



Tres respuestas al antropomorfismo en robótica social: hacia un enfoque crítico, relacional y hermenéutico

Mark Coeckelbergh¹ 

Aceptado: 23 de febrero de
 2021 © El autor(es) 2021

Abstracto

Tanto los diseñadores como los usuarios de robots sociales tienden a antropomorfizar a los robots. Centrándose en la cuestión de cómo conceptualizar la relación entre robots y humanos, este primer artículo esboza dos puntos de vista filosóficos opuestos con respecto a esta relación, que están conectados con varias respuestas normativas al antropomorfismo y la antropomorfización. Luego aboga por una tercera visión: navegando entre lo que llama "instrumentalismo ingenuo" y "posthumanismo acrítico", desarrolla un enfoque hermenéutico, relacional y crítico. Paradójicamente, al analizar la dimensión humana de la robótica en su uso y desarrollo, esta visión permite una discusión crítica sobre la antropomorfización de los robots. Al mismo tiempo, y de nuevo algo paradójicamente, evita una posición instrumentalista ingenua al tomar en serio el papel de los robots como instrumento en una contratecnología más amplia. Como tal, la tercera visión cuestiona el dualismo asumido en el debate. Luego, el artículo explora lo que esto significa para el campo de la robótica social y la educación de científicos e ingenieros informáticos. Propone una reforma basada en una comprensión relacional del propio campo y ofrece sugerencias para el papel de los usuarios-ciudadanos.

Palabras clave Robótica social Relaciones humano-robot Antropomorfismo Antropomorfización Instrumentalismo Posthumanismo · Hermenéutica · Enfoque relacional · Poder

1 Introducción: Antropomorfismo y antropomorfización en lo social Robótica

Los especialistas en robótica social desarrollan robots destinados a funcionar de forma más natural en situaciones sociales que las máquinas del pasado. Para ello, se inspiran en los seres humanos y otros primates superiores, que a su vez tienden a entenderse mediante la metáfora de una especie de máquina social, aunque se desarrolla y ha evolucionado. Con la ayuda de literatura de la psicología, las ciencias cognitivas, la neurociencia y otros campos, los investigadores en robótica social e interacción entre humanos y robots (HRI) identifican características que se consideran necesarias para la interacción social y luego las implementan en el robot. Por ejemplo, Dautenhahn se refiere al trabajo en HRI (en particular [1]) y a la hipótesis del cerebro social, que dice que la inteligencia de los primates evolucionó debido a la necesidad de lidiar con situaciones sociales cada vez más complejas.

dinámica, para definir una serie de características de interacción social que se implementan en robots sociales. Afirma que «los robots socialmente interactivos presentan las siguientes características: expresan y/o perciben emociones; comunicarse con un diálogo de alto nivel; aprender modelos de otros agentes o reconocerlos; establecer y/o mantener relaciones sociales; utilizar señales naturales (mirada, gestos, etc.); exhibir personalidad y carácter distintivos; y podrá aprender y/o desarrollar competencias sociales.» [2:684; 1:145] La implementación de tales características sociales a menudo lleva a que el robot exhiba una apariencia y un comportamiento antropomórficos. Sin embargo, el propósito no es construir robots parecidos a los humanos como tales (o robots parecidos a los primates). La antropomorfización en este contexto tiene un objetivo más pragmático. En robótica social, es una estrategia para integrar robots en entornos sociales humanos. Es un medio para dar a los robots la "capacidad de participar en una interacción social significativa con las personas" [3:178]. Entonces, tanto los diseñadores como los usuarios tienden a antropomorfizar a dichos robots cuando interactúan con ellos, atribuyéndoles características antropomórficas como personalidad, vitalidad, etc. En las próximas páginas me referiré a "antropomorfismo", "antropomorfismo" y "antropomorfización", respectivamente, como el fenómeno y la actividad de atribuir

* Mark Coeckelbergh
 mark.coeckelbergh@univie.ac.at

¹ Universidad de Viena, Viena, Austria

(por los usuarios cuando interactúan con los robots) y la creación de dichas características por parte de los diseñadores y desarrolladores de robots.

A veces no se pretende el antropomorfismo. Los robóticos sociales son conscientes de esto y por eso intentan controlar este fenómeno. Con su diseño, pueden intentar fomentar o desalentar la antropomorfización por parte del usuario. Por ejemplo, algunos expertos en robótica intentan deliberadamente crear robots que se parezcan lo más posible a los humanos: Ishiguro [4] ha sostenido que los humanoides son interfaces ideales. Otros construyen menos parecidos a los humanos.

robots, por ejemplo robots que parecen animales, o intentan evitar por completo el antropomorfismo. En teoría, el antropomorfismo no parece necesario para la robótica social; Puede ser suficiente que el robot presente las características mencionadas anteriormente, sin parecerse a un humano. Pero en la práctica, muchos robóticos sociales acogen con agrado algún grado o forma de antropomorfismo, ya que, según se argumenta, facilita las interacciones entre humanos y robots. Hoy en día, el antropomorfismo es uno de los factores que se miden para dar forma a la interacción entre humanos y robots (HRI) [5]. Se reconoce que el antropomorfismo presenta desafíos, pero en última instancia el objetivo es "facilitar la integración de máquinas similares a los humanos en el mundo real" [6] y el antropomorfismo ayuda a alcanzar ese objetivo.

Los robots son muy adecuados para este fin [7]; Debido a su apariencia física y presencia, tienen una alta "antropomorfización" [6]. Mientras que ya en los años 90 Reeves y Nass descubrieron que la gente trata a los ordenadores como si fueran personas reales [8], con los robots esto funciona aún mejor, especialmente los robots humanoides [9]. Pero el antropomorfismo no sólo se ve afectado por la apariencia: los robots pueden ofrecer todo tipo de señales durante la interacción [10] y existen otros factores como la autonomía del robot, la previsibilidad, etc. Sin embargo, resulta interesante para los fines de este artículo que el antropomorfismo no sea sólo el resultado de cómo se ve el robot o de lo que hace, sino que también depende de las características del observador, como el origen social y el género [6]. Por ejemplo, las personas aplican categorías sociales, como la pertenencia a un grupo, a los robots.

[11] Esto implica que un ángulo social puede ayudar a comprender mejor, y quizás también evaluar mejor, el antropomorfismo en la interacción entre humanos y robots. También sugiere que, además de las características objetivas del robot, la subjetividad del robot El usuario importa.

Los psicólogos ayudan a los robóticos con este proyecto de utilizar la antropomorfización con el fin de desarrollar robots sociales mediante estudios empíricos sobre la percepción de los robots (dependiendo de sus características), por ejemplo, diseñando cuestionarios y analizando los resultados para estudiar la aceptabilidad [12, 13]. Los robóticos también importan la teoría psicológica sobre la socialidad humana y realizan investigaciones en un contexto interdisciplinario: utilizan paradigmas clave de la psicología y utilizan y replican hallazgos de la interacción entre humanos; a menudo se encuentra con enfoques experimentales [14]. Y además de dar forma a las características de diseño del robot, el encuadre también es un método para

Haga que los humanos antropomorfen la tecnología robótica: a los robots se les puede dar un nombre o una historia de fondo, lo que afecta las respuestas de las personas hacia los robots [15]. Una vez más, el antropomorfismo con respecto a la tecnología no es nuevo; Pero la robótica social, que desarrolla robots antropomórficos a propósito, parece aumentar los efectos.

La antropomorfización de los robots plantea todo tipo de preocupaciones, y algunas de ellas han sido identificadas, reconocidas y estudiadas dentro de las comunidades de robótica social y HRI.

ellos mismos. Por ejemplo, Turkle [16] y Scheutz [17] han expresado su preocupación por las relaciones emocionales que las personas desarrollan con este tipo de robots. También está el famoso problema del "valle inquietante". Desde que el robotista japonés Masahiro Mori [18] introdujo el concepto (freudiano) de lo "sinistro" en la robótica, se ha debatido cómo los robots pueden evocar una sensación siniestra y cómo evitarla, por ejemplo, al construir andróides [19]. Para Freud, lo siniestro es la experiencia psicológica de lo extraño en lo familiar. En su ensayo de 1919 sobre el concepto [20], interpreta una historia de ETA Hoffmann que presenta una muñeca realista y otros elementos "extraños": la gente se siente extraña cuando un objeto que se supone que no tiene vida parece real y se siente amenazada por su propia vida. Impulsos inconscientes y ocultos. Lo extraño en el contexto de la robótica significa que cuando un robot parece humano pero se mueve de maneras extrañas y aterradoras, por ejemplo, esto puede afectar a los usuarios. Se puede intentar dar explicaciones psicológicas a esto (freudianas u otras), como el miedo a la muerte. Y desde una visión antropológica, el fenómeno de la antropomorfización puede conceptualizarse como una forma de animismo tecnológico: el animismo, la atribución de un alma o espíritu a objetos inanimados y seres naturales, no se entiende entonces como algo que pertenezca a una etapa "primitiva". de nuestra evolución social o se limita a fenómenos "naturales", sino como algo que subyace a la construcción de agencia y personalidad [21] y que también ocurre en relación con objetos tecnológicos como los robots. Sin embargo, aquí me centraré en el antropomorfismo y en la antropomorfización específicamente.

Dados los problemas planteados por el antropomorfismo, algunos robóticos han argumentado a favor de no aspirar a crear robots antropomórficos como tales. Goetz y sus colegas, por ejemplo, argumentaron que los robots deberían diseñarse para cumplir con su tarea, en lugar de tomar la semejanza humana como el objetivo principal.

[22] Sin embargo, dado que la construcción de robots antropomórficos facilita la interacción entre humanos y robots y ayuda a lograr sus objetivos, por ejemplo en el cuidado de la salud y el cuidado de personas mayores, los desarrolladores de robots sociales a menudo lo persiguen.

Este artículo pregunta: ¿Cómo se puede responder a la cuestión del antropomorfismo desde el punto de vista de la filosofía de la tecnología? Existen posibilidades interesantes para abordar este tema, algunas de las cuales pueden conectarse con las discusiones mencionadas en robótica social. En este artículo, abordo el tema preguntándome cómo enmarcar filosóficamente

La relación entre robots y humanos. Primero establezco dos puntos de vista filosóficos opuestos sobre esta cuestión, que llamaré "instrumentalismo ingenuo" y "posthumanismo acrítico". Tal como están formuladas, pueden parecer visiones de hombre de paja, pero, como mostraré, hay investigadores que sostienen al menos versiones más suaves de ellas. Mostraré cómo cada uno de estos puntos de vista puede conectarse con diferentes normativas.

Posiciones sobre el antropomorfismo y la antropomorfización: posiciones que no se refieren a la descripción e interpretación de estos fenómenos, sino a la cuestión de si debemos antropomorfizar robots y crear robots antropomorfos.

Esbozar estos extremos opuestos me ayuda a desarrollar una tercera visión alternativa, impulsada por un enfoque hermenéutico, relacional y crítico.

Si bien soy consciente de que esta discusión también se vincula con la ética. Si analizamos la ética de la robótica y los debates sobre el estatus moral de los robots, no abordaré aquí directamente estas cuestiones normativas particulares y, en cambio, me centraré en el tema del antropomorfismo y la problemática relacionada con las relaciones entre el hombre y la tecnología, tan relevantes para la robótica social y la HRI. Esto no significa que no se puedan extraer conclusiones normativamente relevantes para la robótica social a partir de las posiciones esbozadas, lo cual haré, pero el enfoque de este artículo no está en el estatus moral. Al final del artículo, exploraré las implicaciones de este enfoque para la robótica social y la educación de científicos e ingenieros informáticos.

Si bien el artículo se refiere al trabajo en robótica social y HRI, dada la audiencia principal de esta revista, vale la pena señalar explícitamente que este artículo presenta una discusión filosófica de varios puntos de vista sobre las relaciones entre humanos y robots y posiciones normativas hacia el antropomorfismo y el antropomorfismo; No es una revisión de la literatura ni un estudio empírico. Sin embargo, espero que el artículo pueda estimular un diálogo fructífero entre los robóticos sociales y los filósofos. Además, la discusión de puntos de vista sobre la relación entre robots y humanos también es relevante para pensar en la relación entre la tecnología y los humanos en general y, por lo tanto, puede atraer a los filósofos de la tecnología.

2 Dos visiones opuestas y problemáticas: Instrumentalismo ingenuo y acrítico Posthumanismo

Hay al menos dos puntos de vista opuestos sobre la relación entre robots y humanos, cada uno de los cuales puede conectarse con diversas respuestas normativas al antropomorfismo y la antropomorfización:

2.1 Instrumentalismo ingenuo

Una es insistir en que los robots son máquinas y meras instrucciones para fines humanos. Según esta opinión, que yo

Llamamos "instrumentalismo ingenuo", antropomorfizar a los robots es una especie de sesgo psicológico. Nosotros (los científicos) sabemos que el robot es sólo una herramienta, pero sin embargo, cuando interactuamos con el robot, nuestra psicología (la psicología de los usuarios) nos lleva a percibir al robot como una especie de persona. Los usuarios piensan que el robot es parecido a un humano, cuando en realidad no lo es. (Nosotros los científicos sabemos que) los robots son sólo máquinas. La base filosófica de esta visión es doble. En primer lugar, supone una visión dualista del mundo en términos de realidad versus apariencia (una visión que a veces se atribuye a Platón): supone que en realidad el robot es un instrumento, mientras que en apariencia se parece más a una persona. El instrumentalismo ingenuo también supone una visión dualista de humanos versus no humanos. Los no humanos son categorías mutuamente excluyentes y existe una profunda división ontológica entre ellos. Para pensar en la tecnología, esto significa que la tecnología se encuentra del lado de los no humanos (por ejemplo, las cosas, los objetos), mientras que los humanos son completamente diferentes (personas, sujetos). Los robots y los humanos forman parte de categorías ontológicas completamente diferentes. En segundo lugar, supone una versión del realismo metafísico y epistemológico: los objetos existen independientemente de nuestros conceptos y percepciones, y podemos describirlos de manera objetiva. Los usuarios se dejan engañar por las apariencias: en realidad, el robot es sólo una máquina. Los científicos pueden describir esta realidad en términos objetivos y estudiar cómo nuestras mentes se dejan engañar por las características antropomórficas del robot, la forma de interacción y nuestro entorno social y cultural.

Esta visión puede conducir a al menos dos posiciones normativas con respecto al antropomorfismo y la tarea de la robótica social. Una es que los investigadores en robótica social pueden aprovechar este sesgo diseñando robots con apariencia humana para mejorar la interacción entre humanos y robots y lograr mejor sus objetivos. No sólo el robot es una herramienta; La morfología en sí misma es también un instrumento antropológico que puede usarse y explotarse. Ésta es la postura que esbocé al principio de este artículo. Implica una distribución desigual del conocimiento sobre el estado real de las cosas: el diseñador del robot conoce la realidad, mientras que el usuario es engañado, al menos temporalmente, haciéndole creer que el robot es una persona. Se podría decir que el diseñador del robot actúa como una especie de mago [23] que crea la ilusión de que el robot es humano. Como ha señalado el filósofo del diseño Flusser [24], los propios términos diseño y máquina están etimológicamente relacionados con la astucia y el engaño. Los robots están diseñados para producir una antropomorfización por parte de los usuarios, a quienes se les engaña haciéndoles creer que el robot es un ser humano o (teniendo en cuenta el estado actual de la robótica) están al menos dispuestos a suspender temporalmente la incredulidad y abrazar la ilusión. Según la visión instrumentalista, esta magia y engaño de la antropomorfización no es problemático si se logra el objetivo: si el antropomorfismo del robot y la antropomorfización por parte de los usuarios conducen a una mejor interacción entre humanos y robots, y si esto a su vez logra los objetivos que los humanos querían. lograr (por ejemplo, un objetivo de salud específico)

tarea de cuidado), entonces este engaño está permitido e incluso recomendado. Éste es nuevamente el instrumentalismo en acción.

Tenga en cuenta que, por el mismo razonamiento, si los objetivos no son logrado por el antropomorfismo, el instrumentista no lo utilizará. Si un robot, por el motivo que sea, no cumple su función, entonces el razonamiento instrumentalista lo desaconsejará.

Otra posición normativa es utilizar el instrumentalismo para argumentar que, independientemente del funcionamiento y efectividad del robot, es altamente problemático antropomorfizar las máquinas y desarrollar dichas máquinas en general y, en principio, así se cree y afirma, los robots son más herramientas. Por lo tanto, se argumenta, los especialistas en robótica social deberían dejar de diseñarlos o al menos asegurarse de que los usuarios sean conscientes de que son meras herramientas. Consideremos, por ejemplo, la posición de Bryson: los robots son herramientas que utilizamos para lograr nuestros propios objetivos [25]. Hablar de ellos de una manera que sugiera que son personas es engañoso.

Además, puede haber razones independientes por las que el antropomorfismo sea problemático. Por ejemplo, se podría argumentar que construir tales máquinas tampoco es deseable debido a los posibles problemas éticos mencionados anteriormente: las personas podrían apegarse emocionalmente a las máquinas y, como han argumentado Sparrow y Sparrow, dichos robots podrían desconectarlas de la realidad [26]. Por ejemplo, un robot social en el sector sanitario puede inducir a error a los usuarios haciéndoles pensar que está realmente vivo o que realmente puede ser un amigo. Si bien el estado actual de la técnica no ofrece esto para todos los usuarios y durante mucho tiempo, los avances en la robótica social parecen apuntar a esta ilusión, por ejemplo con robots como el robot cría de foca Paro, que se utiliza en hospitales y enfermería. hogares. Turkle [16] ha sostenido que los robots como Paro sólo proporcionan la ilusión de una relación.

Desde una perspectiva instrumentalista, se podrían apoyar tales críticas y agregar que los especialistas en robótica social deberían desarrollar robots que eviten esta ilusión y funcionen simplemente como una herramienta. No se deberían construir robots que "pretendan" ser más que una herramienta. Si se implementa, esto bien podría ser el fin del proyecto de robótica social en el sentido descrito anteriormente, ya que si el robot ya no tiene ninguna característica que dé lugar a la antropomorfización (o si al usuario se le recuerda constantemente que es un mera máquina), está claro cómo puede crear una interacción social "natural" con los humanos, o incluso una interacción "social" en absoluto. Desde este punto de vista, entonces, deberíamos dejar de desarrollar y utilizar robots sociales que inviten a la antropomorfización. Si eso significa que la interacción entre humanos y robots se vuelve menos "natural" o menos "social", que así sea.

Pero cualquiera que sea la posición normativa que se adopte respecto de la antropomorfización (ya sea que debamos antropomorfizar o no), ambas posiciones asumen una visión instrumentalista de la relación entre robots y humanos. Ambos piensan que los robots son sólo herramientas; sólo difieren cuando se trata de evaluar el antropomorfismo y la antropomorfización. Además, calificar este instrumentalismo de "ingenuo" de ninguna manera

se refiere a la ignorancia sobre el antropomorfismo, y mucho menos sobre la robótica. Ambas posiciones normativas son muy conscientes de lo que es y hace el antropomorfismo (su naturaleza y sus efectos). El término "ingenuo" en el "instrumentalismo ingenuo" sólo se refiere a la ignorancia sobre la dimensión no instrumental de la tecnología. Lo que esto significa quedará claro en las próximas secciones, especialmente cuando analice la tercera posición.

2.2 Posthumanismo acrítico

Otra visión, en el otro extremo del espectro y tentativamente llamada "posthumanismo acrítico", es abrazar totalmente a los robots sociales como cuasipersonas y "otros". Déjame desempacar esto. El posthumanismo puede tener varios significados, pero aquí me refiero a una teoría que critica las visiones humanistas tradicionales del mundo que colocan a los humanos en el centro del mundo (antropocentrismo) y, en cambio, expande el círculo de preocupación ontológica y moral hacia los no humanos. Podrían tratarse de animales no humanos, pero también, por ejemplo, de (algunos) robots.

En contraste con la posición instrumentalista, aquí los robots sociales son bienvenidos como parte de una ecología o red posthumanista de humanos y no humanos. En lugar de una visión dualista del mundo que opone humanos a cosas, según esta visión los humanos y los no humanos son parte de la misma red o ecología y están entrelazados de diversas maneras.

Los posthumanistas nos animan a cruzar las fronteras y buscar dualismos y binarios. Esto incluye la tecnología humana.

Nología binaria: Según esta visión, los humanos son seres tecnológicos y las tecnologías son humanas: los humanos siempre han usado tecnologías, los humanos crean tecnologías y las tecnologías son parte de nuestro mundo en formas que no son sólo instrumentales. Por ejemplo, podemos hablar de la tecnología como si fuera una cosa, podemos esperar más de la tecnología de lo que puede ofrecer (algo que sucede a menudo en la robótica y también, por ejemplo, en la IA), podemos adorar la tecnología, etc. Para pensar en robots, esto significa que, según esta visión, los robots no deben verse como meros instrumentos.

y cosas si eso significa que la única relación que tenemos con ellas es instrumental; En cambio, podemos reconocerlos como parte de nuestro mundo humano y cultural e incluso "encontrarlos" como otros: no como otros humanos sino también como entidades sociales. vínculos, entidades que comparten nuestro mundo. Además, en lugar de una visión realista del mundo y de cómo conocemos el mundo, los posthumanistas tienden a ser no realistas: creen que no podemos tener una visión objetiva de la "realidad" (como si pudiéramos conocer una realidad independiente de nuestra visión). formas de saber) y que las creencias científicas (por ejemplo, sobre los robots) son una construcción social. Entonces, lo que "son" los robots no está descrito exhaustivamente por la ciencia y no puede saberse independientemente de la subjetividad y la cultura humanas. Consideremos, por ejemplo, los propios términos "robot" y "máquina": el significado de estos términos está subdeterminado por las definiciones científicas y de ingeniería, ya que tienen su propia historia cultural.

y dependen del uso contemporáneo. Lo que significa "robot" puede depender, por ejemplo, del contexto social y del uso del lenguaje relacionado. Por ejemplo, generalmente no llamamos robots a los automóviles autónomos (en la mayoría de los contextos los llamamos automóviles), mientras que según la mayoría de las definiciones técnicas ciertamente lo son.

Un enfoque posthumanista y constructivista social de este tipo podría estar inspirado en Haraway, quien en su 'A Cyborg Manifesto' [27] y en trabajos posteriores ha defendido el cruce de fronteras mediante la inclusión de máquinas y animales en lo político, o en Latour, cuyo enfoque no modélico podría inspirarse en Haraway. El enfoque moderno incluye a los no humanos en lo social [28].

El punto de Haraway no se refiere sólo a la creación de cyborgs literales (fusiones de humanos y máquinas); Su objetivo es cambiar nuestra forma de pensar sobre (la importancia de) los límites entre humanos y máquinas, metafísicamente pero también políticamente. Se podría decir, por ejemplo, que los robots deberían incluirse en la sociedad. De manera similar, el pensamiento de Latour vería a los robots como no humanos que deben ser incluidos en el

colectivo social y político; excluirlos sería mantener una frontera estricta entre naturaleza y cultura, y entre humanos y no humanos. Un enfoque no moderno va más allá de esos binarios. O se podrían interpretar las investigaciones de Gunkel que utilizan a Levinas para explorar si los robots pueden ser considerados como "Otros" [29, 30] como no idénticas (su punto de vista es más crítico), pero que potencialmente conducen a tal punto de vista. En la fenomenología tradicional, el otro (a veces escrito: "Otro") se opone al yo y suele referirse a otros seres humanos. El filósofo Emmanuel Levinas añadió a esto un uso específico del término: el 'Otro' es la contraparte radical del yo, que es absolutamente otro y estamos llamados a responder éticamente a ese Otro. Sin embargo, Levinas todavía estaba preocupado por los seres humanos.

seres. Desde un punto de vista posthumanista, cabe entonces preguntarse: ¿podría este otro u Otro ser también un animal no humano? o un robot?

Según una visión posthumanista, el proyecto de la robótica social no es necesariamente problemático e incluso puede ser adoptado, ya que estamos invitados a incluir a los no humanos en la esfera de lo social, y estas antropomorfizaciones pueden ayudar con eso. Deberíamos ir más allá del antropocentrismo de la posición instrumentalista: lo humano no debería ser el centro. La robótica social puede ayudar con el proyecto de descentrar al ser humano. Esta posición también respaldaría el antropomorfismo en los robots sociales: está bien y quizás incluso deseable construir robots que se parezcan a los humanos (antropomorfismo) y que sean construidos y percibidos como tales (antropomorfización), no porque sean instrumentos para nuestros propósitos, sino porque tales robots ofrecernos un "otro" con el que podamos interactuar y que podamos incluir en lo social. El poshumanismo rechaza el antropocentrismo: rechaza la idea de que los humanos deberían ser el centro del mundo y el motivo de preocupación moral. Humano

Los propósitos, por lo tanto, no son los únicos que cuentan, y es posible que también queramos crear robots sociales que no sean antropomorfos en absoluto.

Esta última posición no necesita rechazar el antropomorfismo como tal; simplemente señala que hay muchas otras posibilidades. Sin embargo, basándose en la visión de Gunkel [29, 30] y adoptando un enfoque posmoderno, también se podría adoptar otra postura normativa menos favorable hacia el antropomorfismo: se podría argumentar que, visto desde un enfoque poshumanista orientado a la diferencia o la alteridad, la antropomorfización en El sentido de percibir a los robots como humanos es problemático en la medida en que no respeta la diferencia o la alteridad del robot, entendido como otro artificial. En lugar de proyectar nuestra humanidad en los robots (cómo somos como humanos y cómo nos vemos a nosotros mismos como humanos) y así (a)usarlos como meras pantallas de proyección de nuestra propia imagen, deberíamos tratarlos como otros por derecho propio y respetar su diferencia. Si no somos capaces de hacer esto, no deberíamos construir robots antropomórficos (pero podemos construir otros).

Por tanto, esta postura normativa también relaciona el antropomorfismo con el antropocentrismo, pero rechaza ambos.

2.3 Problemas

Tanto el instrumentalismo ingenuo como el posthumanismo acrítico son problemáticos. El primero es ingenuo si no comprende plenamente que los robots no son meras herramientas, sino que también tienen efectos no deseados. secuencias y están vinculados con los humanos a través de la experiencia, el lenguaje, las relaciones sociales, las narrativas, etc. Su visión dualista y realista del mundo, que configura la relación entre humanos y tecnología como externa, nos impide ver que los robots no son meras cosas (que son resultado de la construcción técnica y de la escritura) sino que también son al mismo tiempo humanamente. y socialmente construido. El segundo no es crítico si olvida que los robots son máquinas hechas por humanos, que bien podrían enfrentarnos como "cuasi-otros" [31] pero nunca son totalmente otros o completamente no humanos, ya que son creados, interpretados y dotados de significado por humanos. .

Si bien pueden enfrentarnos como un otro externo, nunca son completamente externos a nosotros. El enfoque del posthumanismo en la alteridad y la construcción sociocultural de los robots conduce hasta ignorar su origen en prácticas humanas y materiales. Los robots también son máquinas. Y estas máquinas están hechas por humanos. En otras palabras, ambos puntos de vista no reconocen suficientemente cómo los robots sociales están intrínsecamente y (como lo expondré más adelante) relacionados internamente con los humanos de diversas maneras.

Expreso estas críticas en términos de falta de comprensión total y de olvido, ya que ambos bien podrían tener algún conocimiento de las consecuencias no deseadas respectivamente y del origen de la tecnología, o podrían en principio adquirir este conocimiento. Por lo tanto, para ser más precisos: la cuestión no es que los instrumentistas ingenuos no sepan en principio que su tecnología puede tener consecuencias no deseadas o que

Los posthumanistas acrílicos no saben en teoría que la tecnología se crea. Saben que existen estas relaciones o pueden adquirir ese conocimiento. Más bien, el problema es que normalmente no logran comprender las implicaciones más profundas de este conocimiento para (pensar sobre) la relación entre humanos y robots y, por tanto, para el problema del antropomorfismo.

En particular, ambos puntos de vista no logran comprender cuán fuerte es realmente el entrelazamiento de humanos y robots, es decir, que se trata de una relación interna (ver más abajo).

Sin embargo, existe una posición alternativa, la "tercera", que soluciona este problema pero aún nos permite ser críticos con el antropomorfismo y la antropomorfización (y, de hecho, con el proyecto de la robótica social). ¿Cómo podemos llegar a tal punto de vista? Una forma de hacerlo sería tratar de encontrar una posición "intermedia" entre los dos extremos: los robots sociales no son ni enteramente instrumentales ni completamente distintos. Pero no está claro qué significa esto. ¿Cuál es el aspecto "no instrumental" del robot? ¿Y qué significa que un robot pueda ser "otro" hasta cierto punto? Por lo tanto, para encontrar una alternativa real propongo que cuestionemos y cambiemos los supuestos que respaldan ambas posiciones. En particular, propongo abandonar el supuesto dualista de que la relación entre humanos y robots es externa y mostrar cómo los robots humanos sociales y los robots sociales la robótica lo es. Esto crea una posición nueva, completamente diferente, que no es una posición intermedia sino que escapa del campo definido por las opiniones opuestas aquí esbozadas.

3 Hacia una tercera visión: crítica, relacional y hermenéutica

Permítanme esbozar esta posición, que es aplicable a los robots en general, pero que se formula en respuesta a las posiciones opuestas que surgieron en respuesta al antropomorfismo y la antropomorfización en la robótica social. Para constituir una alternativa real tal como se define anteriormente, debería constar de al menos los siguientes elementos:

3.1 Los robots son humanos, pero no humanos: se crean y se utilizan, y dan forma a nuestros objetivos

En primer lugar, los robots sociales no surgen simplemente de la naturaleza o de la nada (como si fueran parte de la naturaleza "no humana" o como si aparecieran como extraterrestres del espacio exterior), sino que son diseñados y fabricados por seres humanos. Esto implica que la "herramienta" o el "otro" presentado en las dos posiciones que consideramos no sólo son no humanos sino también humanos al mismo tiempo. Los robots no sólo están relacionados con nuestros objetivos e intenciones, como si fueran cosas externas que simplemente nos confrontan desde afuera, por así decirlo; ellos también son creados por nosotros en prácticas tecnocientíficas concretas y son utilizados por nosotros. En este sentido, son humanos. Y esto también significa que nunca podrán ser totalmente "otros". También hay

igualdad. Incluso los robots que no son antropomórficos tienen un aspecto humano, ya que nosotros los creamos y los utilizamos. No necesitan ser bienvenidos y llevados a "nuestra" esfera; Ya son parte de él desde y a medida que los usamos. Y desde que los creamos, nunca son meras herramientas, sino que están impregnados de nuestros fines, nuestros significados y nuestros valores. Son instrumentos, pero son nuestros instrumentos. Y con sus efectos no deseados, también dan forma a nuestros objetivos. Por tanto, el instrumento y el objetivo son interdependientes.

Por ejemplo, a un robot se le pueden dar características antropomórficas para que funcione en un texto de atención médica en particular. El objetivo tiene que ver con la atención sanitaria, por ejemplo proporcionando una forma de comunicarse con los pacientes. Pero al hacer que los robots se comuniquen con los pacientes, el contenido de este objetivo cambia: lo que significa la comunicación con los pacientes cambia de un significado (donde los humanos son quienes se comunican como cuidadores) a otro (la comunicación entre cuidadores y pacientes es algo eso también se puede hacer con una máquina). Además, si nosotros (como pacientes, como cuidadores) olvidamos que los robots son creados por humanos y que les damos significado, esto puede llevar a una adopción acrítica del robot y al cierre de posibilidades de cambio. El robot y la interacción no deben tomarse como algo dado sino como algo modificable. Permítanme decir más sobre este otorgamiento de significado en la interacción.

3.2 La construcción lingüística y social de los robots

En segundo lugar, incluso si tomamos el robot como algo dado, es decir, tal como aparece ante nosotros y como interactuamos con él como usuarios (y los diseñadores también son un tipo de usuarios, aunque una categoría especial de usuarios), deberíamos cuestionar nuevamente la Suposición de que existe una relación externa entre humanos y robots.

Esto puede hacerse considerando lo que propongo llamar antropocentrismo epistemológico. Significa que lo que decimos que "es" un robot, lo que sabemos del robot y cómo nos relacionamos con el robot siempre dependerá de la subjetividad humana, la creación de significado, las narrativas, el lenguaje, las metáforas, etc. Ya sea que podamos alejarnos o no del antropocentrismo en un sentido moral o metafísico, adoptar un enfoque filosófico crítico (después de Kant, la fenomenología y gran parte de la filosofía del siglo XX) significa reconocer que siempre tenemos que "pasar por lo humano" cuando hablar, pensar e interactuar con robots sociales. Lo que "es" el robot no es independiente de lo que decimos sobre él y de cómo respondemos al robot en interacción. Los humanos no sólo crean robots materialmente sino que también (durante el desarrollo, el uso y la interacción) los "construyen" mediante el lenguaje y en las relaciones sociales, lo que debe presuponerse cuando pensamos en estos robots e interactuamos con ellos. Esto puede revelarse mediante esfuerzos críticos filosóficos y sociocientíficos que incluyan una perspectiva temporal e histórica y, de hecho, el papel del lenguaje y la metáfora [32-34]. Consideremos el papel del lenguaje: si le doy un nombre personal a un robot social, entonces

Ya he construido el robot como compañero, otro, como si tuviera un género concreto, etc. [35]. El robot no necesariamente deja de funcionar como instrumento y puede ser visto como instrumento en otros momentos o desde una perspectiva diferente. Pero el nombre dado, junto con otros factores como la apariencia del robot, configura el significado del robot.

Además, este significado está relacionado con el entorno sociocultural en el que está inmerso el robot. Al darle un nombre particular, los usuarios también pueden acceder a toda una cultura de nombres y género. Por ejemplo, darle un nombre femenino a un robot social que sirve en el hogar (en respuesta a características del robot que invitan a este comportamiento social) no sólo "convierte" a ese robot en un tipo de persona "con quien" interactuaré. de maneras que sean aceptadas en mi comunidad y sociedad; también refleja, confirma y prolifera una cultura en la que se supone que las mujeres deben asumir ese papel. En este sentido, los robots sociales no son "meras máquinas" ni totalmente "otros"; Están vinculados con la sociedad y la cultura humanas de diversas maneras y son posibles gracias a esa sociedad y esa cultura. En este sentido, los robots no son "externos" a los humanos y sus objetivos sino que existe una relación interna entre humanos y robots.

3.3 Relacionalidad y creación de significado: los robots están integrados en totalidades culturales

En tercer lugar, y relacionado con los puntos anteriores: una visión alternativa debe ser relacional en todos sus aspectos. "Relacional" debe entenderse aquí en los siguientes sentidos. Una positividad alternativa debería reconocer que la relacionalidad de los robots no sólo está integrada en sistemas técnicos más amplios, sino también en sistemas sociotécnicos más amplios: al igual que otras tecnologías, los robots están "entrelazados con las prácticas sociales y los sistemas de significado de los seres humanos" [36:195]. Por tanto, están "relacionados" con tales prácticas y sistemas de significado. También se podría decir que están vinculados a relaciones sociales [32] o que están integrados en conjuntos culturales.

En un trabajo reciente, he expresado este arraigo sociocultural de los robots utilizando los conceptos de Wittgenstein de "juegos de lenguaje" y "forma de vida": al igual que el uso de palabras, el uso de cosas como los robots también está integrado en los juegos (tecnológicos) y una forma de vida [33, 37]. Utilizo el término "forma de vida" no en un sentido biológico, sino que lo interpreto (usando a Wittgenstein) como juegos y prácticas culturalmente definidos que constituyen un todo o todos, que dan significado a nuestras actividades y tecnologías. Estos todos nos permiten dar sentido a los robots e interactuar con ellos, y los robots a su vez contribuyen o son parte de nuestras actividades de creación de significado. No sólo contamos cosas de tecnología; Las tecnologías también contribuyen activamente a la creación de significado. Esa red u horizonte de significado, a su vez, nos permite dar sentido a experiencias y cosas concretas. Este punto también puede plantearse haciendo referencia al término "narrativa".

Utilizando la teoría narrativa de Ricoeur, Wessel Reijers y yo [38] hemos argumentado que, al igual que el texto, las tecnologías co-constituyen narrativas al configurar personajes y eventos en un todo significativo. Por ejemplo, una interacción particular entre humano y robot puede ubicarse dentro de una narrativa más amplia de construcción de relaciones de amistad con los robots: esto se puede hacer contando historias sobre el robot, pero la cuestión es que el propio robot también contribuye activamente a dar forma a esta narrativa. . Esto se puede hacer, por ejemplo, mediante la forma antropomórfica de un robot, que crea personajes y eventos (por ejemplo, dos amigos que se encuentran y preguntan cómo van las cosas).

De manera más general, se puede decir que las tecnologías tienen una función hermenéutica ya que contribuyen a la creación de significado (el término "hermenéutica" se utiliza aquí para referirse al significado y la interpretación). Esto también sucede cuando interactuamos con robots y a través de nuestras interacciones con robots. No son sólo el objeto de nuestra creación de significados e interpretaciones (por ejemplo, interpreto al robot como una "mera máquina"), sino que también moldean la forma en que damos sentido al mundo. Por ejemplo, un robot doméstico que interactúa con humanos como un sirviente podría transmitir el significado de que el mundo social es uno de amos y sirvientes y que el usuario es un amo.

Por lo tanto, los robots no son meros instrumentos si eso significa que estarían fuera del ámbito de la cultura humana; Más bien, se incluyen en nuestras actividades hermenéuticas, de manera pasiva (hablamos de ellos y les damos sentido) y de manera activa (contribuyen a dar forma al significado de nuestras historias).

Para la base filosófica subyacente, este enfoque relacional y hermenéutico implica que el lenguaje ontológico "es" (afirmaciones filosóficas o preguntas sobre qué es algo o alguien, por ejemplo, ¿qué es un robot?), utilizado por realistas e instrumentalistas, se vuelve problemático, especialmente si este "es" no se entiende de manera relacional. No existe una "cosa" o "herramienta" en sí misma ni un "otro" que no esté relacionado con los humanos.

Todas las tecnologías, tal como las crean y utilizan los humanos, están relacionadas con los humanos. No sólo los llamados robots "sociales", sino que todos los robots son ya sociales y relacionales en este sentido, incluso sin tener o antes de tener características consideradas "sociales". Por ejemplo, un robot industrial está integrado en juegos de lenguaje y tecnología que pertenecen a un contexto social-industrial particular, en el que los trabajadores humanos han sido reemplazados por máquinas en condiciones capitalistas. Pero en el caso de los robots sociales, el carácter social relativo de los robots es aún más evidente; ellos, por así decirlo, exhiben nuestro entorno social y nuestra cultura. Por ejemplo, un robot social habilitado para IA con interfaz de voz nos "hablará" de maneras (y nosotros hablaremos con él de maneras) que están conectadas con cómo "nosotros" (en nuestra familia, en nuestra comunidad, nuestro país, nuestra comunidad lingüística, etc.) hablar. Si logra ser humorístico, por ejemplo, entonces sólo será posible si se aprovecha una base (humana) compartida de conocimiento, experiencia y habilidad. Y si lograba hacer una broma mediante el aprendizaje autom

El robot habría aprendido a jugar nuestro juego. El éxito de su actuación depende enteramente de un todo relacional, de una forma de vida, que debe presuponerse para que el robot sea divertido y, en general, tenga sentido. Los especialistas en robótica social pueden hacer que el robot imite esta creación de sentido y creación de significado (aquí usaré sentido y significado indistintamente) permitiéndole aprender del material que encuentra en la red mundial, donde los corpus de texto materializan parte del significado que se encuentra en la web. alrededor. En la investigación sobre inteligencia artificial ya existen investigaciones sobre aspectos éticamente problemáticos del uso de corpus lingüísticos de Internet; por ejemplo, existe un sesgo de género en estos corpus y en nuestro lenguaje ordinario, que tiende a ser replicado por la IA [39]. Cuando la tecnología de inteligencia artificial se implemente cada vez más en los robots, estos problemas pueden proliferar. Pero considerada en el nivel de la interacción humano-robot, que por definición involucra a humanos y, por lo tanto, también se conecta con juegos y horizontes más amplios de significado, la interacción ya es siempre una actividad (en tiempo real) de creación de significado. Una vez más, en muchos sentidos el robot es más que una mera herramienta o instrumento: es un elemento de actividades y juegos que crean significado y dependen de totalidades más amplias para su significado. Además, el robot no es totalmente "otro" en el sentido de que, en la medida en que es realmente social, adquiere mucha similitud a través de sus relaciones con la sociedad y la cultura humanas. Sólo puede tener éxito, es decir, funcionar con éxito como robot social, si hay suficiente similitud en términos de proporcionar una base para un significado compartido. Un robot extraterrestre "social" no sería considerado social por nosotros, los humanos, ya que carecería de todas estas relaciones con el significado humano.

3.4 Falta de control hermenéutico

En cuarto lugar, sin embargo, decir que estas actividades de creación de material y de significado son humanas (son el tipo de cosas que hacen los humanos), no significa que estén totalmente bajo nuestro control. No controlamos completamente el significado, tampoco en situaciones "puramente" humanas (que rara vez son "puras" ya que generalmente— si no siempre, mediado por otras tecnologías y cosas).

También surge un significado cuando interactuamos con robots y hay encuentros y eventos. A veces surgen cosas malas que no esperábamos;

Éstas son emergencias involuntarias de significado. Por ejemplo, podemos reírnos o sentir miedo cuando de repente un robot suena o parece muy humano.

Como muestra una investigación del HRI, los robots suelen sorprender a la gente. En este sentido y en tales casos, hay una alteridad en ellos:

O al menos, a la interacción y al robot en relación.

Los desarrolladores y diseñadores de tales robots y los usuarios pueden intentar controlar el significado (por ejemplo, un diseñador podría intentar evitar efectos "extraños"), pero el control hermenéutico nunca es absoluto.

Cuando los humanos se relacionan con el mundo y entre sí, siempre hay lugar para sorpresas, para significados que no esperábamos. También en ese sentido los robots nunca son más máquinas, si "máquinas" evoca el significado de "mecánico" y si ese

se interpreta en el sentido de "predecible". La interacción humano-robot, como todas las interacciones, puede producir significados no deseados y nada excepcionales. La cuestión no es sólo que existen comportamientos no exigidos en HRI, como se ha documentado (p. ej., [40]); los humanos pueden y también entenderán la interacción en

maneras inesperadas. Por ejemplo, un gesto del robot que debería ser amistoso podría interpretarse repentinamente como hostil.

Este tipo de interpretaciones inesperadas son inevitables y forman parte de cómo existimos en el mundo. Sin embargo, tales significados emergentes tampoco pueden conceptualizarse como constitutivos de alteridad total, ya que la aparición de significados (por ejemplo, significados no intencionados como en el ejemplo) depende nuevamente por completo de la creación y experiencia de significado humanas. Siempre podremos comparar lo que hace el robot con algo que conocemos del (resto de) el mundo humano y de nuestros propios entornos sociales, con algo que sea igual.

y eso se relaciona con uno mismo. Nunca es completamente diferente u otro.

3.5 Poder

Finalmente, la robótica social entendida como interacción y como una actividad humana integral de creación de significado, uso social y construcción material, siempre está entrelazada con el poder.

Los robots sociales en uso e interacción no son sólo herramientas o actividades puramente "técnicas", sino que tienen significados y efectos sociales y políticos, y esto incluye un aspecto de poder. Por ejemplo, el lenguaje de "esclavos" utilizado por Bryson [25] es profundamente problemático [41], como ha reconocido recientemente el autor.

Y los robots sociales en un contexto de atención médica pueden plantear preguntas sobre las políticas de la atención médica sin robots: ¿cómo se hacen las cosas ahora y cómo se trata a las personas ahora, y cuál es la calidad y justificación de esa forma de hacer y de tratar a las personas? ¿A qué intereses se beneficia la introducción de estos robots sociales en ese contexto particular? Energía y tecnología están conectadas. El instrumentalismo ingenuo pasa por alto este punto, porque tiende a separar la tecnología y lo humano.

el social. Al colocar a los robots sociales en la categoría de "tecnología", se excluyen involuntariamente las cuestiones "humanas" relativas a la sociedad y el poder. (Y a veces el talismo instrumental no es ingenuo y nos ciega a propósito ante las relaciones sociales y las relaciones de poder.) Pero los robots sociales tampoco son "otros" o "no humanos" si eso significa que son meros socios en agradables relaciones de poder neutrales. juego cultural y material posmoderno (una dirección que, creo, Gunkel evita pero que podría ser un efecto de al menos una interpretación de Gunkel o del posmodernismo al estilo Haraway). Tomando en serio el lado político de Haraway, pero también inspirado por Marx [42] y Foucault [43], deberíamos señalar el peligro de que las fantasías posthumanas sobre la interacción con robots como alteridades puedan enmascarar cómo el proyecto de la robótica social puede integrarse en y contribuyendo a las narrativas de dominación y desigualdad de género, los sistemas socioeconómicos capitalistas y las narrativas neoliberales, y los juegos de poder micro y meso a nivel

nivel de interacciones y organizaciones humanas. Por ejemplo, los robots pueden utilizarse para recopilar datos que se venden a terceros, su uso puede eliminar puestos de trabajo, pueden proliferar formas problemáticas de tratar a las mujeres, su producción puede implicar la explotación de personas y los robots pueden utilizarse para demostrar el poder de una gran empresa o estado. En todo caso, estos problemas relevantes para el poder muestran la realidad y la posibilidad de una distopía poshumanista en lugar de una utopía. La robótica social bien puede presentar a los robots como "otros" y "amigos"; pero detrás de la cortina (y en realidad no tan bien escondida), puede haber manipulación, explotación y disciplina. También se podrían considerar las relaciones laborales que hacen posibles los robots sociales. Las celebraciones de la otredad y la alteridad y los argumentos a favor de la inclusión de los no humanos en lo social pueden dejar de lado el contexto socioeconómico más amplio y ocultar que, como sostiene Foucault [44], el poder está presente en todo tipo de relaciones, incluidas las que [tenemos](#) con otras personas al desarrollar y utilizar robots sociales. Centrarse en el robot como "otro" puede distraer la atención de los humanos que utilizan robots para extraer datos de los humanos y afectar los "cuerpos y almas, pensamientos, conducta y forma de ser" de otros humanos [45:18]. Por ejemplo, los robots sociales pueden usarse para manipular a las personas para que compren ciertos bienes y para estimular otros comportamientos, lo que redundaría en interés de una parte que no es visible en la interacción del robot humano en sí. Detrás de la máscara antropomórfica del robot, presentada como una invitación lúdica a una interacción agradable y útil del poder neutral con un cuasi-otro, se esconde la cara seria del poder humano y de las relaciones de poder humano. La relación no es sólo diversión y los juegos no se tratan sólo de jugar; Esta perspectiva de la robótica social (y de hecho del ser humano) también abre una caja de Pandora de problemas relacionados con el poder. La robótica social trata de juegos (en el sentido de Wittgenstein, véase más arriba), y algunos de ellos son, de hecho, juegos muy serios: juegos de poder. El hecho de que esto también sea cierto para todo tipo de otras herramientas digitales no es una excusa para apartar la vista de la dimensión de poder de estos juegos cuando se trata de robótica social.

3.6 Lo que ofrece la tercera visión: los robots como instrumentos en relación

Paradójicamente, al analizar la dimensión humana de la robótica en su uso y desarrollo, esta posición nos permite ser críticos con la antropomorfización de los robots. Precisamente porque los robots sociales están profundamente relacionados con lo humano y lo social, podemos señalar la antropomorfización problemática (o buena) y explicar por qué es problemática (o buena) señalando el gran todo en el que está inmerso el robot, por ejemplo. Por ejemplo, un juego particular sobre género. Al mismo tiempo, y nuevamente de manera un tanto paradójica, la posición evita una posición instrumentalista ingenua al tomar en serio el papel de los robots como un instrumento de creación de significado y un instrumento que es parte de, y co-constituye, un contexto más amplio: los robots.

son parte de lo que he llamado, en analogía con el contexto, una 'con-tecnología' [37: 263]: así como los textos son parte de un gran contexto que da forma a su significado, (otras) tecnologías también son parte de un contexto humano más amplio. entorno social, significativo y relacional en el que están insertos y que da forma a su significado. Los robots son instrumentos; pero son instrumentos en relación: siempre están conectados con los humanos y los campos socioculturales en los que operan. El instrumentalismo ingenuo sugiere una relación externa y superficial entre los humanos y su herramienta. El enfoque que desarrollé rechaza esta suposición. Mi deconstrucción del debate sobre el antropomorfismo se ha basado en una crítica de la supuesta relación externa entre humanos y robots. Los humanos y los robots están entrelazados de las muchas maneras descritas; Están interna y profundamente relacionados.

4 implicaciones para afrontar el fenómeno de la antropomorfización y para las relaciones sociales

La robótica como proyecto y campo.

¿Qué implica este enfoque y esta "tercera" visión de la relación entre humanos y robots para evaluar el antropomorfismo en la robótica social? ¿Qué tipo de posición normativa sobre estos fenómenos y prácticas se desprende de la discusión crítica anterior?

4.1 Los robots no son otros ni meras máquinas

En primer lugar, sobre la base de la visión alternativa desarrollada aquí, parece que tenemos que rechazar no sólo las dos visiones iniciales (instrumentalismo ingenuo y posthumanismo acrítico), sino también las posiciones normativas relacionadas con ellas. Sugiero que el antropomorfismo y la antropomorfización en la robótica social no se condenen por ir más allá de lo que debería ser una máquina (una mera máquina) ni se utilicen (mal) como una oportunidad para proyectar la alteridad en la máquina. En los muchos sentidos esbozados anteriormente, los robots sociales (incluso aquellos que son antropomórficos) pueden ser experimentados como otros, pero no son Otros con "O" mayúscula, como si misteriosamente estuvieran fuera del mundo humano. Los robots sociales están hechos por humanos y son parte de esfuerzos más amplios de creación de significado humano. Por tanto, no pueden construirse como totalmente otros. Hay mucha similitud allí, debido al vínculo entre el robot y su contexto de desarrollo y el uso de contratecnología. Podemos percibirlos como otros, pero no debemos olvidar que fueron creados por e

¹ Podría agregar que tampoco son "otros" en el sentido de amigos o compañeros, ya que lo único que sabemos sobre la amistad y el compañerismo humanos es que su valor reside, entre otras cosas, en que el amigo o compañero no es creado por nosotros y, de hecho, en su alteridad. Esto no se puede crear, aunque puede surgir cierto grado de alteridad en la interacción. Sin embargo, no desarrollaré este punto aquí.

Pero tampoco los robots sociales son más que máquinas. La antropomorfización es una de las cosas que hacemos como humanos creadores de significado: incluso si los robots no están diseñados para invitar a la antropomorfización, siempre habrá "antropomorfización", en el sentido de que el significado de los robots, incluidos los robots sociales, cambiará. siempre estará vinculado a una práctica humana más amplia de creación de significado y a horizontes más amplios de significado, una "forma de vida" (para usar nuevamente el concepto wittgensteiniano). El enfoque presentado no nos permite decidir a favor o en contra de una antropomorfización específica por parte de los desarrolladores/ diseñadores o usuarios. Pero sí nos ayuda a comprender que cualquier discusión sobre un caso particular tendrá que evocar significados más allá de la interacción específica: significados extraídos de nuestra sociedad y nuestra cultura, con todas sus dimensiones, incluida una dimensión de poder. Esto implica que la posición proantropomorfismo dentro del instrumentalismo también es insostenible si y en la medida en que los investigadores que sostienen esta posición no estén dispuestos a considerar esta relacionalidad, significado y normatividad más amplia del trabajo de antropomorfización que realizan.

4.2 Redefinición de los objetivos de la robótica social

Además, el enfoque relacional y hermenéutico esbozado también debería aplicarse a nuestra comprensión de la robótica social como campo. La robótica social debe entenderse como un campo interdisciplinario y multidisciplinario que no sólo se ocupa de la creación de conjuntos de artefactos y software reales llamados "robots" (posición instrumentalista) o de la creación de "otros artificiales" (posición posthumanista), sino que también siempre al mismo tiempo con la creación de significados humanos, sociales, culturales y políticos (y la afirmación misma de que el robot es un "otro" es uno de esos significados que tiene significado político y que debe ser problematizado).

Así, si bien he indicado cómo puntos de vista específicos y sus posiciones normativas relacionadas son problemáticos, el presente artículo no conduce a una conclusión normativa específica sobre si los robóticos deberían o no construir robots antropomórficos. A lo sumo, uno podría derivar una advertencia de esto: si diseñas robots de este tipo, entonces sé consciente de que es más probable que te unas a esta actividad más amplia de creación de significado social con todos sus aspectos de poder, o al menos te unas a esta tecnología más amplia de estafa para en mayor medida. Por esta razón, algunos pueden concluir que es más seguro no desarrollar este tipo de robots. Otros persistirán en crear estos robots, pero luego tienen la responsabilidad de hacerlo de manera que sea consciente y tenga en cuenta los significados y consecuencias socioculturales. Esto me lleva al siguiente punto sobre la responsabilidad de los robóticos.

4.3 La responsabilidad de los robóticos, educadores, y Usuarios-Ciudadanos

Normativamente, reconocer esta dimensión relacional, hermenéutica y política de los robots sociales y de la robótica en general implica que los robóticos, informáticos, ingenieros, psicólogos y todos los demás involucrados en el campo también son responsables de los significados que generan y de las consecuencias relacionales sociales. ellos producen. Esto no debería verse como algo marginal sino como parte del negocio principal de la robótica social. Por ejemplo, tanto un diseñador de hardware de robot como un desarrollador de software de reconocimiento de voz que se utilizará en un robot deben entenderse a sí mismos no sólo como creadores de un producto tecnológico que tiene algunas funciones previstas (por ejemplo, poder mantener una conversación con un humanos en un contexto doméstico), sino también como actores sociales que, a través del robot, intervienen en un campo de juego social que incluye desafíos que no son meramente técnicos (por ejemplo, evitar o disminuir los prejuicios de género). La responsabilidad de los robóticos sociales significa entonces la responsabilidad de garantizar que la interacción robot y humano-robot logre sus objetivos, pero también de hacerlo de una manera que ejerza responsabilidad ante otras personas: significa responder a los usuarios y otras partes interesadas que puedan estar afectado por las consecuencias no deseadas del diseño, código, etc., incluido el potencial impacto hermenéutico/semiótico. Significa responder a los problemas sociales además de los técnicos, y descubrir cómo los cambios en el aspecto técnico pueden contribuir a mitigar lo que sucede a nivel social.

No es fácil definir el alcance preciso de la responsabilidad por el potencial impacto hermenéutico/semiótico y social de los robots. Por ejemplo, es imposible predecir completamente el futuro y sigue sin estar claro hasta qué punto uno es responsable de significados intencionados sorprendentes e inesperados. Como ya hemos dicho, no existe un control hermenéutico pleno y absoluto, ni en la robótica ni en ningún otro lugar de la vida y la cultura humanas. Pero uno debería aumentar tanto como sea posible el nivel de control hermenéutico desde los significados intencionados hasta los significados no intencionados y potenciales (futuros). Se podría esperar de los investigadores y diseñadores de robótica (y de otras personas en las empresas y organizaciones para las que trabajan) que al menos intenten imaginar y evaluar los impactos previstos y potenciales. Investigadores de otras disciplinas, también de las humanidades y las ciencias sociales, podrían ayudar desarrollando y transfiriendo metodologías para ello, por ejemplo métodos para crear escenarios y métodos de evaluación ética.

Dado el vínculo con el nivel social, la responsabilidad en y para la robótica social incluye asumir la responsabilidad de los significados y consecuencias de la robótica social en términos de poder. Los investigadores técnicos en robótica tal vez no consideren que tengan mucho que ver con el poder, porque generalmente se entiende que el poder tiene que ver con las relaciones entre

autoridades políticas y ciudadanos. Una comprensión foucaultiana, sin embargo, nos dirige a cuestiones de poder en todo tipo de lugares y niveles del tejido social. El laboratorio de robótica no está excluido de esta omnipresencia del poder. Junto con otros que toman decisiones sobre robótica y emplean o utilizan robots en diversos contextos, quienes desarrollan y diseñan robots tienen cierto poder para dar forma al significado de las prácticas, las historias de las personas y las metas que nos fijamos como humanos. Para el desarrollo de robots antropomórficos, este poder es quizás más significativo que en el caso de otros tipos de robots, ya que el antropomorfismo sustenta los aspectos sociales de los robots, incluido el poder. Este poder conlleva responsabilidad.

El poder no sólo es negativo: puede ser restrictivo y opresivo, sino que también puede ser habilitante. El poder puede tomar la forma de empoderamiento: se puede dar poder a las personas para que hagan cosas y esto puede aumentar su potencial. En el contexto de la robótica social, la cuestión es entonces cómo empoderar a las profesiones técnicas para que ejerzan la responsabilidad de sus creaciones. Desafortunadamente, tradicionalmente los ingenieros y científicos carecen de todas las herramientas para hacerlo, ya que la atención se centra en los conocimientos y las habilidades técnicas. Es necesario cambiar esto. Metodológicamente, se necesita un enfoque multidisciplinario más amplio, pero también interdisciplinario y transdisciplinario. Esto podría significar, por ejemplo, que además de los lenguajes informáticos y las herramientas de ingeniería, también se utilicen instrumentos de las humanidades y las ciencias sociales. El proyecto mismo de una "robótica social" debe concebirse de manera relacional y vincularse al proyecto de evaluación crítica de la cultura y la sociedad humanas. Esto es necesario ya que, como he sugerido, el laboratorio del robotista está conectado a un ámbito social y entornos, al lenguaje, al poder cultural, etc., y la fabricación y el uso de robots al mismo tiempo moldean todo este campo relacional en la misma medida en que él mismo está moldeado por él. Si los robóticos asumen este desafío y esta responsabilidad, y redefinen su lugar dentro de este todo más amplio, entonces podrán desempeñar un papel importante en la configuración del futuro de nuestros vínculos sociales, algo que los ingenieros siempre hicieron, desde puentes y ferrocarriles hasta electricidad y telefonía móvil. teléfonos móviles, de una forma ética y socialmente responsable.

Estas implicaciones normativas y metodológicas del enfoque descrito anteriormente exigen transdisciplinariedad y una reforma de la educación de los ingenieros en robótica y los científicos informáticos. Cuando ingresan a su vida profesional, no solo deben estar dispuestos, sino también preparados y capaces de conectar su práctica y sus decisiones de desarrollo y diseño con un debate social y un discurso cultural más amplio, en el que deberían poder participar críticamente mediante el uso de palabras. y cosas (robots). Pero la responsabilidad no debería recaer únicamente en los especialistas en robótica. Los usuarios-ciudadanos, por su parte, también deben asumir su responsabilidad por el futuro de la robótica social y, más en general, por el futuro de la tecnología y el futuro de la sociedad. Necesitan entender (y ser

educados para comprender) que los robots sociales que pueden encontrar en algunos de los contextos y contratecnologías de su vida diaria no son "otros" ni meras máquinas o instrumentos, sino herramientas que son parte de su sociedad y cultura, para las cuales también llevar responsabilidad. Si vemos a los robots no como opuestos a la esfera humana sino como una parte crucial de ella, entonces los robots se vuelven interesantes en formas que quizás no hubiéramos imaginado antes. Conectar personas y lenguas de diferentes mundos a través de la educación transdisciplinaria puede conducir a una comprensión transformada de lo que la tecnología puede ser y hacer por la sociedad y, en última instancia, a un cambio profundo de ambos. Esto es para lo que deben estar preparados los estudiantes: estudiantes de ciencias e ingeniería, pero también estudiantes de humanidades y ciencias sociales.

4.4 Conclusión general

En ese metasentido, entonces, sugiero que necesitamos más antropomorfización en lugar de menos. No necesariamente la antropomorfización debida a las características de diseño de los robots y las interacciones que estos permiten, que mejora y aumenta las relaciones ya inevitables y existentes entre el robot y su entorno social y cultural, y que por lo tanto puede ser a la vez emocionante y muy problemático (por ejemplo, (por ejemplo, porque proliferan determinadas relaciones de poder que creemos que son indeseables o inaceptables). Esto necesita un escrutinio y un debate por parte de los diseñadores y los usuarios-ciudadanos, quienes deberían asumir la corresponsabilidad del diseño de dichos robots.

Pero necesitamos especialmente una mayor antropomorfización de los robots sociales en el sentido de un mayor reconocimiento de la percepción misma de que existe ese vínculo entre los robots y su hermenéu.

En un entorno ético, existe esa profunda relacionalidad social y semántica: la comprensión (tanto en el laboratorio como en la sociedad en general) de que los robots sociales, como cualquier otra tecnología, son en verdad muy humanos. Además de todo lo demás que hace, el antropomorfismo en la robótica social muestra esa relacionalidad y esa humanidad: nos recuerda que la tecnología siempre ha sido humana.

Financiación Financiación de acceso abierto proporcionada por la Universidad de Viena.

Declaraciones

Conflicto de intereses El autor declara no tener ningún conflicto de intereses.

Acceso Abierto Este artículo está bajo una Licencia Internacional Creative Commons Attribution 4.0, que permite el uso, compartir, adaptación, distribución y reproducción en cualquier medio o formato, siempre y cuando se dé el crédito apropiado a los autores originales y a la fuente. proporcione un enlace a la licencia Creative Commons e indique si se realizaron cambios. Las imágenes u otro material de terceros en este artículo están incluidos en la licencia Creative Commons del artículo, a menos que se indique lo contrario.

de lo contrario en una línea de crédito al material. Si el material no está incluido en la licencia Creative Commons del artículo y su uso previsto no está permitido por la normativa legal o excede el uso permitido, deberá obtener permiso directamente del titular de los derechos de autor. Para ver una copia de esta licencia, visite <http://creativecommons.org/licenses/by/4.0/>.

Referencias

- Fong T, Nourbakhsh I, Dautenhahn K (2003) Una encuesta sobre robots socialmente interactivos. *Sistema de robot automático* 42:143–166
- Dautenhahn K (2007) Robots socialmente inteligentes: dimensiones de la interacción humano-robot. *Philos Trans R Soc Lond B Biol Sci* 362:679–704
- Duffy BR (2003) Antropomorfismo y robot social. *Sistema de autonomía del robot* 42:177–190
- Ishiguro H (2006) Humanoides y androides interactivos como interfaces ideales para los humanos. En *Actas de la undécima conferencia internacional sobre interfaces de usuario inteligentes (IUI '06)*. Association for Computing Machinery, Nueva York, NY, EE. UU., 2–9
- Bartneck C, Kulk D, Croft E (2008) Instrumentos de medición para el antropomorfismo, la animación, la simpatía, la inteligencia percibida y la seguridad percibida de los robots. *Robot internacional J Soc* 1:71–81
- Zlotowski J, Proudfoot D, Yogeeswaran K et al (2015) Antropomorfismo: oportunidades y desafíos en la interacción humano-robot. *Int J de Soc Robotics* 7:347–360. <https://doi.org/10.1007/s12369-014-0267-6>
- Kiesler S, Hinds P (2004) Introducción a este número especial sobre la interacción entre humanos y robots. *Interacción Hum-Comput* 19(1):1–8
- Reeves B, Nass C (1996) La ecuación de los medios: cómo la gente trata las computadoras, la televisión y los nuevos medios como personas y lugares reales. Prensa de la Universidad de Cambridge, Cambridge
- Kiesler S, Powers A, Fussell SR, Torrey C (2008) Interacciones antropomorfas con un robot y un agente similar a un robot. *Soc. Cogn* 26(2):169–181
- Hegel F, Gieselmann S, Peters A, Holthaus P, Wrede B (2011) Hacia una tipología de señales y señales significativas en robótica social. En: *Actas del taller internacional IEEE sobre comunicación interactiva humana y robótica*, págs. 72–78
- Kuchenbrandt D, Eyssel F, Bobinger S, Neufeld M (2013) Cuando la pertenencia al grupo de un robot importa. *Robot internacional J Soc* 5(3):409–417
- Krägeloh CU, Bharatharaj J, Kutty SKS, Irmala PR, Huang L (2019) Cuestionarios para medir la aceptabilidad de los robots sociales: una revisión crítica. *Robótica* 8(4):88. <https://doi.org/10.3390/robotics8040088>
- Peca A, Coeckelbergh M, Simut R, Costescu C, Pintea S, David D, Vanderborght B (2016) Terapia mejorada con robots para niños con trastornos de autismo: medición de la aceptabilidad ética. *IEEE Technol Soc Mag* 35(2):54–66
- Eyssel F (2017) Una perspectiva psicológica experimental sobre la robótica social. *Sistema autónomo de robots* 87:363–371
- Darling K (2017) ¿Quién es Johnny? Encuadre antropológico en la interacción, integración y políticas entre humanos y robots sociales. En: Lin P, Jenkins R, Abney K (eds) *Ética del robot 2.0*. Prensa de la Universidad de Oxford, Nueva York
- Turkle S (2011) Solos y juntos: por qué esperamos más de la tecnología y menos unos de otros. Libros básicos, Nueva York
- Scheutz M (2012) Los peligros inherentes de los vínculos emocionales unidireccionales entre humanos y robots sociales. En: Lin P, Abney K, Bekey G (eds) *Ética del robot*. MIT Press, Cambridge, MA, págs. 205–222
- Mori M (1970) El valle inquietante. *Energía* 7:33–35
- MacDorman KF, Ishiguro H (2006) La asombrosa ventaja del uso de androides en la investigación de ciencias sociales y cognitivas. *Stud interactivo* 7:297–337
- Freud S (2003) Lo siniestro. Clásicos del pingüino, Londres
- Richardson K (2018) Animación tecnológica. En: Swancutt K, Mazard M (eds) *Animismo más allá del alma: ontología, reflexividad y creación de conocimiento antropológico*. Berghahn, Nueva York, págs. 110–128
- Goetz J, Kiesler S, Powers A (2003) Emparejar la apariencia y el comportamiento del robot con las tareas para mejorar la cooperación entre humanos y robots. En: *Actas del 12º taller internacional IEEE sobre comunicación interactiva humana y robótica (ROMAN 2003)*, 2003, págs. 55–60
- Coeckelbergh M (2018) Cómo describir y evaluar los fenómenos del “engaño”: reformular la metafísica, la ética y la política de las TIC en términos de magia y desempeño y tomar un giro relacional y narrativo. *Ética Inf Technol* 20(2):71–85
- Flusser V (1999). *Forma de las cosas: una filosofía del diseño*. reaccionar ción libros, Londres
- Bryson J (2010) Los robots deberían ser esclavos. En: Wilks Y (ed) *Estrecho compromiso con compañeros artificiales*. John Benjamins, Presa de Ámsterdam, págs. 63–74
- Sparrow R, Sparrow L (2006) ¿En manos de las máquinas? El futuro del cuidado de personas mayores. *Mentes Mach* 16:141–161
- Haraway D (2000) Un manifiesto cyborg. En: David B, Kennedy BM (eds) *El lector de ciberculturas*. Routledge, Londres, págs. 291–324
- Latour B (1993) *Nunca hemos sido modernos*. Harvard University Press, Cambridge, MA, Catherine Porter (traducción)
- Gunkel DJ (2007) *Pensar de otra manera: filosofía, comunicación*. Purdue University Press, West Lafayette, Tecnología
- Gunkel DJ (2017) La otra pregunta: ¿pueden y deben los robots tener ¿derechos? *Ética Inf Technol* 20:87–99
- Ihde D (1990) *Tecnología y mundo de la vida*. Universidad de Indiana Prensa, Bloomington
- Coeckelbergh M (2012) *Relaciones morales crecientes: crítica de la adscripción de estatus moral*. Palgrave Macmillan, Nueva York
- Coeckelbergh M (2017) *Uso de palabras y cosas: lenguaje y filosofía de la tecnología*. Routledge, Nueva York
- Coeckelbergh M (2019) *Movidos por las máquinas: metáforas del rendimiento y filosofía de la tecnología*. Routledge, Nueva York
- Coeckelbergh M (2011) ‘Tú, robot: sobre la construcción lingüística de otros artificiales’ en: *AI & Society* 26(1): 61–69
- Johnson DG (2006) *Sistemas informáticos: entidades morales pero no agentes morales*. *Ética Inf Technol* 8:195–204
- Coeckelbergh M (2018) *Juegos tecnológicos: uso de Wittgenstein para comprender y evaluar la tecnología*. *Ética en ciencias e ingeniería* 245: 1503–1519
- Coeckelbergh M, Reijers W (2016) *Tecnologías narrativas: una investigación filosófica de las capacidades narrativas de las tecnologías mediante el uso de la teoría narrativa de Ricoeur*. *Hum Stud* 39:325–346
- Caliskan A, Bryson JJ, Narayanan A (2017) La semántica se deriva automáticamente de corpus de lenguaje que contienen sesgos similares a los humanos. *Ciencia* 356 (6334): 183–186
- Lemaignan S, Fink J, Mondada F, P Dillenbourg (2015) ¿Lo estás haciendo mal! Estudiar comportamientos inesperados en la interacción niño-robot. En Tapus A., André E., Martin JC., Ferland F., Ammi M. (eds) *Robótica social. ICSR 2015*. Apuntes de conferencias sobre informática, vol 9388. Springer, Cham.
- Coeckelbergh M (2015) La tragedia del maestro: automatización, vulnerabilidad y distancia. *Ética Inf Technol* 173: 219–229
- Marx K (1867/1976) *El capital* (Vol. I). Penguin Books, Londres (reimpresión en 1990).
- Foucault M (1980) *Poder/conocimiento: entrevistas seleccionadas y otros escritos 1972–1977*. Colin Gordon ed. Libros Pantheon, Nueva York

44. Foucault M (1980) Poder/conocimiento. Gordon C (ed) Gordon C, Marshall L (trad.). Libros Pantheon, Nueva York
45. Foucault M (1988) Tecnologías del yo: un seminario con michel foucault. Martin LH, Gutman H, Hutton PH (eds). Prensa de la Universidad de Massachusetts, Amherst

Nota del editor Springer Nature se mantiene neutral con respecto a reclamos jurisdiccionales en mapas publicados y afiliaciones institucionales.