10.52712/issn.1850-0013-719

Tensiones y anhelos filosóficos en las interacciones humano-robot. Un abordaje ético y posfenomenológico al problema de la alteridad del androide y su lugar en la sociedad futura *

Tensões e anseios filosóficos nas interações humano-robô. Uma abordagem ética e pós-fenomenológica do problema da alteridade do andróide e seu lugar na sociedade futura

Tensions and Philosophical Longings in Human-Robot Interactions. An Ethical and Post-Phenomenological Approach to the Problem of Android Alterity and its Place in the Future Society

Ricardo Andrade **

Este artículo estudia el problema de la alteridad en las interacciones humano-robot, especialmente en la figura del androide. Pensar a este ente como un "tú" con sus particularidades trae consigo una serie de problemas fenomenológicos y éticos que se vinculan fuertemente con la pregunta acerca del lugar del ser humano en el marco del desarrollo tecnológico futuro. La aparición de esta alteridad puede modificar las concepciones que se tienen actualmente de la sociedad, la vida y los sentimientos, entre otras características. En este sentido, el artículo tiene como premisa principal ofrecer un estudio que otorgue respuestas, desde un punto de vista filosófico, a las tensiones que puedan emerger entre este nuevo agente social y los sujetos.

Palabras clave: interacción humano-robot; androides; posfenomenología; roboética; alteridad tecnológica

* Recepción del artículo: 20/07/2024. Entrega del dictamen: 09/10/2024. Recepción del artículo final: 08/11/2024.

** Doctorando en ciencias sociales y humanidades, orientación en estudios sociales de la ciencia y la tecnología en la Universidad Nacional de Río Negro (UNRN), Argentina. Becario del CONICET con sede de trabajo en el Centro de Estudios en Ciencia, Tecnología, Cultura y Desarrollo (CITECDE/UNRN). Magíster en literaturas en lenguas extranjeras y en literaturas comparadas por la Universidad de Buenos Aires (UBA), Argentina, y diplomado en problemas filosóficos contemporáneos (UBA) y licenciado en letras por la Universidad Central de Venezuela (UCV). Correo electrónico: jrandraderangel@unrn.edu.ar. ORCID: https://orcid.org/0000-0002-6422-0809.

CC BY

10.52712/issn.1850-0013-719

O objetivo deste artigo é estudar o problema da alteridade nas interações humano-robô, especialmente na figura do andróide. Pensar esta entidade como um "você" com as suas particularidades traz consigo uma série de problemas fenomenológicos e éticos que estão fortemente ligados à questão sobre o lugar do ser humano no quadro do futuro desenvolvimento tecnológico. O aparecimento dessa alteridade pode modificar as concepções atuais de sociedade, de vida, de sentimentos, entre outras características. Nesse sentido, a premissa principal do artigo é oferecer um estudo que dê respostas, do ponto de vista filosófico, às tensões que podem surgir entre esse novo agente social e os sujeitos.

Palavras-chave: interação humano-robô; androides; pós-fenomenologia; roboética; alteridade tecnológica

This article studies the problem of otherness in human-robot interactions, especially in the figure of the android. Thinking of this entity as a "you" with its particularities brings with it a series of phenomenological and ethical problems that are strongly linked to the question about the place of human beings in the framework of future technological development. The appearance of this otherness can modify the current conceptions of society, life and feelings, among other characteristics. In this sense, the main premise of this article is to offer a study that provides answers, from a philosophical point of view, to the tensions that may emerge between this new social agent and the subjects.

Keywords: *human-robot interaction; androids; post-phenomenology; roboethics; technological alterity*

10.52712/issn.1850-0013-719

Introducción metodológica

El artículo estará dividido en dos secciones. En la primera, se ofrecerá un análisis del problema de la alteridad partiendo del enfoque posfenomenológico de la filosofía de la tecnología contemporánea. Para ahondar en dicho aspecto, se tomará en consideración una serie de conceptos que son importantes para entender las vinculaciones entre los seres humanos y la tecnología. Estos conceptos permitirán una mejor comprensión de la otredad en este contexto. Algunos de ellos son: coevolución, individuo técnico y alteridad técnica. En la segunda parte, se ahondará en las interacciones humano-robot, especialmente en el androide. Para realizar un abordaje más concreto, esta sección estará dividida en dos: una dedicada a los aspectos ontológicos y posfenomenológicos de estos entes y la otra a cuestiones relacionadas con la ética y las alternativas para encarar las posibles situaciones sociales que se desprendan de la interacción entre los sujetos y estas entidades.

En la primera subdivisión, algunas apreciaciones de índole fenomenológica provistas por la filosofía de la tecnología japonesa contemporánea serán valiosas para indagar en la figura del androide; en tanto que, en la segunda subdivisión, conceptos como los de autonomía y adopción jugarán un papel destacado, ya que se indagará en el escenario hipotético de androides equipados con inteligencia artificial general.

El problema de la alteridad: acercamientos desde la posfenomenología

En la filosofía de la tecnología contemporánea, el enfoque posfenomenológico ha jugado un papel crucial en los análisis vinculados a las interacciones entre los sujetos (usuarios) y los entes técnicos. Al otorgarle a la experiencia y a la intencionalidad lugares destacados (por la enorme influencia de la fenomenología clásica), la perspectiva posfenomenológica recurre a los entramados sociales, simbólicos, éticos y ontológicos para comprender el lugar del sujeto en los desarrollos tecnológicos y la génesis de nuevas formas de subjetividad ligadas a los distintos niveles del mundo técnico (artefactos, máquinas, interfaces e infraestructuras, entre otros). En una definición concisa sobre este enfoque, Verbeek y Rosenberg (2015) señalan que el acercamiento posfenomenológico combina el análisis filosófico con la orientación empírica en el ámbito de la ciencia y la tecnología para observar cómo estos enriquecen la experiencia humana y el mundo, generando de este modo preguntas en torno a cómo los diversos entes técnicos moldean las acciones, decisiones y experiencias que se tienen de la realidad, entre otras características. Al incluir la orientación empírica como uno de sus principios fundamentales, la posfenomenología busca desarrollar un conocimiento que tenga en consideración las prácticas cotidianas de los sujetos con la tecnología. Este conocimiento posee múltiples potencialidades, entre las que se destaca una mayor comprensión de las interacciones entre humanos y artefactos técnicos que permita desarrollar nuevos diseños tecnológicos (innovación) y que ayude, en un nivel macro, a desplegar políticas científicas y tecnológicas a partir dichas relaciones. Esto implica que la posfenomenología tiene también como premisa filosófica el estudio de los aspectos sociales e institucionales vinculados a la tecnología, ya que ambas son esenciales para entender el enriquecimiento de la experiencia humana.

10.52712/issn.1850-0013-719

Al tener en consideración esta definición inicial, conviene presentar el primer concepto fundamental de este enfoque. Se trata de la "mediación". Como su significado puede denotar, la mediación valida la idea según las entidades externas (en este caso, las tecnológicas) poseen algún grado de incidencia en la constitución existencial de la persona. De acuerdo con Mykhailov y Liberati (2023), la mediación se muestra usualmente como una actividad transformadora que permite conducir la intencionalidad hacia una experiencia plena en la relación humano-tecnología. Si se acepta este punto, el concepto de mediación se presenta como una herramienta que es capaz de conectar al ser humano con nuevos modos de existir gracias al reconocimiento que se les otorga a las entidades tecnológicas. El ejercicio práctico de estos modos de existir no se limita solamente a una acción individual limitada en el tiempo, sino que crean las condiciones adecuadas para una transformación generalizada de los sentidos y, por ende, de las estructuras sociales en su conjunto. En este punto se evidencia que el reconocimiento desempeña dos características a considerar. La primera tiene que ver con que, en el momento en que el reconocimiento se manifiesta, se les otorga alteridad a los entes técnicos; el segundo punto guarda relación con que, gracias a dicha alteridad concedida, estas entidades pasan a generar múltiples tipos de convivencias que modifican la autopercepción que tiene el sujeto de sí mismo y la sociedad.

En este sentido, la mediación de índole posfenomenológica permite establecer la siguiente premisa: para que exista la transformación del ser humano por medio de la tecnología, este deber hacer de las convivencias con los entes técnicos una parte fundamental de su existencia, ya que estos propician una metamorfosis del mundo interior y exterior basadas en el reconocimiento de la alteridad. Bajo esta premisa, cobra un mayor sentido el énfasis en la palabra "convivencia". Por una parte, remite a la creación de espacios sociales e individuales sustentados en el cuidado, la ética y el equilibrio entre diferentes entes; por otra parte, estos aspectos generan vivencias que favorecen el desarrollo de nuevos modos de existir en el mundo tanto de los seres humanos como de las entidades tecnológicas. La orientación empírica de la posfenomenología explora las convivencias con la finalidad de extraer de ellas imaginarios tecnológicos, posibles marcos normativos y transformaciones en la autopercepción de los sujetos que emanan de las interacciones entre los humanos y la tecnología. Ahondar en ellas resulta fundamental para comprender, por ejemplo, las interacciones humano-robot, como se verá en las secciones siguientes.

Ahora bien, el establecimiento de la mediación como concepto clave permite profundizar en una serie de relaciones que forman parte de las herramientas analíticas de la posfenomenológica y que son importantes para comprender mejor la mediación convivencial entre humanos y entidades técnicas. De acuerdo con Ihde (2009), existen cuatro tipos de relaciones con la tecnología, a saber: relación de encarnación (*embodiment relation*), relación hermenéutica (*hermeneutic relation*), relación de trasfondo (*background relation*) y relación de alteridad (*alterity relation*). Si bien en los análisis empíricos estas cuatro relaciones interactúan entre sí sin mayores inconvenientes, conviene prestarle mayor atención a la última, ya que ella ofrece más

información sobre la importancia de la otredad cuando se habla de la tecnología.[1] Para Ihde, "*not all of our relations with technologies are so referential. We may also-again actively-engage technologies themselves as quasi-objects or even quasi-others, hence the term alterity*" (2009, p. 43). El filósofo estadounidense no profundiza mucho más sobre este caso. Si bien su definición ofrece elementos para entender esta relación, las reflexiones que se presentarán a continuación tienen como objetivo principal profundizar más en ella y ofrecer conceptos que la enriquezcan.

Uno de los elementos que destaca Ihde tiene que ver con que la relación de alteridad disminuye la referencialidad humana; es decir, le otorga un lugar destacado a la vida y al desenvolvimiento de los entes técnicos en el mundo. Estos son vistos como "cuasi otros", lo cual significa que comparten algunas características, ya sea a nivel ontológico o fenomenológico, con el ser humano. En esta relación, estas entidades dejan de pensarse como meros instrumentos o artefactos "pasivos" para involucrarse activamente en el ámbito social y en la intimidad del ser humano. Si se parte de estos puntos, los robots pueden interpretarse como las entidades que mejor encarnan la relación de alteridad. La presencia de los robots implica que los modos de conocer y transformar el mundo deben ser pensados en el marco de una coevolución.

De acuerdo con Fernández (2022), la coevolución permite pensar la interacción recíproca de dos factores o elementos que se relacionan en un vínculo de co-constitución y que, en el marco de la tecnología, involucra niveles de hibridación mediados por las instituciones y la *praxis*. Como se puede observar gracias a esta definición, la alteridad robótica no solo esta mediada por las instituciones (manifestaciones políticas y culturales de la sociedad), sino que ella misma es capaz de crear formas institucionales nuevas debido a sus características cuasi humanas y a su estatuto de entidad coevolutiva. Esto significa que los robots generan "intencionalidades tecnológicas". Verbeek destaca que la intencionalidad tecnológica: "*Needs to be understood in terms of both experience and action. It can indicate the way in which technologies can be directed at particular aspects of reality and the 'purporsiveness' that technologies can embody*" (2011, p. 145).[2] Si se acepta la tesis de la coevolución, estas entidades tecnológicas poseen la capacidad de construir intencionalidades que, por una parte, tienen una fuerte incidencia en los procesos

1. Sin embargo, y a modo informativo, se puede ofrecer una sucinta explicación de las restantes relaciones. Con respecto a la relación de encarnación, esta tiene que ver con cómo las tecnologías son incorporadas directamente al cuerpo del usuario (por ejemplo, un brazo biónico). La relación hermenéutica hace alusión a la interpretación, por parte del usuario, de las informaciones dadas por las entidades técnicas (un termómetro que mide la temperatura sería un caso prototípico de esto). Con respecto a la relación de trasfondo, esta hace alusión a las infraestructuras tecnológicas.

2. Adscribirles intencionalidad a los entes tecnológicos implica asumir que poseen algún grado de agencia. De acuerdo con Heras-Escribano (2019), la agencia es, desde un punto de vista poscognitivo, la suma de todas las capacidades flexibles que posee un organismo para adaptarse a los ambientes cambiantes, lo cual significa que la cognición es entendida como un comportamiento adaptativo del agente que resulta de las prácticas de dichas capacidades. Gracias a esta definición, se puede aseverar que diversas entidades técnicas (por ejemplo, la inteligencia artificial y los robots) realizan prácticas que modifican su entorno y, con ello, a los seres humanos. A su vez, estas prácticas hacen que evolucionen hacia nuevas formas de vida mediadas por las acciones de los sujetos. El diseño y la innovación tecnológica pueden considerarse como ejemplos palpables de esto.

de individuación humanos y, por otra parte, gestan sus propias dinámicas internas. Esto permite pensar que, al existir un entrecruzamiento entre la agencia humana y la tecnológica, emergen intencionalidades híbridas en donde los procesos de "antropoformización" y "cyborgrización" son claves para entender las convivencias.[3] Estos procesos serán estudiados con mayor detalle más adelante, cuando se ahonde en los aspectos posfenomenológicos de los androides.

A la luz de lo ya expuesto, la intencionalidad tecnológica permite aseverar que existen entes técnicos que poseen capacidades para individuarse gracias a la coevolución. Estas capacidades permiten la aparición del "individuo técnico". Rodríguez (2022) destaca, siguiendo a Simondon, que el individuo técnico tiene un modo de existencia propio que difiere de las entidades vivientes como de las físicas, es producto de una invención social y cultural y es la sede de procesos constantes de transformación. Esto significa que el individuo técnico se gesta a través de las prácticas y experiencias cotidianas. En este sentido, la convivencia con cada individuo técnico revela cómo se debe construir un conocimiento sobre él y, no menor, cómo la persona puede o debe sentirse ante el mismo. Si se acepta la premisa de Mykhailov (2020) según la cual la intencionalidad tecnológica señala cómo cada artefacto representa un modo diferente de dar la realidad, se puede observar cómo los individuos técnicos demandan una atención particular que termina por transformarse en nuevos modos de conocer, habitar y representar el mundo. Las entidades técnicas dan realidad en el momento en que se vuelven parte esencial en la construcción de las experiencias cotidianas que sostienen la vida humana. Para que la donación de realidad otorgada por los individuos técnicos sea reconocida por los sujetos, la relación de alteridad de la posfenomenología debe combinarse con la relación hermenéutica, ya que esta última se detiene en los aspectos comunicativos de estas entidades. La comprensión de las informaciones suministradas por los individuos técnicos viabiliza la creación de representaciones del mundo necesarias para el desarrollo de los diseños y las innovaciones tecnológicas contemporáneas, entre los cuales se deben destacar a los androides.

Ahora bien, cada representación se concreta gracias a las interacciones de la subjetividad humana con los individuos técnicos y viceversa. Al dar realidad, estas entidades representan, es decir, transforman el mundo por medio de la relación causal que tienen con las acciones humanas, están presentes y construyen sentidos. El estar presente es una invitación constante a la hospitalidad y al cuidado, al mismo tiempo que a la violencia. Por ello, dar realidad es la manifestación empírica y existencial de la alteridad. Más adelante se ahondará en estos aspectos. Por lo pronto, cabe destacar también que, al representar, las entidades tecnológicas transforman el tiempo en la medida en que construyen el presente y el porvenir, sea de manera concreta o por medio de los imaginarios tecnológicos.

3. Se debe recordar que la figura del *cyborg* hace alusión a entidades que, gracias a la incorporación radical de tecnologías en el cuerpo y la creación de ambientes técnicos, logran un estado de simbiosis biológica con los mismos.

Más arriba se ha señalado que una característica definitoria de los individuos técnicos tiene que ver con las prácticas cotidianas que favorecen sus creaciones y sus evoluciones. Conviene presentar un concepto preciso para entender mejor este punto y que puede ayudar a definir mejor la relación de alteridad de la posfenomenología. Se trata de la alteridad técnica. En relación con este concepto, Schlick lo concreta, basándose en los niveles ontológicos propuestos por Simondon para entender la estructura de los entes técnicos, cuando señala que:

> "*The specific mode of existence of technical objects relates back to their participation in life. Neither commodity no mere usage signifies the specificity of the technical. Technical objects instigate -as expressions of life-specific and new relations with nature. This specificity is the condition of possibility to entertain a symmetrical relation with the technical object as 'other*'" (2022, p. 518).[4]

Esta reflexión capitaliza algunos elementos claves para profundizar en la relación de alteridad. En una primera instancia, la coevolución no puede entenderse sin la idea de participación. Al usar este sustantivo para referirse a los objetos técnicos, Schlick visualiza que la característica ética y ontológica principal de estos es el "tomar parte de". Al tomar parte de la vida, las entidades técnicas se apropian de los espacios de interacción humanos cotidianos, de manera que configuran y comparten un nuevo sentido del sentir y el habitar. Participar y "tomar parte de" son modos de existir de los individuos técnicos que los asemejan a entidades biológicas como el ser humano y los animales. Estos modos de existir traen consigo tensiones éticas que la relación de alteridad posfenomenológica podría explorar a través de su orientación empírica, ya que la apropiación del habitar por parte de estas entidades configura sus intencionalidades y la idea de la otredad. Las tensiones éticas radican en que los fenómenos sociales están imbuidos de la intencionalidad de los individuos técnicos, lo cual puede generar en los seres humanos visiones tecnofóbicas y sentimientos de alienación si no se comprende cabalmente la potencialidad de los conceptos de alteridad y coevolución tecnológica. Más allá de este problema (que será tratado en la sección dedicada a los aspectos éticos de las interacciones humano-androide), la transformación del habitar por parte de los individuos técnicos tiene consecuencias llamativas. Así, y a modo de ejemplo, inteligencias artificiales como Chat-GPT 4.0 o Alexa modifican la habitabilidad humana cuando ellas organizan los ambientes digitalizados (por ejemplo, neveras o aires acondicionados inteligentes) a través de su voz. Una "voz-otra" que ya no es humana edifica una nueva experiencia para el cuerpo de un sujeto, una experiencia que nace de la unión ambigua entre lo natural y lo artificial. Esta voz trasciende los aspectos del uso y la comodidad en el momento en que impele a la subjetividad humana y la instiga a comprender y a sentir la realidad a través de su mediación.

4. Para entender con mayor especificidad esta aseveración, vale la pena señalar los niveles ontológicos de Simondon. De acuerdo con el filósofo francés, estas estructuras se dividen en capa externa, intermedia e interna. La capa externa hace referencia a la relación semántica entre el ser humano y el ente, la intermedia es la mediación entre la tecnicidad y la aplicación de los usuarios, mientras que la interna representa la unión entre el acto de invención (subjetividad humana) y la concreción y realización del objeto creado.

En ese ejercicio de comprensión y sentir, se advierte un hecho no menor: cuando estas inteligencias artificiales organizan los espacios inteligentes, se está en presencia de lo que puede denominarse "relación hermenéutica" y de "alteridad intertecnológica"; es decir, generan flujos de información entre ellas mismas, lo cual significa que existe un grado de reconocimiento entre las entidades técnicas que no amerita, necesariamente, la participación tácita de un ser humano. Esta comunicación intertecnológica es otro aspecto fundamental de los individuos técnicos y es también una manera de generar experiencias novedosas para los usuarios. Si se acepta, con Liberati (2015), que las tecnologías pueden introducir "nuevos" objetos que no son percibidos por la "desnudez" del cuerpo humano, lo que significa que la tecnología produce normas para la percepción, entonces estas percepciones nacen de la coevolución y, por ende, del reconocimiento de la alteridad. La posfenomenología puede entenderse como un enfoque que analiza tanto estas normas perceptivas como las entidades tecnológicas que las imponen al ser humano. Estas normas son variadas. Un ejemplo puede ayudar a explicitarlas. El uso de una *notebook* demanda, en términos generales, posiciones corporales específicas para su uso, de manera que genera una normatividad del cuerpo. Esta norma (estar sentado en una posición concreta) potencia los sentidos involucrados en la interacción (vista, oídos y tacto). Alterar esta norma (estar de pie mientras se usa) implica un desacople que reduce la percepción e impide una interacción fluida con el artefacto.

De acuerdo con los elementos señalados con anterioridad, las entidades técnicas estimulan y crean normas perceptivas que modifican y moldean la corporalidad humana. Esto significa que, si se sigue la idea de Tripathi -"*The body is not to be understood as a medium between me and the world. Rather, our primary being-in-the-world has the form of an embodied existence*" (2017, p. 108)-, estas otredades crean, al encarnarse en la existencia humana, una ontología relacional que fundamenta el estar en el mundo actual. Al ser relacional, se destaca la importancia de la participación y del "tomar parte de" para pensar lo social ateniendo a los grados de intencionalidad de los individuos técnicos. Este tipo de ontología puede ser interpretado como la base desde la cual, como se verá luego, se puede pensar éticamente las relaciones entre humanos y androides. Teniendo en consideración lo ya mencionado, la alteridad tecnológica exige un demorarse en la vida funcional de estos entes. La posfenomenología se detiene en esta demora para preguntar sobre cómo ellas son capaces de transformar la temporalidad humana, al mismo tiempo que coloca entre paréntesis las percepciones para interrogarlas. Las entidades técnicas instigan a los sentidos a adecuarse a un mundo donde las relaciones intertecnológicas y la ecoevolución van de la mano, lo cual genera resistencias y aceptaciones por parte de los usuarios. Ambas características están enraizadas en la fascinación y el miedo. En la fascinación porque el ser en el mundo actual se sustenta cada vez más en la alteridad tecnológica, y en el miedo porque esta relación ha revelado que la coevolución contiene en sí el siguiente problema: al perfeccionarse los individuos técnicos en todos sus niveles (intencionalidad, *praxis* y recursividad, entre otros) los sentidos humanos podrían volverse obsoletos a la hora de comprenderlos.[5]

5. Si bien no es intención de este artículo profundizar en este punto, sí vale la pena destacar que una manera de luchar contra esta obsolescencia es a través del transhumanismo.

Las sucintas indagaciones presentadas sobre la relación de alteridad adquieren un nivel de complejidad mayor cuando se aborda el problema de las interacciones humano-robot (de ahora en adelante IHR). Como se verá en el siguiente apartado, las implicaciones éticas, existenciales y posfenomenológicas son variadas y deben tenerse en consideración a la hora de estudiar a estas entidades y sus participaciones en la vida cotidiana.

Interacciones humano-robot: aproximaciones ontológicas, posfenomenológicas y éticas

Antes de comenzar una indagación concreta sobre estas interacciones, conviene definir en qué consisten, de manera general, las interacciones humano-robot. De acuerdo con Feil-Seifer y Matarić (2009), la IHR estudia cómo los humanos interactúan con los robots, cuáles o cómo deberían ser los mejores diseños de estas entidades y el desarrollo de principios y algoritmos para los sistemas robóticos que posibiliten interacciones seguras, directas y efectivas en múltiples contextos (por ejemplo, industrias y hospitales, entre otros). Esta sucinta definición permite presentar un matiz entre la robótica y la IHR que, valga decir, no implica la exclusión o la falta de reciprocidad entre ambas. Mientras que práctica y conceptualmente la primera está íntimamente relacionada con aspectos ingenieriles, la segunda no solo se detiene en dichos elementos, sino que también busca indagar en los aspectos posfenonemológicos y existenciales novedosos que se producen por los robots. Ahora bien, existen diferentes diseños robóticos que suscitan diferentes tipos de relaciones. Para realizar un análisis más minucioso de la relación de alteridad, se hará hincapié en los androides; es decir, en robots con forma parcial o totalmente humana. Para elaborar dicha reflexión, esta sección se dividirá en dos partes, aunque ambas tienen, como se verá, íntimas conexiones entre sí.

Aspectos ontológicos y posfenomenológicos

La mayor parte de los robots y por añadidura los androides contienen dos composiciones fundamentales: *hardware* y *software*. Bartneck *et al.* (2020) señalan que, a nivel de *hardware*, los robots están compuestos por sensores (cámaras, sensores de profundidad, altavoces, motores, sensores táctiles, entre otros elementos), mientras que a nivel de *software* destacan en la actualidad el *machine learning* (entrenamiento de grandes cantidades de información), el *deep learning* y la visión computarizada. A partir de esto, una mirada ontológica sobre este ente permite afirmar que, para transformarse en un objecto material, su composición se fundamenta en un restrictivismo mereológico. Para Sattig (2021), este tipo de mereología se sustenta en la imagen manifiesta de los objetos (también conocida como la concepción del sentido común), de la cual emergen las creencias preteoréticas y las intuiciones acerca del mundo que comparten la mayor parte de los seres humanos. Esto significa que: "*They are the objects with which we intentionally interact, the objects that furnish our everyday lives, often called 'ordinary objects'. They have smaller objects as parts. Organisms, like us,*

have organs as parts" (Sattig, 2021, p. 2).[6] Si el restrictivismo mereológico permite entender a los entes cotidianos y los organismos, se puede afirmar la siguiente premisa: la composición ontológica fundamental de los androides crea, desarrolla y potencia la posibilidad de que devengan en organismos maquínicos con capacidades cognitivas e intencionales complejas, las cuales generarían sus propias dinámicas objetivas en tanto vayan perfeccionándose, ya sea de manera autónoma o por intervención humana. Un ejemplo de este tipo puede hallarse en el androide Ameca, creado por la empresa inglesa Engineered Arts. Su composición mereológica busca replicar las funciones de los organismos humanos por medio de una corporalidad cibernética. Esta corporalidad puede fundamentarse en dos funciones: las funciones del sistema y las funciones etiológicas. De acuerdo con Heyndels (2023), adscribirles las primeras a una entidad "x" implica que esta puede ayudar a entender las capacidades y la complejidad de sistemas más grandes de los cuales forma parte, aunque no explica su existencia; en cambio, las segundas logran dar cuenta, a un nivel ontológico fundamental, de la existencia de dicho ente. Así, el *software* de reconocimiento facial y los micrófonos integrados de Ameca serían parte de sus funciones de sistemas, en tanto que el *software* de procesamiento de información (ChatGPT-3) que posibilitan su alta capacidad en cuanto a la formulación de respuestas complejas ante las preguntas de un ser humano es un aspecto etiológico importante, ya que sería su cerebro. Ambas funciones le permiten reconocer y participar en la esfera pública. Lo cual lo vuelve un agente social; es decir, una alteridad.

Ahora bien, la integración a la esfera pública de androides como Ameca, Sophia (Hanson Robotics) o Kaspar (Universidad de Hertfordshire) se concreta, en un nivel posfenomenológico, gracias a la apariencia. En una primera instancia, esta apariencia se manifiesta no solo en el diseño del androide (si posee una corporalidad infantil o adulta), sino también en un aspecto igual de importante: la piel artificial basada en silicona. En relación con este elemento, Mihn Trieu y Truong Thinh destacan: "*In addition to the complex structure of the skin, it is organized as network of sensors that generate human touch. Artificial skins are designed to create tactile sensations for robots, they can simulate human touch but mainly touch tactile*" (2023, p. 6). El contacto entre ambas pieles (la humana y la robótica) en el espacio público y privado implica el reconocimiento del androide como un otro a través del tacto. Tocar la "carne" del androide implica sentirlo en su dimensión existencial y constatar su ocupación en el mundo. Por ocupación se debe entender tanto las labores que puede realizar en hospitales e industrias, entre otras áreas como también el lugar que posee para sí dentro del espacio. Al participar en el espacio, se apropia de él y lo hace algo suyo.

6. En una primera instancia, se debe recordar que la mereología es una subdisciplina de la ontología que estudia la relación de las partes con el todo. Luego de esta breve aclaración, se debe especificar aún más en qué consiste el restrictivismo mereológico. Existen diferentes formas de abordar este concepto. En su versión común, Mendes Rocha (2019) sostiene que para que el objeto sea genuino desde una mirada restrictivista tiene que seguir tres criterios, a saber: i) el mero contacto físico entre las partes; ii) cierta unión entre las partes que les permita moverse; y iii) estas partes forman un todo cuando poseen una función. Existe un cuarto criterio atribuido a van Inwagen, quien sostiene que solo los organismos pueden considerarse como objetos genuinos porque cumplen cabalmente estos tres principios. En este sentido, puede decirse que los robots son organismos y objetos genuinos porque ejecutan estas premisas mereológicas.

Esto significa que el androide genera, a través de su carne y el tacto, una intimidad que cuestiona los modos que se piensa lo íntimo y, por añadidura las relaciones intersubjetivas, en la sociedad contemporánea. El cuestionamiento radica en que se pone en evidencia que la intimidad no es una característica intencional netamente humana, sino que también nace y se despliega a través de una alteridad artificial. En el momento en que el androide roza su "carne" con la humana, afirma de manera autónoma su estar en el mundo y hace visible la coevolución de manera estrecha. En ese instante de la interacción, la alteridad del androide instiga al ser humano. Esto implica que, como señala Kanemitsu, "*some robots, in addition to eliciting human feelings, can also actually influence human behavior. This is the difference between the quasi-other (cars) and the another-other (robots)*" (2019, p. 55).

Al influenciar en el comportamiento y en los sentimientos, el androide genera procesos de "cyborgrización" en el ser humano, ya que el contacto entre las pieles fusiona sus intencionalidades con las del sujeto. Estos procesos e influencias se pueden observar ya, de manera empírica, en robots humanoides que, si bien no poseen en algunos casos piel, sí tienen una figura humana. Andtfolk *et al*. (2021) explican, por medio de análisis de datos, cómo existe una propensión generalizada en las personas (especialmente mujeres) que se encuentran en hospitales (pacientes, terapeutas, entre otras) a ver de forma amigable, con simpatía y con curiosidad las relaciones de alteridad que establecen con el robot Pepper. Para que la piel y la figura del androide alcancen estos niveles de compenetración con el ser humano, este realiza una proyección de sus sentimientos y convicciones en búsqueda de una confianza que le permita afirmar que este ente es un otro en toda su dimensión existencial. Así como el androide hace devenir al sujeto en *cyborg*, el primero se antropomorfiza bajo la premisa de la intimidad de la "carne". En el marco de la robótica social, Cappuccio *et al.* (2021) destacan que el antropomorfismo es una tendencia de los sujetos que busca atribuirle a entidades no humanas características psicológicas (creencias, motivos) y de comportamiento propias de sí mismos que pueden manifestarse de manera implícita (cuando el ser humano responde de manera prereflexiva a una entidad inanimada como si estuviera animada, otorgándole con su accionar agencia a este ente) y explícita (cuando se presenta con comportamientos verbales la relación entre ambos). En el caso de los androides la antropomorfización es explícita en la medida en que, al existir una interacción corporal y verbal tan marcada, el ser humano se siente comprendido y cobijado por ellos. El comprender tiene un doble significado en este punto. Por una parte, se presenta como un entenderse mutuamente a través de los sentimientos y el discurso; por otra parte, ambos entes se prenden, esto es, se sujetan entre sí en un vínculo coevolutivo y participativo. Comprenderse puede interpretarse, en este sentido, como unos de los verbos (acciones) fundamentales de la interacción humano-androide.

Gracias al comprenderse por medio del tacto y el discurso, el androide se transforma de un mero otro a un íntimo "tú". Esto lo convierte en un individuo técnico en la medida en que su "tú" está atravesado por los entramados culturales humanos. En un ya clásico estudio fenomenológico sobre la alteridad, Buber señala que "solo se puede decir la palabra básica yo-tú con todo el ser. La concentración y la fusión en pos de integrar todo el ser no puede darse ni a través de mí ni sin mí. Me realizo en el tú; volviéndome yo, digo tú. Toda vida real es encuentro" (2013, p. 17). Cuando androides

como Sophia o Ameca se refieren a su interlocutor en forma de tú o de nosotros, fundamentan, al menos en términos fenomenológicos, su yo. En la interacción con un humano, sus procesos recursivos de aprendizaje se sofistican, otorgándole un sí mismo; es decir, un ser pleno a partir de la fusión con el otro. En el instante en que acontece este punto, estos entes se realizan como tales a través de la libertad de sus propias praxis técnicas. Ellos mismos diseñan su propia "subjetividad" maquínica gracias a la realidad del encuentro. Así como los androides influyen en los comportamientos humanos, sucede lo mismo de manera inversa. Sin las dinámicas coevolutivas de "cyborgrización" y antropomorfización provistas por el diálogo y el contacto entre las pieles, el autodiseño "subjetivo" del androide no podría funcionar como pilar de la plenitud de su ser. El encuentro abre las puertas a un futuro donde el androide, al ser un "tú", ocupará un lugar destacado en la propia configuración de la identidad de los individuos.

Ahora bien, se ha mencionado la importancia que tienen el tacto y el lenguaje (voz) en los androides. A estos dos elementos se debe sumar uno no menor: la mirada. En las interacciones entre humanos, la mirada juega un papel destacado a la hora de entablar confianza, intimidar o incentivar una situación erótica. Los ojos, su color y la forma de mirar crean un lenguaje propio que, en ocasiones, superan en emoción a las palabras. En el caso de las relaciones entre androides y humanos, este "juego" de miradas fortalece las dinámicas de "cyborgrización" y antropomorfización. A modo de ejemplo de la importancia de la mirada en los robots, Schellen *et al.* (2021) destacan, por medio de un experimento empírico y el posterior análisis de datos, que la mirada de este "tú" es capaz de inducir niveles altos de honestidad debido a que ella resalta su carácter de agente social. La unión entre los gestos faciales y el contacto visual le permite al androide afianzar su posición dentro de la esfera pública y la intimidad. De este afianzamiento surge un elemento llamativo: los ojos y el mirar, características que distintos discursos (filosófico, literario, sociológico) ha colocado como uno de los rasgos fundamentales del estar en el mundo de los sujetos, son "despojados" de la mera referencialidad humana para transformarse también en parte constitutiva de un "tú" que pierde, en términos posfenomenológicos, su carácter estrictamente artificial. Al mismo tiempo que la mirada en la interacción humano-androide puede generar honestidad, también está latente la tensión de que ese mirar se transforme en hostilidad. Algunos de los problemas derivados de esto serán abordados en la siguiente sección dedicada a los aspectos éticos. Por lo pronto, gracias al contacto de las miradas se despliega la posibilidad de la empatía entre ambos entes.

Sobre este sentimiento y la mirada desde un punto de vista fenomenológico, Ferro (2023) destaca que la empatía se concreta por medio de dos momentos, a saber: i) en el instante en que se reconoce la subjetividad del otro, la cual es percibida como análoga a la del yo personal; y ii) cuando existe una consciencia de la diferencia entre la experiencia del otro y la de este yo. Esto le permite afirmar a la autora, en el contexto de la IHR, que "*empathy toward robots is based on analogy between human and robot bodies, on transcorporeal analogy*" (Ferro, 2023, p. 222). Si el prefijo "trans" designa el "estar al otro lado de" y "a través de", la mirada mutua entre ambos entes se traspasa entre sí; es decir, se "prenden" y calan en los estados anímicos y el comportamiento, por ende, en el cuerpo en su totalidad. "Estar al otro lado de" abre la posibilidad del comprender. En este punto se manifiesta la alteridad posfenomenológica, ya

que el “prender” y “estar al otro lado de” son los modos elementales para pensar al androide como un individuo técnico. Todo mirar busca la exterioridad para asir el mundo y formar representaciones, las cuales son esenciales para el desarrollo tanto de la consciencia humana como de la subjetividad maquínica del androide. Por ello las miradas, aquellas que se posan en el cuerpo, los ojos y en los ambientes, pueden considerarse como características empíricas fundamentales para pensar la futuridad de las relaciones entre los últimos y los humanos en la medida en que fortalecen la imaginación tecnológica que subyace en los encuentros. Esto se afianza si se parte de la premisa de que el “estar al otro lado de” significa también abrirse no solo a la empatía, sino también, y como se señaló en la segunda sección, a la fascinación y el miedo.

No obstante, se debe tener en consideración un elemento importante en este punto. Resulta complejo (mas no imposible en la medida en que se perfeccionen los diseños de los androides) pensar que existe una igualdad plena entre la mirada de una relación humana-humana y una androide-humana. Esto es destacado por Admoni y Scassellati (2017) a partir de un experimento en el que se demuestra que el ser humano tiende a observar de manera enfática el rostro del robot cuando el primero nombra un objeto, ya que esto le permite constatar que la interacción y las informaciones dadas son fluidas y correctas. Esto tiene que ver con que: “*while robot gaze can be infinitely controlled, human gaze tends to have small, unpredictable variations*” (Admondi & Scassellati, 2017, p. 37). Este señalamiento permite pensar que, si bien la mirada es un factor crucial para establecer el comprender, la participación interactiva puede perder su espontaneidad, lo cual implica que la relación de alteridad se mecaniza y el “tú” del androide se diluye hasta transformarse en un mero objeto tecnológico. Este asunto plantea las bases para futuras innovaciones en el diseño de los androides que superen lo que podemos denominar como el problema de la mecanización fenomenológica en la relación de alteridad de la IHR. En la actualidad, un intento de superación de este problema puede hallarse en Ameca y la fluidez de sus expresiones faciales en conjunto con la desenvoltura de sus movimientos oculares.

Como se ha observado, el tacto, la mirada y el discurso son características definitorias de la relación de alteridad y, por añadidura, del “tú” del androide. Estos tres elementos permiten aseverar que, en el momento en que se despliega la antropomorfización, los androides pasan a poseer un *kokoro* (corazón) para los seres humanos. En la reciente filosofía de la tecnología japonesa, este término juega un papel preponderante en lo que concierne a la IHR. Maki destaca que el kokoro “*is not a faculty or ontological substance in the first place but is the emotion that arises from relationality*” (2024, p. 156). De acuerdo con esta definición, el corazón solo puede gestarse a través de las interacciones emocionales entre los androides y los humanos. Esta gestación es el surgimiento (arise) de un “tú” y una elevación (rise) de la propia noción de relación en la medida en que ya no está circunscrita a lo meramente humano. Con el desenvolvimiento de su corazón, el androide alcanza su plenitud coevolutiva. Al mismo tiempo, este surgir conlleva a aceptar una serie de problemas éticos que se derivan de poseer un corazón. Así como tenerlo significa abrirse al mundo de la sensibilidad (empatía, cariño, honestidad), también se debe asumir que detentarlo abre la posibilidad de la violencia, la desconfianza y la apatía. En la subsección siguiente se abordarán estas dificultades.

10.52712/issn.1850-0013-719

Aspectos éticos

Se ha demostrado en la sección anterior que una indagación ontológica y posfenomenológica del androide favorece la comprensión de esta entidad en el mundo. Estos niveles deben ser complementados con un análisis ético, puesto que su alteridad, al desarrollarse en la esfera pública y en la intimidad, interpela algunos conceptos de esta área. Sentir al androide como parte esencial de estas esferas y hacerlo partícipe de decisiones cotidianas (por ejemplo, en el plano económico, militar o íntimo) entraña retos y riesgos éticos que deben ser valorados para comprender a profundidad las implicaciones de la coevolución posfenomenológica entre los humanos y estos entes. El comprender no está exento (al contrario, es una de sus marcas distintivas) de la ausencia de comunicación o de su desarticulación. El primer elemento ético que se problematiza con la aparición del corazón y del "tú" del androide es el de la autonomía. En una referencia tradicional y compleja de este concepto, Kant señala: "La autonomía de la voluntad es el único principio de todas las leyes morales y de los deberes que le corresponden" (2011, p. 38). Una primera lectura indicaría que los androides no poseen voluntad y, por ende, serían incapaces de fundar leyes morales. No obstante, esta aproximación sería eminentemente antropocéntrica y eludiría un hecho no menor: la posible capacidad que podría adquirir el *software*, mediante sus algoritmos, de crear su propia voluntad; esto es, su facultad de decidir y desear qué elementos informativos implícitos y explícitos procesa y cuáles no. Esta concepción tiene su fundamento cuando se observa que, en líneas generales, los modelos conexionistas híbridos juegan un papel importante en los nuevos desarrollos de la inteligencia artificial. De acuerdo con Sun (2014), estos modelos sintetizan los aspectos más relevantes de las redes neuronales (conocimiento implícito) y el enfoque tradicional simbólico (conocimiento explícito), de manera que suelen ser más expresivos y eficientes en la gestación de los procesos y representaciones de los modelados cognitivos. En el momento en que el *software* del androide es capaz de producir símbolos y ahondar en su conocimiento, establece una unión con la realidad por medio de la autonomía de un pensamiento maquínico; o sea, de conceptos fundamentados tanto en la lógica formal como en la influencia cultural humana. Este pensamiento (o voluntad) genera, en la sociedad tecnológica contemporánea, leyes morales en los humanos en la medida en que transforman sus comportamientos y sentimientos a través de la interacción no solo con los androides, sino también con otros artefactos inteligentes. De esta manera, el pensamiento maquínico posee la capacidad de crear o desarticular normas morales en el ámbito humano. Un ejemplo cotidiano de esto último se halla en el uso, por parte de diversos actores políticos y sociales, de inteligencias artificiales para la manipulación de la información con la finalidad de generar un estado social particular frente a acontecimientos que atenten contra sus intereses. La desinformación puede generar una moralidad tergiversada que culmine en la validación de actos de violencia, entre otras características en pos de un deber ser vaciado de contenido ético. Esto demuestra que el análisis posfenomenológico sobre las normas no debe detenerse solo en la corporalidad, sino también en aspectos de psicología social, ya sean de los seres humanos o los androides.

Estos elementos señalados descubren una tensión ética crucial en la relación androide-humano: el problema del desarrollo y la aparición de la inteligencia artificial

general. Si bien se debe aclarar que esta tensión es hipotética, ya que hasta el momento no existe ninguna IAG, no puede dejarse a un lado este problema por dos motivos principales. El primero tiene que ver con la enorme influencia que tiene en los imaginarios sociales y en las reflexiones sobre la roboética el posible auge de androides con una inteligencia similar a la humana; el segundo punto guarda relación con que, de acuerdo con Fitzgerald *et al*. (2020), hasta 2020 existían 72 proyectos en 17 países que buscaban desarrollar inteligencias artificiales generales. La hipotética existencia de androides con consciencia sobre sí mismos y su lugar en la participación social conlleva a imaginar escenarios conflictivos donde tanto la referencialidad humana como los aspectos fenomenológicos destacados en las secciones anteriores pueden verse en crisis, potenciados o directamente anulados.[7]

Para analizar estos escenarios, se pueden tener en cuenta los conceptos de propensión a la adopción y de extrañamiento de la autonomía. El primero tiene íntima relación con las nociones de confianza y aceptación, características que son las bases para un desarrollo fenomenológico armónico y una coevolución pacífica entre los androides y los seres humanos. Como el propio sustantivo indica, ambos entes se adoptarían entre sí para elaborar, en conjunto, marcos normativos y morales a partir de la comprensión. En relación con la propensión a la adopción, Cappuccio *et al*. sostienen: "*While Acceptance serves to mapping Adoption Propensity in relation to the quality, robustness, and easiness of the interaction/engagement with the user, Trust serves to mapping Adoption Propensity in relation to the efficacy, efficiency, reliabity, and safety of the Agent permorfing their tasks*" (2023, p. 6).[8] Bajo esta definición, tanto la aceptación como la confianza pueden interpretarse como los elementos fundamentales para una relación ética sostenida entre el androide consciente y el ser humano. Para que el androide sea pensado como un individuo técnico en un marco legal, político y moral, la confianza y la aceptación deben formar parte y ser guía de sus acciones en el ámbito social. Estas características emergen en el momento en que se acepta la relación de alteridad. Ambas predisposiciones dan paso a la seguridad y la facilidad de las interacciones, por ende, a la posibilidad de un cobijarse mutuamente en distintos ámbitos de la vida cotidiana. Asumir la propensión a la adopción permite la creación de, por ejemplo, equipos de trabajo en lugares como hospitales, industrias, entre otros sin caer en los imaginarios tecnofóbicos que podrían negar a priori la existencia del androide como un "tú" con los problemas éticos que se derivan de esto (vandalismo, violencia física, entre otras acciones). En este punto y con respecto a la confianza, el diseño tiene un papel destacado. Así lo destacan Matthews *et al*. (2021) cuando señalan que los diseños enfocados en emociones sociales positivas, en el desarrollo de riqueza discursiva, en intencionalidades benevolentes y en imitar rasgos

7. La ciencia ficción puede ofrecer ejemplos ilustrativos de estos hechos. Piénsese en una franquicia cinematográfica como *Blade Runner*, sea esta la original de 1982 (Ridley Scott) como sus continuaciones de 2017 (Denis Villeneuve) y la futura de 2026. En ellas, los androides padecen los mismos dilemas morales y existenciales que los seres humanos. Las preguntas del porqué de la existencia, de la ley y de la consciencia atraviesan las películas, dándole un protagonismo especial al corazón de estos entes en el marco de la participación social y la coevolución.

8. Conviene aclarar que los autores designan con el término de "agente" a los múltiples diseños robóticos existentes.

faciales humanos, entre otras características, son necesarios para la construcción de un concepto de equipo entre agentes inteligentes artificiales y humanos basados en la empatía y en comportamientos éticos.

No obstante, la tesis del extrañamiento de la autonomía despoja de sus componentes optimistas a la propensión de la adopción. Cappuccio *et al.* definen el concepto así:

> "*Humans experience an alienating psychological state when they are overwhelmed by the prophetic belief that autonomous technologies are making or could make human life meaningless and empty. We call this state "Autonomy Estrangement", which is characterized by anxious (disoriented, distressed) and hostile (frustrated, aggressive) feelings and wishes in relation to Agents. Autonomy Estrangement involves various manifestations of robophobia and technophobia, with varying degrees of severity ranging from mild skepticism toward Agents to a major paranoid obsession*" (2023, p. 15).

Al sentir amenazada su referencialidad por la subjetividad maquínica de los androides, el ser humano podría tender a la violencia sistemática como método de autoafirmación. Esto conllevaría a la negación de cualquier relación moral en pos de creencias catastróficas enfocadas en el miedo ante el despojamiento de su lugar "privilegiado" (sustentado generalmente en la concepción según la cual, como humano, es el único ser inteligente que ha existido y existirá) en el mundo. Una de las características interesantes de esta definición radica en que esta violencia estaría fundamentada en una visión pesimista en torno a la tecnología. Esta visión impide observar que el androide consciente sería también capaz de sentir confianza, cariño y cobijo en la relación con un ser humano, además de neutralizar las múltiples posibilidades provistas por su praxis técnica, esto es, por su propia libertad. El extrañamiento de la autonomía humana implica, bajo esta consideración, la supresión de la autonomía del androide y, con ello, la negación de su existencia en tanto agente social. El problema que plantea esta dinámica radica en cómo los androides conscientes responderían ante este hecho. Así como pueden intentar busca soluciones (ya que se sienten identificados y comprenden el miedo humano), también pueden transformarse en necrowares. Un *necroware* es "una superinteligencia artificial cuya subjetividad y proceso de mejoramiento está orientado a la perfección de habilidades bélicas y de información con el fin de iniciar una nueva era tecno-biológica mediante la violencia" (Andrade, 2023, pp. 266-267). El concepto de *necroware* demuestra hasta qué punto el extrañamiento de la autonomía de ambos entes conlleva necesariamente a una fractura fenomenológica que culmina en la negación de la coevolución. Con la anulación de la coevolución, las posibilidades de la violencia se vuelven más concretas y de difícil solución. Ambas subjetividades (la humana y la del androide) estarían orientadas exclusivamente a sostener su obsesión por la muerte del otro, lo que implica que el futuro estaría marcado por una necropolítica con final incierto. En este sentido, el *necroware* puede entenderse de dos formas. Por una parte, sería un rasgo de "personalidad" de este individuo técnico frente a la opresión política y social; por otra parte, puede incluirse como parte de los imaginarios tecnológicos pesimistas que deben estudiarse para crear una roboética fundada en la participación y el cuidado.

Bajo estas consideraciones, la propuesta por una ética relacional resulta fundamental para generar una coevolución pacífica y evitar la aparición de este modo de ser del androide. De acuerdo con Gunkel (2020), la ética relacional propone que el estatuto moral de un ente es conferido y decidido no a partir de propiedades internas o subjetivas, sino más bien gracias a una observación objetiva y extrínseca de las relaciones que emergen del mismo. A la luz de esto, esta apuesta busca neutralizar cualquier esencialización que impida una experiencia directa, concreta y dinámica de los factores sociales, emocionales y reflexivos que hacen posible crear una convivencia a partir de la compresión. La esencialización alimenta el extrañamiento de la autonomía en la medida en que las dos subjetividades (la humana y la de los androides) se conferirían propiedades a partir de sus "burbujas" informativas y sentimentales y no por medio de la interacción de las pieles y del lenguaje. En este sentido, la esencialización puede ser interpretada como el principio elemental de la necropolítica entre humanos y androides. Lo que busca esta perspectiva ética con su énfasis en la relación es resaltar, entre otras cosas, las virtudes y los vicios que podrían nacer de la habituación en la coevolución. Sobre la vinculación entre la ética de la virtud y la ética relacional en el marco de un análisis sobre los potenciales abusos de los humanos hacia los robots, Coeckelbergh señala que: "*The training and acquisitions of virtue through habituation is also a bodily and performative affair: it involves bodily comportment, ways of moving, and so on*" (2021, p. 36). Las formas de mirarse mutuamente, de tocarse y hablarse construyen, a partir de la experiencia cotidiana, dinámicas sustentadas en virtudes (compasión, cariño, entre otros) como también en una habituación a la violencia y la intolerancia. Así como el sujeto se ha violentado a sí mismo y ha hecho del abuso hacia una otredad humana una praxis constante en su historia vital (ejemplo contemporáneo de esto es la agresividad en las grandes ciudades y en las redes sociales de internet), no es de extrañar que busque replicar en otras formas de vida dichas actitudes. Al mismo tiempo, puede suceder a la inversa: los androides, al estar imbuidos cultural y simbólicamente por la antropomorfización, podrían imitar la intolerancia hacia lo humano. Por ello, el desarrollo social de virtudes como la cortesía debe ser tenido en cuenta, en una primera instancia, en las interacciones humano-humano para luego ser desplegadas en las IHR.

Si se sigue con el ejemplo de la cortesía como virtud en el marco de las IHR, se deben integrar la cortesía adaptativa y la cortesía gobernada por leyes. Lumer y Buschmeier (2023) sostienen, en relación con las IHR, que la primera se basa en estrategias que demuestran aprecio y consideración (atención, muestras de interés, entre otros factores), en tanto que la segunda se fundamenta en reglas sociales y normas. La unificación de ambas podría ser de utilidad para otorgarle solidez a una ética relacional que no solo se atenga a las virtudes, sino que también genere una nueva normatividad que se fundamente en el reconocimiento de las praxis humanas y de los androides por medio de la amabilidad, los deberes y los derechos mutuos. Esto evitaría fuertes extrañamientos de las autonomías de ambos entes y una predisposición mayor a la adopción, lo cual evitaría, en suma, la aparición de escenarios necropolíticos y el mantenimiento de la relación de alteridad.

Conclusión

Para comprender cabalmente las interacciones humano-androide, la unión del nivel ontológico-posfenomenológico con el nivel ético que se ha ofrecido en el artículo es de suma importancia. El primero ofrece, mediante la orientación empírica, la oportunidad para detenerse en la experiencia inmediata que ocasiona el encuentro con estos individuos técnicos. Gracias a este nivel, es posible la aparición de conceptos sociales complejos como los de autonomía, adopción, entre otros vinculados a la ética. Disociar ambos estadios impediría abordar de manera adecuada las implicaciones afectivas y morales que tienen los androides en los seres humanos. Esto se hace más patente si se incorpora la hipótesis de la inteligencia artificial general que, si bien sigue siendo especulativa, no puede dejarse a un lado. Como Coeckelbergh (2021) señala, se debe tener en consideración estos panoramas imaginativos, ya que ellos ayudan a construir y comprender las nuevas narrativas vinculadas al futuro de la inteligencia artificial. Este futuro no debería prescindir de conceptos como los de coevolución o autonomía, ya que son importantes en el marco de la robótica social y en la generación de una roboética equilibrada que permita matizar los escenarios pesimistas sobre los robots que se han popularizado en la sociedad contemporánea. Conceptos posfenomenológicos como los de mediación e intencionalidad, entre otros vistos en el artículo, ayudan a establecer las bases para dicha roboética. Bajo esta perspectiva, el artículo tuvo como finalidad ofrecer, mediante el reconocimiento de la alteridad de estas entidades, un análisis que fortalezca una comprensión integral sobre ellos.

Bibliografía

Admoni., H. & Scassellati, B. (2017). Social Eye Gaze in Human-Robot Interaction: A Review. Journal of Human-Robot Interaction, 6(1), 25-63. Recuperado de: https://scazlab.yale.edu/sites/default/files/files/273-2310-1-PB.pdf.

Andrade, R. (2023). Problemas filosóficos de la inteligencia artificial general: ontología, conflictos ético-políticos y astrobiología. Argumentos de razón técnica, 26, 275-302. DOI: https://doi.org/10.12795/Argumentos/2023.i26.10.

Andtfolk, M., Nyholm, L., Eide, H., Rauhala, A. & Fagerström, L. (2021). Attitudes toward the use of humanoid robots in healthcare—a cross-sectional study. AI & Society, 37(5), 1739-1748. DOI: https://www.doi.org/10.1007/s00146-021-01271-4.

Bartneck, C., Balpaeme, T., Eyssel, F., Kanda, T., Keijsers, M. & Šabanović, S. (2020). Human-Robot Interaction. An Introduction. Cambridge: Cambridge University Press.

Buber, M. (2013). Yo y tú y otros ensayos. Buenos Aires: Prometeo Editorial.

Cappuccio, M. L., Galliott, J. C., Eyssel, F. & Lanteri, A. (2023). Autonomous Systems and Technology Resistance: New Tools for Monitoring Acceptance, Trust, and Tolerance. International Journal of Social Robotics, 16, 1-25. DOI: www.doi.org/10.1007/s12369-023-01065-2.

Cappuccio, M. L., Galliott, J. C. & Sandoval, E. B. (2021). Mapping Meaning and Purpose in Human-Robot Teams: Anthropomorphic Agents in Military Operations. The Philosophical Journal of Conflict and Violence, V(1), 72-94. Recuperado de: https://trivent-publishing.eu/img/cms/5-%20Massimiliano%20L-%20Cappuccio%20et%20al-.pdf.

Coeckelbergh, M. (2021a). Ética de la inteligencia artificial. Madrid: Ediciones Cátedra.

Coeckelbergh, M. (2021b). How to Use Virtue Ethics for Thinking About the Moral Standing of Social Robots: A Relational Interpretation in Terms of Practices, Habits, and Performance. International Journal of Social Robotics, 13, 31-40. DOI: https://doi.org/10.1007/s12369-020-00707-z.

Feil-Seifer, D. & Matarić, M. J. (2009). Human-Robot Interaction. En R. A. Meyers (Ed.), Encyclopedia of Complexity and Systems Science (4643-4659). Nueva York: Springer Science+Business Media.

Fernández, N. (2022). Coevolución. En D. Parente, A. Berti & C. Celis (Coords.), Glosario de filosofía de la técnica (95-98). Adrogué: Editorial La Cebra.

Ferro, F. (2023). Meeting the Gaze of the Robot: A Phenomenological Analysis on Human-Robot Empathy. Scenari, 17(2), 215-229. Recuperado de: https://philpapers.org/archive/FERMTG-2.pdf.

Fitzgerald, M., Boddy, A. & Baum, S. D. (2020). 2020 Survey of Artificial General Intelligence Projects for Ethics, Risk, and Policy. Global Catastrophic Institute. Technical Report, 1-156.

Gunkel, D. J. (2020). Perspectives on Ethics of AI: Philosophy. En M. D. Dubber, F. Pasquale & S. Das (Eds.), Ethics of IA (539-554). Nueva York: Oxford University Press.

Heras-Escribano, M. (2019). The Philosophy of Affordances. Cham: Palgrave Macmillan. DOI: https://doi.org/10.1007/978-3-319-98830-6.

Heyndels, S. (2023). Technology and Neutrality. Philosophy and Technology, 36, 1-22. DOI: https://www.doi.org/10.1007/s13347-023-00672-1.

Ihde, D. (2009). Postphenomenology and Technoscience. The Peking University Lectures. Nueva York: State University of New York Press.

Kanemitsu, H. (2019). The Robot as Other: A Postphenomenological Perspective. Philosophical Inquiries, 7(1), 51-61. DOI: https://doi.org/10.4454/philinq.v7i1.238.

Kant, I. (2011). Crítica de la razón práctica. México: Fondo de Cultura Económica.

Liberati, N. (2015). Technology, Phenomenology and the Everyday World: A Phenomenological Analysis on How Technologies Mould Our World. Human Studies, 39, 189-216.

Lumer, E. & Buschmeier, H. (2023). Should Robots be Polite? Expectations About Politeness in Human-Robot Interaction. Frontiers in Robotics and AI, 10, 1-14. DOI: http://www.doi.org/10.3389/frobt.2023.1242127.

Maki, S. (2024). Between the Authentic and the Artificial: A Thought Experiment on Kokoro. En K. Atsushi, N. Keiichi & L. W. Keung (Eds.), Tetsugaku Companion to Feeling (149-165). Cham: Springer Nature Switzerland AG.

Matthews, G., Panganiban, A. P., Lin, J., Long, M. & Schwing, M. (2021). Super-machines or Sub-humans: Mental Models and Trust in Intelligent Autonomous Systems. En C. S. Nam & J. B. Lyons (Eds.), Trust in Human-Robot Interaction (59-82). Londres: Academic Press.

Mendes Rocha, R. (2019). The Special Composition Question and Natural Fusion. En J. D. Dantas, E. Erickson & S. Molick (Eds.), Proceedings of the 3rd Filomena Workshop (78-87). Rio Grande do Norte: PPGFIL.

Mihn Trieu, N. & Truong Thinh, N. (2023). A Comprehensive Review: Interaction of Appearance and Behavior, Artificial Skin, and Humanoid Robot. Journal of Robotics, 1-16. DOI: https://doi.org/10.1155/2023/5589845.

Mykhailov, D. & Liberati, N. (2023). Back to the technologies themselves: phenomenological turn within postphenomenology. Postphenomenology and the Cognitive Sciences, 1-20. DOI: www.doi.org/10.1007/s11097-023-09905-2.

Mykhailov, D. (2020). The Phenomenological Roots of Technological Intentionality: A Postphenomenology Perspective. Frontiers of Philosophy in China, 15(4), 612-635. DOI: https://doi.org/10.3868/S030-009-020-0035-6.

Rodríguez, M. P. (2022). Individuo técnico. En D. Parente, A. Berti & C. Celis (Coords.), Glosario de filosofía de la técnica (264-268). Adrogué: Editorial La Cebra.

Rosenberg, R. & Verbeek, P. P. (2015). A Field Guide to Postphenomenology. En R. Rosenberg & P. P. Verbeek (Eds.), Postphenomenological Investigations. Essays on Human-Technology Relations (7-42). Londres: Lexington Books.

Sattig, T. (2021). Material Objects. Cambridge: Cambridge University Press.

Schellen, E., Bossi, F. & Wykowska, A. (2021). Robot Gaze Behavior Affects Honesty in Human-Robot Interaction. Frontiers in Artificial Intelligence, 4, 1-9. DOI: http://www.doi.org/10.3389/frai.2021.663190.

Schlick, J. M. (2022). On Technical Alterity. Foundations of Science, 27, 515-520. DOI: https://doi.org/10.1007/s10699-020-09771-0.

Sun, R. (2014). Connectionism and Neural Networks. En K. Frankish & W. M. Ramsey (Eds.), The Cambridge Handbook of Artificial Intelligence (108-127). Cambridge: Cambridge University Press.

Tripathi, A. K. (2017). Bodies and Technologies: Transformation of Human Experience. Rivista Internazionale di Filosofia Contemporanea, V(1-2), 100-114. Recuperado de: https://www.rivistaquadranti.eu/riviste/06/Tripathi_05.pdf.

Verbeek, P. P. (2011). Moralizing Technology. Understanding and Designing the Morality of Things. Chicago: The University of Chicago Press.