



No tan humanos

Acerca de la antropomorfización de los Agentes No Humanos de Comunicación

JULIÁN TAGNIN (UNPAZ)
16 DE AGOSTO DE 2024

Introducción

La elucidación fenomenológica de conceptos como los de intencionalidad, libertad y conciencia aplicados a los Agentes No Humanos de Comunicación (ANHC) es una tarea necesaria para el cabal entendimiento de estas entidades, especialmente si queremos salir del antropocentrismo reinante en las ciencias sociales. Sin embargo, con esto surge la pregunta acerca de si es válido y necesario asignarle conceptos que utilizamos para el entendimiento de los humanos a otro tipo de agentes, o esto no será una nueva trampa del antropocentrismo del que quisiera que escape la teoría social: de tomarnos como la única medida del mundo, o al menos la preferencial.

En paralelo a esto me parece importante decir que, aunque el lenguaje como producto humano contenga tantísimas huellas que se refieren a nosotros, humanos, y a nuestra historia, no quisiera que este problema filosófico se vea solo como un malentendido sobre el lenguaje. Por el contrario creo, como Ludwig Wittgenstein y la filosofía analítica británica que lo sigue, que al resolver confusiones lingüísticas podemos deshacernos de muchos problemas filosóficos, por eso mento la elucidación como método. Justamente porque las reglas de uso de estas palabras se modifican al mismo tiempo que cambian las formas de vida compartidas que juegan con ellas: nosotros convivimos y nos comunicamos con potencias generativas no humanas.

Este trabajo no se propone resolver estas cuestiones sino un problema particular subyacente. Intentaremos analizar el estado actual respecto de la antropomorfización de los ANHC para señalar sus límites, beneficios y amenazas desde una consideración ontológica y epistémica. Esto es, intentaremos mostrar por qué están mal planteadas algunas preguntas que circulan en medios de comunicación respecto de su naturaleza y rol moral, y por qué otras requieren una atención urgente.

La filosofía tiene un papel crucial en este proceso, no solo al clarificar el lenguaje, sino al proporcionar marcos conceptuales que nos permitan entender mejor la naturaleza y las implicaciones de los ANHC. Este esfuerzo debe ser interdisciplinario, integrando conocimientos de las ciencias cognitivas, la ingeniería, las ciencias sociales y los estudios culturales, entre otros campos disciplinares. Los problemas actuales evidencian la necesidad de terminar con los compartimentos estancos de la academia. Este enfoque busca no solo enriquecer la teoría social, sino también promover una interacción más ética y eficaz con las tecnologías emergentes.

El rol de la metáfora: los humanos y las computadoras

Las metáforas no son meras decoraciones del lenguaje; son herramientas cognitivas que moldean nuestra percepción y comprensión del mundo. Al utilizar una metáfora estamos eligiendo un marco específico para interpretar y conceptualizar un tema. Por ejemplo, considerar la mente como una computadora influye en cómo pensamos sobre los

procesos mentales, nos sugiere ideas de procesamiento de información, almacenamiento y programación.

La antropomorfización de los ANHC plantea un desafío crítico tanto para nuestra comprensión de estos agentes como para su integración en nuestras estructuras sociales y tecnológicas. La tendencia a atribuir características humanas a entidades no humanas no solo refleja un sesgo cognitivo inherente, sino también una limitación epistemológica que puede distorsionar nuestra percepción y evaluación de los mismos. Para abordar la antropomorfización, es esencial desarrollar nuevas metáforas y conceptos que reflejen más fielmente la naturaleza de los ANHC. En lugar de ver a estos agentes como humanos en potencia o juzgarlos por la perfección de sus imitaciones, deberíamos conceptualizarlos en términos de sus roles y funciones dentro de redes de interacción más amplias. Esto puede implicar el uso de metáforas ecológicas o sistémicas que subrayen la interdependencia y la co-evolución.

La elección de una metáfora implica uno u otro camino para abordar un tema, cada una destaca ciertos aspectos y oculta otros. Por ejemplo, al hablar de la “evolución” de la tecnología, implicamos un proceso natural y gradual de mejora, lo cual puede influir en nuestras expectativas y en cómo interpretamos el desarrollo tecnológico. Conceptualizar las interacciones tecnológicas como parte de un “ecosistema” subraya la interdependencia y la co-evolución de diferentes tecnologías y agentes, tanto humanos como no humanos. Esto nos puede ayudar a entender las dinámicas complejas y a desarrollar políticas y prácticas más integradas y sostenibles.

Para escapar del antropocentrismo es necesario pensar a estos agentes desde una ontología relacional que no reconozca únicamente a los ANHC en términos de lo que carecen en comparación con los humanos (como conciencia, emociones o capacidad sintiente), sino en términos de sus propias capacidades y formas de agencia. Esto implica una reorientación de nuestra mirada hacia la valorización de las interacciones y las relaciones en lugar de las características intrínsecas.

Las metáforas antropomórficas pueden facilitar la interacción y la integración social, pero también puede llevar a malentendidos y expectativas incorrectas sobre las capacidades y responsabilidades de los ANHC. Por ejemplo, podemos pensar en los asistentes

virtuales, como Siri o Alexa, que son diseñados con nombres, voces y personalidades que sugieren una humanidad simulada. Esta antropomorfización puede conducir a expectativas y comportamientos que no se alinean con las capacidades reales de estos sistemas, pueden crear confusión y potencialmente desconfianza cuando los límites de la tecnología se hacen evidentes.

En el ámbito educativo y laboral, los LLMs se presentan principalmente como tutores o asistentes. La percepción de que estos modelos tienen intenciones o comprensiones humanas influye en cómo los usuarios interactúan con ellos, lo que redundará en variaciones de la efectividad de la enseñanza o la productividad en el trabajo. Es crucial remarcar que estos modelos no actúan en un sentido humano para evitar confusiones. No piensan, entienden o comprenden, en el sentido en que nosotros lo hacemos. Sin embargo, esto no significa que no actúen o procesen conocimiento. Y esos conceptos también requieren una elucidación fenomenológica, que tampoco haremos aquí.

Beneficios, amenazas y oportunidades de la antropomorfización

La antropomorfización puede acelerar la integración de los ANHC en la sociedad. Los humanos tienden a relacionarse mejor con entidades que se les parecen y actúan de manera similar. Esta familiaridad puede hacer que la adopción de nuevas tecnologías sea más rápida y menos resistente, permitiendo una transición más suave en diversos contextos, desde el hogar hasta el lugar de trabajo. En terapias, especialmente aquellas dirigidas a niños con autismo o personas mayores, los robots antropomorfizados han demostrado ser útiles. La apariencia y el comportamiento humano de estos robots pueden facilitar una conexión emocional y mejorar la efectividad de las terapias. La antropomorfización aquí actúa como un puente para la interacción, haciendo que las personas se sientan más cómodas y dispuestas a participar.

Este fenómeno también puede inspirar nuevas ideas y aplicaciones de ANHC en diversos campos. Puede usarse para crear experiencias educativas más atractivas y efectivas o permitir una interacción más personalizada y adaptada.

Pero antropomorfizar ANHC puede llevar a confusiones sobre su verdadera naturaleza y capacidades. Si un robot o sistema es percibido como consciente, las expectativas y demandas sociales pueden ser inapropiadas. Esto no solo afecta la interacción cotidiana sino también la formulación de políticas y regulaciones, que podrían basarse en supuestos incorrectos sobre la autonomía y la agencia de estos sistemas. Este tema amerita una discusión más extensa que la que podemos dar aquí.

La antropomorfización es una barrera epistémica ya que podemos proyectar nuestros propios comportamientos, emociones y motivaciones en ellos, lo que puede distorsionar nuestra comprensión de cómo realmente funcionan y operan. Asumir que los ANHC entienden y reaccionan como humanos puede crear una falsa sensación de seguridad, o incluso derivar en situaciones peligrosas si los usuarios confían demasiado en su palabra.

Los derechos animales a menudo se basan en la experiencia humana y animal que está intrínsecamente ligada a la finitud y la necesidad de preservación, a ello se asocia la capacidad de sentir dolor y placer, lo que no aplica a los ANHC. Los ANHC son sistemas funcionales diseñados para realizar tareas específicas, y su consideración ética debería basarse en su funcionalidad y el impacto de su uso en la sociedad, más que en una imaginaria capacidad sintiente.

Otra amenaza es el desplazamiento o abandono de las interacciones humanas. Las personas podrían incurrir en usos patológicos de los ANHC con consecuencias sociales y psicológicas, como el aislamiento social. Los robots diseñados para ser compañía tienden a provocar vínculos emocionales, lo que plantea preguntas sobre el bienestar emocional de los usuarios, especialmente en niños y ancianos.

Por último, debemos considerar como una amenaza al conflicto generado en el ámbito legal a partir de atribuir características humanas a los ANHC, lo que dificulta la interpretación de sus acciones y la asignación de responsabilidades. Ya exploraremos el avanzado debate en el área puntual de los robots.

Sin embargo, y aunque advierto que no deberíamos avanzar ni un ápice en este sentido, creo también que, dado el estado de situación, si nos quedamos a medio camino de la antropomorfización (al menos con la atribución de funciones cognitivas y el recono-

cimiento de su agencia) se abre una oportunidad importante de salir de la tradicional mirada instrumental hacia los objetos técnicos. Así podemos empezar a ajustar nuestros marcos normativos a los ANHC que ya están entre nosotros, y preparar el terreno para los que vendrán. Si empezamos a verlos como colaboradores más que como herramientas, se puede promover una cultura de la corresponsabilidad entre humanos y potencias cognitivas.

Un caso límite

El interés en analizar casos paradigmáticos es que representan una situación que se encuentra en los bordes de nuestras definiciones y entendimientos actuales. Estos casos nos permiten poner a prueba los límites de nuestros conceptos, teorías y principios.

Shawn Bayern presenta una situación hipotética, plausible actualmente dentro del sistema legal norteamericano en donde un sistema avanzado y autónomo podría tomar decisiones estratégicas y operativas por sí mismo con efectos legales dentro de la sociedad.¹ En teoría solo se requieren tres pasos:

1. La creación de una sociedad de responsabilidad limitada (LLC, *Limited Liability Company*).
2. La declaración del fundador respecto de que la LLC adopta un acuerdo operativo que indica que tomará acciones según lo determinado por un sistema autónomo, especificando términos o condiciones.
3. La transferencia, por parte del fundador, de sus derechos propietarios a la misma LLC, junto con la propiedad intelectual correspondiente, y su posterior renuncia.

Pero de hecho no es necesario ese marco legal para pensar en las responsabilidades de un ANHC: ya muchos de ellos toman decisiones financieras, económicas y operativas relevantes en empresas. Los QTA (*Quantitative Trading Algorithms*) son utilizados por

¹ Bayern, Sh. (2019). Are Autonomous Entities Possible? 114 Nw. U. L. Rev. Online 23. Recuperado de: https://scholarlycommons.law.northwestern.edu/cgi/viewcontent.cgi?article=1270&context=nulr_online&preview_mode=1&z=1561384769

fondos de cobertura y bancos de inversión para tomar decisiones de compra y venta de activos financieros basándose en modelos matemáticos y análisis de datos en tiempo real. IBM Watson, por ejemplo, opera en diversas industrias y ejecuta acciones profesionales basadas en análisis de datos avanzados, desde diagnósticos médicos hasta estrategias de marketing, que reemplazan o complementan el trabajo humano.

Lo que nos interesa es que el concepto de una empresa de responsabilidad limitada dirigida, o al menos influenciada significativamente por un ANHC, plantea cuestiones críticas sobre la responsabilidad, la agencia e implicaciones éticas reservadas históricamente para los seres humanos. Es cierto que para que un ANHC maneje una SRL, las leyes tendrían que adaptarse para reconocer la posibilidad de entidades no humanas en roles directivos. Hay quienes dicen que el ejemplo de Bayern podría ser visto como un caso de abuso del sistema, y vetado. Adaptar las leyes incluiría definir la responsabilidad legal y los límites de la autoridad del ANHC. Pero en un orden más general implica hablar de derechos y deberes para estas entidades.

El fango de los derechos robóticos

Ya desde la etimología de la palabra robot se plantea el problema del rol moral de los ANHC. La palabra tiene sus orígenes en la obra de teatro “R.U.R.” (Rossum’s Universal Robots) escrita por el dramaturgo checo Karel Čapek en 1920 y proviene del término checo “robota”, que significa “trabajo forzado” o “servidumbre”. En “R.U.R.”, los robots son seres artificiales con una apariencia y comportamiento humanoide, creados para realizar trabajos repetitivos y serviles que los humanos no quieren hacer, que terminan con un reclamo por su derecho a la libertad y reconocimiento moral.

Desde su creación, entonces, la noción de robot implica una forma de antropomorfización de los ANHC. Y no solo por su apariencia humanoide o las funciones para las cuales se diseña, sino incluso por la concepción misma de su relevancia moral. De hecho, desde hace décadas la discusión sobre sus derechos y deberes se asemeja a una bizantina especulación comparable a los debates sobre el sexo de los ángeles. Hay quienes conside-

ran que los robots deberían tener derechos cuando adquirieran cierto grado de conciencia o cuando pueda demostrarse que son seres sintientes, como hemos discutido.

Por otro lado, es muy atendible la posición de autores como Kate Darling, que defienden la idea respecto a la cual extender la protección legal a estos agentes no es tanto un interés en favor de los propios ANHC como del nuestro, porque cometer prácticas insidiosas que no aceptamos entre los humanos contra cualquier otra entidad, especialmente una con la que compartimos atributos, nos envilece.²

David Gunkel se propuso rediscutir la ontología moral y legal que separa a las cosas, las personas y los robots con la misma perspectiva que defiende: la que intenta rediseñar el lugar de la humanidad dentro de un mundo complejo y con múltiples agencias. Allí propone que pensemos cómo decidimos responder y tomar responsabilidad por nuestra realidad social con otros y otras clases de otredades. Gunkel explora los límites ontológicos de los robots para ceñirse a estas categorías, y las problemáticas epistémicas que esto ocasiona.³ Pero el interés principal del libro reside en que invita a pensar el marco legal a construir para convivir con estas entidades desde la consideración de cómo los robots afectan nuestras vidas. Esto nos permite guiar el desarrollo y el uso de estos ANHC en pos de asegurarnos que sean beneficiosos para todos.

Y como me gusta trabajar desde las cosas, y no al revés, quisiera terminar el artículo con el análisis de un escenario en donde se contemplan derechos y deberes específicos para robots.

Imaginemos que estamos ante un robot cuidador, como el actualmente existente Pepper de SoftBank Robotics. Pepper es un robot semi-humanoide con la capacidad de leer emociones, cuyas réplicas están en miles de hogares japoneses. También se usa en otros países como recepcionista y orientador.

En su uso doméstico, podemos definir claramente los ámbitos en los que el robot tiene autoridad para tomar decisiones en nombre del usuario, como la administración de

2 Darling, K. (2016). Extending Legal Protection to Social Robots: The Effects of Anthropomorphism, Empathy, and Violent Behavior toward Robotic Objects. En Calo, R.; Froomkin, M. y Kerr, I. (2016). *Robot Law*, pp. 213-231. Edward Elgar: Massachusetts.

3 Gunkel, D. (2023). *Person, thing, robot*. MIT Press: Massachusetts.

medicamentos, de la agenda, la compra de suministros esenciales, o incluso la gestión de cuentas bancarias. Estas tareas prácticas bien pueden considerarse derechos delegados a Pepper, lo mismo que podemos también pensar un marco legal para garantizar mantenimiento y actualización al robot, o protección contra abusos para evitar que sea utilizado indebidamente por terceros.

Dentro de las obligaciones podemos pensar que deben informar a los usuarios y operadores sobre sus capacidades y limitaciones, así como cualquier error o fallo en el sistema. También sus deberes de cuidado o de protección de datos personales. Lo más espinoso quizá sea pensar un régimen de corresponsabilidad. En ese sentido podemos implementar un sistema de registro y monitoreo que documente todas las acciones y decisiones tomadas por Pepper. En caso de ser el responsable moral de una acción cuestionable, podríamos auditar su código, modificarlo y hacer una rehabilitación funcional. También parece sensato aprovechar el aumento de la productividad que genera la automatización para crear un fondo de compensación que permita reparar los daños ocasionados.

Los robots no son ni personas ni cosas en el sentido tradicional, por lo que podría ser útil crear una categoría legal especial para ellos. En todo caso, como propone Gunkel, no tiene sentido juzgarlos desde esas categorías.