XXI Simposio CEA de Control Inteligente
24-26 de junio de 2026, Salamanca

Gestión de visitas guiadas reales con un robot interactivo que saluda

Barranco, M.[a,1], Fernández, O.[a], Zhang, Z.[a], Galán, D.[a], de la Puente, P.[a,*]

[a] *Centre for Automation and Robotics (UPM-CSIC), Universidad Politécnica de Madrid. C/ José Gutiérrez Abascal, nº2, 28006, Madrid, España.*
[b] *Dronomy, Paseo de la Castellana, nº 40, planta 8, Madrid, España.*

To cite this article: Barranco, M., Fernández, O., Zhang, Z., Galán, D., de la Puente, P. 2026. Management of Real-World Guided Tours Using an Interactive Robot. XXI Simposio CEA de Control Inteligente. 2026.

Resumen

Este trabajo proporciona una integración práctica de detección de objetos y personas basada en YOLO en un robot de servicio real, añadiendo funcionalidades propias para la interacción mediante saludo y la recogida de datos durante las visitas. El sistema resultante permite desplegar visitas guiadas autónomas más interactivas y ha sido utilizado con éxito en el desarrollo de visitas guiadas con público real en un laboratorio de investigación.

Palabras clave: Percepción, Navegación autónoma, Automatización centrada en los humanos, Interacción humano-robot, Robots dotados de IA

Management of Real-World Guided Tours Using an Interactive Robot

Abstract

This work provides a practical integration of object and human detection based on YOLO in a real service robot, incorporating additional functionalities for interaction through greetings and data collection during guided tours. The resulting system enables the deployment of more interactive autonomous guided visits and has been successfully used in guided tours with real visitors in a research lab environment.

Keywords: Robot perception and sensing, Autonomous navigation, Human centered automation, Human-robot interaction, AI-powered robotics

1. Introducción

En los últimos años, la robótica de servicio ha experimentado un gran crecimiento gracias a distintos avances en inteligencia artificial (IA), percepción y navegación autónoma. Esto ha permitido que los robots interactúen de forma más natural con las personas, abriendo las puertas a la posibilidad de desempeñar tareas de interacción con visitantes en espacios públicos o privados Garcia-Fernandez (2024); Barranco (2025).

El presente trabajo se enmarca en esta línea, integrando en un sistema de visitas guiadas autónomas la detección de objetos y personas mediante el algoritmo YOLO en el robot móvil interactivo TIAGo. Con esta combinación, el robot no solo guía a los visitantes por un recorrido predefinido, aportando información relevante en los puntos de interés, sino que también es capaz de identificar elementos del entorno y reaccionar ante ellos en tiempo real, enriqueciendo así la experiencia.

2. Estado del Arte

El desarrollo de robots guía ha experimentado notables avances recientemente Alcantara et al. (2024). En el ámbito de los museos, varios trabajos han mostrado su potencial para mejorar la experiencia de los visitantes mediante asistencia contextual y comportamiento autónomo Rosa et al. (2024); Nasiri and Saleem (2025); de Mello et al. (2024). Estas soluciones combinan sensores, conectividad avanzada (por ejemplo, 5G) y estrategias de planificación para operar eficazmente en entornos dinámicos.

En paralelo, la navegación basada en visión y el uso de técnicas de inteligencia artificial han permitido mejorar la autonomía y robustez de los robots guía en distintos escenarios,

*Autor para correspondencia: paloma.delapuente@upm.es

	“morning”	“afternoon”	“evening”
Español	“Hola, buenos días”	“Hola, buenas tardes”	“Hola, buenas tardes”
Inglés	“Hello, good morning”	“Hello, good afternoon”	“Hello, good evening”

Tabla 1: Texto enviado para saludar dependiendo del idioma y del período del día.

incluyendo entornos interiores complejos Todkari et al. (2025). Estas aproximaciones facilitan una interacción continua con los usuarios durante el desplazamiento.

Por otro lado, estudios recientes ponen de manifiesto que los comportamientos de saludo influyen significativamente en la calidad de la interacción Avelino et al. (2025); Rudaz et al. (2023); Øllgaard et al. (2024); Scheideger et al. (2025). Asimismo, aspectos como el momento del saludo, el movimiento corporal y la sincronización del habla resultan clave para captar la atención de los usuarios Enoki et al. (2022).

En conjunto, la combinación de navegación autónoma y comportamientos sociales, como el saludo proactivo a las personas en el entorno, constituye un elemento fundamental para el diseño de robots guía más naturales, accesibles y efectivos.

3. Descripción del Sistema

3.1. Sistema inicial de visitas

En un fichero externo en formato XML se definen las salas y los lugares de interés dentro de cada una de ellas. Para generar o editar dicho fichero pueden utilizarse herramientas con interfaces gráficas de usuario previamente desarrolladas[de la Puente et al. (2018)]. En otro fichero se define el recorrido a realizar en cada visita concreta dentro de ese mismo entorno.

El sistema se ha diseñado e implementado sobre el Sistema Operativo Robótico (ROS). La Fig. 1 muestra la arquitectura de nodos y topics. El nodo principal es el nodo *visita*, que se encarga de leer los destinos sucesivos, solicitando a places_interpretation que obtenga las coordenadas métricas del siguiente punto a alcanzar. Estas coordenadas son enviadas al nodo sending_goals, el cual, mediante un Action Client, solicita al nodo move_base del robot Tiago que emprenda la tarea de navegación autónoma hacia dicha pose de destino. Este mismo nodo gestiona el feedback de la navegación e informa al nodo visita cuando el robot llega satisfactoriamente al punto deseado. Entonces el nodo visita accede al texto para el diálogo correspondiente a dicho punto y envía la información para ser reproducida por voz mediante el sistema tts del robot Tiago.

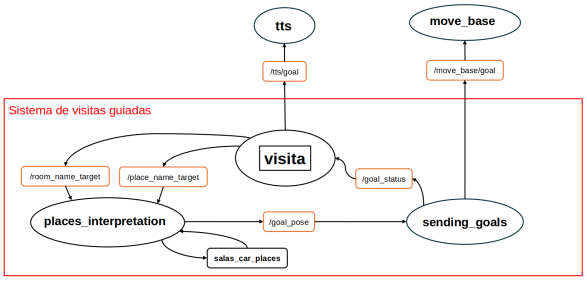


Figura 1: Arquitectura ROS para la gestión de las visitas guiadas .

Sobre dicha estructura inicial, se han integrado las siguientes capacidades:

1. Configuración dinámica del idioma (variable idioma_visita = “es” o “en”)
2. Carpetas independientes para almacenar diálogos en español e inglés
3. Se ha ajustado el tiempo de pausa tras alcanzar un punto de destino (proporcional a la longitud del diálogo reproducido)
4. Se han añadido dos nuevos topics en los que publica el nodo:
   - /visita/language : se publica el idioma seleccionado (“es” o “en”)
   - /visita/status : se publica el inicio (“started”) y fin (“finished”) de la visita
5. Se impide que el texto de despedida interrumpa otros diálogos (suscripción al topic /tts/status)

3.2. Detección de objetos y personas e introducción de saludos

Respecto a la integración del sistema de detección de objetos y personas, debido a la existencia de latencias por falta de recursos computacionales para desplegar algoritmos muy demandantes, se optó por utilizar modelos ligeros con el menor número posible de clases de detección. De este modo, se utilizó el conjunto de paquetes Yolov8_ros Zheng (n.d.), integrando el nodo yolov8_ros_compressed con imágenes comprimidas en formato JPEG, y se decidió reducir el número de clases mediante el uso de modelos YOLO-World Ultralytics (n.d.) customizados (entrenados con el dataset COCO). El modelo utilizado se ha configurado para que detecte exclusivamente 3 clases de las 80 disponibles: “person”, “chair” y “dining table”. La Fig. 2 muestra un ejemplo de detección en el laboratorio con dicho modelo custmizado. La Fig. 3 muestra un ejemplo de detección de persona. La detección de mesas y sillas está disponible para nuevos desarrollos pero actualmente no se está utilizando para modificar el comportamiento del robot.

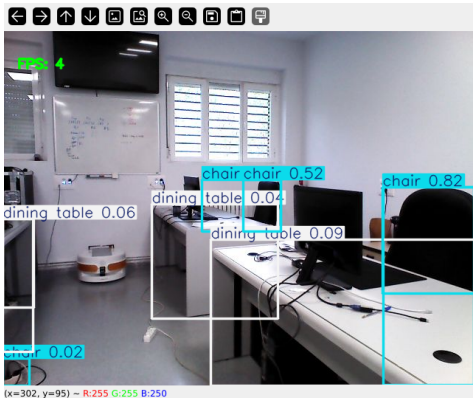


Figura 2: Ejemplo de resultado de detección de mesas y sillas con el modelo customizado.

A partir de la salida del detector, el nodo *object_detector* envía la reproducción de un saludo cuando se detecta a una persona que cumple una serie de condiciones de clase, tamaño, posición y probabilidad.

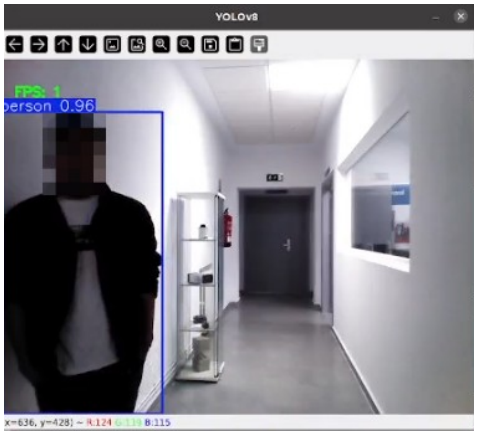


Figura 3: Ejemplo de resultado de detección de persona

El texto de saludo varía en función del período del día y del idioma (ver Tabla 1). El período se establece según los siguientes criterios:

- Si la hora del reloj ≤ 13:00 (hasta 12:59), periodo = "morning".
- Si 13:00 < hora del reloj < 18:00, periodo = "afternoon".
- Si la hora del reloj ≥ 18:00, periodo = "evening".

El robot no va a saludar si se detectan al mismo tiempo más de 3 personas con probabilidad ≥ 50 % y altura de caja ≥ 200 px. Además, el nodo no publica mensajes de saludo en mientras se esté reproduciendo otro texto por los altavoces del robot. Por último, el robot solo va a saludar cuando se encuentra en desplazamiento.

Tras reproducir un saludo, se desactiva temporalmente esta funcionalidad para evitar que se esté saludando repetidamente a una misma persona.

4. Pruebas y resultados

4.1. Robot y entorno de las visitas

El robot utilizado es el robot TIAGo, de la empresa PAL Robotics . TIAGo es un robot móvil de servicio con aspecto humanoide y diseñado para trabajar en entornos interiores. Combina capacidades de movilidad, percepción, manipulación e interacción. Además, es configurable a las necesidades del usuario, tanto a nivel de componentes como a nivel de software.

La base móvil de TIAGo cuenta con un mecanismo diferencial de movimiento, lo cual proporciona movilidad al robot con una velocidad máxima de 1 m/s. Además del mecanismo diferencial, la base contiene un ordenador de a bordo, la batería, un conector de corriente, un sensor LIDAR, 3 sensores sonar, un panel de usuario, un panel de servicio y conexión WiFi.

La cabeza de TIAGo tiene dos grados de libertad (giro e inclinación). En la cabeza se localiza el altavoz, los micrófonos y una cámara RGB-D. La imagen que captura la cámara RGB de TIAGo tiene una resolución de 480×640 píxeles.

El sistema se ha desarrollado principalmente para realizar visitas guiadas en el edificio del Centro de Automática y Robótica (CAR) de la Escuela Técnica Superior de Ingenieros Industriales (ETSII) de la Universidad Politécnica de Madrid (UPM). El trabajo se centra en recorrer la planta inferior, donde se encuentran los laboratorios. La distribución de la planta inferior consiste en : 4 laboratorios, 1 sala de prácticas, 3 despachos y los baños. El objetivo, es poder visitar los distintos laboratorios aportando información sobre ellos y proyectos llevados a cabo. La Fig. 4 muestra el mapa del laboratorio construido utilizando SLAM y el etiquetado de salas y puntos de interés para la realización de las visitas guiadas.

4.2. Pruebas iniciales

4.2.1. Reproducción de saludos

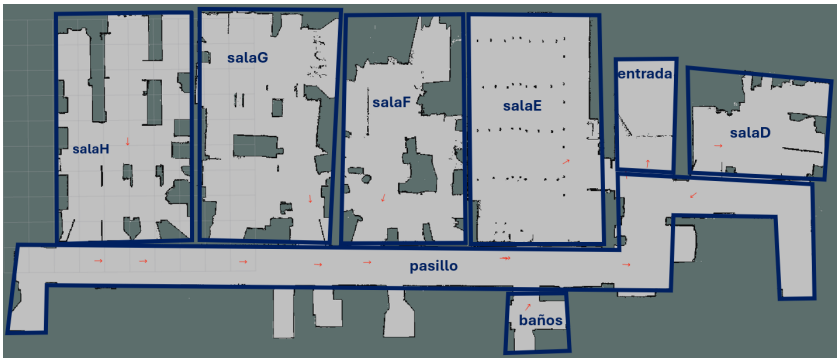
Se realizaron 4 visitas de prueba para 3 valores distintos del parámetro que establece la mínima altura en píxeles de la detección de personas. Los valores seleccionados para las pruebas se eligieron considerando la existencia de reflejos de las personas en los cristales, la presencia de figuras en algunos carteles de las salas y teniendo en cuenta valores de distancia de interacción realistas. En todas las pruebas se repitió el mismo recorrido: centroH → entradaH → vitrina3 → entradaH → centroH. Dos personas se posicionaron en varias ubicaciones a lo largo del trayecto, con el objetivo de verificar si se reproducía correctamente el saludo al detectarlas. Esto permitió observar y comparar el comportamiento del sistema en las mismas condiciones, con distintos valors del parámetro. Durante estas pruebas también se verificó que el robot no reproduce ningún saludo si no se está desplazando.

En la Fig. 5 se muestra el mapa con las ubicaciones en las que se situaron las dos personas durante las pruebas.



Figura 5: Posición de las personas durante las pruebas. Las letras que acompañan a cada punto indican el orden de paso del robot por dichas posiciones. Además, se ha utilizado un código de colores para diferenciar los trayectos de ida y vuelta: el color negro representa los encuentros que se producen en el trayecto de ida desde centroH hasta vitrina3, mientras que el color rojo corresponde a los encuentros durante el trayecto de regreso a centroH.

La tabla 2 muestra el comportamiento del sistema en las 4 pruebas realizadas para un valor mínimo de altura de detección igual a 320 píxeles. Poco antes de llegar al punto B, el robot se cruzó con una persona ajena a las pruebas, a la que saludó. Posteriormente pasó junto a una de las personas objetivo sin reproducir un nuevo saludo, al no haber transcurrido el tiempo suficiente para la reactivación de la funcionalidad. Se observó que con este valor el robot saluda a las personas detectadas desde



(a) División del mapa del entorno en diferentes salas.



1: centroH 2: entradaH 3: vitrina4 4: vitrina3 5: centroG 6: entradaG 7: entradaF 8: centroF 9: vitrina2 10: vitrina1
11: baños 12: escalera 13: centroE 14: entradaE 15: entradaprincipal 16: entradaD 17: centroD

(b) Etiquetado de los puntos de interés de las distintas salas.

Figura 4: Etiquetado del mapa para definir defierentes salas y puntos de interés en ellas.

distancias demasiado lejanas, lo que provoca que puedan darse casos en los que se salude 2 veces a un mismo individuo que permanece inmóvil en una posición. Por tanto, se considera que 320 píxeles es un umbral de altura demasiado pequeño.

Tabla 2: Resultados de reproducción de saludos para un valor mínimo de altura de detección igual a 320 píxeles.

	A	B	C	D	E
1	Sí (2 veces)	Sí	Sí	Sí	Sí
2	Sí	Sí	Sí	Sí	Sí
3	Sí	No	Sí	Sí	Sí
4	Sí (2 veces)	Sí	Sí	Sí	Sí

La tabla 3 contiene los resultados de las 4 pruebas realizadas con un umbral mínimo de altura de detección de 410 píxeles. Se ha observado que un 40 % (8/20) de los saludos programados no se reproducen debido a que la detección no alcanza la altura mínima establecida.

Tabla 3: Resultados de reproducción de saludos para un valor mínimo de altura de detección igual a 410 píxeles.

	A	B	C	D	E
1	Sí	No	No	Sí	No
2	Sí	No	Sí	Sí	No
3	No	Sí	Sí	Sí	No
4	Sí	No	Sí	Sí	Sí

En la tabla 4 se puede observar que para un umbral mínimo de altura de detección de 370 píxeles el robot ha ejecutado satisfactoriamente la gran mayoría de los saludos programados en las ubicaciones previstas, durante las 4 pruebas realizadas.

Tabla 4: Resultados de reproducción de saludos para un valor mínimo de altura de detección igual a 370 píxeles.

	A	B	C	D	E
1	Sí	Sí	Sí	Sí	Sí
2	Sí	Sí	Sí	Sí	Sí
3	Sí	Sí	No	Sí	Sí
4	Sí	No	Sí	Sí	Sí

Durante la primera prueba, mientras el robot se encontraba en el punto entradaH reproduciendo el diálogo correspondiente, una persona ajena a la prueba pasó por delante de él. Aunque se cumplieron las condiciones, no se inició ningún texto de saludo y el diálogo finalizó por completo sin interrumpirse, tal y como estaba previsto.

En la tercera prueba el robot saludó a una persona ajena a la prueba poco antes de cruzarse con la persona de la prueba situada en la ubicación C, por lo que se considera que dicho saludo es válido para ambas personas y no supone un error de comportamiento.

Durante la cuarta prueba, el robot se encontró con una persona ajena a la prueba en el pasillo, justo antes de llegar al punto B. Saludó correctamente a dicha persona y, posteriormente, no reprodujo ningún saludo al pasar junto a B, por lo que no se considera fallo. A continuación, esta misma persona se quedó

a observar el robot y se posicionó muy cerca del punto D. El robot la saludó correctamente al pasar a su lado (en este caso, sí transcurrió tiempo suficiente desde C para reactivar el saludo) y, tras detenerse en entradaH y reproducir el diálogo correspondiente, saludó de nuevo al reanudar la marcha y pasar junto a D.

Con este valor del parámetro, las condiciones para saludar se cumplen a una distancia óptima entre el robot y las personas detectadas, de modo que los saludos se reproducen aproximadamente en el momento en que el robot pasa junto a la persona, lo que mejora la naturalidad de la interacción. Además, la distancia a la que saluda el robot es menor que con 320 px, lo que contribuye a evitar que el robot, durante su desplazamiento, salude dos veces consecutivas a una misma persona que permanece inmóvil.

Tras evaluar diferentes configuraciones, los parámetros utilizados en las pruebas finales para activar los saludos son los siguientes:

- Class = “person”
- (ymax – ymin) ≥ 370 px
- xmax ≤ 240 px ó xmin ≥ 400 px
- probability ≥ 0,50

4.2.2. *Realización de visitas*

Antes de efectuar las visitas reales con público externo se llevaron a cabo diferentes pruebas y ensayos, en los que se registraron algunas medidas relevantes referentes a las mismas, para diferentes recorridos (ver Tabla 5. La primera fila corresponde a las pruebas de saludo presentadas en la sección anterior, la segunda fila corresponde a un ensayo para las visitas reales de la subsección 4.3.1 y la tercera fila corresponde a un ensayo de la visita real presentada en la subsección 4.3.2.

Tabla 5: Medidas tomadas durante diferentes recorridos de visita.

	Tiempo total (min)	Tiempo en movimiento (min)	Distancia (m)	Velocidad media (m/s)
1	4.02	1.20	23.34	0.32
2	3.38	0.99	22.49	0.38
3	5.42	1.43	32.18	0.38

4.3. *Operación real*

A continuación se describen las experiencias correspondientes a visitas con público real.

4.3.1. *Jornadas de Puertas abiertas*

Durante dos días de Puertas Abiertas en la Universidad, se realizaron 4 visitas guiadas. Todas las visitas fueron en español y el público asistente estuvo compuesto por estudiantes de bachillerato y sus familiares. El primer día se organizaron 2 visitas, cuyo recorrido total se dividió en dos partes:

- 1ª parte: entradaprincipal → entradaD → centroD
- 2ª parte: vitrina1 → vitrina3 → entradaH → centroH



Figura 6: Realización de una visita guiada en las jornadas de puertas abiertas.

El segundo día se organizaron otras 2 visitas, utilizando los mismos diálogos y configuraciones que el día anterior, pero con un recorrido mucho más corto:

- Recorrido del segundo día: entradaprincipal → entradaD → centroD

Al igual que en la jornada anterior, las visitas se desarrollaron sin ningún tipo de incidencia.

En esta ocasión, los grupos asistentes fueron más reducidos (5-10 personas por grupo). Esto contribuyó a que la experiencia fuera más fluida y agradable, tanto para los responsables de la actividad como para el público.

4.3.2. *Visita del personal de una empresa*

Se organizó una visita guiada breve para algunos integrantes de una gran empresa relacionada con el sector de la robótica y la automatización. Los visitantes recorrieron una parte de las instalaciones del laboratorio acompañados por el robot TIAGo, que actuó como guía durante el trayecto. La visita se realizó en inglés para un grupo de 7 asistentes, y el recorrido fue el siguiente:

- Recorrido de la visita de empresa: escalera → vitrina1 → vitrina3 → entradaH → centroH



Figura 7: Realización de una visita guiada con personal de una empresa del sector.

En todas las visitas realizadas se observó una percepción positiva de la experiencia, caracterizada por la curiosidad y el interés de los asistentes hacia el robot. Se ha puesto de manifiesto que una gran cantidad de personas responden activamente al saludo del robot, mostrando notable interés y fascinación antes su presencia. No se detectaron signos de rechazo o incomodidad en ningún momento durante las visitas.

5. Conclusiones y trabajos futuros

Este trabajo ha conseguido integrar componentes avanzados de navegación, reconocimiento de objetos e interacción social para la realización de visitas guiadas reales en el entorno de un laboratorio de investigación. La realización de visitas guiadas a un laboratorio con un robot interactivo que saluda ha mejorado notablemente la experiencia de los asistentes, aumentando su interés y confianza en los trabajos de investigación realizados en el mismo. Se recomienda realizar las visitas con grupos de entre 5-10 personas, con el fin de garantizar una experiencia más fluida y agradable para los asistentes. Adicionalmente, esta medida contribuye a evitar aglomeraciones, mejora la interacción durante el recorrido y ayuda a prevenir que los visitantes obstruyan la ruta del robot durante sus desplazamientos entre los puntos de destino.

Como próximos trabajos futuros se plantea la evaluación extensiva del sistema de visitas, incorporando nuevas métricas y cuestionarios registrados en un mayor número de visitas reales. También resulta de interés añadir lenguaje corporal las explicaciones y saludos del robot. Otra funcionalidad muy útil sería que el robot reconociera cuándo su camino está bloqueado por un objeto o una persona y reaccionara de forma autónoma, solicitando ayuda para que se retire el obstáculo o pidiendo amablemente a la persona que se aparte. Para ello, se podría aprovechar la detección de objetos ya implementada, añadiendo gestos como mover la cabeza hacia el obstáculo (persona u objeto) o hacia atrás para pedir ayuda.

Agradecimientos

Esta publicación es parte del proyecto de I+D+i “BRAINS: Building Real-world Applications with Intelligent, Neural, and Sustainable approaches” PID2024-155948OB-C54 financiado por MICIU /AEI /10.13039/501100011033 / FEDER, UE. Esta actuación también ha sido financiada mediante el programa de actividades de I+D con referencia TEC-2024/TEC-62 y acrónimo iRoboCity2030-CM, concedido por la Comunidad de Madrid a través de la Dirección General de Investigación e Innovación Tecnológica a través de la Orden 5696/2024. Zhang Zhenhao agradece la financiación del China Scholarship Council (CSC), Grant No.202408320082.

Referencias

Alcantara, C. B., Júnior, A. O., Costa, J., Ricken, L., Mendes, A., Leitao, P., 2024. Artificial intelligence-based user interaction module for autonomous mobile service robots. In: 2024 IEEE 3rd Industrial Electronics Society Annual On-Line Conference (ONCON).

Avelino, J., Moreno, P., Correia, F., et al., 2025. Evaluating greeting models for human-robot engagement during first encounters. International Journal of Social Robotics 17, 297–314.
URL: https://doi.org/10.1007/s12369-025-01213-w

Barranco, M., 2025. Integración de módulos de reconocimiento y gestión de visitas guiadas con un robot móvil interactivo. Trabajo Fin de Grado. Grado en Ingeniería en Tecnologías Industriales (GITI).

de la Puente, P., Bajones, M., Reuther, C., Wolf, D., Fischinger, D., Vincze, M., 2018. Robot navigation in domestic environments: experiences using RGB-D sensors in real homes. Journal of Intelligent & Robotic Systems.
URL: https://doi.org/10.1007/s10846-018-0885-6

de Mello, J. C., Secci, G. P. N., Ribeiro, P. H., 2024. Robotics and ai in museums – the future of the present. Robotic Systems and Applications 4 (2), 44–58.

Enoki, M., Inaishi, T., Noguchi, H., 2022. Extraction and evaluation of greeting speech-timing and characteristic upper body motion for robots to gain attention of older adults. Journal of Robotics and Mechatronics 34 (6), 1338–1347.
URL: https://doi.org/10.20965/jrm.2022.p1338

Garcia-Fernandez, O., 2024. Desarrollo de visitas guiadas con un robot interactivo. Trabajo Fin de Grado. Grado en Ingeniería en Tecnologías Industriales (GITI).

Nasiri, N., Saleem, A., 2025. The museum guide robot: Enhancing visitor experience through autonomous robotics. In: 2025 International Conference for Artificial Intelligence, Applications, Innovation and Ethics (AI2E).

Øllgaard, J. A., Bundgaard, K., Sorag Kjaer, M., Bodenhagen, L., Palinko, O., 2024. Utilizing a social robot as a greeter at a children’s hospital. In: Social Robotics. ICSR + AI 2024. Springer Nature Singapore.

Rosa, S., Randazzo, M., Landini, E., Bernagozzi, S., Sacco, G., Piccinino, M., Natale, L., 2024. Tour guide robot: a 5g-enabled robot museum guide. Frontiers in Robotics and AI Volume 10 - 2023.

Rudaz, D., Tatarian, K., Stower, R., Licoppe, C., 2023. From inanimate object to agent: Impact of pre-beginnings on the emergence of greetings with a robot. J. Hum.-Robot Interact. 12 (3).

Scheideger, C. E., Fisher, E. L., Wince, K., Pond, E., 2025. Design of a greeting robot. In: 2025 ASEE North Central Section (NCS) Annual Conference.

Todkari, V. C., Kaldate, A. P., Kolhar, S., Jagtap, A., Sable, N. P., 2025. Ai-driven visual navigation for smart lab tour guide robot. Journal Européen des Systèmes Automatisés 58 (12).

Ultralytics, n.d. Yolo-world model. Documentación oficial.
URL: https://docs.ultralytics.com/es/models/yolo-world/

Zheng, Z., n.d. Yolov8_ros. Repositorio de GitHub.
URL: https://github.com/qq44642754a/Yolov8_ros