

Máquinas sintientes, sueños co-constitutivos: interacción humano-robots, ética relacional tecnológica y posthumanismo.

Andrade, Ricardo.

Cita:

Andrade, Ricardo (2024). *Máquinas sintientes, sueños co-constitutivos: interacción humano-robots, ética relacional tecnológica y posthumanismo*. III Congreso Internacional de Ciencias Humanas. Escuela de Humanidades, Universidad Nacional de San Martín, Gral. San Martín.

Dirección estable: <https://www.aacademica.org/3.congreso.eh.unsam/356>

ARK: <https://n2t.net/ark:/13683/esz9/Wam>



Esta obra está bajo una licencia de Creative Commons.
Para ver una copia de esta licencia, visite
<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.es>.

Acta Académica es un proyecto académico sin fines de lucro enmarcado en la iniciativa de acceso abierto. Acta Académica fue creado para facilitar a investigadores de todo el mundo el compartir su producción académica. Para crear un perfil gratuitamente o acceder a otros trabajos visite: <https://www.aacademica.org>.

Máquinas sintientes, sueños co-constitutivos: interacción humano-robot, ética relacional tecnológica y posthumanismo

Ricardo Andrade
CITECDE/UNRN/CONICET
jandraderangel@unrn.edu.ar

Resumen ampliado

Palabras clave: Filosofía de la tecnología; Inteligencia artificial general; Roboética; Hibridación agencial; Estudios del futuro

El desarrollo creciente de sistemas inteligentes autónomos en distintos ámbitos de la vida social (por ejemplo, a nivel militar e industrial) han fortalecido la necesidad de pensar en las implicaciones que podrían tener, a nivel filosófico, las apariciones de inteligencias artificiales generales en el futuro. Si bien estas entidades son aún hipotéticas, una de las múltiples tareas de la filosofía de la tecnología es ahondar en los escenarios especulativos en torno a los avances tecnológicos para, de esta manera, preguntar sobre el lugar del ser humano en un mundo donde la agencia técnica está neutralizando, de forma progresiva, el pensamiento antropocéntrico que ha sustentado la modernidad, al menos en Occidente. Estos escenarios adquieren mayor relevancia cuando se piensa en las progresivas avances de la robótica social, especialmente en las interacciones humano-robot. Dos conceptos deben ser definidos antes de continuar. El primero es el de interacción humano-robot (desde ahora IHR). De acuerdo a 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, hospitales, entre otros). El segundo es el de inteligencia artificial general. Bubeck *et al.* Señalan que “Usamos IAG para referirnos a sistemas que demuestran amplias capacidades de inteligencia, incluido el razonamiento, la planificación, y la capacidad de aprender de la experiencia, y con estas capacidades al nivel humano o por encima de él (2023, p. 4).

Estas dos definiciones permiten realizar la siguiente pregunta: ¿cómo establecer interacciones seguras y efectivas con entidades robóticas que poseerían rasgos reflexivos similares a la de los seres humanos? En una primera instancia, se debe asumir que estos entes serían agentes sociales que desempeñarían un papel crucial en el desarrollo de los entramados productivos de la sociedad futura, al mismo tiempo que transformarían las concepciones que se tienen sobre la afectividad, ya que los seres humanos deberán lidiar con las emociones que se despliegan de los robots, sean estas negativas o positivas. Al tener una fuerte influencia en la producción y en los estados anímicos, los robots se transformarían en *co-actores*, lo cual significa que sus propias experiencias, intencionalidades y deseos los llevarían a construir su propia noción de futuro. En este punto se presenta un primer foco de tensión, ya que se hace patente que puede existir una pugna en torno a

la *futuridad*. Por una parte, el porvenir pensado e imaginado por el ser humano en donde la visión antropocéntrica de la existencia seguiría teniendo un peso crucial; por otra parte, el futuro concebido por los robots, en donde la existencia humana puede derivar en tres escenarios: el de la total indiferencia, el del mundo compartido bajo la empatía recíproca y el de la progresiva violencia política. Estos escenarios demanda un análisis ético que permita comprender, de manera especulativa, cómo se debe actuar teniendo en consideración un elemento no menor: el de pleno reconocimiento de la alteridad (y por añadidura, la autonomía) del robot.

Dentro de la filosofía de la tecnología contemporánea, la ética relacional ha desempeñado un importante papel para comprender las dinámicas entre los desarrollos tecnológicos de diferentes índoles (artefactos, inteligencia artificial, entre otros) y su impacto en la vida humana. En un ya famoso trabajo sobre roboética y ética relacional, Coeckelbergh (2010) señala que para comprender la dimensión moral de los robots se debe pensar en una ecología social, ya que esta tiene como objetivo analizar las relaciones entre agentes humanos y no humanos que son inter-dependientes y se adaptan entre sí, lo cual hace que las consideraciones morales que nacen de estas relaciones se sustenten principalmente en los contextos situados y en las interacciones concretas. Si se sigue esta apreciación, el vínculo moral entre un humano y un robot no solo se gestaría a partir de una reflexión pormenorizada sobre las intencionalidades de ambos, sino también, y de manera más inmediata, sobre los gestos, las miradas o los “roces” corporales de ambas entidades. Este aspecto fenomenológico es crucial en la medida en que él fundamenta la intimidad o el desencuentro definitivo.