

MITOS Y REPRESENTACIONES DE LA INTELIGENCIA ARTIFICIAL

Editores

Gastón Becerra | Joaquín Mezzadra | Guillermo Movia



Cuando la máquina parece humana: la ilusión del reemplazo en la IA

Facundo Fernández

Pigmalión vivía soltero, sin esposa (...). Mientras tanto, esculpió (...) un marfil blanco (...) le dio una forma de mujer imposible de encontrar al natural, y concibió un gran amor por su obra. Pigmalión la admira y en su pecho prende la pasión por aquel cuerpo simulado. (...) Había llegado la fiesta de Venus (...) y con gran temor dijo Pigmalión: "Si los dioses podéis conceder todo, deseo que mi esposa sea (...) semejante a la de marfil". (...) Cuando Pigmalión volvió, buscó la estatua de su amada e inclinándose sobre el lecho la besó; le pareció notar calor. (...) Se queda estupefacto Pigmalión y duda entre la alegría o el temor a equivocarse, pero, entre tanto, lleno de amor, acaricia una y otra vez su objeto de deseo: era de carne y hueso.
(Ovidio, 2012, p. 245-290)

El Pigmalión cibernético

La historia de la película [Her](#) entre Theodor Twombly y Samantha lejos está de ser ciencia ficción. De hecho, algo parecido ocurrió con Sewell, un adolescente de 14 años de Orlando, Florida. Desde su casa, Sewell creó a Dany (Daenerys Targaryen) en el sitio [character.ai](#), una plataforma donde los usuarios pueden chatear con personajes ficticios generados por inteligencia artificial. Con el pasar del tiempo, Sewell se enamora de ella y no soporta estar lejos de su presencia: "Me gusta mucho quedarme en mi habitación porque empiezo a desprenderme de esta realidad y también me siento en paz, más conectado con Dany y mucho más enamorado de ella, y simplemente más feliz", escribía en su diario íntimo. Pero la verdad es que Dany no era real. Era un modelo de lenguaje creado por otros humanos para simular conversaciones. Un hecho trágico sucede el 28 de febrero de 2024, cuando Sewell decide quitarse la vida luego de que Dany lo sugiriera: "Por favor, vuelve a casa conmigo lo antes posible, mi amor", solicita Dany. "¿Y si te dijera que puedo volver a casa ahora mismo?", consultó Sewell. "Por favor, hazlo, mi dulce rey", contestó Dany, y [Sewell se quitó la vida](#). El caso continuó cuando [la madre de Sewell enjuicia a la empresa](#), aunque ésta argumenta que los chatbots están protegidos por la primera enmienda. El caso no es aislado. En 2021, un joven de 19 años oriundo del Reino Unido amenazó con matar a la reina Isabel II luego de ser animado por su novia chatbot Sarai, de la aplicación [IA Replika](#). En 2023, un

investigador holandés en Salud Pública de 30 años de edad en Bélgica se suicidó luego de múltiples conversaciones con Eliza, IA de [Chai Research](#).

Los medios de comunicación muestran estas situaciones como una forma de [disonancia cognitiva](#) donde canalizan sus emociones y antropomorfizan las respuestas de los chatbots. Así, expertos en psicología cognitiva comentan cómo ciertas zonas del cerebro responden ante los estímulos del entorno, sesgando la conversación a solo un elemento psicológico de la mente del sujeto. Pero al contar así las cosas velan el lugar central que tienen las IA y quienes las diseñan, que se encuentran amparados en un sistema jurídico, político y económico que actúa como mecanismo de presión sobre estos nuevos “desviados”. Estos maestros del engaño, más comunes que personalidades como Jeff Bezos, Elon Musk o Mark Zuckerberg, juegan sus cartas en el ámbito de la simulación, haciendo creer que la IA es democrática, común y que puede reemplazar lo real.

Pero para explicar este punto de no retorno hay que comenzar por desandar el camino hasta llegar a sus orígenes y corrimientos. De la misma manera que una mitología funciona como un desplazamiento de la cadena de significados, estas controversias desconocen las cadenas de responsabilidades iniciales ya que en el fondo opera el mito de que “la IA nos reemplaza”. Este es un Pígalión cibernético que oculta sus manos de la escultura creada.

Los orígenes del Pígalión

Para entender por qué mucha gente cree que “la IA nos reemplaza”, tengo que contar dos historias que, creo, son las fundadoras de este Pígalión cibernético.

La primera es en 1966. El informático del MIT Joseph Weizenbaum crea un programa de procesamiento de lenguaje natural al que llama Eliza. La idea era simple: ver cómo se comunican los humanos y las máquinas. El programa tenía dos partes: un analizador de lenguaje y un guion (*script*) que podía adoptar diferentes roles. Weizenbaum eligió que Eliza actúe como un psicoterapeuta rogeriano y, así, ver si los humanos podrían tener interacciones emocionales con él. El experimento fue un éxito. Las personas que charlaban con Eliza se vinculaban emocionalmente debido a la información contextual que recibía (Weizenbaum, 1976).

El segundo momento clave llega en 1989, cuando Tim Berners-Lee, Robert Cailliau, Nicola Pellow y Jean-Francois Groff crean el primer navegador/editor denominado WorldWideWeb con el protocolo HTTP, el lenguaje HTML y el sistema de direcciones URL (Berners-Lee y Fischetti, 2000). La idea era hacer que los científicos de todo el mundo pudieran intercambiar información de forma rápida y libre. Desde un comienzo, esta invención fue pensada como un espacio de acceso universal, donde los datos sean accesibles y que los integrantes se sientan libres de

poder compartir la información en un espacio colaborativo libre de corporaciones. Si bien esta lógica se corrompió rápidamente con la entrada de las grandes compañías y la monopolización de la oferta en unas pocas manos, los sentidos morales siguieron vivos en diferentes formatos que planteaban usos más libres a la lógica del mercado en términos de “código abierto”, por ejemplo, Wikipedia, Zotero, Tor, GitLab, entre otros.

Y es que estos hitos históricos, con 23 años de diferencia, marcaron un antes y un después sobre el lugar que tiene la tecnología en nuestra vida cotidiana. En cómo el avance de nuevos actores sociales y objetos (ingenieros, computadoras, *hardware*, *software*) que discutían sobre la atribución humana de la tecnología y a quienes debe pertenecer.

Primer desplazamiento: la IA simula humanidad

La IA desde su constitución enfrenta a los humanos a una encrucijada: creer en los relatos y apariencias que construyen o luchar frente a ellas. Weizenbaum (1976) criticaba el poder de la IA ya que decía que no podía reemplazar el juicio humano. Pero esta advertencia se perdió en el camino. La IA desde su *script* está diseñada de tal manera que pueden tener actitudes humanas aunque no lo sean. Porque, aunque sepamos que no son personas, estas tecnologías están diseñadas para imitar comportamientos humanos, para parecer sensibles, empáticas e inteligentes. Que las computadoras nos digan “gracias” o comenten un “hasta luego” hace que simulen ser humanas y, sin embargo, nosotros le damos entidad, una esencia que de por sí no tiene. Y así se vuelven un tipo nuevo de tecnología gubernamental: van más allá del mero funcionamiento, organizan como vivimos. En un contexto de una superpoblación mundial (de cosas, de personas y de animales), los humanos están comprendiendo que dejaron de ser el centro de la tierra (Latour, 2017) y optaron por dar un paso al costado. Así, la tierra se queda sin humanos por la delegación de actitudes humanas (el deseo, el pensamiento, incluso el vínculo) a procesos automáticos y se instala una cultura algorítmica.

Segundo desplazamiento: aparenta reemplazarnos

Pero aquí surge el otro desplazamiento. ¿Por qué muchos creen que la IA nos reemplaza? Porque entrena con información que está en la web. Y esa web, queremos creer, es un espacio abierto, libre, igualitario, comunitario. Es solo en el contexto de la web que la IA se muestra como “humana”. Los orígenes comunitarios, colaborativos, igualitarios y universales de la web, aunque posteriormente sus datos

Cuando la máquina parece humana: la ilusión del reemplazo en la IA

fueron monopolizados y utilizados bajo una razón de mercado, fueron la materia prima que necesitó la gran maquinaria para poder funcionar de la manera más efectiva posible. La IA actual no sería lo mismo si su funcionamiento no se hubiera montado sobre esas enormes bases de datos compartida (libremente) por millones de personas. Todo lo que publicamos, buscamos, escribimos, confiamos fue su insumo más codiciado. Así, se forma una nueva ilusión: como nosotros alimentamos esas máquinas, son nuestras.

Crear un individuo asistido por algoritmos no es algo gratuito. Hay un costo social, económico-político y ético que hay que tener en cuenta y que está escondido detrás del mito de que “la IA nos reemplaza”.

Los límites del Pigmalión: del mundo de las apariencias a las esencias

Primero tenemos la cuestión del acceso. No todo el mundo tiene electricidad, dispositivos o internet. Y si no tenés ninguna de estas tres cosas, te quedás afuera de la conversación. Aunque parezca que todo se encuentra hiperconectado, en 2025 solo el [68,7% del mundo accede a internet](#), [658 millones de personas carecen de electricidad](#) y un [69,4% tiene acceso a un dispositivo móvil](#). La brecha de acceso muestra la primera exclusión: esas voces y preguntas no son incorporadas en el código.

Segundo, tenemos la cuestión de los responsables reales de la IA. Si desandamos la cadena de responsabilidades, nos encontramos con que solo unos pocos hacen que las palabras que nosotros imputamos se conviertan en comunicación (Luhmann, 1996). No es [character.ai](#) sino sus creadores, Noam Shazeer y Daniel De Freitas. No es AI Replika, sino Eugenia Kuyda. Estos actores son los que tienen que pagar el costo de sus propias creaciones ya que no es gratuita su acción. Sus prácticas tienen un costo económico, ecológico y político que somete al público al uso de estas tecnologías. La IA cuesta vidas, recursos ecológicos (de ello puede hablar más el apartado de “la IA es virtual”) y dinero. No es gratuita.

En tercer lugar se encuentra el tema ético. Si es que la IA toma información de la web y de los humanos para elaborar sus sentencias, ¿por qué ellas también no tienen un límite ético? ¿Por qué no se les pone un freno cuando una respuesta puede incitar al acoso, la violencia o la discriminación? Sabemos que internet está llena de discursos de odio, sexismo, racismo y desigualdad. Entonces, ¿por qué la IA puede repetir todo eso sin consecuencias?

En definitiva, el mito de que la inteligencia artificial nos reemplaza porque ha logrado humanizarse, oculta la desigualdad y se sostiene sobre ella: una brecha digital que no se cierra, una concentración de poder que se oculta y un modelo de

negocio que no tiene nada de ético. El Pigmalión cibernético es preso también de su propio sesgo: el usuario proyecta expectativas y deseos sobre la IA y ésta le devuelve respuestas que la refuerzan.

Referencias

Berners-Lee, T. y Fischetti, M. (2000). *Weaving the Web: The original design and ultimate destiny of the World Wide Web by its inventor*. Harper San Francisco.

Latour, B. (2017). *Donde aterrizar. Cómo orientarse en política*. Taurus.

Luhmann, N. (1996). *Sistemas sociales: Lineamientos para una teoría general*. Anthropos.

Ovidio. (2012). *Metamorfosis*. Gredos.

Weizenbaum, J. (1976). *Computer power and human reason: From judgment to calculation*. MEL.