

La metáfora del shoggoth en la inteligencia artificial

The shoggoth metaphor in the Artificial Intelligence

Tadeo Masís González

Universidad de Costa Rica, Costa Rica.

tadeo.masis@ucr.ac.cr

ORCID: <https://orcid.org/0000-0003-2867-9011>

Fecha de recibido: 14-06-2024

Fecha de aceptación: 24-10-24

Resumen

Mediante la metáfora del *shoggoth*, la comunidad alrededor del desarrollo de los grandes modelos de lenguaje en la inteligencia artificial describe características particulares de esta tecnología. Dadas sus condiciones de opacidad y aparente inconmensurabilidad, estos modelos algorítmicos llaman a cuestionar algunas propiedades epistemológicas de la tecnología y su rol en el proyecto moderno. El presente texto tiene como propósito exponer reflexiones derivadas del uso de esta metáfora en el contexto de la inteligencia artificial, de manera que sea posible señalar paradojas para la sociedad moderna, así como realizar un bosquejo de lo que podrían ser nuevas propiedades de la tecnología.

Palabras clave: Inteligencia artificial, filosofía de la tecnología, metáfora tecnológica, H.P Lovecraft, epistemología.

Abstract

Through the shoggoth metaphor, the community around the development of large language models in artificial intelligence describes particular characteristics of this technology. Given their conditions of opacity and apparent incommensurability, these algorithmic models call to question some epistemological properties of technology and its role in the modern project. The purpose of this text is to present reflections extracted from the use of this metaphor in the context of artificial intelligence, so that it is possible to point out paradoxes for modern society and draw what could be new properties of technology.

Key words: Artificial intelligence, Philosophy of technology, technological metaphor, H.P Lovecraft, epistemology.

1. Introducción

Durante el año 2023, dentro de la comunidad dedicada al desarrollo de la inteligencia artificial, se ha popularizado la caricatura de un monstruo del universo literario de H.P Lovecraft denominado *shoggoth*. El propósito de esta ilustración difundida en redes sociales es representar las cualidades que se le atribuyen a la inteligencia artificial (Nicholson,

2023), especialmente el alto grado de desconcierto que experimentan los desarrolladores de estas tecnologías con sus creaciones (O'Connor, 2024). Esta criatura ficticia es descrita en la obra literaria de Lovecraft (2020) de la siguiente manera: «Era un algo indescriptible, terrible, mayor que cualquier tren subterráneo, un conjunto informe de protoplasma burbujeante, sutilmente luminoso y con miríadas de efímeros ojos que se formaban y desvanecían conti-

nuamente como pústulas de luz vercosa» (p.193). De este modo, la idea de *shoggoth* alude a lo terrorífico, carente de forma e indescifrable. Al mismo tiempo, dentro del canon lovecraftiano, también se plantea que en algún momento de la historia estas criaturas se rebelaron contra sus creadores, lo que añade significado a esta referencia dentro del ámbito de la inteligencia artificial (Calia, 2023).

Dicha metáfora se emplea específicamente para referirse a los grandes modelos de lenguaje (LLM por sus siglas en inglés), que en el contexto de la inteligencia artificial son potentes algoritmos —como el *GEMINI* desarrollado por Google o las diferentes versiones del GPT desarrolladas por *OpenAI*— que derivan en productos como el ampliamente conocido *Chat GPT* (Bowman, 2023). Precisamente, la comparación de estos modelos con la noción de *shoggoth*, busca condensar algunas preocupaciones compartidas por quienes desarrollan estas tecnologías. Así, esta criatura monstruosa representa el carácter terrorífico y potencialmente incontrolable que atribuyen a esta clase de algoritmos, cuyos procesos de razonamiento son en algunos casos enigmáticos y parcialmente desconocidos para sus creadores (Farrell & Shalizi, 2023; Roose, 2023). A su vez, se ha planteado que la ilustración del *shoggoth* refleja también la necesidad de problematizar el sentido de otredad que genera la interacción humano/máquina en el contexto de nuevos desarrollos de la inteligencia artificial (Grimmelman, 2024; Milička, 2024).

Aunque pueda parecer una metáfora trivial, la comparación entre un *shoggoth* y los modelos de lenguaje de la inteligencia artificial puede ser muy sugerente para discutir nuevas propiedades de la tecnología. En primer lugar, el atribuir imprevisibilidad e insubordinación a un producto de la ciencia moderna, resulta altamente paradójico y marca una ruptura importante con el estatus epistemológico que tradicionalmente se le adjudica a la tecnología. También, si tomamos como referencia los imperativos de racionalidad de la sociedad moderna, no podría concebirse un escenario donde los desarrolladores de una tecnología desconocen aspectos sobre la manera en que funcionan sus creaciones. Es

decir, esta metáfora no corresponde con la idea de un mundo racional, librado de pensamiento mágico y enfocado en la previsión por medio de los mismos productos de la razón.

En este marco, también cabe señalar que en obras de la filosofía contemporánea se han descrito características de la obra literaria de Lovecraft muy pertinentes para aproximarse a esta idea de algoritmos como *shoggoths*. En primer lugar, Graham Harman (2012) señala que la obra de Lovecraft viene a establecer una brecha entre los objetos y sus cualidades, o bien entre la realidad y su accesibilidad. Es decir, las extrañas criaturas cósmicas de este universo literario darían cuenta de aquello que puede ser real, pero imperceptible. Por otra parte, se ha argumentado también que este tipo de ficción pone de manifiesto lo *extraña* que puede resultar la realidad y lo familiar que puede resultar lo sobrenatural (Fisher, 2018). Por ejemplo, un agujero negro —propio de la naturaleza— resulta ajeno a la experiencia cotidiana de los seres humanos, mientras que la idea de un vampiro o un hombre lobo —seres sobrenaturales— es bastante común y fácil de asimilar. Así, un punto de partida para abordar la relación entre *shoggoth* y tecnología es considerar la obra lovecraftiana como un medio para reflexionar sobre la inconmensurabilidad de lo real.

El presente trabajo consta de tres apartados. En el primero de ellos, se señala la situación paradójica en que uno de los desarrollos tecnológicos más recientes de la sociedad moderna se percibe como terrorífico e inquietante por parte de las personas a cargo de su producción. En un segundo apartado se señalan las grandes implicaciones epistemológicas y ontológicas de una tecnología a la que se atribuyen las cualidades de un *shoggoth*. Finalmente, en un tercer apartado se exponen conclusiones alrededor de la metáfora en los modelos de lenguaje de la inteligencia artificial y se aportan ideas generales sobre lo que podrían ser algunas cualidades incipientes de la tecnología.

2. Un paradoja para el proyecto moderno

Si consideramos los grandes modelos de lenguaje como uno de los más recientes productos de la razón tecno-científica —derivada del proyecto moderno—, su comparación con una criatura monstruosa e indescifrable puede implicar una gran paradoja. El surgimiento de la sociedad moderna supuso el *desencantamiento del mundo*; una racionalización de la vida orientada al control, la medición y el cálculo (Weber, 2004). Este proceso derivó en la secularización de las sociedades occidentales, el progresivo abandono del pensamiento mágico y la emergencia de la tecno-ciencia como eje fundacional de la modernidad. También, considerando el desarrollo de las tecnologías contemporáneas, el presente puede caracterizarse como una época marcada por la transparencia, donde los crecientes flujos de datos e información contribuyen a la erosión de cualquier rasgo de negatividad, o dicho de otra manera, de aquello con propiedades ocultas o inconmensurables (Han, 2013).

En este marco, la atribución de esta clase de opacidad a los modelos de la inteligencia artificial implica tanto una contradicción como una paradoja con respecto al imperativo moderno de previsión y transparencia. En primer lugar, resulta altamente contradictorio que aquella tecnología llamada a emplear los flujos de información para profundizar la medición y el control de las actividades humanas, sea asociada con el terror y lo oculto. En este caso, la impronta positivista que fundamenta la construcción de modelos algorítmicos deviene en negatividad; los procesos de razonamiento de estas tecnologías se vuelven indescifrables aún para sus creadores. En consecuencia, también resulta paradójico que la infraestructura diseñada para procesar grandes cantidades de información, genere más bien un clima de ansiedad e incertidumbre para una comunidad tecno-científica.

Si bien se ha reconocido la incertidumbre como un rasgo creciente y de gran complejidad en el ámbito de las decisiones técnicas (Callon et al., 2009), la

ciencia y la tecnología persisten como medios para la consecución de los imperativos del proyecto moderno. Para evidenciar este rasgo de las sociedades contemporáneas es pertinente acudir a la noción de *hipermodernidad* acuñada por el sociólogo francés Gilles Lipovestky (2018), quien plantea que nos encontramos en una sociedad *hipermoderna*, donde la dimensión científica de la modernidad se impone a los proyectos utópicos de índole política: “No se ha destruido la fuerza del futuro: lo que ocurre es que éste ya no es ideológico-político, sino que se sustenta en la dinámica técnica y científica” (p.71). De esta manera, el eje de la modernidad se sitúa más que nunca en la promesa de los desarrollos técnicos para alcanzar el mito del progreso.

Esta lectura de la modernidad evidencia la magnitud de la paradoja que supone una representación de la tecnología —por parte de sus mismos creadores— como un *shoggoth*. Al verse desprovista de grandes utopías políticas, la fase actual de la modernidad refleja algún grado de incertidumbre ontológica en las sociedades occidentales. Sin embargo, su confianza en la tecno-ciencia es continua, el desarrollo de tecnologías se valora positivamente e incluso ha motivado el desarrollo de nuevas ideas políticas (Avanessian y Reis, 2017). La inconmensurabilidad que se le atribuye a la inteligencia artificial es sumamente paradójica si consideramos que la búsqueda de certeza y cálculo racional por medio de algoritmos es uno de los más recientes instrumentos de la sociedad *hipermoderna* en su búsqueda por la utopía del progreso. Sin embargo, el hecho de que aquellas personas dedicadas a la producción de modelos de inteligencia artificial carezcan de certeza sobre el alcance de sus creaciones, marca un punto de ruptura; la fe en los avances técnicos como medios para obtener control y transparencia, deriva en este caso en profundas incógnitas y *cajas negras*.

Evidentemente esto no implica que hasta el momento no hayan existido críticas o temores en torno a los productos de la modernidad. Por ejemplo, los teóricos de la escuela de Frankfurt mostraban ya en la primera mitad del siglo XX gran consternación por las transformaciones culturales de la

sociedad moderna (Adorno & Horkheimer, 1970) y por la percepción del mito del progreso como una fuerza incontenible (Benjamin, 2008). De igual forma, el sociólogo Bauman (2008) expone la dimensión destructiva de la modernidad por medio de su descripción del Holocausto como un producto de la razón moderna, apoyada en la industria química y el diseño de eficientes procesos operativos:

Llegó [...] sobre un vehículo construido en una fábrica, empuñando armas que solo la ciencia más avanzada podía proporcionar y siguiendo un itinerario trazado por una organización científicamente dirigida. [...] Fue el mundo racional de la civilización moderna el que hizo que el Holocausto pudiera concebirse. (p.34)

En este marco, la paradoja del *shoggoth* y la sociedad moderna adquiere una dimensión más profunda. En contraste con estas visiones alrededor de la modernidad, la preocupación en torno al carácter *monstruoso* de la inteligencia artificial expresa más bien el temor a que los productos de la racionalidad moderna no funcionen como tal. Así, lo que genera ansiedad entre las personas que trabajan con estos complejos algoritmos no es la lógica racional de previsibilidad, medición y control, sino la ausencia de esta. Es decir, el *shoggoth* no representa un temor o preocupación sobre las consecuencias del mundo moderno, sino sobre la pérdida de su vigencia; más tecnología implica en este caso menor certeza.

3. Propiedades ocultas de la máquina

El uso recurrente de esta metáfora alrededor de los grandes modelos de lenguaje de la inteligencia artificial implica también una ruptura con las premisas epistemológicas que usualmente se le atribuyen a la tecnología. La idea de una *tecnología-shoggoth* conduce a cuestionar características como la precisión, el control y la ausencia de secretos en las máquinas. El filósofo Evandro Agazzi (1998) plantea que en tanto la tecnología permite constituir el objeto de estudio de cualquier ciencia empírica, esta es un elemento constitutivo de la ciencia moderna. En concreto, la tecnología es descrita como un ámbito

necesariamente transparente y confiable para la observación de los fenómenos naturales y sociales:

¿Qué representa la máquina? La máquina posee una enorme ventaja: dentro de ella no hay secretos, todo es conocido, quizás no para mí pero sí para el técnico o el especialista. Si se estropea, se puede arreglar, pues está construida según un proyecto que permite saber cómo funciona y devolverla a su funcionamiento. Antes de construirla sabemos cómo funcionará y por qué funcionará así y esto no se conoce por la experiencia propiamente dicha, sino que se sabe antes de que exista. (Agazzi, 1998, p.29)

Desde esta óptica, la tecnología sería la consecución del ideal mecanicista: un sistema sin secretos, fácilmente descifrable y absolutamente previsible. En un plano similar, el filósofo Byung-Chul Han (2013) plantea que: «Solo la máquina es transparente» (p.14), para indicar que en contraste con los seres humanos, las máquinas carecen de propiedades ocultas —o negatividad—. De este modo, la metáfora del *shoggoth* vendría a plantear un desafío a estas perspectivas; la opacidad no sería un monopolio de la condición humana, sino que la tecnología podría tener secretos, incluso para sus mismos creadores. En particular, el factor de previsibilidad atribuido a las máquinas —que asume un conocimiento pleno sobre sus funciones— es superado; los temores cristalizados en la representación *monstruosa* de los modelos de lenguaje, evidencian secretos e incertidumbre para las mismas personas que conocen y diseñan estas tecnologías.

La atribución de opacidad a la inteligencia artificial implica también situar a las máquinas al mismo nivel de los seres humanos. Esta concepción sobre la tecnología supone no solo la imposibilidad de conocer plenamente a las máquinas, sino la imposibilidad de conocer por medio de estas. Es decir, si es gracias a la fiabilidad de las máquinas que la tecnología es una condición *sine qua non* del desarrollo de la ciencia moderna, la posibilidad de no conocer todos los procesos que esta desarrolla, genera también obstáculos para medir y controlar los diversos ámbitos que el ser humano quiera abordar por medio

de estas herramientas.

Una de las ansiedades más extremas en torno a estas ideas es la posibilidad de que la máquina mida y controle al ser humano por su cuenta. Desde este lugar pueden emanar diversas ideas en torno a la singularidad tecnológica —o una *rebelión de las máquinas*—, tópico recurrente en la ciencia ficción y en los escritos del filósofo Nick Land (2017) de la última década del siglo XX: «Nada humano consigue escapar del futuro próximo» (p.51). Pero más allá del potencial especulativo de esta situación, la relevancia de la noción de *shoggoth* reside en este caso en la pérdida de confianza absoluta en la tecnología como medio para aproximarse a la realidad.

Al mismo tiempo, esta aparente pérdida de control sobre las máquinas rompe la idea —al menos en el plano epistemológico— de que la tecnología sea: «la construcción de *otro mundo* al lado del mundo de la naturaleza» (Agazzi, 1998, p.29). En otras palabras, el caso de los modelos de lenguaje, muestra que las propiedades de imprevisibilidad y ausencia de control que se le atribuyen a la naturaleza no son diferentes de las que pueden manifestarse en la tecnología. Es decir, la *emancipación* que puede atribuirse a estos algoritmos no está en función de emular de la capacidad humana de pensar, como podría derivarse del *Imitation Game* de Alan Turing (1950) o del *experimento de la habitación china* de John Searle (1980). Más bien, la ruptura que se plantea tiene un corte epistemológico con respecto a la posibilidad de conocer a las máquinas y por medio de estas.

Por tanto, la inteligencia artificial vista desde esta perspectiva puede concebirse como algo distinto a las máquinas tradicionales, para ello podría ser pertinente la noción de cyborg: «criatura híbrida, compuesta de organismo y de máquina» (Haraway, 1995, p.62). Así, en tanto este tipo de tecnologías constituyen una suerte de cyborgs, la criatura indescifrable con la que se compara a los modelos de lenguaje de la inteligencia artificial integra propiedades de máquina y de naturaleza, no a un nivel físico, sino a un nivel semiótico y epistemológico.

De este modo, el carácter terrorífico atribuido a la inteligencia artificial radica en su naturaleza híbrida. La experiencia de la comunidad dedicada al desarrollo de estas tecnologías, evidencia una brecha entre la caracterización moderna de la tecnología y su incipiente mezcla de transparencia y opacidad.

4. Conclusión: las propiedades incipientes de la tecnología

Es posible plantear que en su búsqueda por mayor certeza y control sobre las actividades humanas, las personas dedicadas a los modelos de inteligencia artificial han terminado por *invocar shoggoths*; desde su propia perspectiva, han topado con que sus creaciones son *monstruos* inconmensurables e impredecibles. La opacidad de los grandes modelos de lenguaje sitúan metafóricamente a programadores y otros tecnólogos en una especie de retorno a la alquimia: su ámbito de conocimiento integra hechos tecno-científicos con terrores esotéricos. En este marco, el desarrollo de estas reflexiones resulta tanto sugerente como necesario para observar pequeños pero importantes cambios en la forma en que se percibe a la tecnología.

En este breve artículo se han explorado las implicaciones ontológicas y epistemológicas del uso de la metáfora del *shoggoth* en la inteligencia artificial. En primer lugar, esta comparación entre tecnología y una criatura inconmensurable muestra una gran contradicción en las pretensiones del proyecto moderno de control y previsión. La imprevisibilidad que los tecnólogos han detectado en el ámbito de la inteligencia artificial, indica que el imperativo moderno de racionalización posee límites dentro de su propio desarrollo tecnológico, a su vez se muestra que una tecnología más avanzada no se traduce en mayor control y fiabilidad. Es paradójicamente en el desarrollo de tecnologías donde podemos encontrar preocupaciones por la falta de previsión y transparencia.

En segundo lugar, este caso permite observar lo que podrían ser algunas *cualidades incipientes de la tecnología*. La opacidad de los grandes modelos de

lenguaje invita a considerar la existencia de tecnologías que denominamos *híbridas*, que funcionan entre el control y la incertidumbre; mezclan propiedades de las máquinas y la naturaleza como parte de sus operaciones regulares. También, este tipo de tecnologías incorporan rasgos de insubordinación, no en términos de una *emancipación de las máquinas*, pero sí en un sentido epistemológico. La *monstruosidad* que se le adjudica a estas aplicaciones de la inteligencia artificial podría ser síntoma de un cambio en la confianza con la que el ser humano aborda la realidad a través de las máquinas, o al menos puede sugerir una nueva acepción alrededor de la noción de tecnología.

5. Referencias bibliográficas

- Adorno, T & Horkheimer, M. (1971). *Dialéctica del iluminismo*. Sur.
- Agazzi, E. (1998). El impacto epistemológico de la tecnología. *Argumentos de Razón Técnica*, (1) 17-31.
- Avanessian, A. y Reis, M. (2017). Introducción. En A. Avanessian & M. Reis (Comps.), *Aceleracionismo: Estrategias para una transición hacia el poscapitalismo* (pp. 9-31). Caja Negra.
- Bauman, Z. (2008). *Modernidad y Holocausto*. Sequitur.
- Benjamin, W. (2008). *Tesis sobre la historia y otros fragmentos*. Ítaca.
- Bowman, S. R. (2023). Eight things to know about large language models. *arXiv preprint*. <https://doi.org/10.48550/arXiv.2304.00612>
- Calia, M. (12 de junio de 2023). The world's top H.P. Lovecraft expert weighs in on a monstrous viral meme in the A.I. world. *CNBC*. <https://www.cnbc.com/2023/06/12/lovecraft-joshi-shoggoth-ai-meme.html>
- Callon, M., Lascoumes, P. & Barthe, Y. (2009). *Acting in an Uncertain World. An essay on Technical Democracy*. MIT Press.
- Farrell, H. & Shalizi, C. (24 de junio de 2023). Artificial intelligence is a familiar-looking monster. *The Economist*. <https://www.economist.com/by-invitation/2023/06/21/artificial-intelligence-is-a-familiar-looking-monster-say-henry-farrell-and-cosma-shalizi>
- Fisher, M. (2018). *Lo raro y lo espeluznante*. Alpha Decay.
- Grimmelmann, J. (2024). The Defamation Machine. SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4914458
- Han, B-C. (2013). *La sociedad de la transparencia*. Herder.
- Haraway, D. J. (1995). *Ciencia, cyborgs y mujeres: La reinención de la naturaleza*. Cátedra.
- Harman, G. (2012). *Weird Realism: Lovecraft and Philosophy*. Zero Books.
- Land, N. (2017). Colapso. En A. Avanessian & M. Reis (Comps.), *Aceleracionismo: Estrategias para una transición hacia el poscapitalismo* (pp. 49-64). Caja Negra.
- Lipovetsky, G. (2018). *Los tiempos hipermodernos*. Anagrama.
- Lovecraft, H. P. (2020). *Antología de Relatos de Terror*. Plutón.
- Milička, J. (2024). Theoretical and Methodological Framework for Studying Texts Produced by Large Language Models. *arXiv preprint*. <https://doi.org/10.48550/arXiv.2408.16740>
- Nicholson, K. (1 de junio de 2023). This Terrifying Meme Is Capturing All Of Our Deepest Fears About AI. *HuffPost*. https://www.huffingtonpost.co.uk/entry/artificial-intelligence-shoggoth-fears_uk_64789526e4b045ce2486feff
- O'Connor, M. (2024). In the Craftsman's Garden: AI, Alan Turing, and Stanley Cavell. *Minds & Machines* 34, 22. <https://doi.org/10.1007/s11023-024->

[09676-y](#)

Roose, K. (23 de mayo de 2023). Why an Octopus-like Creature Has Come to Symbolize the State of A.I. *The New York Times*. <https://www.nytimes.com/2023/05/30/technology/shoggoth-meme-ai.html>

Searle, J. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3), 417–424. <https://doi.org/10.1017/S0140525X00005756>

Turing, A. (1950). Computing machinery and intelligence. *Mind* 59(236), 433–460. <https://doi.org/10.1093/mind/LIX.236.433>

Weber, M. (2004). *La ética protestante y el Espíritu del Capitalismo*. Fondo de cultura económica.