



La heteronomía en los robots sociales asistenciales. Una experiencia piloto con un robot para dar de comer en el hospital

Vallès-Peris, Núria^a, Barrué, Cristian^b, Alenyà, Guillem^b

^a Institut d'Investigació en Intel·ligència Artificial, Consejo Superior de Investigaciones Científicas (IIIA-CSIC), Campus de la UAB, 08193 Bellaterra (Barcelona), email: nuria.valles@iia.csic.es;

^b Institut de Robòtica i Informàtica Industrial, CSIC-UPC, Carrer Llorens i Artigas 4-6, 08028 Barcelona, email: cbarrue@iri.upc.edu; galenya@iri.upc.edu.

Palabras clave: robótica social asistencial, autonomía, heteronomía, estudios de la ciencia y la tecnología

Grupo de trabajo: CT10-Sociología de la salud

1) Introducción: presentación y objetivos

En un contexto de creciente tensión y presión asistencial en los sistemas sanitarios a nivel global (Topol, 2019), la inteligencia artificial y la robótica aplicadas a la salud han recibido un importante impulso. Se han intensificado las investigaciones y desarrollo de la robótica social asistencial, un ámbito de la robótica que trabaja en el diseño de artefactos para acompañar y apoyar en la asistencia y los cuidados. La emergencia de este tipo de artefactos plantea una serie de controversias para el ámbito de la salud, entre ellas el debate sobre el potencial efecto de los robots en nuestra autonomía (Borenstein & Arkin, 2016).

Utilizando el enfoque interdisciplinar de los estudios de la ciencia y la tecnología (STS), en este estudio analizamos de qué manera un robot social (en concreto un robot para dar de comer, MAR), puede contribuir (o no) a mejorar el bienestar y autonomía de los pacientes hospitalizados. A partir de un estudio etnográfico de un robot *in the wild* (fuera del laboratorio) mostramos cómo el efecto de estos artefactos no se explica por su eficiente ejecución de las funcionalidades, sino por lo que llamamos su capacidad de heteronomía.

2) Planteamiento teórico-metodológico

El estudio presentado se basa en la realización de un estudio piloto experimental, llevado a cabo por un equipo multidisciplinar (ingeniería, ciencias sociales y salud), de utilización de un MAR con personas con severos problemas de movilidad. La experiencia piloto se llevó a cabo en el hospital de cuidados intermedios Parc Sanitari

Pere Virgili de Barcelona, un hospital público especializado en geriatría y cuidados paliativos.

Para analizar si el diferente nivel de autonomía del robot condicionaba la valoración o experiencia de los pacientes se diseñaron 3 versiones diferentes del MAR con diferentes capacidades de interacción, en las que el paciente controla el robot a través de gestos (Barrué et al., 2024). Se utilizó el robot con un total de 8 pacientes con graves problemas de movilidad en la parte superior del tronco, para dar un yogur a la hora de la merienda. Cada variación del robot se utilizó con 3-2 pacientes, durante 5 días con cada paciente, utilizándose un total de 40 días. Se realizó una etnografía focalizada (Knoblauch, 2005) de la experiencia piloto, a partir de la recopilación de los siguientes datos:

- Datos de uso del robot durante las comidas
- Compilación manual de datos cuantitativos sobre la interacción del robot con el paciente
- Entrevistas semi estructuradas cara a cara con los pacientes
- Observación no-participante de una merienda con cada paciente

3) Resultados: principales aportaciones, resultados y conclusiones

El concepto de autonomía, al igual que muchos otros conceptos morales, como la justicia o la igualdad, deriva su significado en perspectivas teóricas particulares y a menudo contradictorias. Desde una perspectiva dominante la autonomía se entiende como capacidad de autodeterminación y/o autogobierno (Darwall, 2006). En el ámbito de la robótica la autonomía puede entenderse como la capacidad de un artefacto computarizado para seguir un algoritmo complejo en respuesta a los datos del entorno, independientemente de la intervención humana en tiempo real (Etzioni y Etzioni, 2016). Aunque se refieran a los humanos o a los robots, esta manera de comprender la autonomía se basa en la idea de autonomía individual, es decir, en que las personas o los artefactos tiene independencia, y para que esta sea posible deben estar libres de coacción u otras interferencias. Este es el modelo integrado en la conceptualización de la Interacción Humano-Robot (HRI), basada en una relación diádica entre dos actores racionales y autónomos.

La relación entre la autonomía de los robots sociales y la autonomía de los humanos se ha entendido tradicionalmente desde la bioética y la robótica desde dos enfoques diferentes: (a) como un juego de suma cero: más autonomía para los robots equivale a menos autonomía para los humanos, porque se delegan las decisiones en los robots (Floridi y Cowsls, 2019) o; (b) al contrario, que al mejorar la autonomía de los robots sociales se puede mejorar la autonomía humana, ya que estos artefactos permiten conseguir fines más valiosos, competencias mejoradas y elecciones más auténticas (Formosa, 2021). Nuestra investigación con MARs, sin embargo, muestra que la relación entre los robots sociales y la autonomía es más variada y compleja que lo que sugieren estos enfoques basados en la idea de autonomía individual.

Analizando la autonomía en la relación entre los pacientes y los robots desde la perspectiva de la HRI, los resultados de nuestro estudio reportan tanto una evaluación positiva de la eficiencia de la autonomía a través de los registros de uso del robot, como en la valoración de los pacientes. Los resultados sobre las tres variaciones del robot con diferentes capacidades según diversos niveles de autonomía, no se identificaron como relevantes. Tampoco los diversos problemas para poder calibrar los robots cada día con los pacientes o las numerosas incidencias acontecidas durante la hora de la merienda tuvieron ninguna influencia negativa en la valoración de la experiencia por parte de los

pacientes. Desde esta perspectiva estos resultados tan positivos aparentemente ajenos a cualquier variación del artefacto, de su funcionamiento o de sus múltiples problemáticas, solamente se podrían explicar por el efecto experimental, el conocido como efecto Hawthorne.

Sin embargo, en lugar de analizar los resultados desde la lógica de la HRI tomamos como referencia la noción de autonomía relacional y su integración con la perspectiva teórica de los STS y la mediación tecnológica (Latour, 1998). La autonomía relacional se entiende como la necesidad y capacidad permanente de establecer redes de dependencia e interdependencia entre humanos con otros humanos y con tecnologías y otros agentes. Desde la idea de mediación, entonces, el análisis se centra en la red de relaciones asistenciales en las que participa el robot, nos referimos a esta perspectiva como del *Robot Embedded in a Network* (REN) (Vallès-Peris y Domènech, 2023).

Desde la perspectiva REN podemos identificar una serie de relaciones que se producen cuando se introduce el MAR en el hospital, que no se limitan a la interacción del robot con el paciente. Agrupamos estas relaciones como la ‘capacidad de heteronomía’ del robot. En este caso, la heteronomía no como carencia de autonomía, sino como la “ley del otro” (Levinas, 2000).

A partir del estudio realizamos mostramos como la capacidad de heteronomía del robot está conformada por dos dimensiones:

- a) La capacidad relacional del robot, es decir, de establecer relaciones con otros agentes de la red que sostienen los cuidados de ese paciente (lo otro de las relaciones que el robot establece con el paciente): las interacciones del robot con las técnicas auxiliares sanitarias, la posibilidad de adaptarse al mobiliario de la habitación, la relación entre el artefacto y el técnico que lo acompaña, las interacciones con familiares u otras visitas, etc.
- b) La potencia de interpelación del robot, es decir, de hacer emerger e intensificar aquellas procesos o relaciones de asistencia y cuidados del paciente (lo otro de las relaciones en las que participa el robot): la intensificación de las visitas de los familiares, la atracción de fondos del hospital, la flexibilización de las tareas del personal sanitario, la relación directa de la dirección del centro con los pacientes, etc.

Desde esta perspectiva proponemos situar la idea de heteronomía en el centro del debate sobre las relaciones de cuidado con robots, planteando nuevos retos en el diseño de estos artefactos; así como también subrayando la necesaria vinculación de la robótica con programas y políticas de salud y atención, que permitan desarrollar desde una perspectiva interdisciplinar la capacidad relacional y de interpelación de estos artefactos.

Referencias bibliográficas

Barrué, C; Suárez, A; Inzitari, M; Ribera, A & Alenyà, G. “NYAM: the role of configurable engagement strategies in robotic-assisted feeding” in ACM/IEEE International Conference on Human-Robot Interaction, 2024.

Borenstein, J., & Arkin, R. (2016). Robotic nudges: The ethics of engineering a more socially just human being. *Science and Engineering Ethics*, 22(1), 31–46.

Darwall, S. (2006). The value of autonomy and autonomy of the will. *Ethics*, 116, 263–284.

Etzioni, A., & Etzioni, O. (2016). AI assisted ethics. *Ethics and Information Technology*, 18(2), 149–156.

Floridi, L., & Cows, J. (2019). A Unified Framework of Five Principles for AI in Society. *Harvard Data Science Review*.

Formosa, P. (2021). Robot Autonomy vs. Human Autonomy: Social Robots, Artificial Intelligence (AI), and the Nature of Autonomy. *Minds and Machines*, 31(4), 595–616.

Knoblauch, H. (2005). Focused ethnography. *Forum Qualitative Sozialforschung*, 6(3).

Latour, B. (1998). De la mediación técnica: filosofía, sociología, genealogía. In M. Domenech & F. J. Tirado (Eds.), *Sociología simétrica. Ensayos sobre ciencia, tecnología y sociedad* (pp. 249–302). Gedisa.

Levinas, (2000). *La huella del otro*. Taurus.

Topol, E. J. (2019). High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine*, 25(1), 44–56.

Vallès-Peris, N., & Domènech, M. (2023). Caring in the in-between: a proposal to introduce responsible AI and robotics to healthcare. *AI and Society*, 38(4), 1685–1695.