

Culturally Sensitive Social Robotics for Africa

D2.1 Use Case Scenario Definition

Due date: 31/10/2023 Submission Date: 07/11/2023 Revision Date: 10/03/2025

Start date of project: 01/07/2023

Duration: 36 months

Lead organisation for this deliverable: Carnegie Mellon University Africa

Responsible Person: **D. Vernon**

Revision: 1.3

Project funded by the African Engineering and Technology Network (Afretec) Inclusive Digital Transformation Research Grant Programme				
Dissemination Level				
PU	Public	PU		
РР	Restricted to other programme participants (including Afretec Administration)			
RE	Restricted to a group specified by the consortium (including Afretec Administration)			
СО	Confidential, only for members of the consortium (including Afretec Administration)			

Executive Summary

This deliverable represents the outcome of Task 2.1. It presents a detailed scenario definition for the two use cases: lab tour guide and receptionist. It describes the aim of the use case, the setting, the procedure, and the measurable variables. It provides the basis for the robot behavior specification (D2.2) and the visitor behavior specification (D2.3). The report includes a walk-through of the scenario, providing a decomposition into a time sequence of elementary robot actions. For each action, the deliverable specifies the set of triggers for the action, e.g., input from visitor using speech, spatial movement, or the tablet PC on the robot, and the sequence of movements, expressions, or vocal output that constitute the robot actions.

The report does not detail the layout of the environment in which the scenarios are set as this will depend on the local situation while will be mapped automatically by the Pepper robot during a set-up phase.

Contents

1.	Use-Case Scenario Methodology	4
2.	Use-Case Scenario Specification	4
3.	Use-Case Scenario 1: Lab Tour	5
4.	Use-Case Scenario 2: Receptionist	7
Princi	ipal Contributors	9
Document History		10

CSSR for

1. Use-Case Scenario Methodology

This document defines the two use cases in which the culturally sensitive social robot behavior will be demonstrated: lab tour guide and receptionist.

It does this by setting out the aim of the use case scenario, the setting, and the procedure using a structured walk-through of all the interactions that instantiate the scenario. The purpose of the walk-through is to unwrap the interaction in each use case into micro-steps of elementary robot perceptions and actions and actions by the interaction partner, i.e., the visitor.

On the basis of this unwrapped timeline of elemental perceptions and actions, we then identify in Deliverable D2.3 the measurable sensory indicators required to parameterize and quantify the information about the visitor that is necessary to allow the robot to interact effectively (i.e., in a culturally sensitive manner) with her or him, e.g., locating the position of the visitor, their face, and eyes.

The main form of interaction by the human will be through limited spoken requests and instructions, implemented with automated speech recognition, and, if necessary, the tablet PC on the robot. This unwrapped walk-through provides the baseline data for tasks T2.2 and T2.3.

2. Use-Case Scenario Specification

The next two sections set out a detailed decomposition of each use-case scenario. These decompositions are relatively straightforward but do contain a lot of information. To make it easier to read, the parts of the decomposition – actions, movements and sensory cues, sensory-motor processes, and comments – are colour-coded.

The actions (in green) define the interaction tasks in a relatively abstract and intuitive manner.

The component movements and sensory cues (in black) make explicit all of the constituents of each action. The component movements and sensory cues always refer to the robot's perspective, i.e., they define what the robot does and what the robot sees and hears.

The sensory-motor processes (in red) provide the essential input for the definition of robot behaviour specification (Deliverable D2.2) and the visitor behaviour specification (Deliverable D2.3).

The comments (in blue) add simple explanations of what is happening in the task at that point.

Note that no cultural knowledge is embedded in either use-case scenario specification. This knowledge will be included after it has been identified in Tasks 1.1 and 1.2, formalized in Tasks 1.3 and 1.4, and encapsulated in a knowledge base in Task 5.4.1.



3. Use-Case Scenario 1: Lab Tour

The robot stands at the entrance to the Robotics Laboratory, actively waiting to engage with a visitor. In this scenario, we assume the visitor approaches the robot.

Launch the application

REPEAT	
REPEAT	
The robot makes natural movements	
Adjust gaze and sway gently, flexing hands slightly	Animate behaviour
The robot looks for a person approaching	
Try to detect a face	Face detection
IF a face is detected	
Determine the position of the visitor	Face localization
Look at the visitor	Rotate head to centre gaze on the visitor
Adjust the body posture to face the visitor	Rotate torso to face visitor and adjust gaze
UNTIL a visitor is standing in front of the robot	Face distance estimation
The robot engages the visitor	
Make a welcome gesture	Iconic gesture
Say" Hello! Welcome to the Robotics Lab!"	Speech synthesis
Say "I'm Pepper, your friendly robot tour guide."	Speech synthesis
Check to see that the visitor is looking at the robot REPEAT	
Look at the visitor	Rotate head to centre gaze on the visitor
UNTIL the visitor is looking at the robot	Detect mutual gaze (*)
Say "Would you like a tour?"	Speech synthesis
IF automatic speech recognition (ASR) enabled	
REPEAT	
Listen to the visitor and wait for a response	
Say "I only understand "Yes" or "No"". Pause for a few seconds.	Speech synthesis
UNTIL the visitor says "Yes" or "No", OR the robot has prompted three times	Speech event
ELSE	
REPEAT	
The robot prompts the visitor and waits	
Display the words "Yes" and "No" on the robot's tablet PC Say "Please press "Yes" or "No" on my screen." Pause for a few seconds.	Display dialogue Speech synthesis

UNTIL the visitor taps "Yes" or "No", OR the robot has prompted three times

Tablet PC event

(*) When detecting mutual gaze, the robot only has to determine whether or not the visitor looks at the robot's head, but not necessarily at the robot's eyes. The gaze has to be held for a short period.

ENDIF

CSSR for

D2.1 Use Case Scenario Definition

The robot decides whether to give the tour or look for another visitor

IF the visitor taps or says "No", OR does not tap the screen or say anything

Say "Maybe another time"

ELSE

The robot starts the tour and visits each exhibit in turn

REPEAT

The robot goes to a tour landmark location Select a landmark location on the tour Look up the location on a map Say "Please follow me" Navigate to the location Rotate to face the visitor

Check to see that the visitor is looking at the robot REPEAT Look for a face UNTIL the visitor's face is detected

Determine the location of the visitor Move head to look at the visitor Adjust body posture to face the visitor

REPEAT Look at the visitor UNTIL the visitor is looking at the robot

Point to the featured exhibit (e.g., a lab demo) Identify the exhibit location Point to the exhibit Look at the exhibit Look at the visitor Describe the exhibit

UNTIL all landmark locations have been visited

The robot escorts the visitor back to the entrance

Say "That was the last exhibit on the tour" Say "I hope you enjoyed your visit" Say "Let me show you to way out" The robot looks up the location of the entrance on a Lab map The robot navigates to the location The robot rotates to face the visitor

Check to see that the visitor is looking at the robot **REPEAT**

Look for a face UNTIL the visitor's face is detected

Determine the location of the visitor Move head to look at the visitor Adjust body posture to face the visitor

REPEAT Look at the visitor UNTIL the visitor is looking at the robot

Say "Goodbye!" Wave goodbye

ENDIF UNTIL the application is terminated Look up knowledge-base Look up knowledge-base Speech synthesis Navigation Locomotion

Speech synthesis

Face detection

Face localization Rotate head to centre gaze on the visitor Rotate torso to face visitor and adjust gaze

Face localization Detect mutual gaze (*)

Look up knowledge-base Deictic gesture Rotate head to centre gaze on the visitor Rotate head to centre gaze on the visitor Speech synthesis

Look up knowledge-base

Speech synthesis Speech synthesis Speech synthesis Look up knowledge-base Navigation Locomotion

Face detection

Face localization Rotate head to centre gaze on the visitor Rotate torso to face visitor and adjust gaze

Face localization Detect mutual gaze

Speech synthesis Iconic gesture



4. Use-Case Scenario 2: Receptionist

The robot stands at the entrance to the building, actively waiting to engage with a visitor. In this scenario, we assume the robot approaches the visitor.

Launch the application

REPEAT	
REPEAT	
The robot makes natural movements	
Adjust gaze and sway gently, flexing hands slightly	Animate behaviour
The robot looks for a person close by and listens for a greeting	
Try to detect a face	Face detection
Try to detect a spoken sound	Sound detection
IF a face, body, or spoken sound is detected	
Determine the position of the visitor	Face, sound localization
Look at the visitor	Rotate head to centre gaze on the visitor
Adjust the body posture to face the visitor	Rotate torso to face visitor and adjust gaze
UNTIL the visitor in the centre of the robot's field of view	
Check to see that the visitor is looking at the robot	
Look at the visitor	Face localization
UNTIL the visitor is looking at the robot	Detect mutual gaze (*)
The robot attempts to engage with the visitor.	
In a realistic setting it would only do this if it recognized that the person does	not work or study in the building.
The robot moves toward the visitor	
REPEAT	
Rotate to face the visitor	Locomotion
Move towards the visitor	Locomotion
UNTIL the robot is close to the visitor	Face distance estimation
The robot engages the visitor	
Make a welcome gesture	Iconic gesture
Say "Hello, I'm Pepper, your friendly robot receptionist."	Speech synthesis
Say "I can give you directions to anyone who works here."	Speech synthesis
Check to see that the visitor is still looking at the robot	
REPEAT	
Look at the visitor	Face localization
UNTIL the visitor is looking at the robot	Detect mutual gaze (*)
IF automatic speech recognition (ASR) enabled REPEAT	
Listen to the visitor and wait for a response	
Say "Please say the name of the person you wish to meet"	Speech synthesis
Pause for a few seconds.	
UNTIL the visitor says a name OR the robot has prompted three times	Speech event
ELSE	
REPEAT	
The robot prompts the visitor and waits	
Say "Please type the name of the person you wish to meet"	Speech synthesis
Display a keyboard on the robot's tablet PC	Display dialogue
Pause for a rew seconds.	Tablet DC event
ENDIE	Tablet PC event
EINDIF	

(*) When detecting mutual gaze, the robot only has to determine whether or not the visitor looks at the robot's head, but not necessarily at the robot's eyes. The gaze has to be held for a short period.



The robot finds out whether or not the visitor wants assistance

IF the visitor does not type or say a name

Say "Maybe another time. I'll try to assist someone else"

ELSE

The robot gives directions

The robot looks up the location of the person on a map Determine location of worker's office Look up the location on a map

The robot gestures in the direction of the person's office Say "The person you want to meet is in Room XYZ" Point in the direction of the room Look in the direction of the room Look at the visitor Say "It is located in that direction" The robot gestures goodbye Say "Goodbye!" Wave goodbye

ENDIF

UNTIL the application is terminated

Speech synthesis

Look up knowledge-base Look up knowledge-base

Speech synthesis Deictic gesture Rotate head toward the office Rotate head to centre gaze on the visitor Speech synthesis

Speech synthesis Iconic gesture

Tablet PC event



Principal Contributors

The main authors of this deliverable are as follows (in alphabetical order).

David Vernon, Carnegie Mellon University Africa.



Document History

Version 1.0

First draft. David Vernon. 7 November 2023.

Version 1.1

Fixed problem with header. David Vernon. 16 November 2023.

Version 1.2

Removed "Stop pointing" from Use-Case Scenario 1: Lab Tour. David Vernon. 9 September 2024.

Version 1.3

Removed need for person detection in Use-Case Scenarios 1 and 2. We rely instead on face detection and localization. David Vernon. 10 March 2025.