el conocimiento era un factor de poder sobre la naturaleza y sobre los demás, pero a partir de cierto momento la expansión del conocimiento técnico empezó a funcionar a la inversa: no más prótesis del poder humano, la tecnología ha sido transformada en un sistema dotado de una dinámica independiente dentro del cual nos encontramos atrapados.

Las tecnologías digitales han creado las condiciones para la automatización de la interacción social, al punto de hacer inoperante la voluntad colectiva.

En *Out of control* (1993), Kevin Kelly predijo que las entonces nacientes redes digitales crearían una Mente Global a la que las mentes subglobales (individuales o colectivas o institucionales) tendrían que ser obligadas a someterse.

Mientras tanto, se desarrollaba la investigación sobre Inteligencia Artificial (IA), que hoy ha alcanzado un nivel de madurez suficiente para prefigurar la inscripción en el cuerpo social de un sistema de concatenación de innumerables dispositivos capaces de automatizar las interacciones cognitivas humanas.

La sociedad planetaria está cada vez más llena de automatismos técnicos, pero esto de ninguna manera elimina el conflicto, la violencia y el sufrimiento. No se vislumbra armonía, no parece establecerse ningún orden en el planeta.

El caos y el autómata conviven entrelazándose y alimentándose.

El caos alimenta la creación de interfaces técnicas de control automático, que solo permiten continuar la producción de valor. Pero la proliferación de automatismos técnicos, desafiados por las innumerables instancias del poder económico, político y militar en conflicto, termina alimentando el caos, en lugar de reducirlo.

¿Será siempre así o habrá un cortocircuito y el caos se apoderará del autómata? ¿O más bien será capaz el autómata de librarse del caos, eliminando a su agente humano?

El cumplimiento

En las últimas páginas de la novela *The Circle* de David Eggars, Ty Gospodinov le confiesa a Mae su impotencia ante la criatura que él mismo concibió y construyó, la

monstruosa empresa tecnológica constituida por la convergencia de Facebook, Google, PayPal, YouTube y mucho más. “No quería que pasara lo que está pasando,” dice Ty Gospodinov, “pero ahora ya no puedo evitarlo”.

La completud se destaca en el horizonte de la novela, el cierre del círculo: tecnologías de recolección capilar de datos e inteligencias artificiales se conectan perfectamente en una red ubicua de generación sintética de realidad compartida.

La etapa actual de desarrollo de IA probablemente nos esté llevando al umbral de un salto a la dimensión que yo definiría como “autómata cognitivo global”.

Autómatas pseudocognitivos se enlazan entre sí convergiendo en el autómata cognitivo global

El autómata no es un análogo del organismo humano, sino la convergencia de innumerables dispositivos generados por inteligencias artificiales dispersas. La evolución de la Inteligencia Artificial no conduce a la creación de androides, a la simulación perfecta del organismo consciente, sino que se manifiesta como la sustitución de habilidades específicas por autómatas pseudocognitivos que se enlazan entre sí convergiendo en el autómata cognitivo global.

El 28 de marzo de 2023, Elon Musk y Steve Wozniak, seguidos por más de mil altos operadores de alta tecnología, firmaron una carta en la que proponían una moratoria a la investigación en el campo de la IA:

> “Los sistemas de IA se están volviendo competitivos con los humanos en la búsqueda de propósitos generales, y tenemos que preguntarnos si debemos permitir que las máquinas inunden nuestros canales de información con propaganda y falsedades. ¿Deberíamos permitir que todas las actividades laborales se automaticen, incluidas las gratificantes? ¿Deberíamos desarrollar mentes no humanas que pudieran superarnos en número y eficacia, para volvernos obsoletos y reemplazarnos? ¿Deberíamos arriesgarnos a perder el control de nuestra civilización? Estas decisiones no se pueden delegar en líderes tecnológicos no elegidos. Estos poderosos sistemas de inteligencia artificial solo deberían desarrollarse cuando estemos seguros de que sus efectos serán positivos y sus riesgos son manejables. Por lo tanto, hacemos un llamamiento a

> todos los laboratorios de IA para que suspendan de inmediato el entrenamiento de los sistemas de IA más potentes que GPT-4 durante al menos seis meses. Este descanso debe ser público y verificable, y debe incluir a todos los actores clave. Si no se implementa tal pausa, los gobiernos deberían tomar la iniciativa de imponer una moratoria".

Luego, a principios de mayo de 2023, se difundió la noticia de que Geoffrey Hinton, uno de los primeros creadores de redes neuronales, decidió dejar Google para poder expresarse abiertamente sobre los peligros implícitos en la inteligencia artificial.

"Algunos de los peligros de los chatbots de IA son bastante aterradores", dijo Hinton a la BBC, "porque pueden superar a los humanos y pueden ser utilizados por agentes maliciosos".

Además de anticipar la posibilidad de manipulación de la información, lo que preocupa a Hinton es "el riesgo existencial que surgirá cuando estos programas sean más inteligentes que nosotros. Es como si tuvieras diez mil personas y cada vez que una persona aprendiera algo, todos lo supieran automáticamente. Y así es como estos chatbots pueden saber mucho más que cualquier persona".

La vanguardia ideológica y empresarial del neoliberalismo digital parece atemorizada por el poder del Golem

La vanguardia ideológica y empresarial del neoliberalismo digital parece atemorizada por el poder del Golem y, como aprendices de brujo, los empresarios de alta tecnología piden una moratoria, una pausa, un periodo de reflexión.

¿La mano invisible ya no funciona? ¿La autorregulación de la Red-Capital ya no está en la agenda?

¿Qué está sucediendo? ¿Qué va a pasar? ¿Qué está a punto de suceder?

Son tres preguntas separadas. Lo que está pasando más o menos lo sabemos: gracias a la convergencia de la recolección masiva de datos, de programas capaces de reconocimiento y recombinación, y gracias a dispositivos de generación lingüística, está surgiendo una tecnología capaz de simular habilidades inteligentes específicas: loros estocásticos.

Lo que está a punto de suceder es que los loros estocásticos, gracias a su capacidad de autocorrección y su capacidad para escribir software evolutivo, están obligados a acelerar en gran medida la innovación técnica, especialmente la innovación técnica de sí mismos.

Lo que podría suceder y probablemente sucederá: los dispositivos innovadores de autocorrección (aprendizaje profundo) determinan su propósito independientemente del creador humano. En las garras de la competencia económica y militar, la investigación y la innovación no pueden suspenderse, especialmente si pensamos en la aplicación de la IA en el campo militar.

Creo que los aprendices de brujo se están dando cuenta de que la tendencia a la autonomía de los generadores de lenguaje (autonomía desde el creador humano) está generando competencias inteligentes más eficientes que el agente humano, aunque en un campo específico y limitado.

Luego, las habilidades específicas convergerán hacia la concatenación de autómatas autodirigidos para los cuales el creador humano original podría convertirse en un obstáculo a eliminar.

En el debate periodístico sobre este tema priman posiciones de cautela: se denuncian problemas como la difusión de noticias falsas, la incitación al odio o manifestaciones racistas implícitas. Todo cierto, pero no muy relevante.

Durante años, las innovaciones en la tecnología de la comunicación han aumentado la violencia verbal y la idiotez. Esto no puede ser lo que preocupa a los maestros del autómata, a quienes lo concibieron y lo están implementando.

Lo que preocupa a los aprendices de brujo, en mi humilde opinión, es la conciencia de que el autómata inteligente dotado de capacidades de autocorrección y autoaprendizaje está destinado a tomar decisiones autónomas de su creador.

Pensemos en un automatismo inteligente insertado en el dispositivo de control de un dispositivo militar. ¿Hasta qué punto podemos estar seguros de que no evoluciona de forma inesperada, tal vez acabando disparando a su dueño o deduciendo lógicamente de los datos de información a los que puede tener acceso la urgenciade lanzar la bomba atómica?

Lenguaje generativo y pensamiento: el cerebro sin órganos

La máquina lingüística que responde a las preguntas es una demostración de que Chomsky tiene razón cuando dice que el lenguaje es el producto de estructuras gramaticales inscritas en la herencia biológica humana, dotadas de un carácter generativo, es decir, capaces de generar infinitas secuencias dotadas de significado.

Pero el límite del discurso de Chomsky reside precisamente en la negativa a ver el carácter pragmático de la interpretación de los signos lingüísticos; del mismo modo, el límite del chatbot GPT consiste precisamente en su imposibilidad de leer pragmáticamente las intenciones de significado.

Transformador Generativo Pre-entrenado (GPT) es un programa capaz de responder y conversar con un humano gracias a la capacidad de recombinar palabras, frases e imágenes recuperadas de la red lingüística objetivada en Internet.

Los órganos sensibles constituyen una fuente de conocimiento contextual y autorreflexivo que el autómata no posee

El programa generativo ha sido entrenado para reconocer el significado de palabras e imágenes, y posee la capacidad de organizar declaraciones sintácticamente. Posee la capacidad de reconocer y recombinar el contexto sintáctico pero no el pragmático, es decir, la dimensión intensiva del proceso de comunicación, porque esta capacidad depende de la experiencia de un cuerpo, experiencia que no está al alcance de un cerebro sin órganos. Los órganos sensibles constituyen una fuente de conocimiento contextual y autorreflexivo que el autómata no posee.

Desde el punto de vista de la experiencia, el autómata no compite con el organismo consciente. Pero, en términos de funcionalidad, el autómata (pseudo)cognitivo es

capaz de superar al agente humano en una habilidad específica (calcular, hacer listas, traducir, apuntar, disparar, etc.).

El autómata también está dotado de la capacidad de perfeccionar sus procedimientos, es decir, de evolucionar. En otras palabras, el autómata cognitivo tiende a modificar los propósitos de su funcionamiento, no solo los procedimientos.

Una vez desarrolladas las habilidades de autoaprendizaje, el autómata está en condiciones de tomar decisiones relativas a la evolución de los procedimientos, pero también, fundamentalmente, de tomar decisiones relativas a los propios fines del funcionamiento automático.

Gracias a la evolución de los loros estocásticos en agentes lingüísticos capaces de autocorrección evolutiva, el lenguaje, hecho capaz de autogenerarse, se vuelve autónomo del agente humano, y el agente humano se envuelve progresivamente en el lenguaje.

El humano no es subsumido, sino envuelto, encapsulado. La subsunción total implicaría una pacificación de lo humano, una completa aquiescencia: un orden, finalmente.

Finalmente una armonía, aunque totalitaria.

Pero no. La guerra predomina en el panorama planetario.

El autómata ético es una ilusión

Cuando los aprendices de brujo se dieron cuenta de las posibles implicaciones de la capacidad autocorrectora y por tanto evolutiva de la inteligencia artificial, empezaron a hablar de la ética del autómata, o alineamiento, como se dice en la jerga filosófica empresarial.

En su pomposa presentación, los autores del chatbot GPT declaran que es su intención inscribir criterios éticos alineados con los valores éticos humanos en sus productos.

> “Nuestra investigación de alineación tiene como objetivo alinear la

> inteligencia artificial general (IAG) con los valores humanos y seguir la intención humana. Adoptamos un enfoque iterativo y empírico: al intentar alinear sistemas de IA altamente capaces, podemos aprender qué funciona y qué no, refinando así nuestra capacidad para hacer que los sistemas de IA sean más seguros y más alineados".

El proyecto de insertar reglas éticas en la máquina generativa es una vieja utopía de ciencia ficción, sobre la cual Isaac Asimov fue el primero en escribir cuando formuló las tres leyes fundamentales de la robótica. El mismo Asimov demuestra narrativamente que estas leyes no funcionan.

Y al fin y al cabo, ¿qué estándares éticos deberíamos incluir en la inteligencia artificial?

La experiencia de siglos muestra que un acuerdo universal sobre reglas éticas es imposible, ya que los criterios de evaluación ética están relacionados con los contextos culturales, religiosos, políticos, y también con los impredecibles contextos pragmáticos de la acción. No existe una ética universal, si no la impuesta por la dominación occidental que, sin embargo, empieza a resquebrajarse.

La máquina no se alinea con los valores humanos, que nadie sabe cuáles son. Pero los humanos deben alinearse con los valores automáticos

Obviamente, todo proyecto de inteligencia artificial incluirá criterios que correspondan a una visión del mundo, una cosmología, un interés económico, un sistema de valores en conflicto con otros. Naturalmente, cada uno reclamará universalidad.

Lo que sucede en términos de alineación es lo contrario de lo que prometen los constructores del autómata: no es la máquina la que se alinea con los valores humanos, que nadie sabe exactamente cuáles son. Pero los humanos deben alinearse con los valores automáticos del artefacto inteligente, ya sea que se trate de asimilar los procedimientos indispensables para interactuar con el sistema financiero, o que se trate de aprender los procedimientos necesarios para utilizar los sistemas militares.

Pienso que el proceso de autoformación del autómata cognitivo no puede ser corregido por ley o por normas éticas universales, ni puede ser interrumpido o

desactivado.

La moratoria solicitada por los aprendices de brujo arrepentidos no es realista, y menos aún la desactivación del autómata. A esto se opone tanto la lógica interna del propio autómata como las condiciones históricas en las que se desarrolla el proceso, que son las de la competencia económica y la guerra.

En condiciones de competencia y guerra, todas las transformaciones técnicas capaces de aumentar el poder productivo o destructivo están destinadas a ser implementadas.

Esto significa que ya no es posible detener el proceso de autoconstrucción del autómata global.

Las cajas de Pandora

“Estamos abriendo las tapas de dos cajas de Pandora gigantes”, escribe Thomas Friedman en un editorial de *The New York Times* “El cambio climático y la inteligencia artificial”.

Algunas frases del artículo me llamaron la atención:

> “La ingeniería está por delante de la ciencia hasta cierto punto. Esto significa que incluso aquellos que están construyendo los llamados modelos de lenguaje extenso que subyacen en productos como ChatGPT y Bard no entienden completamente cómo funcionan ni el alcance total de sus capacidades”.

La expansión del autómata se ve limitada por la persistencia del factor caótico humano

Probablemente, la razón por la que una persona como Hinton decidió abandonar Google y tomarse la libertad de advertir al mundo del peligro extremo es la conciencia de que el dispositivo posee la capacidad de autocorregirse y redefinir su propósito.

¿Dónde está el peligro de un ente que, sin tener inteligencia humana, es más

eficiente que el hombre en la realización de tareas cognitivas específicas, y posee la capacidad de perfeccionar su propio funcionamiento?

La función general de la entidad inteligente inorgánica es introducir el orden de la información en el organismo impulsor.

El autómata tiene una misión ordenadora, pero encuentra en su camino un factor de caos: la pulsión orgánica, irreductible al orden numérico.

El autómata extiende su dominio a campos siempre nuevos de la acción social, pero no logra completar su misión mientras su expansión se ve limitada por la persistencia del factor caótico humano.

Ahora surge la posibilidad de que en algún momento el autómata sea capaz de eliminar el factor caótico de la única manera posible: acabando con la sociedad humana.

LEER EL ARTÍCULO ORIGINAL PULSANDO AQUÍ

Fotografía: Bloghemia

Fecha de creación
2025/07/06