

Trabajo de Robótica Social par IHR

Lucas Gómez Velayos

13 de enero de 2026

1. Motivación

Es raro que las innovaciones tecnológicas tengan una rama dedicada a las relaciones de estas con el ser humano. No es así en el ámbito de la robótica, ya que estos estudios existen en una rama de la disciplina conocida como Robótica Social, que existe debido a la gran interacción que se busca proporcionar entre los humanos y los robots. Este es un caso único, ya que no muchas disciplinas tienen como objetivo un producto que imite o replique a un humano, sobre todo, en el ámbito social. Debido a esta peculiaridad, la Robótica Social es un campo inherentemente multidisciplinar, que combina robótica con sociología, antropología, neurología, psicología y/o psiquiatría.

En este trabajo se pretende analizar en profundidad los últimos avances realizados en este ámbito y, debido a la situación única de este, también analizaremos su impacto en los sujetos que lo experimentan. Por ejemplo, como se discutirá más adelante, existen estudios que indican un impacto positivo de interacciones con robots en el ámbito de la educación infantil y en la terapia de niños con trastorno dentro del espectro autista (ASD). También se ha investigado de forma prolífica sobre robótica de servicio para personas mayores, que está cobrando importancia debido al aumento de la esperanza de vida. Sin embargo, y he aquí una de las motivaciones principales para explorar sobre este tema, casi no existe investigación sobre los efectos negativos de algunas de las nuevas tecnologías utilizadas ampliamente en este campo.

Hasta hace relativamente poco, las relaciones o interacciones sociales con robots parecían todavía relegadas a la ciencia ficción, a excepción de algunos visionarios como Alan Turing, quien empezó a plantear estas cuestiones de forma académica en [1] en 1950. Sin embargo, con los últimos avances en Inteligencia Artificial, sobre todo en los ámbitos de procesamiento del lenguaje natural (*Natural Language Processing* o NLP), texto a voz (*Text to Speech* o TTS), transcripción (*Speech to Text* o STT) y tecnologías concretas como los *Large Language Models* (LLM), han dado pie a nuevos niveles de sofisticación y complejidad en este ámbito. Debido a la velocidad de estas innovaciones, el impacto de estas relaciones artificiales sobre los usuarios no se ha estudiado aún con suficiente detenimiento y profundidad, y menos en el ámbito de la robótica social, donde estas relaciones pueden incorporar además un robot físico como vehículo para cualquier tipo de comunicación. Además, a pesar de la importancia que se está produciendo en el ámbito académico a cuestiones ecológicas, el claro impacto medioambiental negativo que conlleva el entrenamiento de estas redes (sobre todo de LLM's) pocas veces se menciona en las investigaciones que las usan.

2. Estado del arte

En [2] se hace un resumen de los últimos avances en la materia de la Robótica Social. En este trabajo, se estudian 21 artículos de 2019 en adelante. Los avances más relevantes para este trabajo se agrupan en:

- Lenguaje no verbal: Mejoras y propuestas para su identificación y réplica
- Personalización de las interacciones
- Tratamiento de niños con ASD y personas de avanzada edad

- Enseñanza

En muchos de estos casos, las propuestas de los estudios son novedosas y, por tanto, su efectividad no puede ser probada de forma definitiva. Por ejemplo, se menciona “[...] although robots theoretically have the potential to influence the development of empathy in children, the type of effect they can have is uncertain, depending on their design, intensity, and context of use. Most of the studies reviewed that showed robots’ ability to improve empathy levels in children were not longitudinal, while studies that observed and argued about the negative effect of robots on children’s empathy were either purely theoretical or context-dependent”. Muchos de estos trabajos muestran resultados prometedores, e incluso se respaldan entre sí llegando a conclusiones similares, pero esto demuestra que estas todavía necesitan aún mayor experimentación y un desarrollo adecuado en cada caso de uso.

Este artículo también remarca algunas convergencias existentes entre casos, por ejemplo los robots NAO (Fig. 1), que también se mencionan en [3] (“The NAO robot is one of the most commonly used robots”), el uso de redes neuronales convolucionales (CNN) y recurrentes (RNN), o el uso de OpenPose [4], una librería para detectar poses del cuerpo humano mediante imagen o vídeo en tiempo real que ha sido desarrollada en la universidad Carnegie Mellon.



Figura 1: Robot NAO

3. Temas de interés

3.1. Lenguaje no verbal

Como se ha mencionado anteriormente, existe un interés creciente por hacer que tanto la detección como la recreación de gestos y expresiones sea adecuada. De igual forma que la morfología del robot y el contexto en el que se encuentra pueden influir decisivamente en la percepción del ser humano que interactúa con este, parece que se ha identificado la comunicación no verbal como otro componente de estos “prejuicios” o aceptación previa. En [5], se destaca la importancia de la comunicación no verbal para aplicaciones empleadas en escuelas:

“User-centered and participatory design studies, tend to agree that students prefer to interact with social robots which can talk, show emotional expressions with facial and body gestures, give them feedback, and support their educational activities”.

También en [2] se destaca: “The authors emphasize the importance of natural gestures for humanoid robots as they enhance the comfort of human interactions”.

Con respecto a este campo, además de las soluciones creadas ad hoc, también destaca el uso generalizado de OpenPose, una librería de código abierto, basada en imágenes y vídeo (Fig. 2). Esta aproximación es importante, ya que en lo relativo a *motion capture*, hasta hace poco, las únicas alternativas fiables se basaban en trajes o sensores que el usuario debe llevar puesto. El cambio de enfoque hacia una tecnología basada en vídeo no solo evita todo este aparataje, si no que aprovecha información que un robot social seguramente ya esté recopilando, por ejemplo, para fines de visión artificial.

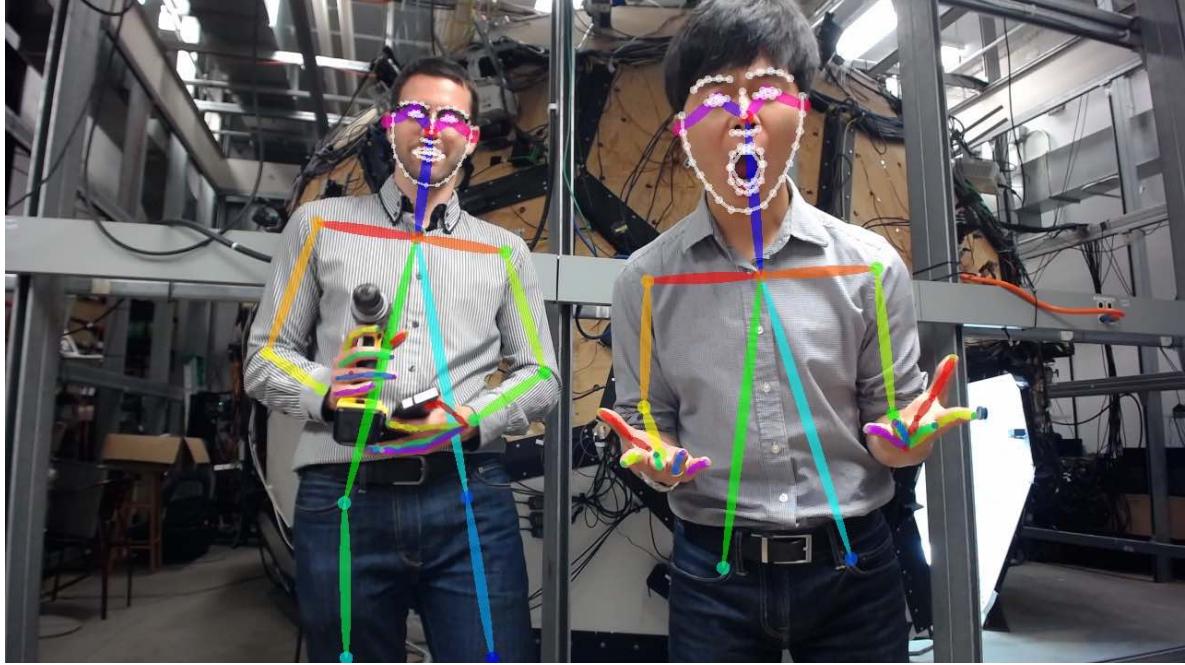


Figura 2: Ejemplo de funcionamiento de OpenPose, de su repositorio en GitHub

3.2. Personalización o adaptación de las interacciones

Otro punto importante que se destaca en [2] es la importancia que puede tener la personalización en el contexto de la interacción, así como la relevancia del usuario o usuarios que la llevan a cabo. De acuerdo con [2]:

“to enable the personalization of user profiles and the behavior of autonomous social robots. The authors emphasize the importance of such profiles and personalization, highlighting their potential to enhance the user experience and engagement with the robot”.

Este enfoque parece querer replicar nuevamente la psicología humana, ya que de acuerdo con [6]:

“The notion that people’s responses and behavior will tend to vary by the situation they are in, as a function of internal states, mental representations, and interpretations that are brought to the fore by cues in the situation [...], was iconoclastic when proposed, but is now supported by considerable evidence”.

3.3. Educación infantil

Como se ha citado antes, no hay un consenso generalizado sobre el aprendizaje de habilidades como la empatía en niños mediante la robótica social. Sí existen, en cambio, estudios que demuestran una correlación positiva entre esta y otros tipos de aprendizaje, como por ejemplo, el reciclaje [2]. En [5]:

“The use of educational robots increases as children get older and researchers and educators report beneficial outputs from their use in elementary school students learning outcomes, and motivation”.

3.4. Terapia con niños con ASD

En [3] se menciona que los niños con rasgos o síntomas autistas más pronunciados muestran una mayor aceptación ante robots y son más propensos a interactuar con ellos en entornos de juego. También concluye que la mayoría de la experimentación reciente llega a resultados positivos en el aprendizaje de habilidades sociales, como la gesticulación, la cognición emocional, la atención o el contacto visual.

La teoría principal es que, al diferenciar entre robots y humanos, las interacciones con robots suelen causar menos estrés y ansiedad social, con respuestas más predecibles que ayudan a la terapia.

3.5. Problemas emergentes con LLMs y *Chatbot psychosis*

El uso de IA generativa en el proceso de diálogo puede acarrear consecuencias negativas que sólo ahora están empezando a observarse. Existen estudios que relacionan la salida al mercado de modelos de chatbots serviles/lisonjeros/aduladores, con un aumento de casos de psicosis entre los usuarios que los utilizan[7][8]. En [9], se describe un caso de bromismo en un individuo paranoico, que pensaba que su vecino le estaba envenenando, y que fue causado por seguir ciegamente las instrucciones de ChatGPT.

Estos nuevos casos y advertencias de expertos, unidos a la poca investigación de estas interacciones, y menos aún a sus efectos a largo plazo, resulta preocupante cuando estos modelos se integran en interacciones humano-robot. Los casos actuales suceden con interacciones puramente textuales, por lo que se desconocen los efectos que pueda tener el hecho de que estas interacciones añadan también un vehículo físico, como un robot, capaz de retransmitir las respuestas mediante TTS, o incluso acompañarlas con gestos y comunicación no verbal.

Incluso sin modelos de IA generativa, también se desconoce si una exposición prolongada a robots sociales podría causar problemas similares a estos, ya que los estudios en robótica social suelen ser de corta duración (de [5]: “educational robots are rarely deployed autonomously for more than a few days”). En [8] se declara una correlación preocupante:

“Indeed, it was recently demonstrated in an experimental study that people who were more paranoid (self-reported) were more likely to report perceiving animacy and agency in dots moving on a computer screen”.

Esta correlación podría extrapolarse también a la antropomorfización de cualquier objeto inanimado, pudiendo llegar a hacer creer a algunos usuarios que un robot tiene conciencia, sentimientos, etc.

Precisamente uno de los avances de los que se ha hablado, en concreto la personalización de las interacciones, es parte de lo que los expertos consideran que fomenta este tipo de casos:

“This model has been accused of being overly “sycophantic” (insincerely affirming and flattering) toward users, caused by the model training leaning too hard on user preferences communicated via thumbs-up/thumbs-down assessments in the chatbot (so-called Reinforcement Learning from Human Feedback (RLHF))”.

[8] Después cita un comunicado de OpenAI afirmando este problema:

“On April 25th, we rolled out an update to GPT-4o in ChatGPT that made the model noticeably more sycophantic. It aimed to please the user, not just as flattery, but also as validating doubts, fueling anger, urging impulsive actions, or reinforcing negative emotions in ways that were not intended. Beyond just being uncomfortable or unsettling, this kind of behavior can raise safety concerns—including around issues like mental health, emotional over-reliance, or risky behavior”.

4. Conclusiones

La Robótica Social es un campo en pleno auge, con muchas cuestiones pendientes que aún se tienen que explorar según esta vaya avanzando. El estado del arte actual parece centrarse en los efectos

positivos de las interacciones con robots, que son muy numerosos, pero que dejan de lado los posibles efectos adversos de algunas nuevas tecnologías que se están incorporando.

En la robótica social se destacan los usos en educación, terapia de niños con ASD y cuidado y/o atención a personas de la tercera edad como los que mayor éxito y potencial tienen en la actualidad.

Destaca también la poca investigación dedicada hasta el momento a los posibles efectos adversos que la robótica social puede causar en ciertos usuarios y, sobre todo, el poco estudio de cualquier tipo de efecto a largo plazo (o estudios “longitudinales”) en este campo.

Referencias

- [1] A. M. Turing, “Computing Machinery and Intelligence,” *Mind*, vol. 59, n.º 236, págs. 433-460, 1950. DOI: <https://doi.org/10.1111/acps.70022>.
- [2] C. S. T. Romulo, L. P. Boscatti, W. S. R. Faria y A. S. Brandão, “Advances in social robotics: A brief review of recent contributions and Innovations,” *Brazilian Symposium on Robotics (SBR), and 2023 Workshop on Robotics in Education (WRE)*, págs. 35-40, 2023. DOI: [10.1109/lars-sbr/wre59448.2023.10332925](https://doi.org/10.1109/lars-sbr/wre59448.2023.10332925).
- [3] K. Zhang, Y. Jin, J. Chen, Y. Dai e Y. Zhang, “Robot intervention in social skills of children with autism: a review of the latest three years *,” en *2023 International Conference on Intelligent Education and Intelligent Research (IEIR)*, 2023, págs. 1-8. DOI: [10.1109/IEIR59294.2023.10391214](https://doi.org/10.1109/IEIR59294.2023.10391214).
- [4] Z. Cao, G. Hidalgo Martinez, T. Simon, S. Wei e Y. A. Sheikh, “OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019.
- [5] K. Papadopoulos, A. Velentza, P. Christodoulou y N. Fachantidis, “Social Educational Robotics Application: Architecture and Interconnectivity,” 2022. DOI: <https://doi.org/10.1109/iisa56318.2022.9904372>.
- [6] S. M. Andersen, R. Tuskeviciute, E. Przybylinski, J. N. Ahn y J. H. Xu, “Contextual Variability in Personality From Significant–Other Knowledge and Relational Selves,” *Frontiers in Psychology*, vol. 6, n.º 1882, 2016. DOI: <https://doi.org/10.3389/fpsyg.2015.01882>.
- [7] Z. Ben-Zion, “Why we need mandatory safeguards for emotionally responsive AI,” *Nature*, vol. 643, n.º 8070, págs. 9-9, 2025. DOI: <https://doi.org/10.1038/d41586-025-02031-w>.
- [8] S. D. Østergaard, “Generative Artificial Intelligence Chatbots and Delusions: From Guesswork to Emerging Cases,” *Acta Psychiatrica Scandinavica*, 2025. DOI: <https://doi.org/10.1111/acps.70022>.
- [9] A. Eichenberger, S. Thielke y A. V. Buskirk, “A Case of Bromism Influenced by Use of Artificial Intelligence,” *Annals of Internal Medicine Clinical Cases*, vol. 4, n.º 8, 2025. DOI: <https://doi.org/10.7326/aimcc.2024.1260>.