

## Desarrollo de un sistema de interacción mediante agentes LLM: Implementación en un robot social NAO

Zamorano, A.<sup>a,\*</sup>, Zalama, E.<sup>a,b</sup>, Gómez-García-Bermejo, J.<sup>a,b</sup>, Duque-Domingo, J.<sup>a</sup>

<sup>a</sup>Departamento de Ingeniería de Sistemas y Automática, ITAP, Universidad de Valladolid, C/Dr Mergelina s/n, 47011, Valladolid, España.

<sup>b</sup>Centro tecnológico CARTIF, Boecillo, 47151, Valladolid, España.

### Resumen

El uso de robots sociales surge como una estrategia prometedora para afrontar los retos del cuidado de personas mayores en una población cada vez más envejecida, particularmente en el ámbito domiciliario. La incorporación de los recientes avances en Inteligencia Artificial, especialmente los modelos de lenguaje de gran escala (LLMs), abre nuevas oportunidades para potenciar las capacidades de la robótica social. Este trabajo presenta el desarrollo e integración de un agente conversacional basado en LLMs en un robot social NAO, con el objetivo de mejorar sus capacidades de interacción humano-robot. El sistema implementa un agente de tipo ReAct, capaz de razonar y actuar mediante herramientas externas, integrando reconocimiento de voz, procesamiento conversacional y síntesis de voz en un flujo completo de diálogo verbal. Se exploran funcionalidades como acceso a APIs externas, ejecución de acciones físicas y adaptación contextual mediante memoria. Los resultados muestran un comportamiento flexible y natural por parte del robot, así como una mejora notable respecto a su sistema original.

*Palabras clave:* Inteligencia Artificial, Robótica Inteligente, Sistemas Robóticos Autónomos, Robotica Social, Salud Urbana

### Development of an agent-based interaction system: Implementation on a NAO social robot

#### Abstract

The use of social robots has emerged as a promising strategy to address the challenges of elderly care in an increasingly aging population, particularly in home-based environments. The incorporation of recent advances in Artificial Intelligence, especially large language models (LLMs), opens new opportunities to enhance the capabilities of social robotics. This work presents the development and integration of a conversational agent based on LLMs into a NAO social robot, aiming to improve its human-robot interaction capabilities. The system implements a ReAct-type agent, capable of reasoning and acting through external tools, integrating speech recognition, conversational processing, and speech synthesis into a complete verbal interaction loop. Features such as access to external APIs, execution of physical actions, and context-aware conversational adaptation are explored. The results show flexible and natural behavior from the robot, along with a significant improvement over its original system.

*Keywords:* Artificial Intelligence, Intelligent Robotics, Autonomous Robotic Systems, Social Robotics, Urban Healthcare.

## 1. Introducción

### 1.1. Una necesidad creciente

El envejecimiento de la población mundial es uno de los cambios demográficos más importantes del siglo XXI. En Europa, más del 21 % de la población supera ya los 65 años, y se prevé que este número aumente considerablemente en la próxima década (Eurostat, 2025). Se observan tendencias similares en todo el mundo, incluyendo China, donde la población de edad avanzada alcanzó los 280 millones en 2022 (GlobalTimes,

2023). En España, y especialmente en regiones rurales como Castilla y León, la situación es aún más crítica. En esta comunidad, el 25 % de los residentes tienen más de 65 años, lo que supone un claro desafío a la sostenibilidad del cuidado de personas mayores (INE, 2024).

Esta presión demográfica subraya la necesidad urgente de soluciones de atención innovadoras, especialmente en zonas desatendidas o despobladas. En los últimos años, el campo de la *Robótica Social Asistencial (SARs)* ha surgido como una vía

\*Autor para correspondencia: alvaro.zamorano@uva.es

prometedora para aliviar la presión sobre los sistemas de salud y reducir los riesgos de aislamiento social entre los mayores.

### 1.2. Asistencia mediante robots

Impulsados por este imperativo demográfico, la cantidad de trabajos de investigación centrados en SARs ha crecido considerablemente. Publicaciones recientes destacan el aumento de estudios conceptuales y prácticos enfocados en entender y mejorar las dinámicas de interacción entre usuarios mayores y robots de servicio (Cesário et al., 2025). Una parte importante de estos esfuerzos se canaliza hacia el diseño de ecosistemas *Ambient Assisted Living (AAL)*, que integran sensores, actuadores y agentes inteligentes en los entornos donde residen personas mayores. Estos entornos permiten que las personas mayores mantengan su independencia mientras aseguran su seguridad y bienestar mediante una monitorización no intrusiva, estimulación cognitiva y compañía, todo ello cada vez más impulsado por sistemas robóticos inteligentes.

Los robots sociales actúan como interfaz de usuario en muchos de estos sistemas, transformando la complejidad del sistema en interacciones accesibles y empáticas. (Cruces et al., 2024; Merino-Fidalgo et al., 2023). La integración de agentes robóticos personalizados ha demostrado un notable potencial en el apoyo a la función cognitiva y la salud emocional de las poblaciones envejecidas. Los estudios muestran que, cuando están adaptados al individuo, los robots sociales pueden ayudar a mitigar los efectos del deterioro cognitivo y mejorar el estado de ánimo (Rincon-Arango et al., 2025). En particular, robots como NAO (ver Figura 1) se han empleado en entornos con usuarios que padecen depresión o demencia, obteniéndose resultados positivos asociados al uso de agentes conversacionales con conciencia emocional (Abdollahi et al., 2023; Macis et al., 2023).

Estos agentes robóticos involucran a los usuarios en diálogos en lenguaje natural que van más allá de respuestas guionizadas, aprovechando los avances en *Large Language Models (LLMs)* para generar interacciones. En particular, los modelos GPT permiten gestionar conversaciones complejas y ofrecer contenido útil y atractivo al usuario, desde noticias diarias hasta ejercicios físicos guiados.

### 1.3. Modelos de Lenguaje: oportunidades y desafíos

La introducción de LLMs en el campo de la Interacción Humano-Robot (HRI) representa un salto tecnológico significativo. Estos modelos ofrecen una comprensión robusta del lenguaje natural, retención de contexto e incluso capacidades de razonamiento de alto nivel y planificación de tareas (Zhang et al., 2023). Permiten que los robots sociales funcionen no solo como acompañantes, sino también como asistentes inteligentes capaces de ejecutar tareas específicas, recabar información desde APIs externas o guiar rutinas cognitivas.

A pesar de estos avances, la implementación de sistemas basados en LLMs no está exenta de desafíos. Aspectos como la gestión de turnos en la conversación, la naturalidad de la síntesis de voz o la variedad en los diálogos siguen siendo áreas de mejora (Irfan et al., 2025). Sin embargo, el rápido progreso de la tecnología ha mitigado significativamente muchas de estas limitaciones, haciendo cada vez más viable su despliegue en entornos reales.

Más allá de la eficacia técnica, la *aceptación* del usuario sigue siendo un factor clave para la adopción exitosa de robots sociales en el cuidado de mayores. Los estudios señalan que la confianza, la utilidad percibida y la facilidad de interacción son los principales impulsores de la aceptación, especialmente entre adultos mayores sin conocimientos tecnológicos (Heerink et al., 2010; Abdi et al., 2018). Las mejoras en la interacción y el comportamiento personalizado —facilitadas por los modelos LLM y sistemas *emotion-aware*— han demostrado aumentar el compromiso de los usuarios y la usabilidad a largo plazo.



Figura 1: Robot humanoide NAO

No obstante, también deben abordarse las preocupaciones éticas. El uso de robots inteligentes en entornos tan sensibles como el de los cuidados, plantea cuestiones sobre privacidad, autonomía y consentimiento informado. Es crucial implementar medidas sólidas de transparencia y protección de datos para asegurar interacciones respetuosas y seguras (Vozna and Costantini, 2025; Dubbaka and Sankar, 2024).

### 1.4. Marco de la investigación

La investigación presentada en este artículo forma parte del proyecto EIAROB, una iniciativa orientada a la creación de un ecosistema ambiental inteligente que fomente la autonomía y mejore la calidad de vida de las personas mayores y dependientes en sus hogares (ver Figura 2). Como componente central de este ecosistema, se incorporan robots sociales para desempeñar funciones tanto asistenciales como de acompañamiento mediante la interacción directa con el usuario. En este contexto, el desarrollo realizado para el robot humanoide NAO se centra en proporcionar apoyo emocional, compañía, estimulación cognitiva y orientación en la realización de actividades físicas estructuradas, como rutinas de ejercicio físico (EIAROB, 2022).

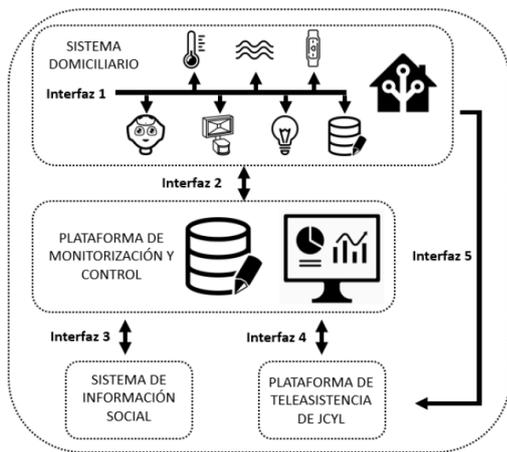


Figura 2: Arquitectura del ecosistema EIAROB

## 2. Evolución de la IA moderna

### 2.1. Fundamentos de los modelos de lenguaje

El panorama actual de la Inteligencia Artificial se basa fundamentalmente en la revolucionaria arquitectura propuesta por Vaswani et al. (2017) en su artículo “*Attention Is All You Need*”. Este trabajo presentó el modelo *Transformer*, una arquitectura centrada en el mecanismo de auto-atención (*self-attention*). Desde entonces, ha constituido la base sobre la cual se construyen los grandes modelos de lenguaje (LLMs) como GPT, BERT y Claude.

Tras este avance, las capacidades de los modelos generativos se expandieron rápidamente a medida que aumentaban de tamaño. Estos modelos comenzaron a desplazar a arquitecturas neuronales concebidas específicamente para algunas tareas, como las redes convolucionales (CNNs) o las redes de memoria corto-largo plazo (LSTMs), gracias a su impresionante capacidad de aprendizaje y generalización (Brown et al., 2020). Un hito importante llegó en 2022 con el lanzamiento de ChatGPT, una interfaz de *chatbot* web impulsada por GPT-3.5 de OpenAI, que democratizó aún más el acceso a los LLMs y dio lugar a una ola de experimentación sin precedentes.

### 2.2. De modelos estáticos a sistemas agénticos

A pesar de su éxito, los primeros LLMs presentaban una limitación crítica: eran sistemas inherentemente estáticos, entrenados sobre un conjunto fijo de datos y aislados de un contexto en tiempo real. Su “conocimiento” quedaba fijado en el momento de entrenamiento y carecían de interacción con el exterior. Una de las primeras mejoras para abordar esta limitación fue la introducción de *memoria* en forma de ventana de contexto, lo que permitió a los modelos mantener la coherencia entre turnos de la conversación. Sin embargo, esta forma de memoria pasiva seguía siendo insuficiente para lograr un comportamiento activo y autónomo.

Un cambio significativo se produjo con el surgimiento del *framework ReAct* (Reasoning + Acting) (Yao et al., 2023). ReAct introdujo una arquitectura en la que los LLMs podían alternar entre pasos de razonamiento y acciones externas, permitiéndoles utilizar herramientas y reflexionar sobre su resultado en un ciclo iterativo. Este ciclo intercalado de razonamiento

y acción sentó las bases conceptuales para los *Agentes de IA* modernos —sistemas que no solo generan respuestas, sino que planifican y ejecutan tareas de forma estratégica.

A partir de ReAct, surgieron más *frameworks* de agentes autónomos. Proyectos como Auto-GPT y BabyAGI llevaron las capacidades de los LLMs al ámbito de la autonomía orientada a objetivos. Auto-GPT se estructura en torno a la ejecución de múltiples pasos pre-planificados, mientras que BabyAGI implementa una gestión dinámica de tareas, generando nuevas sub-tareas basadas en los resultados de pasos anteriores. Estos sistemas superaron algunas de las primeras aplicaciones de LLMs actuando como agentes autónomos.

### 2.3. Arquitectura de los sistemas agénticos

A día de hoy (2025), el campo de la inteligencia artificial agéntica, abarca un espectro de arquitecturas (Anthropic, 2024) en cuyo núcleo se encuentra el denominado *augmented LLM*. Esta arquitectura (ver Figura 3) potencia al modelo base con acceso a memoria, sistemas de *knowledge retrieval* y herramientas externas. Funciona como bloque básico para la construcción de sistemas más complejos (ver Figura 4), permitiendo a los LLMs ir más allá de la generación reactiva de texto. A partir de esta base, surgen paradigmas más sofisticados:

- *AI Workflows*: En estos sistemas, los LLMs se integran en secuencias de ejecución fijas donde cada paso de una tarea está orquestado mediante lógica codificada. El uso de herramientas y la invocación de modelos está predefinido por los desarrolladores.
- *Agentes*: A diferencia de los *workflows*, los agentes tienen autonomía sobre el flujo de ejecución. Deciden qué herramientas usar y cuándo, adaptándose dinámicamente a la situación hasta completar una tarea.

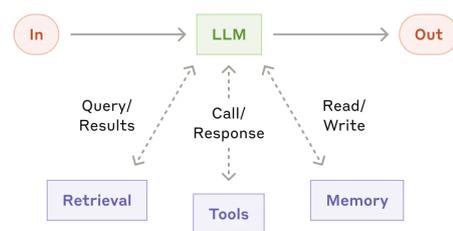


Figura 3: Estructura de un *Augmented LLM*

A medida que los agentes han adquirido la capacidad de controlar dinámicamente la invocación de herramientas y la secuenciación de tareas, han ido asumiendo cada vez más el rol de *orquestadores centrales* en los sistemas de IA. A diferencia de los *workflows*, que siguen cadenas lógicas rígidas, los sistemas agénticos modernos exhiben comportamientos no lineales y realimentados: inician acciones, analizan resultados y adaptan sus estrategias en función de esos resultados. Para describir este nuevo paradigma lógico, investigadores y desarrolladores han recurrido a *frameworks* basados en grafos para modelar sistemas agénticos.

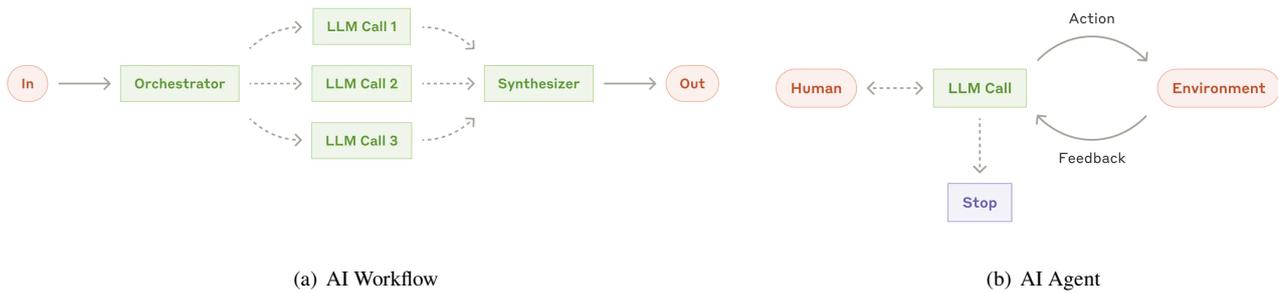


Figura 4: Ejemplos de distintas arquitecturas.

### 3. Descripción del sistema propuesto

El diseño del sistema (ver Figura 5) se basa en la arquitectura de un agente *ReAct* (*Reasoning and Acting*), que actualmente constituye el enfoque por defecto para la implementación de agentes capaces de utilizar herramientas. Este modelo combina el razonamiento paso a paso con la ejecución de acciones, permitiendo a los agentes tomar decisiones informadas y actuar de forma iterativa en función del contexto conversacional. Su popularidad ha llevado a que sirva de inspiración para librerías ampliamente utilizadas como *LangChain* (Chase, 2022), consolidándose como la base conceptual para la mayoría de los sistemas agénticos modernos.

Para construir una experiencia de interacción íntegramente por voz, el agente está precedido por un módulo de reconocimiento de voz (*speech-to-text*, STT), encargado de captar y transcribir las instrucciones verbales del usuario. Posteriormente, las respuestas generadas por el modelo son transformadas nuevamente en audio mediante un sistema de síntesis de voz (*text-to-speech*, TTS), cerrando así el bucle de diálogo oral entre humano y robot.

#### 3.1. Módulo de reconocimiento de voz (STT)

El sistema no emplea el módulo de reconocimiento de voz nativo del robot NAO, ya que este se basa en un enfoque obsoleto de detección mediante palabras clave y frases predefinidas, lo que limita gravemente su flexibilidad. Bajo este sistema, cualquier desviación significativa respecto a las expresiones esperadas impide que el robot identifique correctamente la intención del usuario. Para solventar esta limitación, se ha integrado el proyecto *PepperSpeechRecognition* (Bramauer, 2018), que permite incorporar capacidades de reconocimiento de voz en robots NAO y Pepper mediante la API de *Google Speech Recognition*.

A diferencia de otros modelos de STT como Whisper (Radford et al., 2022), orientados principalmente a la transcripción de audio grabado, la API de Google está optimizada para aplicaciones de reconocimiento en tiempo real, aspecto crucial en entornos de interacción conversacional. Esta API, además, implementa un mecanismo de confianza que asigna un nivel de certeza a cada transcripción, permitiendo al sistema rechazar aquellas que no superen un umbral determinado. Este enfoque resulta preferible a aceptar de forma errónea comandos mal re-

conocidos, ya que evita la propagación de errores hacia etapas posteriores del flujo de la conversación.

#### 3.2. Agente *ReAct*

El núcleo del sistema es un agente conversacional diseñado siguiendo el esquema *ReAct* y alimentado por un modelo de lenguaje GPT proporcionado por la API de OpenAI (utilizando el *endpoint ChatCompletions*). Esta interfaz permite la utilización de toda la gama de modelos disponibles en la plataforma, adaptando el rendimiento del sistema según las necesidades de cada escenario. Durante el desarrollo y pruebas básicas se ha empleado el modelo GPT-4o-mini debido a su eficiencia y bajo coste computacional. Para las pruebas de razonamiento más complejas y evaluación de capacidades de planificación y ejecución de funciones, se utilizó el modelo GPT-4o, más avanzado y preciso.

Dado que este *endpoint* en particular carece de memoria, se implementa un sistema que conserva el historial completo de los mensajes intercambiados durante la sesión. Esto mantiene el contexto de la conversación, permitiendo al modelo razonar en base a información previamente mencionada, evitando redundancias como llamadas duplicadas a una misma función externa para obtener información que ya está disponible. Esta memoria contextual no solo mejora la coherencia de las respuestas sino que también favorece una experiencia de diálogo más fluida y natural, similar a una conversación humana.

Para convertir este modelo de lenguaje en un agente es necesario dotarlo de herramientas (*tools*) que le permitan interactuar con el exterior. Estas herramientas suelen implementarse mediante funciones de código y pueden clasificarse en dos grandes categorías: aquellas orientadas a recabar información (*knowledge retrieval*) y aquellas diseñadas para ejecutar acciones concretas. Con el objetivo de ilustrar ambos tipos de capacidades, se ha proporcionado al agente acceso a las siguientes funciones:

- *get\_location\_info()*: devuelve la ubicación local.
- *get\_weather\_current()*: obtiene el clima actual.
- *get\_weather\_forecast()*: obtiene la previsión del clima.
- *get\_movie\_info()*: consulta una base de datos de películas.
- *robot\_run\_exercise()*: el robot ejecuta una rutina de ejercicios predefinida.

- `robot_goto_posture()`: el robot se posiciona en una postura predefinida.

### 3.3. Módulo de síntesis de voz (TTS)

Para completar el bucle conversacional, se emplea el sistema de síntesis de voz nativo del robot, concretamente el módulo *ALAnimatedSpeech* del framework NAOqi (SO del robot). Esta elección no responde tanto a la calidad de la voz sintética generada —que resulta aceptable pero limitada en comparación con soluciones más modernas—, sino a su capacidad de integración directa con el sistema de animaciones del robot. Este módulo permite que, durante la locución, el robot acompañe el discurso con gestos sincronizados, como movimientos de brazos o inclinaciones de cabeza, activados automáticamente en función de palabras clave presentes en el texto a pronunciar. Esta coordinación entre voz y movimiento incrementa significativamente el grado de expresividad y naturalidad percibida en la interacción.

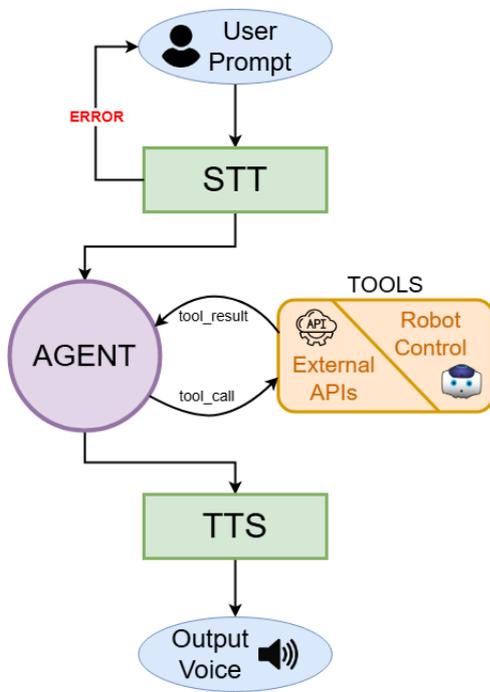


Figura 5: Estructura del sistema

## 4. Resultados

El agente propuesto en el presente artículo, ha demostrado una serie de comportamientos complejos que reflejan un uso inteligente y contextual de las funciones (*tools*) proporcionadas, así como una adecuada comprensión conversacional y gestión del flujo de tareas.

Uno de los aspectos más destacables es la capacidad del robot para discernir cuándo ejecutar funciones externas. Esta habilidad sugiere una comprensión de la intencionalidad del usuario y de las herramientas disponibles, permitiéndole decidir cuándo basta con una respuesta textual generada y cuándo es más apropiado delegar en funciones específicas.

En particular, el robot ha mostrado un comportamiento notable al emplear la función de geolocalización (ver Figura 6). Al recibir peticiones relacionadas con el clima, el agente identifica de forma autónoma si se trata de una consulta sobre el tiempo “local”. En ese caso, ejecuta primero la función de ubicación para obtener la ciudad del usuario y posteriormente encadena esta información con la llamada a la API del tiempo. El modelo demuestra así su capacidad de razonamiento en múltiples pasos ejecutando funciones de manera secuencial.

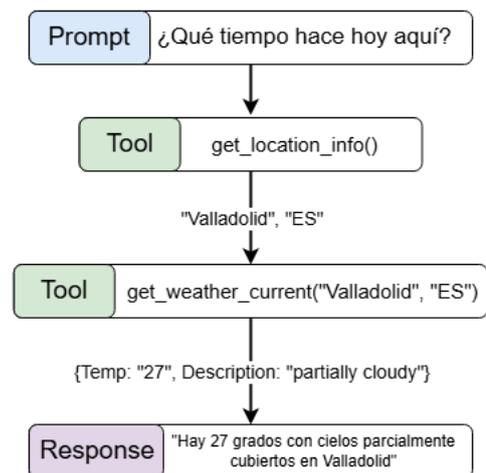


Figura 6: Ejemplo de razonamiento multi-paso

Otro comportamiento destacado es el uso en conjunción del conocimiento interno (pre-entrenado) del modelo con las funciones proporcionadas. Por ejemplo, si el usuario pide información de una película de la cual existen varias versiones y la respuesta inicial de la API no es satisfactoria —debido a que la consulta devuelve únicamente el primer resultado de la búsqueda—, el agente se adapta incluyendo en la petición a la API el año de la versión más reciente de dicha película, obtenido desde su conocimiento interno.

Durante las pruebas, se constató el correcto comportamiento del agente en cuanto al cumplimiento de las instrucciones establecidas en el *prompt* de sistema. El agente respeta de forma consistente los parámetros de comportamiento establecidos, incluyendo el tono, el estilo conversacional y la información del perfil del usuario. Esta memoria contextual se mantiene estable durante la conversación, contribuyendo a una interacción más personalizada y coherente.

## 5. Conclusiones

El presente trabajo se enmarca en un contexto tecnológico emergente, marcado por la reciente irrupción de los modelos de lenguaje a gran escala (LLMs). En este sentido, el desarrollo llevado a cabo debe entenderse como un demostrador tecnológico, cuyo principal objetivo ha sido explorar y ejemplificar el potencial de integrar un agente conversacional avanzado en un robot social como NAO.

Pese a no haberse completado aún la fase de pruebas en entornos reales —como viviendas piloto o residencias—, se ha logrado implementar un conjunto funcional de características

que demuestran una mejora sustancial respecto al sistema original del robot. La experiencia conversacional obtenida supera ampliamente a la ofrecida por defecto, la cual se basaba en estructuras rígidas sin inteligencia artificial.

La integración no ha estado exenta de dificultades, debido principalmente a las limitaciones de la plataforma NAO. Este robot no está originalmente concebido para una reconfiguración tan disruptiva de componentes críticos como el reconocimiento de voz, la toma de decisiones o el sistema de diálogo, lo que ha requerido soluciones creativas y adaptaciones específicas.

Los resultados obtenidos abren interesantes posibilidades de investigación futura:

- Mejora de la capacidad de toma de decisiones del agente combinando el enfoque ReAct con modelos de razonamiento *Chain of Thought (CoT)*.
- Ampliación del contexto del agente incorporando nuevas fuentes de información a través de técnicas más avanzadas de *knowledge retrieval* como RAG (*Retrieval-Augmented Generation*) y bases de datos vectoriales.
- Exploración de un diseño alternativo basado en grafos, representando los posibles estados del robot y sus transiciones como nodos interconectados mediante *frameworks* como LangGraph.

En resumen, el trabajo realizado sienta una base sólida para futuras aplicaciones de robots sociales inteligentes en contextos de asistencia y acompañamiento, alineándose con las necesidades tecnológicas y sociales actuales.

## Agradecimientos

La investigación que se presenta en este trabajo ha recibido financiación del proyecto ROSOGAR PID2021-123020 OB-I00 financiado por MCI-N/AEI/10.13039/501100011033/FEDER, UE, y del proyecto EIAROB Financiado por la Consejería de Familia de la Junta de Castilla y León - Next Generation EU.

## Referencias

Abdi, J., Al-Hindawi, A., Ng, T., Vizcaychipi, M. P., 2018. Scoping review on the use of socially assistive robot technology in elderly care. *BMJ Open* 8 (2).  
DOI: 10.1136/bmjopen-2017-018815

Abdollahi, H., Mahoor, M. H., Zandie, R., Siewierski, J., Qualls, S. H., Jul. 2023. Artificial emotional intelligence in socially assistive robots for older adults: A pilot study. *IEEE Transactions on Affective Computing* 14 (3), 2020–2032.  
DOI: 10.1109/taffc.2022.3143803

Anthropic, 2024. Building effective agents. URL: <https://www.anthropic.com/engineering/building-effective-agents>, accessed: 2025-04-08.

Bramauer, J., 2018. Pepper speech recognition. <https://github.com/JBramauer/pepperspeechrecognition>, accessed: 2025-04-08.

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Nee-lakantan, A., Shyam, P., Sastry, G., Askell, A., et al., 2020. Language models are few-shot learners. *Advances in neural information processing systems* 33, 1877–1901.

Cesário, L. C., Barbosa, P., Miguel, P. A. C., Mendes, G. H., 2025. Service robots in caring for older adults: Uncovering the current conceptual and intellectual structures and future research agenda. *Archives of Gerontology and Geriatrics* 131, 105755.  
DOI: <https://doi.org/10.1016/j.archger.2025.105755>

Chase, H., October 2022. Langchain. Accessed: 2025-04-30.  
URL: <https://github.com/langchain-ai/langchain>

Cruces, A., Jerez, A., Bandera, J. P., Bandera, A., 2024. Socially assistive robots in smart environments to attend elderly people—a survey. *Applied Sciences* 14 (12).  
DOI: 10.3390/app14125287

Dubbaka, S., Sankar, B., 2024. Ethical decision-making for social robots in elderly care scenario: A computational approach. In: *Social Robotics*. Springer Nature Singapore, Singapore, pp. 143–160.

EIAROB, 2022. Ecosistema de inteligencia ambiental para el apoyo a los cuidados de larga duración en el hogar mediante uso de robots sociales. Accessed: 2025-04-08.  
URL: <https://www.itap.uva.es/eiarob/>

Eurostat, 2025. Population structure and ageing. Accessed: 2025-04-08.  
URL: [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Population\\_structure\\_and\\_ageing](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Population_structure_and_ageing)

GlobalTimes, 2023. China's elderly population aged 60 and over reaches 280m. Accessed: 2025-04-08.  
URL: <https://www.globaltimes.cn/page/202312/1303684.shtml>

Heerink, M., Kröse, B., Evers, V., Wielinga, B., 2010. Assessing acceptance of assistive social agent technology by older adults: the almere model. *International Journal of Social Robotics* 2 (4), 361–375, available at: <https://doi.org/10.1007/s12369-010-0068-5>.  
DOI: 10.1007/s12369-010-0068-5

INE, 2024. Proporción de personas mayores de cierta edad por comunidad autónoma. URL: <https://www.ine.es/jaxiT3/Datos.htm?t=1451#!tabs-tabla>, accessed: 2025-04-08.

Irfan, B., Kuoppamäki, S., Hosseini, A., Skantze, G., 03 2025. Between reality and delusion: challenges of applying large language models to companion robots for open-domain dialogues with older adults. *Autonomous Robots* 49.  
DOI: 10.1007/s10514-025-10190-y

Macis, D., Perilli, S., Gena, C., 2023. Employing socially assistive robots in elderly care (longer version).  
URL: <https://arxiv.org/abs/2304.14944>

Merino-Fidalgo, S., Zalama, E., Gómez-García-Bermejo, J. and Duque-Domingo, J., Gómez, R., Viñas, P., García, D., Uruña, H., 7 2023. Sistema de monitorización no intrusiva para vivienda de personas mayores. *Jornadas Nacionales de Robótica y Bioingeniería 2023: Libro de actas*, 115–121.  
DOI: 10.20868/UPM.BOOK.74896

Radford, A., Kim, J. W., Xu, T., Brockman, G., McLeavey, C., Sutskever, I., 2022. Robust speech recognition via large-scale weak supervision.  
URL: <https://arxiv.org/abs/2212.04356>

Rincon-Arango, J. A., Marco-Detchart, C., Julian Inglada, V. J., 2025. Personalized cognitive support via social robots. *Sensors* 25 (3).  
DOI: 10.3390/s25030888

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., Polosukhin, I., 2017. Attention is all you need. *Advances in neural information processing systems* 30.

Vozna, A., Costantini, S., 2025. Ethical, legal, and societal dimensions of ai-driven social robots in elderly healthcare. *Intelligenza Artificiale*.  
DOI: 10.1177/17248035241310192

Yao, S., Zhao, J., Yu, D., Du, N., Shafran, I., Narasimhan, K., Cao, Y., 2023. React: Synergizing reasoning and acting in language models.  
URL: <https://arxiv.org/abs/2210.03629>

Zhang, C., Chen, J., Li, J., Peng, Y., Mao, Z., 2023. Large language models for human–robot interaction: A review. *Biomimetic Intelligence and Robotics* 3 (4), 100131.  
DOI: <https://doi.org/10.1016/j.birob.2023.100131>