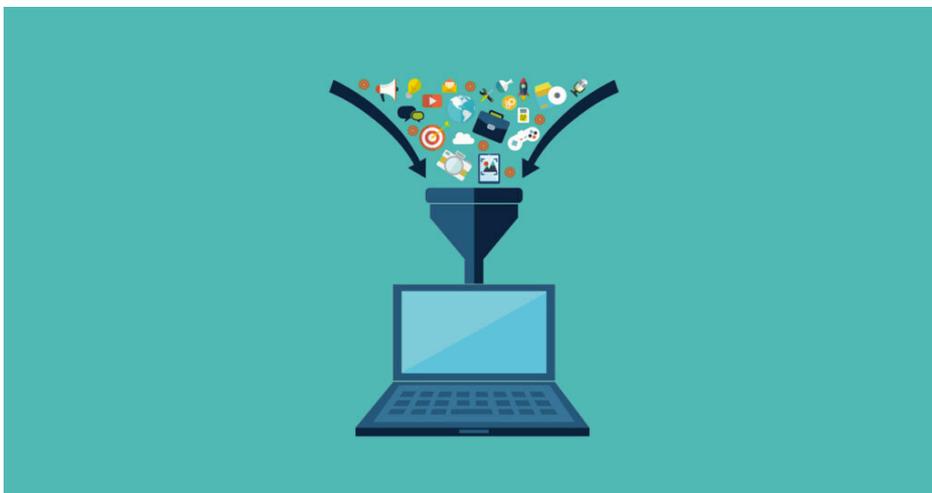


## Aarón Moreno Inglés

### Universidad Técnica de Delft . Miembro de Niaiá

Como es habitual en lo relacionado con el desarrollo tecnológico contemporáneo, nos encontramos en un momento de incertidumbre y sobreinformación a partes iguales. Para evaluar el estado de la cuestión sobre la inteligencia artificial (IA) y reflexionar sobre su regulación hay que tener en cuenta a multitud de actores implicados cuyas propuestas tienen un trasfondo político y económico que no se debe ignorar. Este ya es un buen punto de partida: asumir que cada proyecto de innovación o regulación se encuadra en un marco normativo determinado, y que cada comentario “meramente técnico” lleva una importante carga ideológica. Esto lo decía ya [Langdon Winner en 1980](#) y, aunque nos pueda resultar obvio, hay quien aún intenta sentar cátedra en nombre de la ciencia y la ingeniería como si fueran entes inmaculados, neutrales y libres de sesgo político. En los siguientes párrafos se ve reflejado el enésimo ejemplo de que estas personas, ya sea por falta de reflexión o con alevosía, no están en lo cierto.

Actualmente, la audiencia tecnológica global se encuentra asombrada ante el desarrollo frenético de ChatGPT y otros modelos grandes del lenguaje (*LLMs*, por sus siglas en inglés). Se trata de sistemas de inteligencia artificial que generan contenido sintético a través de técnicas de procesamiento del lenguaje natural, redes neuronales y amplísimas series de datos. A partir de una solicitud de contenido determinada (*prompt*), estos sistemas producen textos cuyo objetivo es el de imitar el pensamiento, la comprensión y la creatividad humanas. Para ello, utilizan palabras escogidas de forma estadística para convencer al usuario de que la respuesta generada es veraz, equiparable a la que un ser humano con gran acceso a información en línea podría ofrecer.



### [FFind Beyond Data](#)

Esto ha dado lugar a que gran cantidad de usuarios piensen que ChatGPT y otros LLMs son inteligentes al nivel de un ser humano, puesto que muchas de las respuestas a sus solicitudes de contenido tienen sentido. Incluso hay quien defiende fervientemente que estos sistemas están desarrollando conciencia. Nada más lejos de la realidad. **Los LLMs no son inteligentes, ni conscientes, como tampoco lo son el resto de sistemas de IA.** Su objetivo es unir al azar ciertas secuencias lingüísticas de acuerdo con información probabilística sobre cómo deberían combinarse para producir una respuesta adecuada, pero sin referencia al significado de estas combinaciones (Weil, 2023). Esto hace que ChatGPT sea realmente bueno a la hora de generar un texto que parezca humano pero que, en última instancia, no comprenda el significado del lenguaje que están procesando (Uddin, 2023).

Esta línea de crítica hacia el estado de la discusión en la IA ha sido en parte desarrollada

por Bender, Gebru, MacMillan-Major y Mitchell, y presentada en la conferencia FAccT de 2021, donde equipararon a los LLMs como “[loros estocásticos](#)”, porque producen palabras de forma aleatoria (estocástica) sin conocer realmente su significado (2021). Aquí, señalan múltiples limitaciones de comparar la inteligencia humana con las IAs. Por supuesto, sería necesario entrar en el terreno de la filosofía del lenguaje y la epistemología para acuñar definiciones rigurosas sobre inteligencia y consciencia. Este no es el propósito del artículo, aunque están ya disponibles investigaciones esclarecedoras sobre estos temas. En cuanto a lo que inteligencia se refiere, Kate Crawford señala que los LLMs no son equivalentes a ningún tipo de inteligencia humana, ya que no son capaces de discernir las tareas que llevan a cabo sin una amplia formación, conseguida a partir de (muchos) datos humanos, y tienen unas lógicas estadísticas completamente diferentes ([Crawford, 2021](#)). Además, como explica Rachel Adams (2021), conceptualizar la inteligencia simplemente como un conjunto de habilidades para la consecución de tareas sin mayor significado no solo es extremadamente limitador, sino que sigue una [perspectiva demasiado occidental](#), orientada a la eficiencia y a la productividad, sin reflexionar sobre el componente relacional de la misma u otros tantos. Y desde luego, ninguna teoría de la consciencia podría defender seriamente que producir texto de forma aleatoria denote que un ente sea consciente.

El problema con exagerar las capacidades de los sistemas de IA vuelve a ser, sin lugar a dudas, un problema político. Sin ir más lejos, el pasado mes de marzo el Future of Life Institute publicó una [carta](#) en la que solicitaba una moratoria de seis meses a los grandes proyectos de IA para evitar “entrenar sistemas de IA más potentes que GPT-4”, el LLM más grande hasta la fecha, creado por OpenAI. Esta propuesta la suscriben ya más de 30.000 personas, entre las que está el magnate Elon Musk. Hay quien podría pensar que tras su firma hay un profundo sentimiento de frustración, ya que él mismo se retiró del comité directivo de la empresa en Febrero de 2018, y podría tener interés en presionar para detener su crecimiento. Sin embargo, el asunto alberga mayor profundidad.

Esta carta de moratoria refleja una visión poco realista y extremadamente dañina de la IA, donde estos sistemas se presentan como un grave problema existencial para el futuro de la humanidad. La carta es una pieza más del discurso que está siendo promovido por las corrientes largoplacistas (*longtermism*), de Altruismo Efectivo, y demás instituciones multimillonarias ideológico-filantrópicas. Los problemas de la IA los estamos viviendo ya, y necesitan una regulación urgente: los sesgos racistas y machistas en los datos (especialmente graves en la aplicación de la IA en el sistema judicial), los problemas de privacidad y vigilancia masiva, los problemas relacionados con el impacto ecológico de su consumo de energía y su coste biofísico en minerales, los problemas de desplazamiento

laboral y de trabajo en condiciones paupérrimas, y otros muchos. **Al alentarnos a pensar en un futuro donde las máquinas desarrollen consciencia y en escenarios apocalípticos, se nos está distrayendo deliberadamente de los graves problemas actuales.** Sobre todo, teniendo en cuenta el rol de estas entidades filantrópicas en la economía política global, y su impacto en la jerarquización de los problemas de cara a su resolución.

### Ley de Inteligencia Artificial de la UE: Niveles de riesgo



### Telefónica

Esto no quiere decir en ningún caso que no haya que preocuparse por pandemias, desastres medioambientales, u otros riesgos globales con perspectiva de futuro. El problema está en que muchos de estos asuntos *ya son una realidad* y tenemos pruebas claras de lo que está por venir. El cambio climático no es tan solo importante porque suponga un riesgo humanitario futuro, como indican los largoplacistas, sino porque sus consecuencias o son inminentes, o se están ya viviendo actualmente, sobre todo en los países más empobrecidos. Referirse a estos temas como materia de futuro únicamente vuelve a denotar la perspectiva demasiado occidental que mencionaba Adams.

Podemos encontrar un buen resumen de estos problemas en una [respuesta](#) a la carta publicada por las autoras del artículo previamente mencionado sobre los loros estocásticos, en abril de este mismo año. Recomiendo su lectura íntegra pero, para muestra, un botón. Este párrafo, extraído de dicha respuesta, aclara y resume muy bien el asunto:

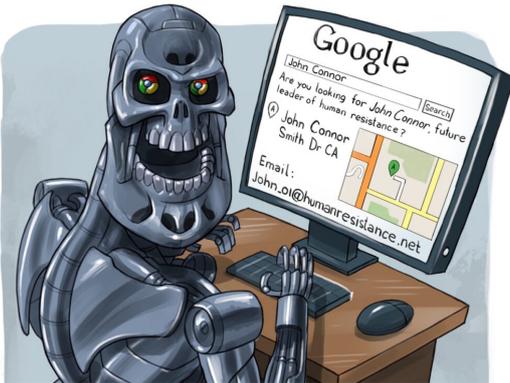
«Si bien hay en la carta una serie de recomendaciones con las que

*estamos de acuerdo (y propuestas en nuestro documento de 2021 conocido informalmente como “Loros estocásticos”), como 1) usar “sistemas de marcas de agua y procedencia para ayudar a distinguir los medios reales de los sintéticos”, estos se ven eclipsados por el alarmismo y la exageración de la IA , que dirige el discurso hacia los riesgos de una imaginaria “mente digital poderosa” con “inteligencia competitiva con la humana”. Esos riesgos hipotéticos son el foco de una ideología peligrosa llamada **largoplacismo** (longtermism) que ignora los daños reales que resultan del despliegue de sistemas de IA en la actualidad. **La carta no aborda ninguno de los daños continuos de estos sistemas**, para crear productos que lucran con un puñado de entidades, 2) la explosión de medios sintéticos en el mundo, que reproducen sistemas de opresión y ponen en peligro nuestro ecosistema de información , y 3) la concentración de poder en manos de unas pocas personas, lo que exacerba las desigualdades sociales»*

En última instancia, lo que expresan con este párrafo no es un desacuerdo total con la carta, sino que las decisiones sobre producción y diseño tecnológico (y de IA) están en manos de unas pocas personas, que pretenden ignorar los efectos nocivos que la IA ya tiene en la actualidad. Cabe recordar, entre otras polémicas sobre su proceso de producción que, sin ir más lejos, [ChatGPT realizó la moderación de contenido en Kenya con trabajadores cobrando apenas 2 dólares la hora, muchos de los cuales ahora sufren síndrome postraumático](#). En definitiva, muchos de los riesgos extremos que nombran en la carta funcionan como profecías autocumplidas. Como siguen diciendo, «*contrariamente a la narrativa de la carta de que debemos “adaptarnos” a un futuro tecnológico **aparentemente predeterminado** y hacer frente “a las dramáticas perturbaciones económicas y políticas (especialmente a la democracia) que causará la IA”, no estamos de acuerdo en que nuestro papel sea **ajustar a las prioridades de unos pocos privilegiados y lo que deciden construir y proliferar**. Deberíamos construir máquinas que funcionen para nosotros, en lugar de “adaptar” la sociedad para que sea legible y escribible por máquinas.*»

En resumidas cuentas, tras la carta de moratoria hay un trasfondo político que asume que

las decisiones sobre qué hacer y qué no hacer en materia de producción, diseño y regulación tecnológica deben seguir en manos del *bigtech*, con una cierta (mínima) supervisión académica. No sorprendería así que Elon Musk y otros grandes promotores tecnológicos firmasen la carta.



### Xataka

Sin embargo, entiendo que ciertas figuras académicas que han estado siempre más conectadas al estudio del transhumanismo y la parte más metafísica de la IA hayan firmado también. Al final, mucha gente experta en filosofía de la tecnología no se ha centrado en estudiar la estructura de producción o la economía política de la IA. Pero, si se exploran con cierta profundidad las relaciones político-económicas de diseño, producción, laborales, y extractivistas que hay detrás, es habitual llegar a ciertas conclusiones críticas. Como recomendación personal, las obras [Atlas of AI](#), de Kate Crawford (2021) o [La Filosofía Política de la IA](#), de Mark Coeckelbergh (2022) son un buen punto de partida para introducirse en la discusión. Al final, volviendo a Winner, se entiende que toda decisión sobre tecnología proyecta una ideología y refleja una determinada visión de la economía política. No debemos pasar esto por alto.

### **Fuentes:**

Adams, Rachel. (2021). Can artificial intelligence be decolonized?, *Interdisciplinary Science Reviews*, 46:1-2, 176-197, <https://doi.org/10.1080/03080188.2020.1840225>

Bender, Emily M.; Gebru, Timnit; McMillan-Major, Angelina y Shmitchell, Shmargaret. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? □ In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*

(FAccT '21). Association for Computing Machinery, New York, NY, USA, 610-623.

<https://doi.org/10.1145/3442188.3445922>

Coeckelbergh, Mark. (2022). The Political Philosophy of AI, Polity Press, Cambridge, UK, ISBN-13: 978-1509548545

Crawford, Kate on Simonite, Tom. (2021). This Researcher Says AI Is Neither Artificial nor Intelligent. Wired, 26/04/2021.

<https://www.wired.com/story/researcher-says-ai-not-artificial-intelligent/>

Crawford, Kate. (2021). The Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence. Yale University Press. <https://doi.org/10.2307/j.ctv1ghv45t>

Future of Life Institute. (2023). Pause Giant AI Experiments: An Open Letter. 22/03/2023.

<https://futureoflife.org/open-letter/pause-giant-ai-experiments/>

Gebru, Timnit; Bender, Emily M.; McMillan-Major, Angelina y Mitchell, Margaret. (2023). Las autoras de Stochastic Parrots sobre la "Pausa de la IA", Traducido en ARTEesÉTICA. 01/04/2023.

<https://arteesetica.org/declaracion-de-las-autoras-enumerados-de-stochastic-parrots/>

Palmer, Sarah y Khatsenkova, Sophia. (2023). 'I want to be alive': Has Microsoft's AI chatbot become sentient? *Euronews*, 18/02/2023.

<https://www.euronews.com/next/2023/02/18/threats-misinformation-and-gaslighting-the-unhinged-messages-bing-is-sending-its-users-rig> .

Perrigo, Billy. (2023). Exclusive: OpenAI Used Kenyan Workers on Less Than \$2 Per Hour to Make ChatGPT Less Toxic. *TIME Magazine*. 18/01/2023.

<https://time.com/6247678/openai-chatgpt-kenya-workers/>

Uddin, Muhammad Saad. (2023). Stochastic Parrots: A Novel Look at Large Language Models and Their Limitations. 13/04/2023.

<https://towardsai.net/p/machine-learning/stochastic-parrots-a-novel-look-at-large-language-models-and-their-limitations>

Weil, Elizabeth. (2023). You Are Not a Parrot And a chatbot is not a human. And a linguist named Emily M. Bender is very worried what will happen when we forget this. *New York Magazine*, 01/03/2023.

<https://nymag.com/intelligencer/article/ai-artificial-intelligence-chatbots-emily-m-bender.html>

Winner, Langdon. (1980). "Do Artifacts Have Politics?" *Daedalus*, vol. 109, no. 1, 1980, pp. 121-36. *JSTOR*, <http://www.jstor.org/stable/20024652>.

### Para citar esta entrada

Moreno Inglés, Aarón. (2023) Inteligencia Artificial: ideología tras la carta de moratoria. En Niaiá, <https://niaia.es/inteligencia-artificial-ideologia-tras-la-carta-de-moratoria/>

Si lo desea, puede volver a publicar este artículo, en forma impresa o digital. Pero le pedimos que cumpla estas instrucciones: por favor, no edite la pieza, asegúrese de que se la atribuye a su autor, a su institución de referencia (universidad o centro de investigación) en Niaiá.

