HRI-RBC 2025 - Final Project

Smart Museum Guide Robot

Najeh Alhalawani (2223737), Yelizaveta Tskhe (2224260)

https://youtu.be/oqimMFCMnOQ

All team members contributed equally to the development and writing of this project.

Introduction

In recent years, the integration of robots into public and educational spaces has seen rapid growth, particularly in the domain of service and assistive robotics. Among these, museum environments offer a unique opportunity for human-robot interaction, where robots can act not only as information providers but also as engaging companions who enhance the overall visitor experience with tailored content and interaction styles.

The motivation for this project stems from our belief that tailored and context-aware interactions can enhance engagement, learning outcomes, and inclusivity in museums and similar public educational spaces. Museums often host diverse audiences with different interests and from various age groups, and a one approach for all visitors limits the potential of robots to provide deeper and more relevant interactions. We view this as a critical problem worth addressing in the evolution of social robotics.

This project proposes the design and simulation of a robot museum guide capable of delivering customized tours and experiences based on the visitor's age, prior visit history, and level of domain knowledge. By leveraging technologies such as facial recognition, user modeling, and adaptive dialogue systems, the robot aims to provide meaningful, engaging, and informative interactions. For example, while children may receive playful and interactive guidance including games, jokes, and quizzes, adult visitors will be offered in-depth, technical explanations of exhibits that match their interests and understanding.

The primary objectives of this project are:

- To simulate a socially interactive robot capable of recognizing and adapting to different user profiles.
- To implement a mechanism for recognizing returning visitors and tailoring recommendations based on previous visits.
- To develop adaptive dialogue strategies that dynamically adjust content based on the user's age and knowledge level.
- To create a virtual prototype of the robot using a suitable platform for demonstration and testing purposes.

After a thorough analysis, Pepper has been selected as the platform for this simulation. The choice is supported by Pepper's expressive communication tools, built-in sensors for facial recognition, and compatibility with virtual development environments. Through this project, we aim to explore how intelligent, personalized robot behavior can enrich museum visits and provide a more inclusive and memorable experience for all visitors.

In the project, we have demonstrated that an adaptive robot museum guide, capable of recognizing visitors through facial recognition and adjusting its interaction strategy based on age and visit history, can significantly enhance engagement, satisfaction, and knowledge retention compared to a standard generic approach. Our simulations showed that playful strategies were more effective for children, while detailed explanations improved the experience for adult visitors. These results highlight the effectiveness of personalized, context-aware robot behaviors in public educational settings, such as museums.

Related Work

The use of robots as museum guides has been explored in various research and real-world contexts.

- Real-World Deployments

One of the examples is The Heinz Nixdorf MuseumsForum [1][2], which offers multiple robots, such as Beppo, Pepper, Aibo, and Cozmo, and a robot Paul that greets the customers, tells stories, suggests taking selfies together and even tries to dance. The robots PETER and PETRA guide visitors to the exhibits while paying great attention to interactions with children.

Another example is the REEM humanoid robot at CosmoCaixa Museum in Barcelona. It was present at the museum for one week, during which it guided visitors through the facilities. Its behavior can be described as a “follow-me” along with face-to-face communication. Robot’s main tasks were guiding visitors and informing them about current exhibitions [3][4]. In both cases, while robots were effective in delivering static content, such systems typically lacked adaptability to individual users or contextual cues.

The first museum that deployed Pepper robots was the Smithsonian Museum in 2018 [5]. It was presented in six spaces in an experimental program to test how interactions with robots can enhance visitors’ experience. Its main focus was on providing a customized visitor engagement with the exhibits.

- Academic Research

Recently, the focus has been on enhancing robot guides with facial recognition and user modeling capabilities. Projects such as FACE (Facial Automaton for Conveying Emotions) explore how facial cues and recognition can be used to adjust robot behavior and tailor interactions. One of the important aspects of human-robot interaction is age-appropriate communication which is particularly beneficial in learning environments.

The article “Interactive Humanoid Robots for a Science Museum” [6] presents a study on deploying humanoid robots in a science museum. Visitors were equipped with RFID tags, enabling the robots to identify them and interact accordingly. Over a two-month trial period, robots received high evaluations from visitors, indicating their effectiveness in the visitors’ museum experience. The study highlights the importance of adaptive interaction strategies in maintaining visitors’ interest and improving learning outcomes.

Another article “Coordination of Verbal and Non-Verbal Actions in Human-Robot Interaction at Museums and Exhibitions” [7] investigates how synchronizing speech with gestures enhances the effectiveness of robots in the museum setting. It was observed that guides often coordinate their speech with non-verbal signals, such as turning their head towards the exhibit while talking about it. Observations indicated that when the robot's speech is synchronized with the gestures, the visitors are more engaged and likely to respond.
Unlike previous implementations that often focus on one aspect of personalization, this project aims to develop a fully adaptive museum guide that dynamically adjusts its behavior based on age group, prior interaction history, and inferred knowledge level. Furthermore, by recognizing returning visitors through

facial recognition, the robot can personalize recommendations and revisit themes of interest, making each museum visit a more meaningful and connected experience.

Solution

HRI

The design of our HRI system focuses on delivering an adaptive, socially engaging, and informative experience tailored to museum visitors of diverse backgrounds and interests. We selected the Pepper robot due to its humanoid features, expressive interaction capabilities, and integration of multiple sensing and communication modalities.

If we had infinite time and unlimited resources, the robot museum guide would be designed as a fully autonomous, hyper-personalized companion equipped with advanced multimodal sensors, including eye-tracking, thermal imaging, emotion recognition, and LIDAR for precise environmental awareness. Its hardware would feature soft robotics for expressive, human-like gestures, adaptive locomotion systems for navigating diverse terrains, and a holographic projection system for immersive storytelling. The software would be powered by a cloud-synced, continually learning AI integrating advanced natural language understanding, real-time emotion modeling, and lifelong user profiling across visits and locations. Functionality-wise, the robot would offer seamless multi-language support, real-time exhibit augmentation using AR/VR, and group interaction management, capable of dynamically shifting tone and content for children, adults, or families in shared tours; transforming the museum into a fully interactive, intelligent learning environment.

1. Hardware Components

Pepper comes equipped with an extensive set of built-in hardware features that support our project objectives:

- Cameras: Two HD RGB cameras (forehead and mouth area) and one 3D depth camera support facial recognition, gesture detection, and visitor tracking.

- Microphones: Four directional microphones enable speech recognition, noise localization, and directional interaction even in moderately noisy museum environments.

- Touch Sensors: Located on the head and hands, these support tactile interaction, useful for playful engagement with children.

- Tablet Interface: A 10.1-inch touchscreen provides a visual medium for displaying rich content such as images, quiz questions, videos, or navigation maps.

- Wheels and Mobility: Pepper's omnidirectional base allows for smooth navigation and "follow-me" interactions, especially for guiding visitors through exhibit paths.

- LEDs and Gesture System: For expressive communication via eye color changes, head movement, and arm gestures, reinforcing Pepper's social presence.

2. Software Architecture

Our robot system will be built upon the NAOqi OS and leverage Choregraphe and QiSDK for programming and behavior design. The architecture includes built-in Pepper capabilities and custom-developed modules :

The built-in components that are gonna be integrated into the system:

- Pepper hardware simulation (via Choregraphe)
- Face detection and recognition module
- Speech recognition and synthesis
- Predefined animation blocks

Custom-developed modules:

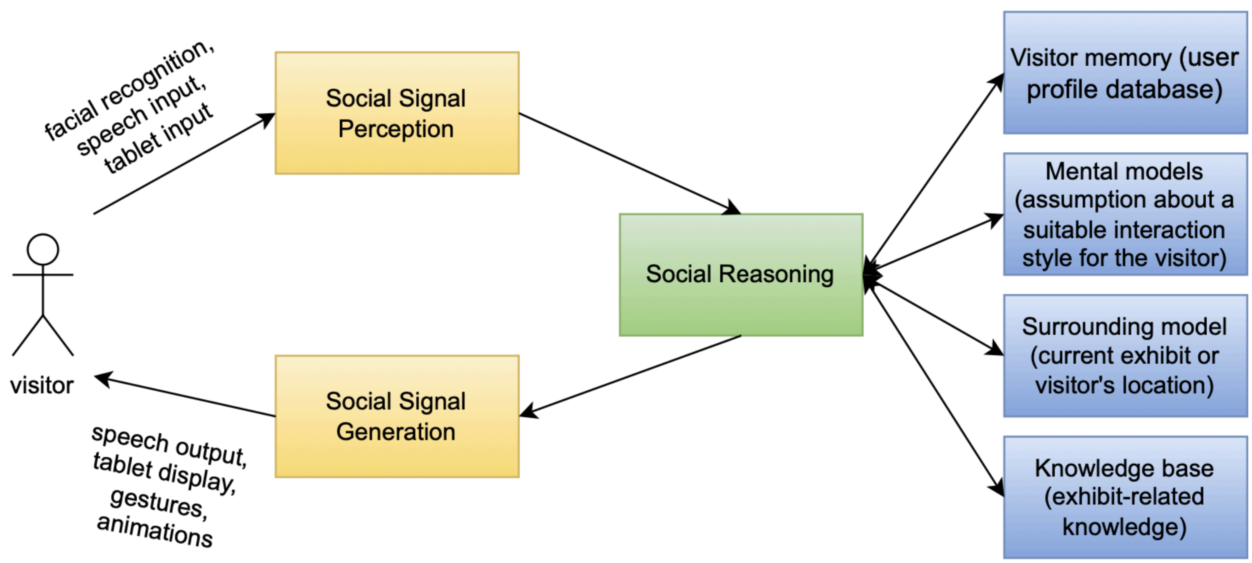
- Visitor Profile Manager
  - Maintains persistent visitor data (age group, visit history, interests)
  - Uses facial recognition to identify returning visitors
  - Stores interaction history to adapt future dialogues and content
- User Modeling Module
  - Dynamically infers the visitor's knowledge level (e.g., beginner, intermediate, advanced) using short quizzes or keyword detection from speech.
  - Categorizes age group (child, teen, adult) using facial analysis and height estimation.
- Dialogue System
  - A hybrid architecture: rule-based for structure + NLP-based for flexibility
  - Integrates speech recognition (ASR), natural language understanding (NLU), and text-to-speech (TTS)
  - Adapts tone, vocabulary, and content depth based on the visitor profile
- Content Delivery Engine
  - Retrieves exhibit-specific content tagged by complexity level
  - Uses the tablet to supplement spoken explanations with visual content
  - Offers interactive options such as quizzes, fun facts, and storytelling for children
- Behavior Engine
  - Controls multimodal outputs: gestures, gaze, facial expressions (LED), speech, and tablet visuals
  - Synchronizes physical behavior with speech to enhance engagement

All modules are orchestrated through Choregraphe's behavior flow, possibly enhanced with Python scripts for data processing and control logic. Data flows from sensors (e.g., face recognition) into the profiling system, which informs the dialogue and interaction flow. Output is returned to the robot through speech and animations.

Functional components:

- Social signal perception
  - Utilizes Pepper's built-in sensors and pre-existing modules to detect facial features, speech input, and tablet touch interactions. This module identifies the visitor (via facial recognition) and collects input signals like spoken questions or tap responses.
- Social reasoning
  - Acts as a central processing unit that integrates sensory input with internal knowledge and context. This component selects appropriate responses, decides the tone (playful vs. informative), and triggers recommendations. It uses visitor memory (user profile database), mental models (assumption about what type of interaction style suits the user), knowledge base (exhibit-related knowledge), and surrounding model (current exhibit or a location of the visitor).
- Social signal generation
  - Translates the reasoning output into multimodal robot behavior. It includes speech generation, tablet output, gestures, and animations.

Functional Architecture diagram



<u>**RBC**</u>

Functionality Benchmarks:

1. Automatic Speech Recognition (ASR)
   a. Metric: recognition accuracy over various voices and ambient noise conditions.

    b. Evaluation: real-world trials with participants giving commands in different accents and distances.
2. Multimodal Interaction Handling
    a. Metric: latency between user input (from button or speech) and robot reaction.
    b. Evaluation: time measurements of end-to-end response delay using a stopwatch or video timestamps.
3. Content Delivery
    a. Metric: completeness and clarity of presented information
    b. Evaluation: human raters assessment.
4. Text-to-speech (TTS)
    a. Metric: clarity, latency, and multilingual accuracy.
    b. Evaluation: simulated delivery of content rated by human evaluators.
5. Sonar-based Obstacle Detection
    a. Metric: obstacle detection accuracy, latency, and robustness across various object types.
    b. Evaluation: virtual navigation tests with static and moving obstacles.
6. Face Recognition
    a. Metric: classification accuracy, latency, and robustness to lighting.
    b. Evaluation: dataset-based evaluation using simulated faces under various lighting conditions.

Task Benchmarks

Task Definition: successfully guide a human visitor through an adaptive museum tour based on their profile (e.g., age), interest (e.g., art or science), and behavior (e.g., asking follow-up questions).

Success Metrics:
- Task Completion: did the user complete the full tour without confusion?
- Adaptive Behavior: did the robot adapt content appropriately for the user?
- Engagement: did the user remain engaged throughout the interaction?

Role of Human Participants: human users simulate museum visitors, including both first-time and returning guests, across various age groups (children and adults). Participants provide natural input via speech and button presses, feedback to test the robot's adaptive strategies, and reactions that allow it to measure user satisfaction.

Variability Across Trials - it allows us to test the robustness and flexibility of the system in realistic scenarios:

- Different user profiles (age, language, familiarity with the robot) - for diversity.
- Various environmental conditions (lighting and background noise).
- Environment unpredictability (users interruption, ignoring the robot).

Success Levels:
- Full Success: the robot completes the full task (delivers a complete tour personalized to the user’s profile).

- Partial Success: the robot manages to engage the user and deliver part of the content but fails to complete all steps (skips quizzes or misses user classification).
- Failure: the robot is unable to initiate or sustain the interaction (no response due to ASR or facial recognition error).

World and Context Modelling Approach

Representation: the robot uses a predefined semantic map of the museum, linked with a lightweight knowledge graph that defines visitors, interaction history, key exhibits, and associated metadata (artwork, name, artist, topic). The knowledge graph enabled personalization, interaction continuity, and dynamic content delivery. The knowledge graph acts as a central context memory, continuously queried by the robot's reasoning module to determine the most appropriate content delivery and adjust interactions. The interface was implemented as a set of rule-based logic conditions that operated on data to generate new interactions.

Knowledge Sources:
- Predefined JSON file containing exhibit information.
- User input (age selection or button responses).
- Sensor inputs (gesture triggers).

Dynamic Updates: the robot dynamically updated its internal state based on the user's selected age and feedback, enabling personalized content delivery and adapted dialogue.

Object Relocation and New User Intentions: the robot did not dynamically remap the physical objects, but it accounted for the intent shift (if the user changed a topic via button inputs, the robot adapted to the new topic) and content redirection (the knowledge graph allowed Pepper to direct the conversation in the flow suitable for the visitor's age and engagement).

Managing Uncertainty: the uncertainty from ASR and face recognition modules was handled by manual typing on the tablet or robot repeating the question.

An example of the simplified knowledge graph illustrating how the system interpreted user profiles and adapted the content based on it:

```
[Visitor_01]
 ├── age_group: "child"
 ├── preferred_language: "English"
 ├── visited: ["Exhibit_Leonardo"]
 ├── interaction_outcome: "quiz_completed"
 └── last_seen: "2025-06-05"

[Exhibit_Leonardo]
 ├── type: "Flying Machine"
 ├── content: ["description_en", "description_it", "quiz_easy"]
```

```
 └── location: "Room_2"

[Reasoning Rule]
 IF Visitor_01.age_group == "child"
 → Use content = "quiz_easy", voice_tone = "cheerful"
```

Another example of a robot's adaptive behavior is when a child user is detected as crying. The knowledge graph retrieves a response suggesting ice cream, while the semantic map identifies the nearest café. The reasoning module prioritizes this suggestion, delivering it via TTS in a cheerful tone, enhancing user comfort.

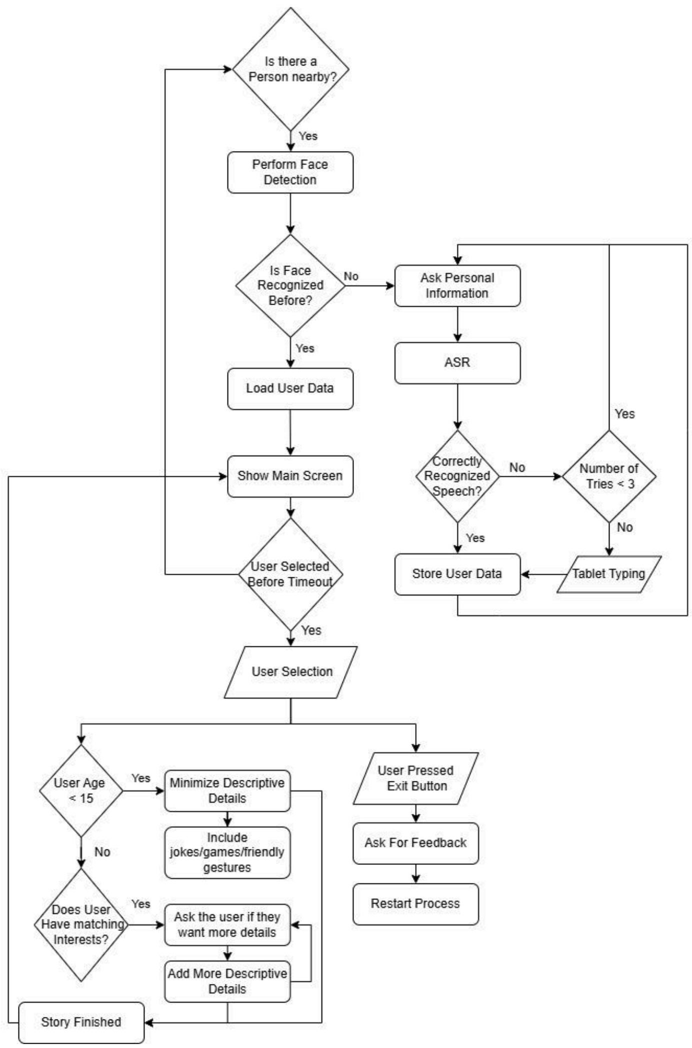
Connection to HRI:

- Speech and button inputs enable intuitive user control and accommodate different preferences.
- Gesture and TTS enhanced the robot's expressiveness and engagement.
- Content adaptation increased user satisfaction by personalizing the interactions.
- Error recovery aims to detect errors and provide the user with quick assistance.

Implementation

The virtual robot system was implemented using Python integrated within the MODIM framework. The system simulates an interactive robot capable of detecting users, collecting visitor profiles, adapting information to user needs, and managing conversation flows.

The architecture follows a perception-reasoning-action pipeline: sensor data is processed to build a visitor profile, which is then used to select appropriate behavior strategies. The behaviors are realized through speech and gesture commands.



The diagram above illustrates the flow between the museum guide robot and a visitor, including personalization, user profile creation, and appropriate content delivery.

Standard Conversation Flow

1. Human Detection

Using Pepper's front sensor ([p1]), the robot checks for a person standing in front of it. It ensures the user remains for at least 2 seconds before starting the interaction.

```
im.robot.startSensorMonitor()
im.robot.normalPosture()
im.robot.headscan()
while not flagP:
    while not detected:
        p = im.robot.sensorvalue() #p is the array with data of all the sensors
        detected = p[1] > 0.0 and p[1] <= 1.0 #p[1] is the Front sonar
    if detected:
        print('*Person Detected*')
        time.sleep(2)
        p = im.robot.sensorvalue()
        detected = p[1] > 0.0 and p[1] <= 1.0
        if detected:
            print('*Person still there*')
            flagP = True
        else:
            print('*Person gone*')
```

2. Language Selection

Once the person is detected, the robot prompts a user to select a language (EN or IT) to proceed in the preferred language.

```
lang_ques = im.ask('lang', timeout=20)
if lang_ques == 'en':
    im.robot.setLanguage('en')
else:
    im.robot.setLanguage('it')
```

3. User Introduction

If the user agrees to an introduction, the robot collects basic information: age group (child or adult), area of interest (e.g. art / robots / history), and for adults, level of knowledge (fun facts or deep details)

4. Dialogue & Exhibit Information

Based on the collected information, the robot adapts its interactions: children receive fun and easy explanations and quizzes, while adults get deeper details or fun facts based on what they chose previously.

```
if age == 'child':
    if interest == 'art':
        im.execute('art_leo_child')
        a0 = im.ask('art_leo_child_ques', timeout=20)
        if a0 == "yes":
            im.execute('art_leo_child_incorrect')
        elif a0 == "no":
            im.execute('art_leo_child_correct')
```

```
# ADULTS
if level == 'fun_facts':
    im.execute('art_leo_fun_facts')
    im.execute('art_leo_fun_facts_random')
elif level == 'deep_details':
    im.execute('art_leo_deep_details')
    a0 = im.ask('art_leo_deep_details_ques', timeout=20)
    if a0 == "bird_wings":
        im.execute('art_leo_deep_details_correct')
    else:
        im.execute('art_leo_deep_details_incorrect')
```

5. Gestures

Pepper performs simple gestures (e.g. hand wave, arm raise) and posture to be more engaged and human-like.

6. Follow-up and Menu Options

After delivering the content, the robot asks if the user wants to continue with another topic or exit. If the user times out, it offers to repeat the menu.

7. Feedback and Goodbye

Upon exit, the robot requests user feedback and concludes with a goodbye message.

```
q = im.ask('menuexist', timeout=4)
if q == 'timeout':
        im.execute('repeat')
```

```
        time.sleep(3)
        q = im.ask('menuexist', timeout=10)
if q == 'main_menu':
        pass
elif q == 'exit':
        feedback = im.ask('feedback', timeout=10)
        im.execute('goodbye')
        time.sleep(1)
```

This structured flow ensures a user-friendly and adaptive experience tailored to different groups of age and interests.

To handle the user interface on the robot tablet, the custom file (qaws.js) provided by MODIM has been used. It establishes a WebSocket connection with the backend and dynamically updates the UI elements (text, images, buttons). Furthermore, it relays user responses from buttons back to the robot for further conversation flow.

Navigation Scenario

In case a visitor needs help with navigation, the robot provides a map of the museum guiding the lost person to the desired location. The examples implemented were "Where is the Restroom?", "Where is the Dinosaur Exhibit?", and "Where is the Impressionist Paintings?". In all cases, it gives clear directions, shows the map, and points to the correct path.

```
if ask == 'restroom':
        im.robot.setPosture([0.00, -0.21,1.55, 0.13, -1.24, -0.52, 0.01,0.3, -1.5, 1.0, 0.05, 0.0,0, 0, 0, 0, 0])
        im.execute('restroom')
elif ask == 'din_ex':
        im.robot.setPosture([0.00, -0.21,1.55, 0.13, -1.24, -0.52, 0.01,0.3, -1.5, 1.0, 0.05, 0.0,0, 0, 0, 0, 0])
        im.execute('din_ex')
elif ask == 'imp_paint':
        im.robot.setPosture([0.00, -0.21,1.55, 0.13, -1.24, -0.52, 0.01,0.3, -1.5, 1.0, 0.05, 0.0,0, 0, 0, 0, 0])
        im.execute('imp_paint')
```

Returning Visitors

Whenever the robot recognizes a returning visitor, it asks if the visitor wants to continue from the last topic or to get new information on a different topic. Then it asks for the level of knowledge the user prefers for the chosen topic (fun facts or deeper details). After that, it provides users with fun facts or deeper details and a short quiz.

```
im.execute('return_welcome_cont')
time.sleep(0.6)
interest = im.ask('return_interest', timeout=20)
level = im.ask('level', timeout=20)
if interest == 'art':
        if level == 'fun_facts':
                    im.execute('art_leo_fun_facts')
                    print('* fun_facts page')
                    time.sleep(3)
                    im.execute('art_leo_fun_facts_random')
         elif level == 'deep_details':
                    im.execute('art_leo_deep_details')
                    print('* deep_details page')
                    a0 = im.ask('art_leo_deep_details_ques', timeout=20)
                    if a0 == "bird_wings":
                        im.execute('art_leo_deep_details_correct')
                    else:
                        im.execute('art_leo_deep_details_incorrect'
```

The “actions” folder contains modular action definitions that the robot executes using MODIM. The dialogue-action files include modalities like IMAGE (image shown on the tablet), TEXT (text shown on the tablet), TTS(text-to-speech for Pepper to say), GESTURE (animation for Pepper), and optionally BUTTONS (to collect user inputs).

As an example, in the file “art_leo_fun_facts”, where Pepper provides information without prompting for user input:

```
IMAGE
<*, *, *, *>:  img/leo.jpg
----
TEXT
<*,*,*,*>: Leonardo's flying machine is called an 'ornithopter.' He studied how birds fly and drew this around the year 1485. It had large wings powered by a human. Even though it never flew, it helped scientists think about flight in new ways.
----
TTS
<*,*,*,*>: Leonardo's flying machine is called an 'ornithopter.' He studied how birds fly and drew this around the year 1485. It had large wings powered by a human. Even though it never flew, it helped scientists think about flight in new ways.
----
GESTURE
<*,*,*,*>: BodyTalk_10
----
```

On the other hand, in some cases Pepper asks a question to maintain the engagement with the user, like in "art_leo_deep_details_ques", where the user has to click one of the buttons:

```
IMAGE
<*, *, *, *>:  img/leo.jpg
----
TEXT
<*,*,*,*>: Time for a quick quiz! What do you think Leonardo used to design his flying machine?
----
GESTURE
<*,*,*,*>: BodyTalk_10
----
TTS
<*,*,*,*>: Time for a quick quiz! What do you think Leonardo used to design his flying machine?
----
ASR
<*,*,*,*>: {'rocket_fuel': ['rocket','fuel','rocket fuel', 'a', '1', 'first', 'first one'], 'bird_wings':
['bird','wings','bird wings', '2', 'b', 'second', 'second one'], 'magic': ['magic', '3', 'c', 'third', 'third one'] }
----
BUTTONS
rocket_fuel
<*,*,*,*>:  Rocket fuel
bird_wings
<*,*,*,*>:  Bird Wings
magic
<*,*,*,*>:  Magic
----
```

- Representation of environment (RBC):

The robot has a general representation, such that it is aware that it is in a museum context and structured to handle interactions around exhibits, greetings, and navigation. It maintains specific representation, such as a persistent profile for each visitor, including their age group, interests, and past visited exhibits. The internal representation is updated during each interaction to reflect new preferences or behaviors. The robot knows exhibit metadata and how to adapt explanations by age and interests. This structure allows a robot to reason about both static environmental and dynamic user data to personalize interactions.

In case the robot detects the sound of a crying baby, it tries to get the child's attention and asks "How about we get some ice cream? Do you like ice cream?" and as the child or their parent chooses "Yes", the robot shows a map and guides visitors to the ice cream place.

For this project with the virtual robot only, the crying baby was simulated, but Pepper's social interaction stayed the same as if the situation happened in a real real-life scenario.

```
ask == 'crying_baby':
    im.execute('navigation_crying_baby')
    time.sleep(0.2)
    icecream =  im.ask('crying_baby_2', timeout=20)
    if icecream == 'yes':
            im.robot.setPosture(turn_pose_1)
            time.sleep(0.3)
            im.robot.setPosture(turn_pose_2)
            time.sleep(0.3)
            im.robot.setPosture(turn_pose_3)
            time.sleep(0.3)
            im.robot.setPosture(turn_pose_4)
            time.sleep(0.3)
            im.robot.setPosture(turn_pose_5)
            time.sleep(0.3)
            im.execute('icecream')
```

- Behavior and Reasoning (RBC):

Behavioral decisions are controlled by rule-based reasoning. When a visitor is recognized, the robot accesses its internal memory and determines what content has already been shown, what type of content matches the visitor's profile, and what strategy to follow. The behavior is computed in real time based on dynamic internal variables. For example, if a user is under the age of 12, the robot prioritizes interactive content, such as jokes or quizzes. Modules are connected via a shared memory. The architecture allows behavior to change as more information is gathered about the visitor.

- Interaction and dialogue design (HRI):

The robot communicates using multi-modal interaction: speech and gestures. Dialogue generation is driven by custom rules, enabling both structured information delivery and dynamic interaction. Children receive short, playful explanations with interactive components, while adults receive detailed content with historical background. Facial recognition is used to detect returning users, enabling the system to refer to past visits and create personalized tours.

- Tools:
- Docker container
- Python and NAOqi SDK for control of robot modules
- File-based persistence for profile memory
- Custom-built modules

Some modules, like facial recognition, were simulated due to the lack of hardware, but still behave as if the robot is reasoning about real-world data.

The robot goes beyond default behavior by creating and using its internal visitor profiles, adapting behavior based on memory, choosing actions based on visitor data, and operating entirely from code bypassing the limitations of Choregraphe's behavior tree.

Dialogue Design and Personalization

Dialogue generation in our system is fully custom-built and rule-based, designed to adapt to each visitor's profile in real-time. The dialogue manager selects and assembles conversations based on visitor attributes, such as group age, interest area, and visit history. It allows robots to behave in more natural, engaging, and context-aware interactions.

- Greeting
    - New Visitor (child):
        - "Welcome to the museum! I'm your robot guide. Let's have some fun together!"
    - New Visitor (adult):
        - "Good day! I'm Pepper, your museum guide. Wishing you an amazing first experience here!"
    - Returning Visitor (any age):
        - "Hey! I remember you! Last time we saw the robot dogs—do you want to see something new today?"
        - "Welcome back! I recall you were interested in AI. There's a new exhibit you might love!"
        - "Great to see you again! Should we continue from where we left off, or start a new adventure?"

- Introduction / Creation of Visitor Profile
        - "To make your visit extra special, can I ask you a few quick questions?"
        - "Can I ask your age group? It'll help me choose the best tour style for you."
        - "What are you most interested in? Robots? History? Art? I can adapt the tour just for you."
        - "Would you prefer basic info, fun facts, or deeper technical details?"

- Exhibit Information
    - Child
        - "This is a flying machine! Can you believe Leonardo da Vinci dreamed it up 500 years ago? It flaps like a bird's wings! Pretty cool, right? Want to hear a fun fact?"
        - "Do you think it could actually fly? Yes, no, or maybe?"
    - Adult

        - “This is Leonardo da Vinci’s concept for a human-powered flying machine. He modeled it on bird wings, and theorized it could take off by flapping. Of course, he didn’t have modern engines—so it never flew—but his vision was extraordinary for the 15th century.”
        - “This design, known as the *ornithopter*, was sketched by Leonardo around 1485. It features a wingspan of over 10 meters and a harness system intended for human-powered flapping flight. Though untested, it reflects early biomimicry principles and significant anatomical study.”

- Knowledge Adaptation
    - Beginner (children)
        - “Leonardo da Vinci had a dream to fly like a bird. He made drawings of a flying machine with big wings that flap, just like a bird’s. But back then, they didn’t have engines, so it didn’t really fly. Still, pretty amazing , right?”
    - Intermediate (curious adults and teens) (fun facts)
        - “Leonardo’s flying machine is called an ‘ornithopter.’ He studied how birds fly and drew this around the year 1485. It had large wings powered by a human. Even though it never flew, it helped scientists think about flight in new ways.”
    - Expert (adults with high domain knowledge)
        - “The ornithopter design reflects Leonardo’s study of avian anatomy and biomechanics. The mechanism involves manual wing articulation, likely via a pulley system. While impractical for sustained flight, the design represents an early instance of biomimicry and engineering foresight.”

- Entertainment
    - Jokes (for children and adults with low engagement)
        - “Want to hear a joke? Why did the robot go on vacation?” *(Pause)* “Because it needed to recharge its batteries!”
        - “Here’s a classic: Why was the museum robot such a great worker?” “Because it always followed the rules… and the route!”
    - Quizzes (to engage playfully)
        - “Time for a quick quiz!
          What do you think Leonardo used to design his flying machine?”
          A) Rocket fuel
          B) Bird wings
          C) Magic *(wait for answer)* “Correct! He studied birds and tried to copy their wings. Great job!”
    - Fun facts (as a surprise or to break up long content)
        - “Here’s a fun fact: Did you know Leonardo da Vinci used to write backwards? Some say it was to keep his ideas secret!”
    - Mini games

        - "Let's play a game! I'll describe something, and you guess the exhibit. It has wings, it doesn't fly, and it was drawn by a genius. What is it?"

- Follow-up topics
    - Recognition of returning visitor
        - "Hi again! I remember you liked the tech exhibits—want to check out something similar?"
    - Continuing from the past topic
        - "You seemed really interested in inventions last time. I've saved a few cool facts just for you!"
    - Avoiding repetitions
        - "You've already seen the flying machine - should we skip it and head to another section instead?"
    - Suggestions based on visit history
        - "You've visited the art exhibit but not the science one yet. Want to go there next?"

- Error handling
        - "I'm sorry, I didn't catch that. Could you please repeat?"
        - "If you prefer, you can also type your question on the screen."
        - "Sorry, my connection is a bit slow. Meanwhile, I can guide you to the nearest exhibit."
        - "Thanks for your patience! How about I simplify the information for you?"

- Navigation Help
        - "The Impressionist Paintings are on the second floor, Gallery B. Follow me, and I'll guide you there!"
        - "You'll find the Dinosaur Exhibit down this hall to the right. If you'd like, I can provide a map of the museum."
        - "The restrooms are just past the main lobby on your left."

- End of Visit
        - "Thank you for visiting! I hope you enjoyed the tour."
        - "Since you enjoyed the art exhibits, don't miss our upcoming special exhibition next month!"
        - "Don't forget to check out our seasonal events."

Evaluation

1. Research Questions and Hypotheses

Research Questions:

- RQ1: How effectively does the Pepper robot's playful interaction strategy (e.g., greetings, quizzes, and jokes for children) increase engagement among new and returning children compared to a generic interaction strategy?
- RQ2: To what extent does the Pepper robot's detailed interaction strategy (e.g., suggesting tech exhibitions with deep explanations for adults) improve knowledge retention and satisfaction among new and returning detail-oriented adult museum visitors?
- RQ3: Does the robot's ability to adapt its interaction strategy based on user profiles (new vs. returning visitors) and age group (children vs. adults) enhance engagement, satisfaction, and learning outcomes compared to a non-adaptive generic strategy?
- RQ4: Which interaction modality (speech, touch, or visual/animation) is most effective for children and detail-oriented adults in terms of engagement and usability, given the robot's age-based reasoning rules?

Hypotheses:

- H1: Children (under 12) will show higher engagement (measured by interaction time and emotional response) with the robot's playful strategies (e.g., greetings, quizzes, and jokes) compared to a generic interaction strategy.
- H2: Detail-oriented adults will demonstrate higher knowledge retention (quiz scores) and satisfaction (self-reported ratings) when the robot uses detailed strategies
(e.g., general introductions for new adults; tech exhibitions with deep explanations for adults who like technology) compared to a generic strategy.
- H3: Adaptive interaction, where the robot tailors its strategy based on user profile (new vs. returning) and age group (playful for children under 12, detailed for adults), will result in higher engagement, satisfaction, and learning outcomes compared to a non-adaptive generic strategy.
- H4: Interaction modality preferences will differ by age group, with children favoring visual/animation and touch (e.g., for quizzes and jokes), and adults preferring speech and text-based visuals (e.g., for deep tech explanations).

Null Hypotheses:

- H0_1: There is no difference in engagement (interaction time, emotional response) for children between the robot's playful strategies and a generic strategy.
- H0_2: There is no difference in knowledge retention or satisfaction for detail-oriented adults between the robot's detailed strategies and a generic strategy.
- H0_3: Adaptive interaction based on user profile and age group does not improve engagement, satisfaction, or learning outcomes compared to a non-adaptive generic strategy.
- H0_4: There is no difference in the effectiveness of interaction modalities (speech, touch, visual/animation) between children and detail-oriented adults in terms of engagement and usability.

2. Variables

Independent Variables:

- Interaction Strategy Type (playful strategy with quizzes and jokes, detailed strategy with in-depth explanations).
- Visitor Profile (new or returning).
- Age group (child or adult).
- Interaction modality (speech, visual).

Dependent Variables:

- Engagement:
  - Interaction time (seconds spent interacting, logged by the robot).
  - Emotional response
    (self-reported on a 5-point scale, e.g., 1 = bored, 5 = excited).
- Satisfaction: Self-reported satisfaction (5-point scale, e.g., enjoyment, clarity).
- Knowledge Retention: Performance on a post-interaction quiz
  (5–10 questions on exhibit content, e.g., space facts for children, tech details for adults, scored as % correct).
- Usability: Self-reported ease of use for each modality
  (5-point scale, e.g., 1 = very difficult, 5 = very easy).

User Study - Evaluation Questionnaire (Godspeed + RoSaS)

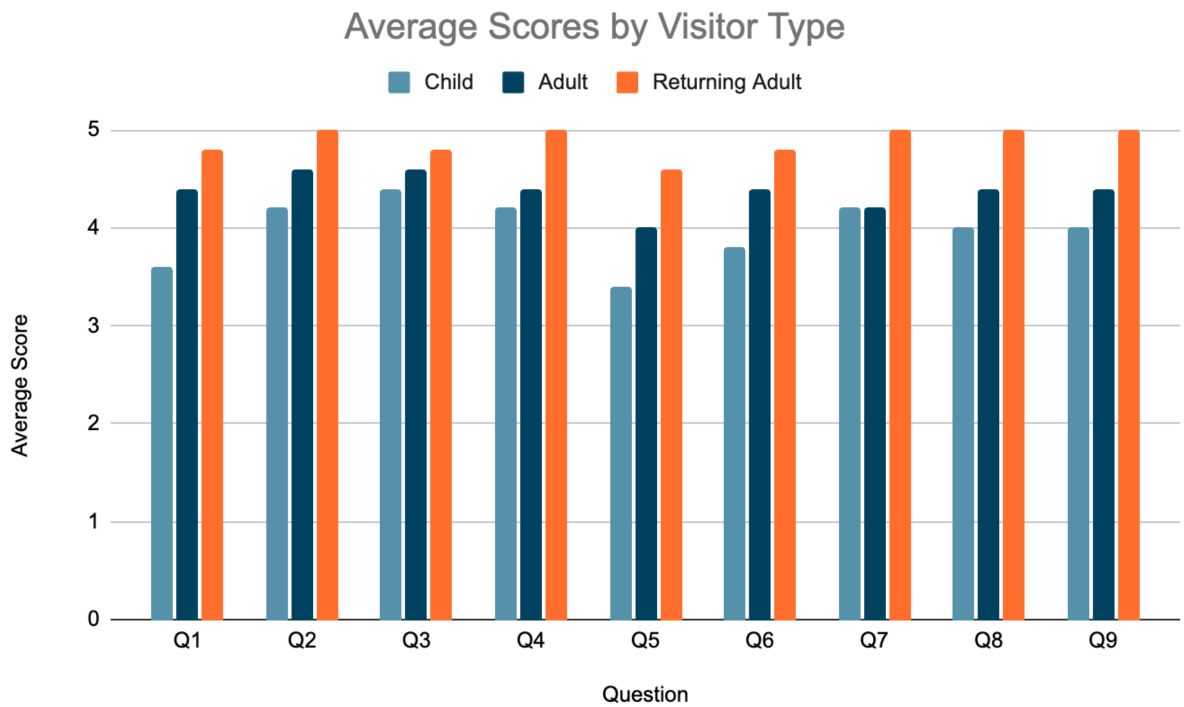
Rate your agreement with the following statements (1 = Strongly Disagree, 5 = Strongly Agree):

1. The robot seemed intelligent. (Godspeed)
2. The robot provided appropriate information about the exhibits. *(*RoSAS*)*
3. The robot was friendly. (Godspeed)
4. I felt comfortable interacting with the robot. (RoSaS)
5. The robot behaved in a human-like manner. (Godspeed)
6. The robot responded appropriately to my behavior. (RoSaS)
7. The robot improved my museum visit. (Overall)
8. The robot adjusted its responses based on my needs. (Functional)
9. The robot personalized the conversation to my interests. (Functional)
10. What would you improve in the robot's behavior? (Open-ended, optional)

Responses:

Visitor	**Category**	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**	**Q8**	**Q9**	**Q10**
Emma	Child	4	5	5	5	3	4	5	4	5	More jokes
Luca	Child	3	4	5	4	3	4	4	3	3	-
Sophie	Child	5	5	5	4	4	4	5	5	5	-
Alex	Child	2	2	3	4	3	2	3	3	2	More games

Visitor	Category	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
Noah	Child	4	5	4	4	4	5	4	5	5	-
Alice	Adult	5	5	4	5	5	5	4	4	4	Increase depth of explanations
Marco	Adult	4	5	5	4	3	4	4	4	5	Speak faster
Fatima	Adult	4	4	4	4	4	4	4	5	4	-
David	Adult	5	5	5	5	5	5	5	5	5	-
Elena	Adult	4	4	5	4	3	4	4	4	4	More historical context
Alex	Returning Adult	5	5	5	5	5	5	5	5	5	-
Beatrice	Returning Adult	5	5	4	5	4	5	5	5	5	More fun facts
Carlos	Returning Adult	5	5	5	5	5	5	5	5	5	Allow me to bookmark favorite exhibits
Dana	Returning Adult	4	5	5	5	4	4	5	5	5	-
Emre	Returning Adult	5	5	5	5	4	5	5	5	5	-



The bar chart shows the average score for questions Q1 to Q9 by category. Children rated the robot positively in friendliness and entertainment but slightly lower in personalization and intelligence. Adults gave favorable ratings, especially valuing robot friendliness and informational content. Returning visitors consistently rated the robot highest across most dimensions, indicating excellent personalization and memory-based interactions.

Question 10 Responses Analysis:

Feedback	Theme
More jokes	Entertainment
More games	Entertainment
Increase depth of explanations	Information depth
Speak faster	Pacing
More historical context	Information depth
More fun facts	Entertainment
Allow me to bookmark favorite exhibits	Personalization

Results

The robot demonstrates adaptive behavior based on its internal reasoning system and structured representation of the environment. It has a persistent user profiling system, so it can distinguish between first-time and returning visitors, and adapt interactions accordingly.

Visitor	Detected Behavior	Selected Strategy
New adult	No profile	General introduction, ask interests
Returning adult (likes technologies)	Profile retrieved	Suggest tech exhibitions with deep explanations
New child	No profile	Playful greeting, introduction quiz
Returning child (likes space)	Profile retrieved	Space jokes, new quizzes

HRI

- Adaptive Interactions:

For example, when a returning visitor is recognized, the robot retrieves their stored profile from internal memory, which includes visited exhibits and interests. It avoids repeating the same information, instead suggesting new unvisited areas of the museum aligned with their preferences. An example of such interaction can be a visitor who previously was interested in the "Space" theme and now is greeted with *"Welcome back! Since you enjoyed the Space section last time, would you like to explore Robotics today?"*. On the other hand, once a new visitor is detected, the robot initializes a blank profile and begins with general introductions, gradually updating the profile based on interactions and engagement.

- Reasoning based on Age:

The robot shows adaptive reasoning based on the age group. If a visitor is a child (under the age of 12), the behavior module switches to a more interactive strategy, and instead of technical facts, it uses simple language, jokes, and quizzes: *Did you know astronauts eat ice cream in space? Want to play a space trivia game?"*. For adult visitors, the robot presents more detailed, technical explanations, following a different path based on their profile.

- Internal Representation:

The robot's reasoning is built on internal rules, such as 'if age < 12, then use playful tone', and these rules dynamically affect speech, gesture, and content selection. The profile is updated in real-time - for instance, if a visitor expresses interest in history, this preference is stored and prioritized in future sessions. The robot's pipeline follows the path: of profile data -> reasoning rules -> selected behavior, showing a high level of personalization and reasoning that goes beyond predefined interactions.

RBC

All benchmarks were conducted in a simulated environment using a virtual Pepper robot, hence no hardware-related issues were encountered. The results reported here are based on controlled test scenarios and represent estimated performance in real-world conditions. Limitations due to simulation fidelity and the absence of real-world noise, sensor variability, and embodiment are acknowledged.

The functional benchmarks that could be evaluated in the simulated environment, such as Multimodal Interaction and Content Delivery got high results. In the case the button was clicked, the robot started the animation/gesture or continued the interaction promptly without any delays. Moreover, the delivered content was clear and complete in all cases for various user profiles.

The task benchmarks acquired high results as well. The simulated tour scenario with full script execution was successful along with the visitor recognition and personalization. In the scenario with the simulated dropout (ASR or button errors), the robot managed to recover the engagement. The quiz flow in the interaction scenarios was smooth and clear to the visitors.

Although no real-world participants were involved, the simulated scenarios aimed to capture the realistic variability by modeling users of different age groups and preferences, simulating interruptions and engagement dropouts.

Conclusion

This project demonstrated the development and evaluation of an adaptive robot museum guide capable of personalizing its interactions based on visitor age, profile (new or returning), and preferences. Through different strategies: playful for children and detailed explanations for adults, the robot was able to engage different visitor groups more effectively than a generic approach. The use of facial recognition and user profiling enabled personalized experiences that increased customer satisfaction and knowledge retention.

Developing this project was a rewarding experience that highlighted the potential of human-robot interaction in real-world settings. Designing adaptive behaviors requires careful consideration of social cues and age-appropriate communication. Throughout the project, we learned valuable lessons about system integration in simulated environments, the limitations of existing robotic platforms, and how contextual adaptation and user modeling can significantly enhance interaction quality.

Future work could focus on enhancing the robot's contextual awareness by incorporating more advanced emotion recognition and natural language understanding. Extending the system to include multi-modal user feedback and adaptive learning from visitor interactions over time could further improve engagement and educational outcomes. Additionally, expanding the robot's knowledge base and interactive content would allow it to cover a wider range of exhibitions and visitor interests. Additional problems worth addressing are accessibility (e.g. people with hearing impairments or users unfamiliar with technology) and group interactions (families or school groups where multiple users interact simultaneously).

From a non-technical perspective, we should consider ethical concerns around face recognition, psychological impacts (over-reliance on robotic explanations or reduced human interactions), and economic accessibility (the need to develop scalable open-source systems).

Overall, this project illustrates the promising role that socially intelligent robots can play in enriching museum experiences, making visits more interactive, personalized, and memorable.

References

1. https://www.discovergermany.com/the-heinz-nixdorf-museumsforum-hnf-experience-technology-hands-on/
2. https://www.hnf.de/en/permanent-exhibition/exhibition-areas/everything-goes-digital/man-robots-living-with-artificial-intelligence-and-robotics.html
3. https://www.researchgate.net/publication/277930820_Evaluating_Group-Robot_Interaction_in_Crowded_Public_Spaces_A_Week-Long_Exploratory_Study_in_the_Wild_with_a_Humanoid_Robot_Guiding_Visitors_Through_a_Science_Museum
4. https://upcommons.upc.edu/bitstream/handle/2117/80877/IJHR-D-13-00060.pdf?isAllowed=y&sequence=1
5. https://www.si.edu/newsdesk/releases/smithsonian-launches-pilot-program-pepper-robots
6. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=7a28fbd549d71abee18010ae1b36855f4f4ff3f3
7. https://www.researchgate.net/profile/Matthew-Burdelski/publication/248452265_Coordination_of_verbal_and_non-verbal_actions_in_human-robot_interaction_at_museums_and_exhibitions/links/5c4301ec458515a4c730eede/Coordination-of-verbal-and-non-verbal-actions-in-human-robot-interaction-at-museums-and-exhibitions.pdf