

Article

Development of an Autonomous and Interactive Robot Guide for Industrial Museum Environments Using IoT and AI Technologies

Andrés Arteaga-Vargas , David Velásquez , Juan Pablo Giraldo-Pérez  and Daniel Sanin-Villa * 

Área de Industria, Materiales y Energía, Universidad EAFIT, Medellín 050022, Colombia; aarteagav@eafit.edu.co (A.A.-V.); dvelas25@eafit.edu.co (D.V.); jgiral95@eafit.edu.co (J.P.G.-P.)

* Correspondence: dsaninv2@eafit.edu.co

Abstract

This paper presents the design of an autonomous robot guide for a museum-like environment in a motorcycle assembly plant. The system integrates Industry 4.0 technologies such as artificial vision, indoor positioning, generative artificial intelligence, and cloud connectivity to enhance the visitor experience. The development follows the Design Inclusive Research (DIR) methodology and the VDI 2206 standard to ensure a structured scientific and engineering process. A key innovation is the integration of mmWave sensors alongside LiDAR and RGB-D cameras, enabling reliable human detection and improved navigation safety in reflective indoor environments, as well as the deployment of an open-source large language model for natural, on-device interaction with visitors. The current results include the complete mechanical, electronic, and software architecture; simulation validation; and a preliminary implementation in the real museum environment, where the system demonstrated consistent autonomous navigation, stable performance, and effective user interaction.

Keywords: Industry 4.0; interactive museums; autonomous navigation; mobile robot; collaborative robotics; generative AI



Academic Editor: Ognjen Arandjelović

Received: 18 October 2025

Revised: 13 November 2025

Accepted: 26 November 2025

Published: 1 December 2025

Citation: Arteaga-Vargas, A.; Velásquez, D.; Giraldo-Pérez, J.P.; Sanin-Villa, D. Development of an Autonomous and Interactive Robot Guide for Industrial Museum Environments Using IoT and AI Technologies. *Sci* **2025**, *7*, 175. <https://doi.org/10.3390/sci7040175>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Recent advances in mobile robotics and Artificial Intelligence (AI) have led to the development of autonomous systems, which are being applied in multiple fields [1], one of them being interactive museums to capture and maintain visitors' attention [2–6]. Although the use of robots in museums is not a recent development, the latest technology has opened up new capabilities in this type of robot [7], allowing them to overcome the main challenges they encounter, such as navigating dynamic obstacles [8] and communicating with visitors [9]. In this context, this research aims to present the development of an autonomous and interactive robot guide for use in a museum setting within a motorcycle assembly plant in Medellín, Colombia. This robot presents a set of particular challenges in its development, the main being precise autonomous navigation in a changing environment, where the presence of people and unforeseen obstacles can affect the robot's performance. To address this challenge, technologies such as navigation sensors (LiDAR and cameras), route-planning algorithms, and artificial vision systems will be integrated. In addition, an Internet of Things (IoT) platform was implemented to monitor and analyze the robot's data. The primary objective is to develop and test a design that achieves a Technology Readiness

Level (TRL) of 6 [10], ensuring that the robot's navigation capabilities, interactions with visitors, and overall system reliability are validated in a real-world environment.

In addition to its technical and research-oriented motivations, this project also addresses a practical challenge faced by the hosting institution. At present, the museum tours rely heavily on the availability and performance of human guides, which introduces significant variability in the visitor experience. The information provided during the tours often depends on each presenter's personal interpretation, level of expertise, and communication style, as there is no standardized script or protocol to ensure uniformity across sessions. This inconsistency can lead to differences in the depth, accuracy, and engagement of the explanations given to visitors. Therefore, the primary goal of the project is to develop an autonomous robotic system capable of delivering guided tours in a standardized, engaging, and reliable manner, reducing reliance on human personnel while ensuring a consistent and high-quality dissemination of information throughout all visits.

Beyond addressing these practical challenges, the project also contributes to advancing the role of interactive robotics in industrial museum environments. The deployment of interactive robotic guides in museum settings represents an emerging, highly relevant field of application for intelligent systems, particularly when embedded in industrial contexts. A review of current literature reveals that, while autonomous robots have been explored for cultural, educational, and social settings, there is a notable absence of documented implementations in industrial museums or similar semi-structured environments. This gap underscores the novelty and importance of the present work, which not only addresses a practical institutional need but also contributes new insights into the integration of autonomous navigation, human–robot interaction, and IoT-based monitoring in industrial spaces. The real-world development within a motorcycle assembly plant not only validates the proposed architecture but also establishes a reference framework for future developments in industrial interactive robotics.

The remainder of this paper is organized as follows: Section 1 reviews the related work, and theoretical background. Section 2 describes the methodology and system design. Section 3 presents the experimental setup, results and discusses the findings with their implications. Finally, Section 4 concludes the paper and outlines directions for future research.

1.1. State of the Art

Autonomous navigation in guide robots for museums has been an active research area in recent decades, with significant advances in the integration of localization, mapping, path planning, and obstacle avoidance technologies [2,5,7,11].

Social navigation is an essential component for museum guide robots. Social navigation involves not only the ability to move autonomously but also the capability to interact safely and respectfully with visitors, adhering to social norms such as personal distance (proxemics). Robots must be able to avoid dynamic obstacles (such as people) and follow predefined routes or generate routes in real-time [5,7]. Moreover, recent work highlights the potential of hybrid human–machine design, where generative AI supports natural and adaptive interaction in social contexts [12]. In parallel, recent reviews on multi-robot systems and soft robotics emphasize the importance of distributed coordination, adaptive control, and compliant interaction in unstructured environments, identifying challenges such as nonlinear dynamics and hyper-redundant configurations [13].

1.1.1. Localization and Mapping

Sensors such as laser scanners (LIDAR) and RGB-D cameras are used for localization and mapping. One of the most widely used algorithms for localization and mapping is

SLAM (Simultaneous Localization and Mapping), especially in some off-the-shelf packages of the ROS (Robot Operating System) framework [14,15].

1.1.2. Path Planning

For path planning, algorithms such as A* (also known as A-star) are used for global planning. At the same time, Social Momentum is applied for human avoidance, and ORCA (Optimal Reciprocal Collision Avoidance) is used when other robots are present. These algorithms enable the robot to navigate safely, avoiding collisions with both static and dynamic objects [15,16].

1.1.3. Obstacle Avoidance

3D obstacle avoidance is crucial to ensuring the safety of both the robot and visitors. The robots use a map generated from RGB-D camera data, which is projected onto a 2D plane or floor and combined with the laser sensor data for path planning. This approach enables the robot to detect obstacles above ground level, such as tables or shelves, which is particularly useful in museum environments where objects may be at different heights [7,17]. Recent work on multimodal object detection using depth and image data for manufacturing parts has shown that combining RGB information with depth cues improves detection robustness in structured industrial environments, which is closely aligned with the use of RGB-D sensing in the proposed system for reliable perception in a factory-museum setting [18].

Millimeter-wave (mmWave) radar has recently emerged as a powerful sensing modality for robotic navigation and obstacle avoidance, offering robustness in environments where optical sensors face limitations such as visual clutter, transparent surfaces, or partial occlusions [19]. Unlike vision-based systems, mmWave provides reliable depth perception and motion tracking through radio frequency reflections, enabling robots to detect and localize obstacles in three dimensions, even through materials like glass or thin walls, with high resilience to noise. As highlighted in Kyle et al. [20], mmWave sensing is increasingly being integrated into mobile robots to complement LiDAR and camera systems, enhancing safety and adaptability in dynamic, cluttered spaces. This makes it a promising technology for next-generation 3D obstacle avoidance frameworks.

1.1.4. LLM in Robotics

Human–robot interaction (HRI) has been extensively studied across diverse domains [21], particularly with the recent breakthroughs in large language models (LLMs), which have greatly expanded the capacity of robots to engage with humans in natural, context-aware, and emotionally responsive ways. Baijun et al. [22] introduced an affection model into a robotic conversational agent by fine-tuning a large-scale pre-trained language model on empathetic dialogue datasets. Their approach enhanced sentiment analysis capabilities, allowing robots not only to generate contextually appropriate responses but also to classify emotions with estimations of arousal and valence levels. Complementing this, Liu et al. [23] explored the integration of LLMs into robotic systems for multimodal perception and decision-making, highlighting their potential to unify natural language understanding with real-world robotic control. Together, these works demonstrate how LLMs are reshaping the state of the art in robotics by bridging affective computing, dialogue systems, and embodied intelligence, paving the way for more adaptive, socially aware, and autonomous robotic agents.

1.1.5. Commercial Robots

Commercial robots designed for guiding functions in museums and exhibitions already exist. One such example is Promobot, a humanoid robot equipped with speech capabilities,

facial expressions, and interactive features. Promobot can autonomously navigate through exhibitions, providing additional information via a touchscreen display on its chest [24].

Although advancements have been made in robotic systems for museums, current applications are limited to controlled settings with predictable layouts and minimal human interaction. Existing solutions frequently show insufficient resilience in dynamic, semi-structured industrial environments like assembly facilities modified for public visits, where human movement and alterations in space often occur. Additionally, only a limited number of systems combine social navigation, multi-modal perception, and IoT-driven monitoring into one operational prototype. This initiative addresses these shortcomings by developing an autonomous robotic guide that features socially mindful navigation, 3D obstacle recognition, and real-time connectivity for data oversight. In the Latin American context, these developments align with the opportunities and challenges of Industry 4.0, where robotics, IoT, and AI are considered key enablers for bridging technological gaps and promoting innovation in cultural and industrial environments [25].

Recent literature also emphasizes that progress in autonomous robotics depends not only on navigation algorithms, but also on the integration of emerging technologies and collaborative innovation ecosystems. These elements enable predictive maintenance and data-driven decision-making in dynamic environments, aspects that are equally relevant in the context of museum guide robots [26].

The innovation of this study lies in the convergence of Industry 4.0 technologies within a single robotic platform operating in a semi-structured industrial museum. The integration of LiDAR, RGB-D, and mmWave sensors provides complementary spatial perception under reflective and dynamic indoor conditions, while the inclusion of a locally deployed large language model enables natural interaction without reliance on cloud services. This combination allows the robot to navigate, communicate, and operate autonomously in real environments characterized by uncertainty and human presence.

2. Materials and Methods

The project will be based on the Design Inclusive Research (DIR) methodology [27], which provides a scientific approach to engineering problems, enabling product-oriented projects to be developed with scientific rigor. The methodology consists of 9 steps, which can be divided into 3 phases. The first phase is considered the exploration phase, where knowledge is gathered, the problem is defined, and the theoretical basis is established. The second phase is the creative phase, where the designs are used as a research tool. For this phase, another methodology will be employed, specifically the VDI 2206 [28,29], a German method for mechatronic product design that focuses on iterative design. Finally, the last phase is an application and evaluation phase, where the prototype is tested, and the results of these tests are examined to draw general conclusions. These findings can then be used for future implementations.

2.1. Design Inclusive Research (DIR)

The Design Inclusive Research (DIR) methodology [30] was employed during the exploratory and hypothesis validation phases to ensure that the robot's design adequately addressed environmental and operational requirements. The process began with systematic observation to identify the main challenges of the deployment context, followed by the formulation of hypotheses concerning autonomous navigation, visitor interaction, and adaptation to dynamic environments. These hypotheses were grounded in theoretical models and principles that provided a conceptual framework for the design. Building on this foundation, initial conceptions of the robot were developed and progressively refined into detailed mechanical, electronic, and software architectures. A functional prototype was

then constructed, enabling experimental proof of the proposed functionalities. Subsequent evaluation involved testing the system in real environments, assessing its performance, and implementing necessary adjustments. Finally, the outcomes were generalized through analysis and reporting, providing insights and guidelines for future implementations.

The main steps of the DIR methodology are shown in Figure 1 as depicted in [27].

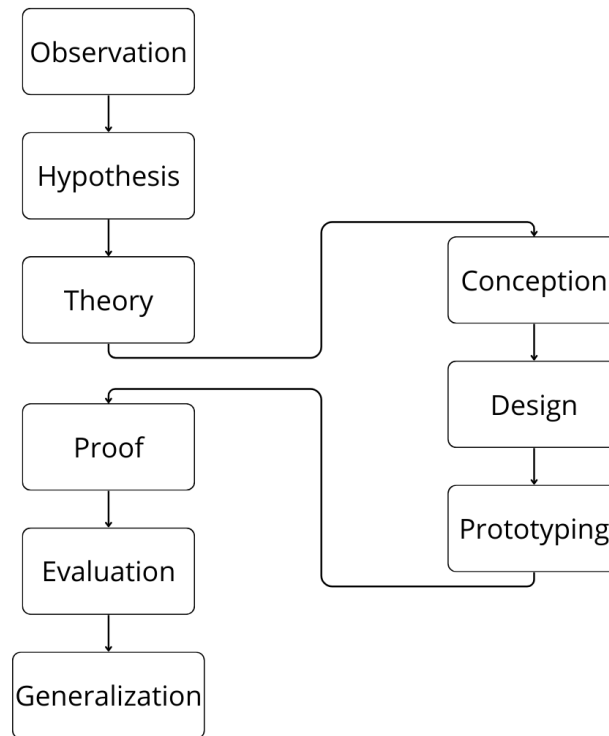


Figure 1. DIR image summary.

2.2. VDI 2206

It will guide the development of the mechatronic system in its implementation, integration, and testing, ensuring a structured and efficient approach [31,32].

The development process of the robotic system followed a structured methodology encompassing several key stages. First, the requirements were established to define the functional and technical specifications that would serve as the foundation for design and evaluation. Based on these specifications, a comprehensive system design was created to ensure effective integration across mechanical, electronic, and software domains. Each sub-system was then addressed through domain-specific design, where engineering principles were applied to optimize functionality and reliability. Following this, system integration was carried out by assembling and testing the individual components as a unified whole, with particular attention to seamless interaction between subsystems. Verification and validation activities were conducted to assess performance against the initial requirements, employing both simulations and real-world testing. Throughout the process, modeling and computational analysis played a central role, enabling iterative refinement of the design and providing insights into system behavior under different conditions. Ultimately, this structured approach culminated in the realization of the final product, representing the synthesis of requirements, design, integration, and validation into a coherent and functional robotic system.

The main steps of the DIR methodology are shown in Figure 2 as depicted in [28].

The combined use of the DIR and VDI 2206 methodologies provided both the scientific rigor and the engineering structure required for the development of a complex mechatronic system such as the autonomous museum guide robot. The DIR methodology was

selected for its robustness in bridging theoretical research and practical design, ensuring that each decision from conceptual definition to prototype validation was supported by empirical observation and iterative evaluation. Meanwhile, the VDI 2206 standard was adopted for its product-oriented nature and its emphasis on requirement driven design, which aligns closely with the expectations of an industrial client. Within this framework, system level design and domain-specific stages guided the detailed development of the mechanical, electronic, and software subsystems, ensuring their subsequent integration into a coherent architecture. Each phase was accompanied by verification and validation activities, promoting an iterative design process where feedback from simulations and physical testing directly informed improvements in component selection, control architecture, and interaction mechanisms. Overall, the joint application of these methodologies enabled a structured and traceable decision-making process that strengthened the scientific foundation and practical feasibility of the project.

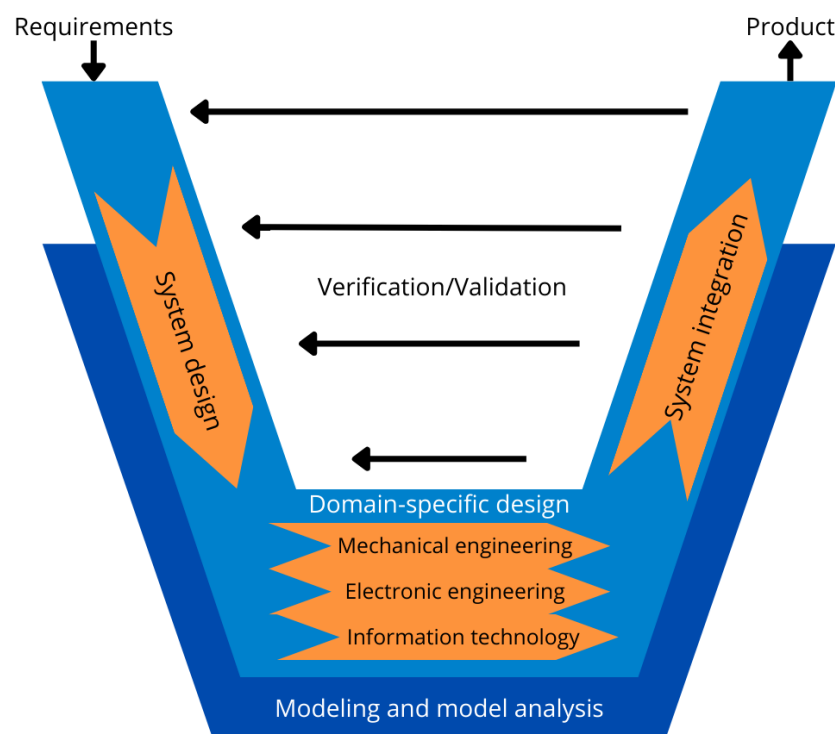


Figure 2. VDI 2206 image summary [28].

2.3. System Design

This section presents the system design of the project following the DIR and VDI2206 methodology. It includes the mechanical, electronic, and software design, as well as preliminary simulation outcomes. This design demonstrates the application of the proposed methodologies and the integration of advanced sensing and navigation technologies to address challenges in dynamic indoor environments.

2.3.1. Mechanical Design

The TurtleBot4 standard version was selected as the basis for the mechanical design, due to its competitive price and performance, as it features a lidar sensor and an RGB-D camera, which is unique in its price segment. On top of the TurtleBot three other floors of components were placed as shown in Figure 3, on the first floor there is the Jetson Orin Nano, which serves as the robot's central processing unit, and a rfid sensor, held by a 3D printing base, which will be used to identify the unique tags that each exhibit has, to then present the information of each one, the second level houses four mmWave sensors

mounted on custom 3D-printed bases, used for accurate human detection, along with LED indicators, on the top floor we have a support to use a tablet to display information on both the museum exhibits and relevant data in real time of the operation of the robot.

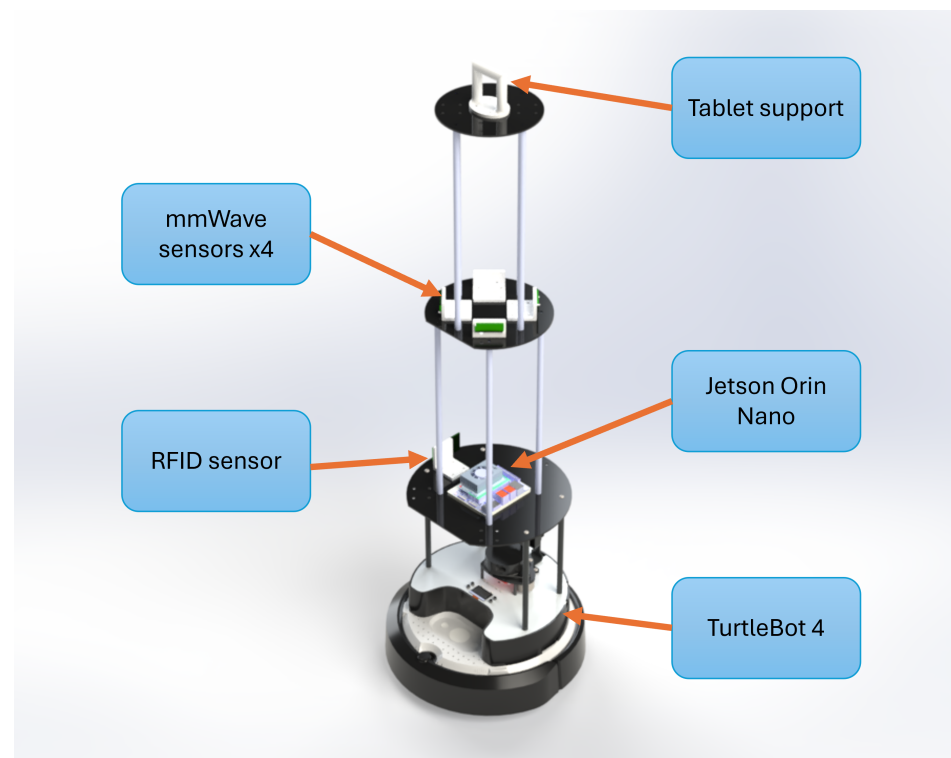


Figure 3. Mechanical design render.

2.3.2. Electronic Design

The electronic diagram is organized into seven main sections as shown in Figure 4. The Power Distribution Section describes how power is sourced from the TurtleBot outputs and distributed to the other subsystems. The mmWave Section illustrates the recommended implementation provided by the component manufacturer; its outputs are routed through a multiplexer to enable the use of a single serial port on the CPU. The LED Indicators Section employs two power MOSFETs to control an RGB LED strip. The Emergency Stop Section details the circuit that disconnects power in the event of an emergency. The TurtleBot4 I/O Section shows all connections between the TurtleBot platform and the auxiliary circuits. The RFID Section explains the power supply configuration and the serial communication interface with the TurtleBot. Finally, the Other Connections Section depicts the HDMI link between the tablet and the main CPU, the associated power supply, and the Ethernet connection with the Raspberry Pi for data communication.

2.3.3. Information Technology Design

The proposed software architecture, illustrated in Figure 5, is organized into three main subsystems: the robot control, the natural language processing (NLP) pipeline, and the cloud services. Within the robot, the architecture is based on the Robot Operating System (ROS), which coordinates various modules for perception, localization, and navigation. Sensor data from devices such as cameras, LiDAR, IMU, encoders, mmWave radar, and RFID readers is first processed by an Orchestrator, which then feeds the data to the SLAM and Navigation Algorithm modules. The Movement Controller handles the resulting navigation commands. Data is stored locally and made available for local visualization, while a Data Transmission module publishes relevant information to the cloud via MQTT. The NLP

pipeline runs on the Jetson platform inside the robot and integrates speech-to-text, a local large language model (LLM), and text-to-speech, enabling real-time natural language interaction with visitors. The tablet operates as a user interface, providing a dashboard for monitoring and modules for voice capture and audio output, connected to the robot through WebRTC. Finally, the cloud services provide remote storage and visualization, allowing administrators to monitor performance and access historical records, ensuring a scalable and distributed operation.

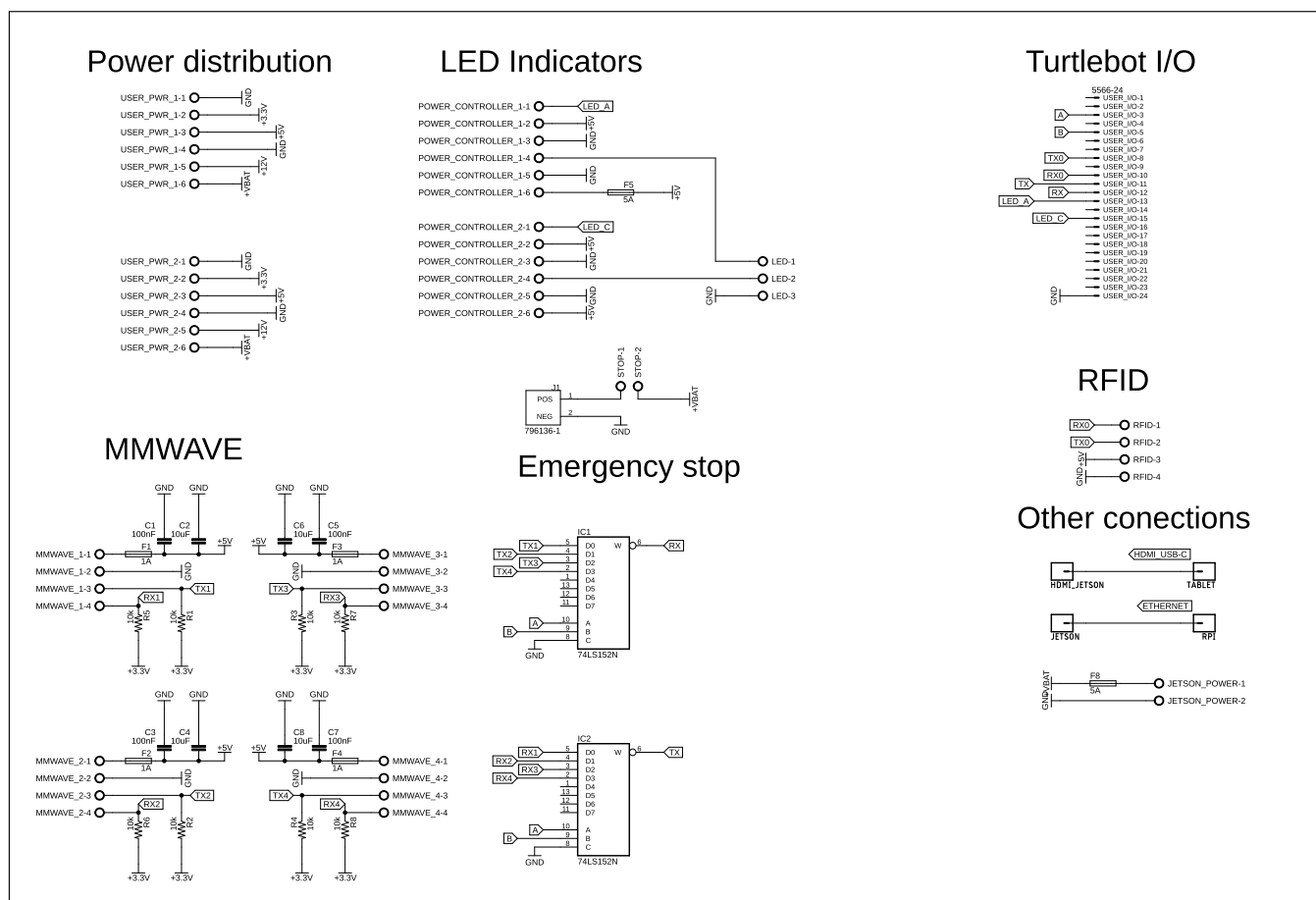


Figure 4. Electronic design.

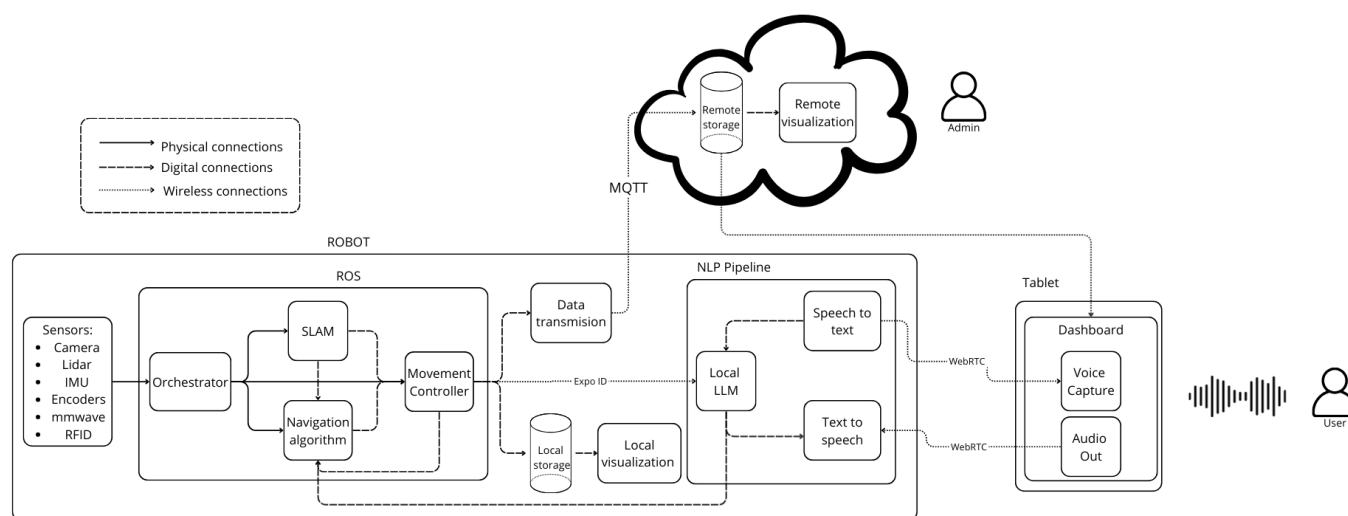


Figure 5. Software architecture.

Robot control: The robot control subsystem is organized around a finite state machine (FSM) as seen in Figure 6 implemented within the ROS 2 Orchestrator. This FSM defines the robot's operational flow through five main states: Idle, Undock, Navigate, Dock, and Done.

- Idle: Initial waiting state where the system decides the next action depending on whether the mission has just started, a navigation goal remains, or all goals have been completed.
- Undock: Triggered at the beginning of a mission, this state commands the robot to disengage from the docking station using a service client. Upon success, the FSM transitions to navigation; on failure, it returns to the docking sequence.
- Navigate: In this state, the robot publishes navigation goals through a dedicated ROS topic. The FSM remains here until feedback is received via the `/goal_status` topic. If the goal succeeds, the FSM resets retries and transitions back to Idle to evaluate the next step. If it fails, the system retries navigation up to a maximum threshold before aborting the mission and returning to Dock. During this state, the FSM invokes the ROS 2 `go_to_pose` routine, which incorporates dynamic obstacle avoidance by recalculating the path whenever new obstacles such as visitors entering the robot's trajectory. This reactive behavior allows the robot to adapt to changing conditions in real time, maintaining both safety and mission continuity during autonomous tours.
- Dock: Commands the robot to return to and connect with its docking station. Successful docking leads to the Done state, while failure also results in mission termination.
- Done: Final state where the FSM halts execution, signaling that the mission has either been completed successfully or aborted due to failure.

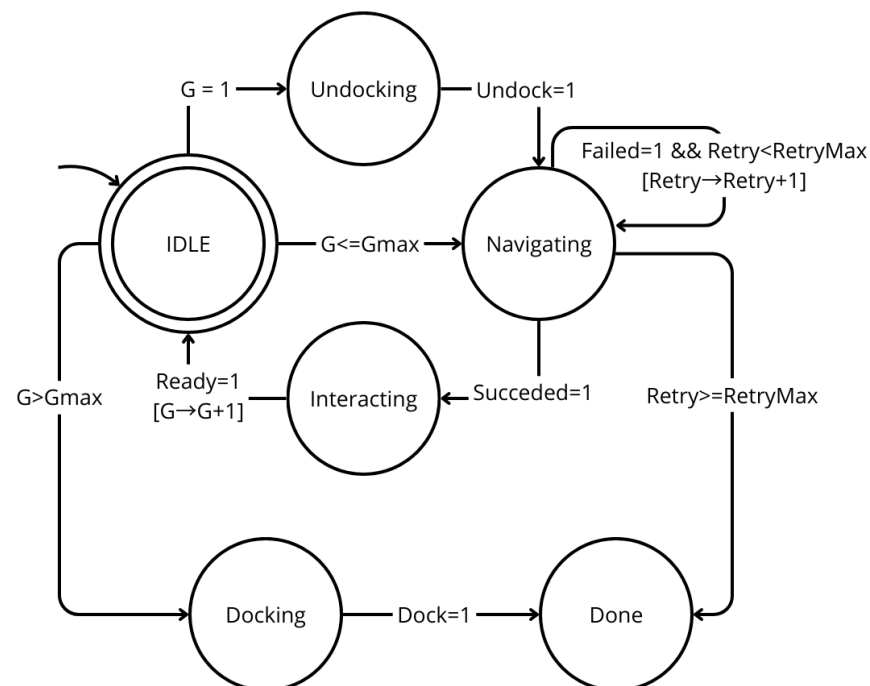


Figure 6. Finite State Machine.

The robot's sensor fusion process operates by combining multiple perception sources to enhance environmental awareness and navigation safety. Sensor data from the LiDAR, depth camera, and mmWave radar are integrated within the ROS 2 navigation stack through the obstacle layer of the local costmap. This fusion process enables the planner to generate and continuously update safe paths by detecting both static and dynamic obstacles. These data streams are time-synchronized through ROS message headers and combined at the costmap level, ensuring coherent spatial awareness for path planning. In contrast, the RFID

reader is not part of the navigation subsystem; it is solely used for exhibit identification, providing the Orchestrator and the language model with contextual information about the robot's current position.

The Orchestrator acts as the core decision-making unit: it runs the FSM, manages retries, and coordinates ROS services and topics. It communicates asynchronously with docking and undocking services, navigation modules, and the MQTT communication layer, ensuring that high-level logic is directly linked to concrete robot actions. This design makes the system modular and fault-tolerant, since each transition explicitly manages success and failure cases, allowing the robot to react adaptively to operational conditions.

Language model: The language model is implemented as a Natural Language Processing (NLP) pipeline running locally on the Jetson platform integrated into the robot. This pipeline is composed of three modules: speech-to-text (STT), a local large language model (LLM), and text-to-speech (TTS). The STT module converts user queries into text, which are then processed by the LLM using both the local knowledge base and exhibition-related information retrieved via the RFID-based Expo ID. The TTS module subsequently generates natural and expressive audio responses.

This configuration ensures low-latency interaction and independence from constant internet connectivity, while also safeguarding sensitive data by avoiding reliance on external cloud services. The tablet operates solely as a user interface, equipped with a dashboard for monitoring as well as voice capture and audio output modules. Communication between the tablet and the robot is established via WebRTC, enabling real-time bidirectional interaction with visitors in the museum environment.

Cloud services: The cloud services integrate monitoring, analytics, and control functions, acting as the central hub for both administrators and visitors. Information transmitted by the robot via MQTT is processed, stored, analyzed, and visualized in real time. As illustrated in Figures 7 and 8, the system provides two complementary dashboards:

- **Overview Dashboard:** Displays the robot's operational status, including battery level, docking state, odometry, velocity, navigation goals, and current location within the museum. It also provides historical data, such as battery trends over time, and highlights waypoint tracking to ensure mission progress.
- **Control Dashboard:** Focused on direct robot management, it allows operators to adjust motion parameters such as speed and direction, select the operating mode, and trigger emergency stop commands. Additional status indicators include bumper contact, charging state, and docking confirmation, offering a clear view of safety and mobility conditions.

These dashboards support two main user roles. For visitors, the tablet interface provides an interactive but restricted mode, where they can access the language model and multimedia content related to exhibitions. For administrators, the cloud platform offers multiple capabilities, enabling remote control of the robot, performance monitoring, and mission management.

By combining local robot intelligence with cloud-based visualization and control, the system ensures a distributed, fault-tolerant, and user-centered architecture. This hybrid approach enhances the visitor experience while providing administrators with insights and reliable supervision tools.

The language model operates as the core of the interaction system, enabling the robot to engage in natural, informative, and context-aware dialogues with visitors. Rather than serving as a general-purpose conversational agent, the model has been adapted to the museum environment, where its responses are grounded in curated exhibition information provided by the company. This ensures that all generated content aligns with institutional narratives and educational objectives. By running entirely on the onboard Jetson platform,

the model processes user queries locally, avoiding the transmission of sensitive information to external cloud services and preserving data privacy. Beyond providing factual answers, the model also contributes to the robot’s personification as a museum guide, shaping its tone and verbal behavior to deliver a coherent and engaging interactive experience for visitors.

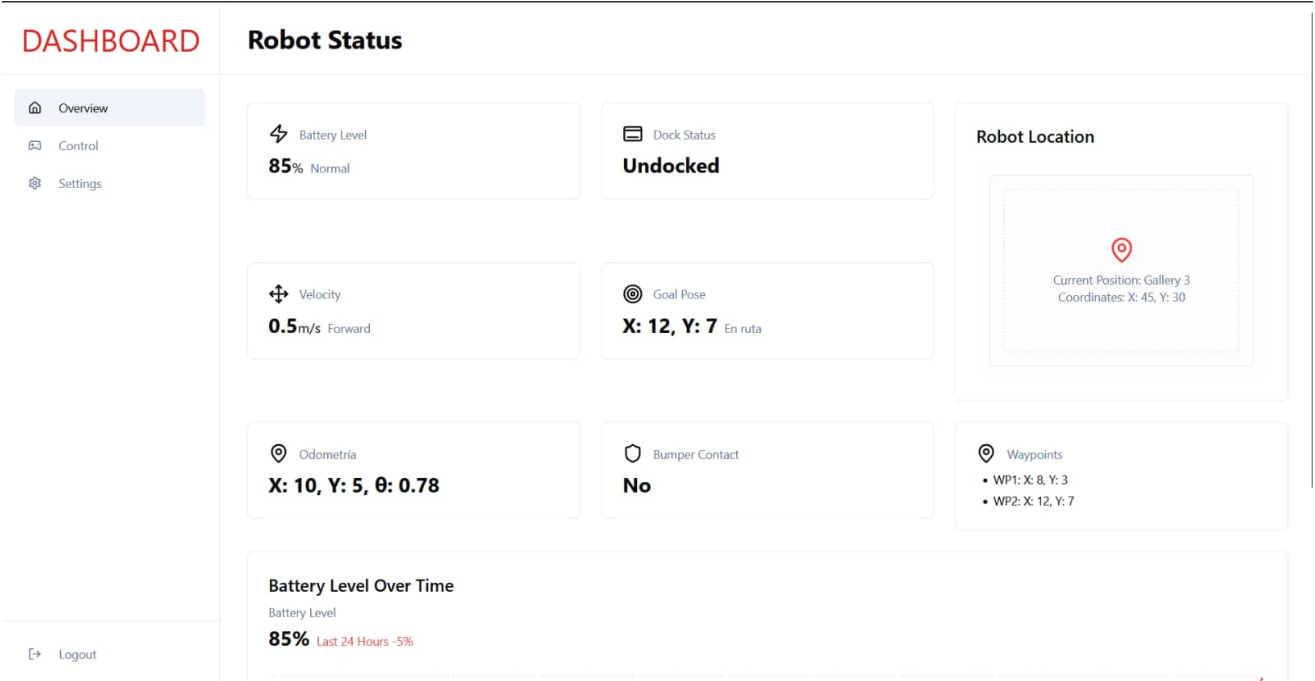


Figure 7. Dashboard overview on cloud services.

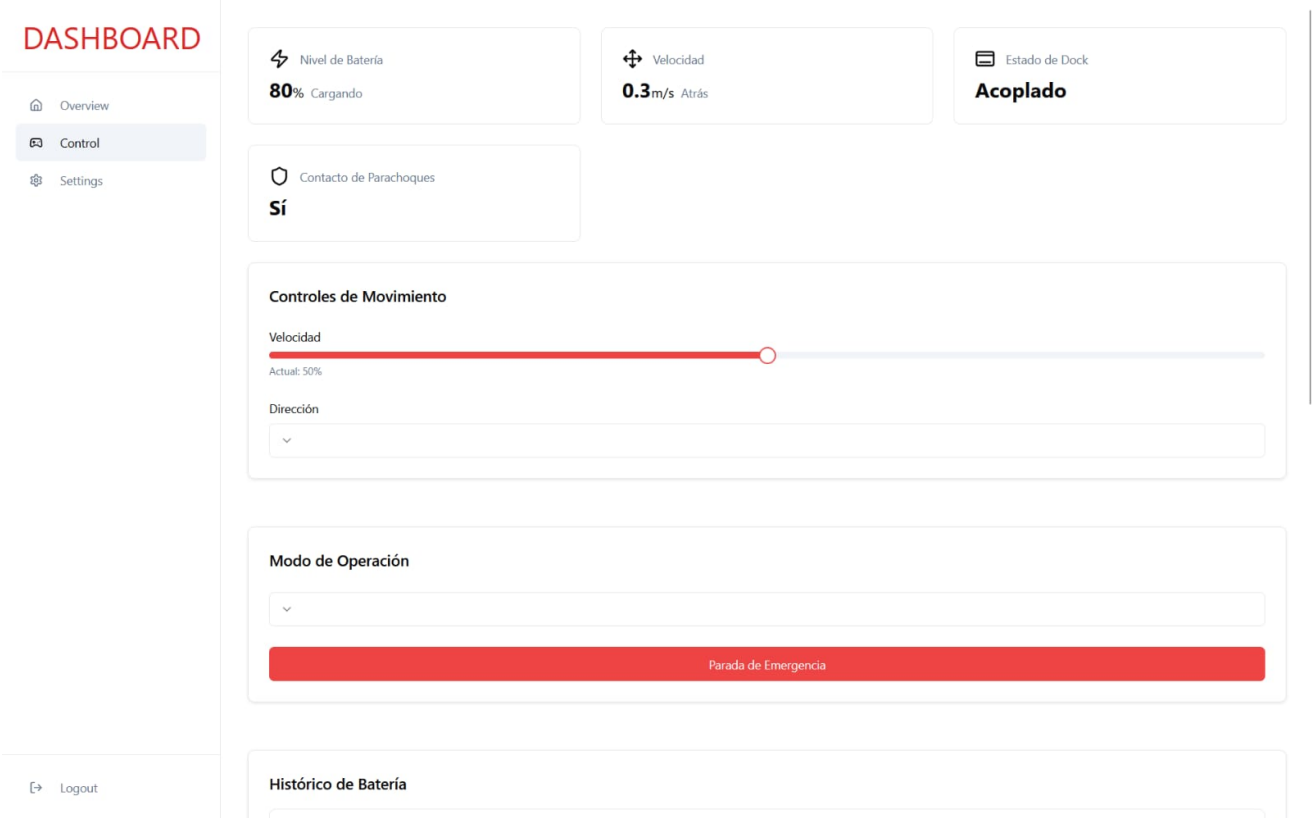


Figure 8. Control Dashboard.

This project follows a hypothesis-driven approach rooted in DIR and VDI 2206 principles. Each stage: requirements definition, domain specific design, integration, and validation; was guided by measurable research questions addressing perception robustness, navigation reliability, and user interaction. This structure elevates the work beyond implementation by ensuring reproducibility, quantification of performance, and scientific traceability.

3. Results and Discussion

This section presents the main findings of the study and examines their implications in relation to the objectives outlined earlier. The results are first introduced through simulation outcomes, which provide a controlled environment to validate the proposed approach. These are then contrasted with experimental observations and practical considerations to highlight both the strengths and limitations of the system. By combining quantitative analysis with qualitative insights, the discussion aims to contextualize the performance of the robot guide within broader scientific and technological frameworks, offering a comprehensive understanding of its effectiveness and potential applications.

3.1. Simulation

To validate the initial system design and navigation capabilities, preliminary simulations were conducted in a virtual industrial environment similar to the real museum. Figure 9 shows the 2D occupancy map used for SLAM and path planning, while Figure 10 illustrates the 3D simulated environment. The robot was deployed using ROS 2 and Gazebo, successfully performing localization, navigation, and obstacle avoidance tasks. These early tests demonstrate the feasibility of the proposed integration and provide a foundation for future work.

The integration of hardware and software components used in this implementation is summarized in Table 1. The selection was guided by the system's functional requirements, such as autonomous navigation, user interaction, and environmental perception. This setup enabled the execution of simulation tests in Gazebo using ROS 2, and demonstrated that the proposed configuration is technically viable for deployment in a real-world museum scenario.

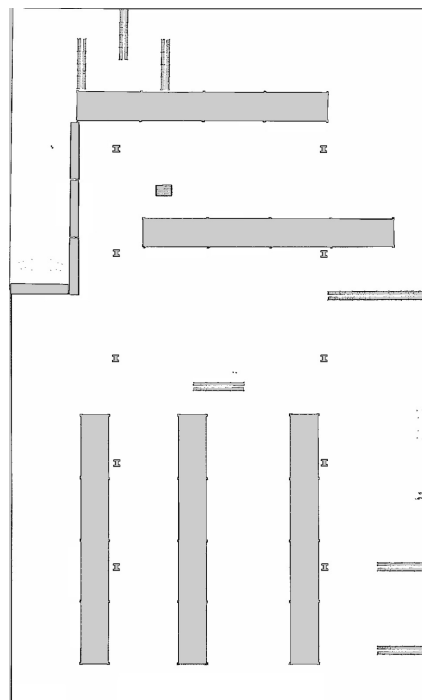


Figure 9. 2D map used for SLAM and navigation in the simulated environment.



Figure 10. Simulated 3D environment with the robot in a virtual industrial scenario.

Table 1. Hardware and Software Components.

Component	Specification/Version
Main Controller	NVIDIA Jetson Orin Nano
LiDAR	RPLIDAR-A1
Depth Camera	OAK-D-PRO
Software Stack	ROS2 Humble Hawksbill
Simulation Environment	Gazebo + RViz
Robotic platform	TurtleBot 4 Standard Version
Wireless Communication	Wi-Fi + MQTT for IoT integration
Visitors interaction	Tablet
Exhibition Identification	RFID Reader M7E-HECTO
Human Detection	Rd-03D mmWave Sensor

3.2. Initial Deployment and Validation

3.2.1. Robot Mapping and Navigation

To validate the proposed architecture under real conditions, an initial deployment was conducted at the museum exhibition space. The objective of this phase was to generate a reliable navigation map of the environment and to test the autonomous navigation capabilities of the robot in a dynamic and structured indoor scenario.

Figure 11 shows the occupancy grid generated using SLAM during the mapping process. This representation provides a digital layout of the museum, enabling localization and global path planning. Once the map was completed, the robot was able to navigate autonomously to assigned waypoints while avoiding static and dynamic obstacles such as exhibition platforms, motorcycles, visitors and staff members.

The physical environment of the museum is illustrated in Figure 12, which highlights the complexity of the space, including narrow corridors, display areas, and open regions. These structural characteristics demand precise localization and adaptive path planning, making the scenario suitable for validating the navigation system.

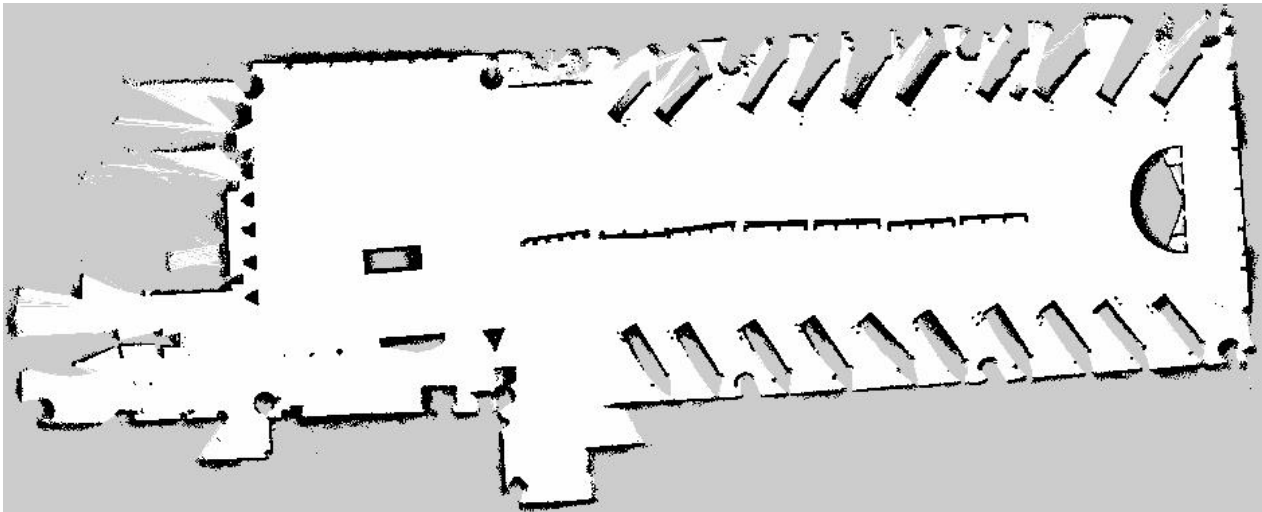


Figure 11. Museum navigation map.

Finally, Figure 13 presents the robot during navigation tasks within the exhibition area. Preliminary trials demonstrated successful undocking, trajectory execution, and docking sequences. Furthermore, the robot achieved consistent navigation performance in the presence of visitors and staff, confirming the feasibility of autonomous tours.

This initial validation represents a key milestone: the system has successfully transitioned from simulation-based testing to real-world operation. Future iterations will focus on improving robustness in dynamic environments by integrating obstacle detection, replanning strategies, and enhanced interaction with visitors.

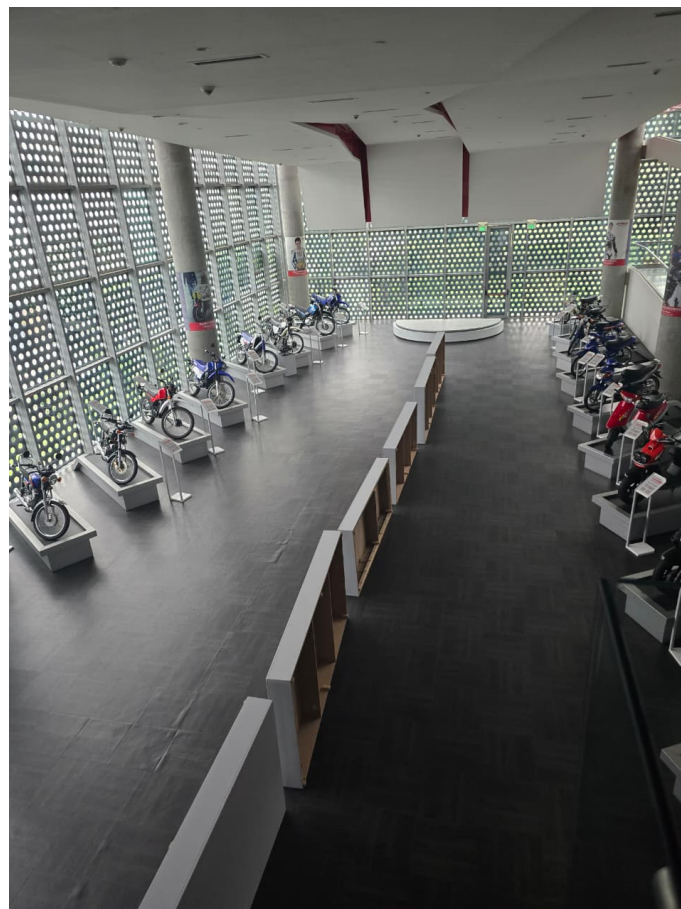


Figure 12. Museum layout.

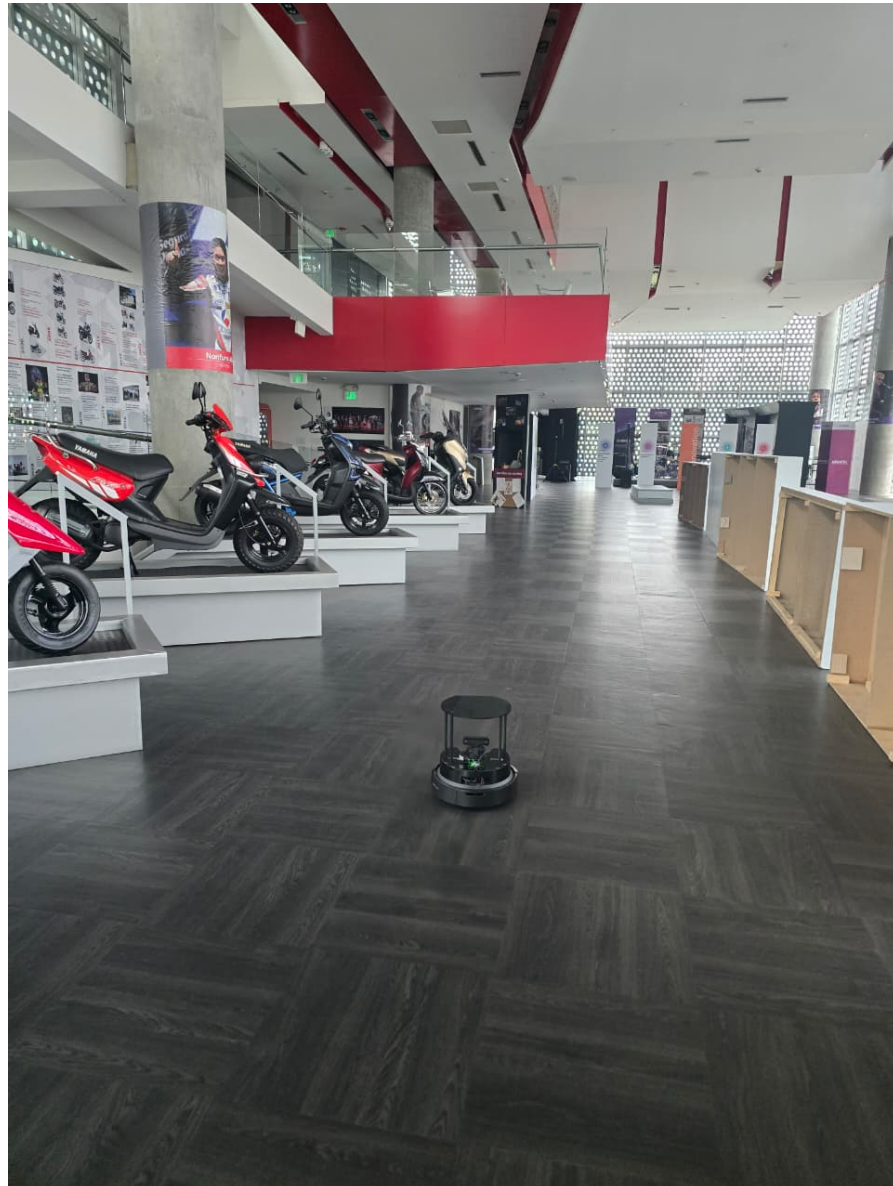


Figure 13. Robot navigating the museum.

3.2.2. mmWave Sensor Evaluation

The mmWave radar technology represents a novel and promising approach for perception in autonomous robots, offering robustness against lighting variations and partial occlusions. In this work, an Ai-Thinker RD-03D mmWave radar module was integrated into the robotic platform to assess its potential for human detection and tracking in indoor environments. With an effective range of up to 5 m, the RD-03D was selected for its compact design and low power consumption, making it suitable for embedded robotic applications. Unlike higher-end mmWave systems, this sensor provides a single detection point per object rather than a dense point cloud, which limits its spatial resolution but simplifies integration and data processing.

Two experimental configurations were conducted to evaluate its performance: static and dynamic tests. In the static setup, the sensor was fixed on a stationary support and connected to a custom ROS 2 node that converts radar data into a `PointCloud2` message for visualization in RViz (see Figure 14). This configuration allowed real-time observation of detected targets, where each object is represented as a single point in the 3D environment.

These tests showed that the RD-03D performs moderately well when both the environment and targets remain static, providing consistent detections and minimal positional noise.

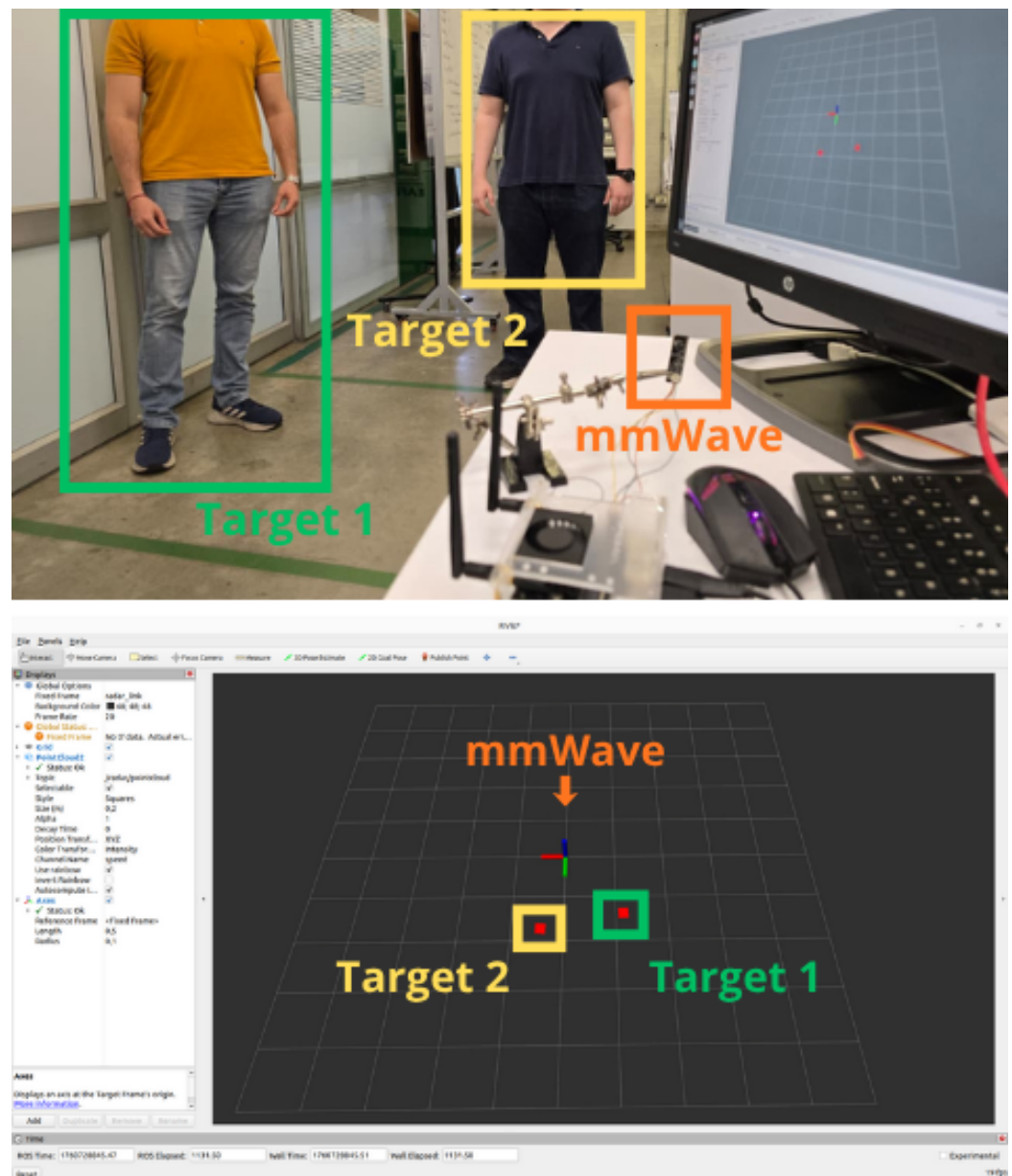


Figure 14. RD-03D mmWave sensor test.

However, in dynamic conditions, when either the sensor or the target was moving, the radar exhibited significant measurement noise and cumulative positional drift. This drift caused detected points to shift gradually over time, even when the target was stationary. In multi-target mode, the sensor can display up to three targets simultaneously, but precision decreases, and objects may intermittently disappear or show oscillatory movement. Conversely, single-target mode yielded greater stability and accuracy, though the limited spatial representation restricts detailed tracking.

3.2.3. Language Model and Robot Body

The Figure 15 illustrates the external appearance of the robot, featuring a humanoid structure equipped with an integrated tablet. This outer design complements the internal mechanical platform and architecture previously described, which houses the Jetson-based computing system, sensors, and supporting electronics. The tablet functions as the main

interface for user interaction, displaying the control and monitoring dashboard. Interaction is enabled through a natural language processing model running locally on the Jetson platform, supported by speech-to-text (STT) and text-to-speech (TTS) technologies. Furthermore, the model is connected to a dedicated database containing detailed information about the exhibitions, allowing the robot to provide dynamic explanations, deliver contextual content, and enhance the visitor experience during autonomous tours. To ensure a standardized and coherent visitor experience, the robot delivers a structured narrative at the beginning of the tour and at each exhibition station, presenting the most relevant information about the displayed items. After each presentation, visitors are encouraged to ask follow up questions, which are processed by the language model to provide detailed and context aware answers. This approach guarantees that every tour maintains consistent educational content while still allowing personalized interaction through the model's generative capabilities.

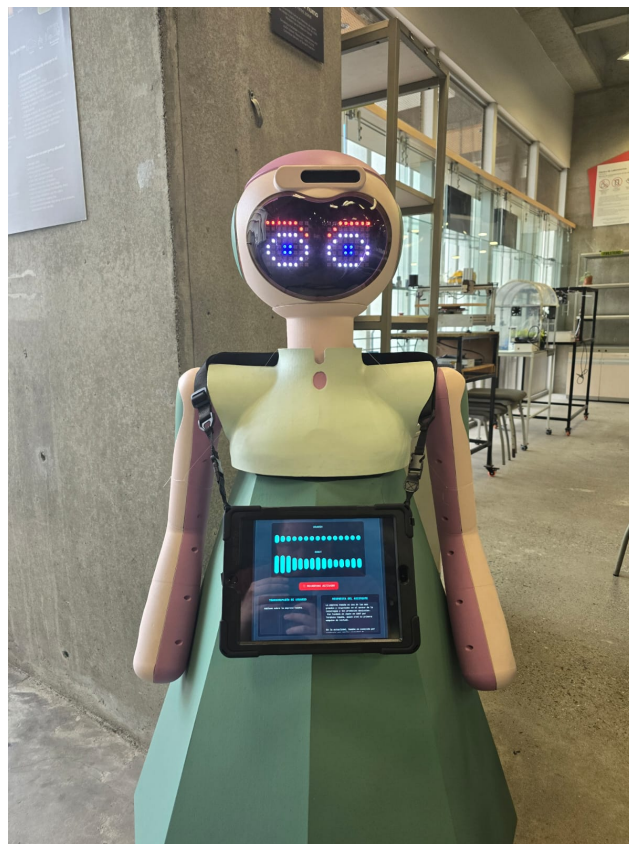


Figure 15. Robot platform body.

Overall, the preliminary implementation successfully demonstrated core functionalities of the proposed system in a real museum environment. The robot achieved autonomous navigation along the predefined exhibition route, performing docking, undocking, and way-point tracking with reliable performance under moderate visitor activity. Throughout the tests, the navigation system consistently executed its routes successfully, maintaining stable localization and obstacle avoidance. The language model operated locally on the Jetson platform and proved functional as an interactive guide, allowing users to ask questions about the exhibitions and receive accurate, context-based answers retrieved from the internal database. Regarding environmental perception, the mmWave radar module is currently in the testing phase. Initial experiments revealed its potential for human detection and obstacle tracking but also highlighted limitations.

4. Conclusions and Future Work

This project achieved the complete design of an autonomous museum guide robot, although only an initial phase of implementation has been carried out. The design process, structured under the Design Inclusive Research (DIR) methodology and the VDI 2206 standard for mechatronic systems, ensured a rigorous and systematic framework, resulting in a solution that is both innovative and adaptable to dynamic indoor environments.

One of the most remarkable contributions of this work is the incorporation of mmWave radar technology into the robot's sensing architecture. Although this technology has been scarcely applied in social and service robotics, its ability to detect and track humans under cluttered or partially occluded conditions makes it a highly promising addition to the perception stack. Despite the current limitations observed during preliminary tests, such as measurement noise and drift, the mmWave radar demonstrates strong potential as a complementary modality to traditional optical sensors like LiDAR and RGB-D cameras. Future efforts will focus on refining signal filtering, mitigating drift through sensor fusion with IMU and odometry data, and integrating the radar-generated point cloud into the robot's obstacle detection layer within the navigation stack to fully validate and exploit its capabilities in dynamic environments.

Another significant outcome is the integration of a natural language interaction system based on generative AI. Running locally on the Jetson platform, this model enables speech-to-text and text-to-speech functionalities, as well as access to a database with exhibition information. This configuration provides a highly interactive experience, improving communication with visitors and reinforcing the robot's role as an engaging museum guide.

In summary, this work demonstrates that the mechatronic design of a museum robot guide using novel technologies has been successfully achieved. These technologies include mmWave sensing for advanced perception and generative AI for natural communication, both of which highlight the originality of the proposed solution.

Looking ahead, the evolution of interactive robotics in museums is expected to move toward more immersive, adaptive, and scalable systems that combine autonomous navigation, generative artificial intelligence, and multimodal perception to deliver personalized experiences. Future museum guide robots could serve not only as information mediators but also as adaptive companions capable of learning from visitor behavior and dynamically adjusting their communication style and tour content. The modular and distributed architecture developed in this project provides a foundation for scalability, allowing the same technological framework comprising sensor fusion, IoT connectivity, and local AI processing to be extended to larger museums or transferred to industrial environments such as smart factories or logistics centers. In such contexts, robots could operate as interactive assistants for maintenance, inspection, or guided tours, demonstrating the broader applicability of the proposed design principles beyond the cultural domain.

In the short term, future work will focus on the complete implementation of the physical prototype inside the industrial museum and on a more systematic evaluation of its performance. This will include the definition and measurement of system-level indicators such as navigation success rate, docking reliability, mission completion time, interaction response time, and operational uptime during extended deployments. These metrics will make it possible to assess robustness and stability under realistic operating conditions. In parallel, user-centered studies will be designed to quantify the contribution of the language model to the visitor experience, through structured questionnaires, observation of interaction patterns, and analysis of anonymized dialogue logs to evaluate perceived usefulness, clarity of explanations, and engagement for different visitor profiles. On the perception side, a dedicated experimental campaign will be conducted to characterize the accuracy and drift of the mmWave radar in static and dynamic scenarios and to compare

the current module with alternative sensors. The resulting models will support improved signal filtering and tighter fusion with LiDAR, RGB-D, and odometry within the navigation stack. Finally, comparative benchmarks will be performed against alternative architectures, such as vision-only perception stacks and cloud-based dialogue systems, to quantify the trade-offs of the proposed multimodal sensing and on-device language model in terms of performance, privacy, and computational cost.

Future developments will also explore the integration of augmented reality (AR) and virtual reality (VR) technologies to enrich the visitor experience and strengthen the educational dimension of the robotic guide. Through AR interfaces, visitors could visualize interactive digital content overlaid on physical exhibits, while VR implementations could allow remote users to experience virtual tours guided by the robot in real time. These approaches align with the project's vision of creating immersive, informative, and inclusive museum environments where human–robot interaction extends beyond physical presence, bridging technology and cultural education.

Author Contributions: Conceptualization, D.V., A.A.-V. and D.S.-V.; methodology, D.V.; software, A.A.-V.; validation, D.V., J.P.G.-P. and D.S.-V.; formal analysis, A.A.-V.; investigation, A.A.-V., D.V., D.S.-V. and J.P.G.-P.; resources, D.V. and D.S.-V.; data curation, A.A.-V.; writing—original draft preparation, A.A.-V. and D.S.-V.; writing—review and editing, D.V. and J.P.G.-P.; visualization, A.A.-V.; supervision, D.S.-V.; project administration, D.V. and D.S.-V. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in the study are included in the article; further inquiries can be directed to the corresponding authors.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

FSM	Finite State Machine
IMU	Inertial Measurement Unit
LiDAR	Light Detection and Ranging
MQTT	Message Queuing Telemetry Transport
RFID	Radio-Frequency Identification
RGBD	Red-Green-Blue-Depth (sensor/camera)
ROS	Robot Operating System
SLAM	Simultaneous Localization and Mapping
STT	Speech-to-Text
TTS	Text-to-Speech

References

1. Hu, K.; Chen, Z.; Kang, H.; Tang, Y. 3D vision technologies for a self-developed structural external crack damage recognition robot. *Autom. Constr.* **2024**, *159*, 105262. [[CrossRef](#)]
2. Gasteiger, N.; Hellou, M.; Ahn, H. Deploying social robots in museum settings: A quasi-systematic review exploring purpose and acceptability. *Int. J. Adv. Robot. Syst.* **2021**, *18*, 1–13. [[CrossRef](#)]
3. Bickmore, T.; Pfeifer, L.; Schulman, D. Relational Agents Improve Engagement and Learning in Science Museum Visitors. In *Proceedings of the Intelligent Virtual Agents*, Reykjavik, Iceland, 15–17 September 2011; Vilhjálmsson, H.H., Kopp, S., Marsella, S., Thórisson, K.R., Eds.; Springer: Berlin/Heidelberg, Germany, 2011; pp. 55–67.

4. Kayukawa, S.; Sato, D.; Murata, M.; Ishihara, T.; Takagi, H.; Morishima, S.; Asakawa, C. Enhancing Blind Visitor's Autonomy in a Science Museum Using an Autonomous Navigation Robot. In Proceedings of the Conference on Human Factors in Computing Systems, Hamburg, Germany, 23–28 April 2023. [\[CrossRef\]](#)
5. Hellou, M.; Lim, J.; Gasteiger, N.; Jang, M.; Ahn, H.S. Technical Methods for Social Robots in Museum Settings: An Overview of the Literature. *Int. J. Soc. Robot.* **2022**, *14*, 1767–1786. [\[CrossRef\]](#)
6. Álvarez, M.; Galán, R.; Matía, F.; Rodríguez-Losada, D.; Jiménez, A. An emotional model for a guide robot. *IEEE Trans. Syst. Man, Cybern. Part A Syst. Humans* **2010**, *40*, 982–992. [\[CrossRef\]](#)
7. Rosa, S.; Randazzo, M.; Landini, E.; Bernagozzi, S.; Sacco, G.; Piccinino, M.; Natale, L. Tour guide robot: A 5G-enabled robot museum guide. *Front. Robot. AI* **2023**, *10*, 1323675. [\[CrossRef\]](#) [\[PubMed\]](#)
8. Daza, M.; Barrios-Aranibar, D.; Diaz-Amado, J.; Cardinale, Y.; Vilasboas, J. An approach of social navigation based on proxemics for crowded environments of humans and robots. *Micromachines* **2021**, *12*, 193. [\[CrossRef\]](#) [\[PubMed\]](#)
9. Duchetto, F.; Baxter, P.; Hanheide, M. Lindsey the Tour Guide Robot—Usage Patterns in a Museum Long-Term Deployment. In Proceedings of the 2019 28th IEEE International Conference on Robot and Human Interactive Communication, RO-MAN 2019, New Delhi, India, 14–18 October 2019. [\[CrossRef\]](#)
10. Manning, C.G. *Technology Readiness Levels*; NASA: Washington, DC, USA, 2023.
11. Yasuda, Y.D.V.; Martins, L.E.G.; Cappabianco, F.A.M. Autonomous Visual Navigation for Mobile Robots: A Systematic Literature Review. *ACM Comput. Surv.* **2020**, *53*, 13. [\[CrossRef\]](#)
12. Pareschi, R. Beyond Human and Machine: An Architecture and Methodology Guideline for Centaurian Design. *Sci* **2024**, *6*, 71. [\[CrossRef\]](#)
13. Tejada, J.C.; Toro-Ossaba, A.; López-Gonzalez, A.; Hernandez-Martinez, E.G.; Sanin-Villa, D. A Review of Multi-Robot Systems and Soft Robotics: Challenges and Opportunities. *Sensors* **2025**, *25*, 1353. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Alletto, S.; Cucchiara, R.; Fiore, G.D.; Mainetti, L.; Mighali, V.; Patrono, L.; Serra, G. An Indoor Location-Aware System for an IoT-Based Smart Museum. *IEEE Internet Things J.* **2016**, *3*, 244–253. [\[CrossRef\]](#)
15. Tufekci, Z.; Erdemir, G. Experimental Comparison of Global Planners for Trajectory Planning of Mobile Robots in an Unknown Environment with Dynamic Obstacles. In Proceedings of the HORA 2023–2023 5th International Congress on Human-Computer Interaction, Optimization and Robotic Applications, Proceedings, Istanbul, Turkey, 8–10 June 2023. [\[CrossRef\]](#)
16. Hussain, S.; Atif, M.; Daniyal, S.; Ahmed, L.; Memon, B.; Pasta, Q. Navigating the Maze Environment in ROS-2: An Experimental Comparison of Global and Local Planners for Dynamic Trajectory Planning in Mobile Robots. In Proceedings of the 6th International Conference on Robotics and Automation in Industry, ICRAI 2024, Rawalpindi, Pakistan, 18–19 December 2024. [\[CrossRef\]](#)
17. Asadi, K.; Suresh, A.K.; Ender, A.; Gotad, S.; Maniyar, S.; Anand, S.; Noghabaei, M.; Han, K.; Lobaton, E.; Wu, T. An integrated UGV-UAV system for construction site data collection. *Autom. Constr.* **2020**, *112*, 103068. [\[CrossRef\]](#)
18. Mahjourian, N.; Nguyen, V. Multimodal object detection using depth and image data for manufacturing parts. In Proceedings of the International Manufacturing Science and Engineering Conference, Greenville, SC, USA, 23–27 June 2025; Volume 89022, p. V002T16A001.
19. Soumya, A.; Mohan, C.K.; Cenkeramaddi, L.R. Recent Advances in mmWave-Radar-Based Sensing, Its Applications, and Machine Learning Techniques: A Review. *Sensors* **2023**, *23*, 8901. [\[CrossRef\]](#) [\[PubMed\]](#)
20. Harlow, K.; Jang, H.; Barfoot, T.D.; Kim, A.; Heckman, C. A new wave in robotics: Survey on recent mmwave radar applications in robotics. *IEEE Trans. Robot.* **2024**, *40*, 4544–4560. [\[CrossRef\]](#)
21. Kazempour, B.; Bhattathiri, S.; Rashedi, E.; Kuhl, M.; Hochgraf, C. Framework for Human-Robot Communication Gesture Design: A Warehouse Case Study. *SSRN Preprint* **2025**. [\[CrossRef\]](#)
22. Xie, B.; Park, C.H. Empathetic Robot with Transformer-Based Dialogue Agent. In Proceedings of the 2021 18th International Conference on Ubiquitous Robots (UR), Gangneung, Republic of Korea, 12–14 July 2021; pp. 290–295. [\[CrossRef\]](#)
23. Liu, Y.; Sun, Q.; Kapadia, D.R. Large Language Models in Robotics: A Survey on Integration and Applications. *Machines* **2023**, *6*, 158. [\[CrossRef\]](#)
24. Promobot. Home. Available online: <https://promo-bot.ai/> (accessed on 1 October 2025).
25. Rueda-Carvajal, G.D.; Tobar-Rosero, O.A.; Sánchez-Zuluaga, G.J.; Candelo-Becerra, J.E.; Flórez-Celis, H.A. Opportunities and Challenges of Industries 4.0 and 5.0 in Latin America. *Sci* **2025**, *7*, 68. [\[CrossRef\]](#)
26. Espina-Romero, L.; Hurtado, H.G.; Parra, D.R.; Pirela, R.A.V.; Talavera-Aguirre, R.; Ochoa-Díaz, A. Challenges and Opportunities in the Implementation of AI in Manufacturing: A Bibliometric Analysis. *Sci* **2024**, *6*, 60. [\[CrossRef\]](#)
27. Horváth, I. Comparison of three methodological approaches of design research. In Proceedings of the ICED 2007, the 16th International Conference on Engineering Design, Paris, France, 28–31 July 2007; Volume DS 42.
28. Graessler, I.; Hentze, J. The new V-Model of VDI 2206 and its validation das Neue V-Modell der VDI 2206 und seine Validierung. *At-Automatisierungstechnik* **2020**, *68*, 312–324. [\[CrossRef\]](#)

29. Vazquez-Santacruz, J.A.; Portillo-Velez, R.; Torres-Figueroa, J.; Marin-Urias, L.F.; Portilla-Flores, E. Towards an integrated design methodology for mechatronic systems. *Res. Eng. Des.* **2023**, *34*, 497–512. [[CrossRef](#)]
30. Horváth, I. Differences between ‘research in design context’ and ‘design inclusive research’ in the domain of industrial design engineering. *J. Des. Res.* **2008**, *7*, 61–83. [[CrossRef](#)]
31. Vega-Rojas, J.K.; Andrade-Miranda, K.S.; Justiniano-Medina, A.; Vasquez, C.A.A.; Beraún-Espíritu, M.M. Design and Implementation of an Automated Robotic Bartender Using VDI 2206 Methodology. *E3S Web Conf.* **2023**, *465*, 02062. [[CrossRef](#)]
32. Graessler, I.; Hentze, J.; Bruckmann, T. V-models for interdisciplinary systems engineering. In Proceedings of the International Design Conference, DESIGN, Dubrovnik, Croatia, 20–23 May 2018; Volume 2, pp. 747–756. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.