



Miraikan
THE NATIONAL MUSEUM OF EMERGING
SCIENCE AND INNOVATION



ChitChatGuide: Conversational Interaction Using Large Language Models for Assisting People with Visual Impairments to Explore a Shopping Mall

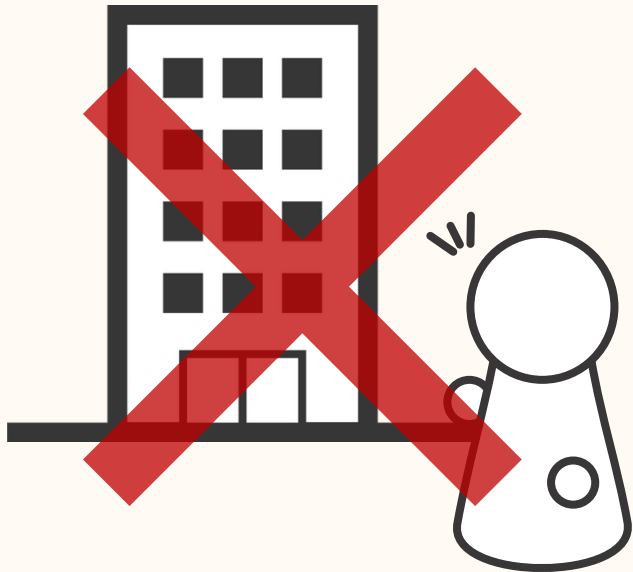
**Yuka Kaniwa^{*1} Masaki Kuribayashi^{*1} Seita Kayukawa² Daisuke Sato³
Hironobu Takagi² Chieko Asakawa^{4,5} Shigeo Morishima⁶**

1. Waseda University 2. IBM Research - Tokyo 3. Carnegie Mellon University
4. Miraikan - The National Museum of Emerging Science and Innovation 5. IBM Research
6. Waseda Research Institute for Science and Engineering (* - equal contribution)

**A system that assists people with visual impairments (PVI)
in exploring a shopping mall through conversational interaction**

It is Challenging for PVI to Enjoy Exploration Independently^[1]

Visit facilities without a specific purpose



Casually explore places that interest them



[1]Kamikubo et al., "We are at the mercy of others' opinion": Supporting Blind People in Recreational Window Shopping with AI-infused Technology

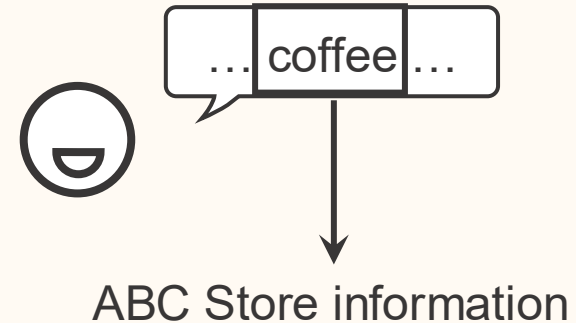
Two Requirements for Providing Information

1. Improve the Autonomy for Exploring Based on Their Interest

List of store names^[2, 3]



Simple voice input^[3]



[2] Kim et al., Navigating visually impaired travelers in a large train station using smartphone and bluetooth low energy

[3] Sato et al., NavCog3 in the wild: Large-scale blind indoor navigation assistant with semantic features

Two Requirements for Providing Information

2. Grasp Surrounding POIs Based on Preferences

**Name and direction
when passing nearby^[3]**



**Brief description of
destination^[4]**



[3] Sato et al., NavCog3 in the wild: Large-scale blind indoor navigation assistant with semantic features

[4] Kayukawa et al., Enhancing Blind Visitor's Autonomy in a Science Museum Using an Autonomous Navigation Robot

Using Conversational Interaction to Fulfill Two Requirements

Conversational interaction allows for flexible information exchange^[5, 6]



We employ **Large language models** to realize conversational interaction

1. Flexibly interact with a deep understanding of conversational context
2. Create personalized sentences in various formats

[5] Gamage et al., What do Blind and Low-Vision People Really Want from Assistive Smart Devices? Comparison of the Literature with a Focus Study

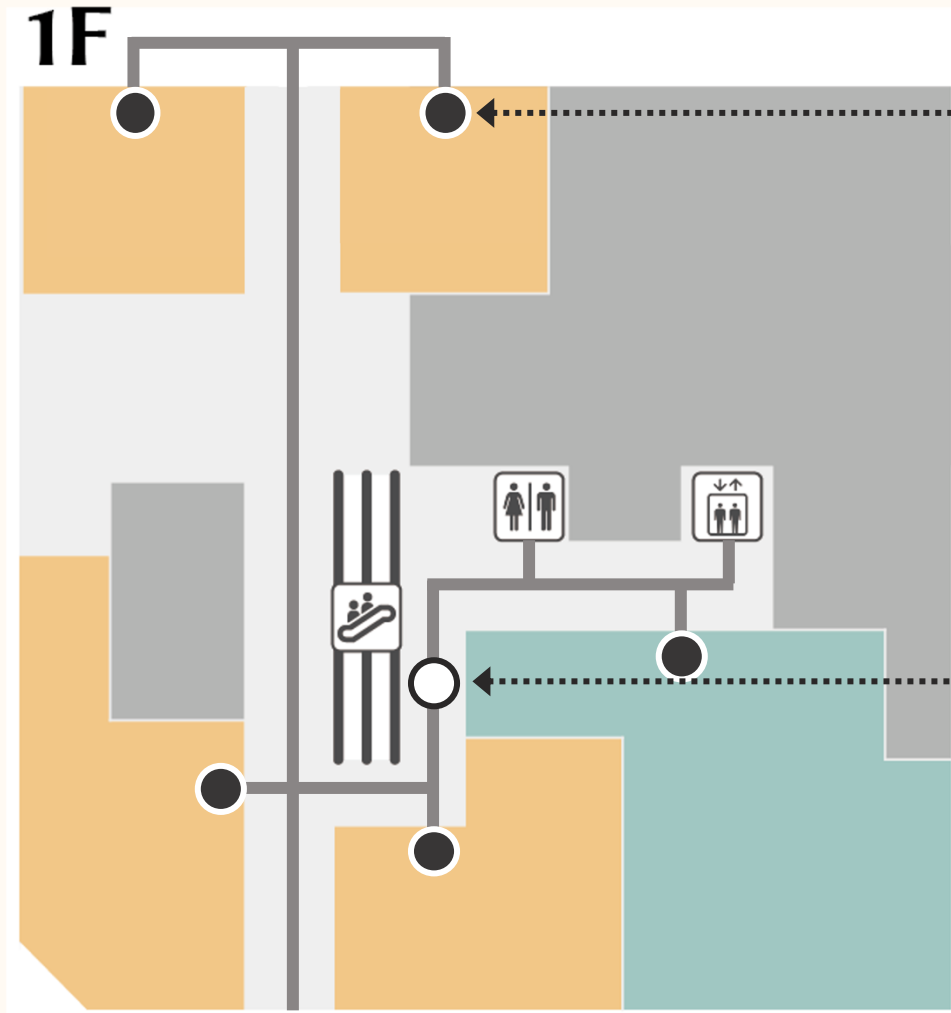
[6] Banovic et al., Uncovering information needs for independent spatial learning for users who are visually impaired

ChitChatGuide:

Assists PVI in exploring a shopping mall
using conversational interaction



Prebuilt Map of a Shopping Mall



POI

- Name
- Detailed Information
 - Shop's concept
 - Recommended items
 - Category
 - Business Hours
 - Phone Number
 - Number of seats etc.

Users' Position

- Floor
- Nearby POIs
- Direction from users

Process Overview

Planning a Tour

Summary

Our shopping mall offers a wide variety of stores. The facility has various shops selling things like bread on the first floor. On the second floor, there are...

Q&A

Tell me two stores you recommend.



The first store is ABC Cafe, which offers handmade scones. The second one is DEF Clothing Store. You can get clothes designed exclusively for this store.

I want to go to the first one.



Of course, ABC Cafe is located on the second floor.

Two Types of Method for Setting Route

Single Destination Route:

Use with a specific destination in mind

Route to a specific destination

e.g., *“I want to go to ABC Cafe.”*

Predefined Route:

Use **without** a specific destination in mind

Manually determined route for describing all the stores on the floor

e.g., *“I want to go through this floor.”*

Overview of Factors Considered when Generating Descriptions

User Preferences

- Identify POIs that need to be described
- Adjust the content of description

Duration of Descriptions

- Adjust the length of description based on walking distance

1) Identify POIs That Need to Be Described

All POIs on the route



Clothing Store



Bank



Shoe Store



Cafe

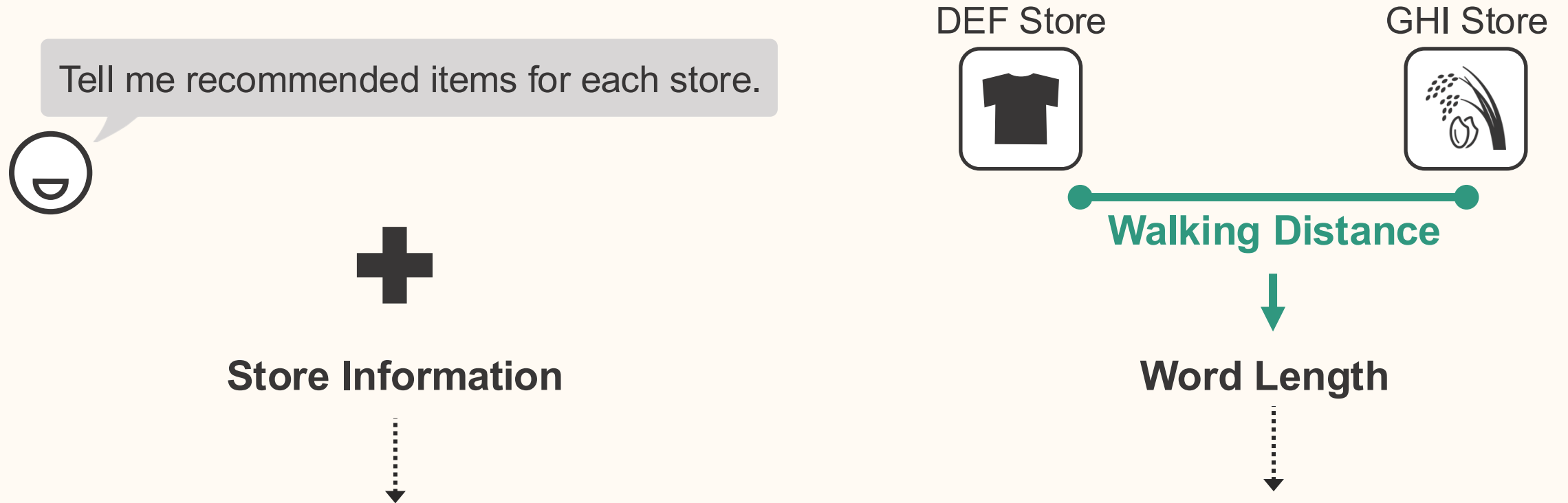
I don't need descriptions related to banks.



POIs need to be described




2) Create Explanation Texts for Each POI with a Specified Length



After a short walk, you will see GHI Store in front of you. This store offers items that allow you to enjoy the taste of natural food restaurants at home. The recommended item is fish broth.



User Study



In-the-wild study with 11 blind participants

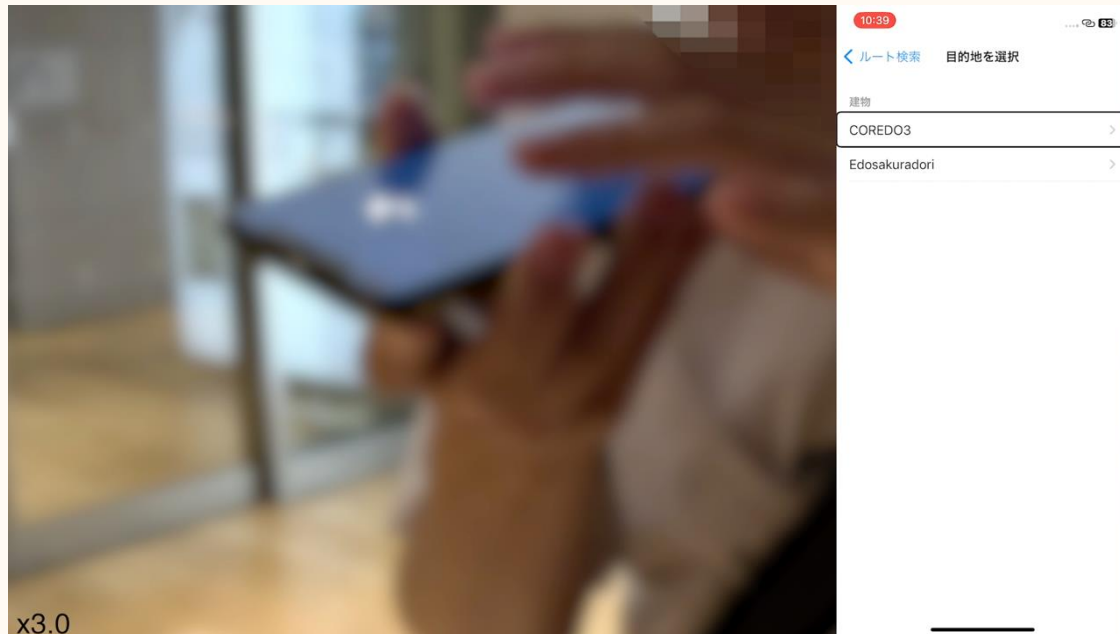
Task: Freely explore the shopping mall for a specified time

Baseline System

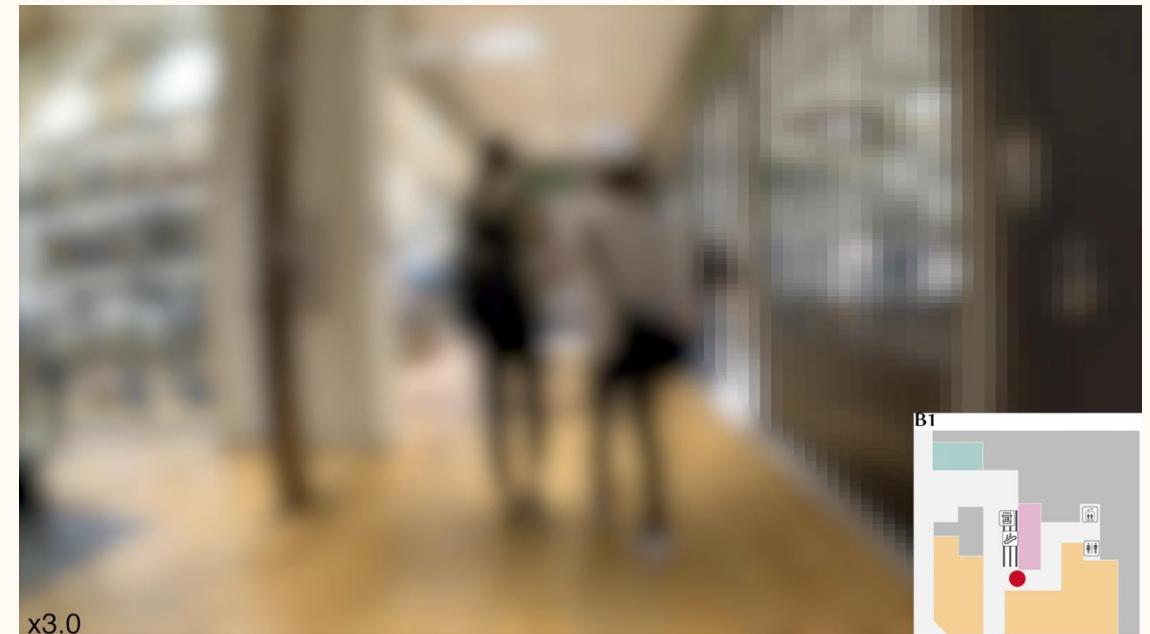
Inclusive Navi^[8]:

A navigation system for PVI publicly available in the shopping mall

Select a destination from a list of store names



Perceive store names and their directions

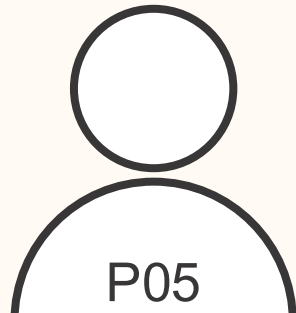
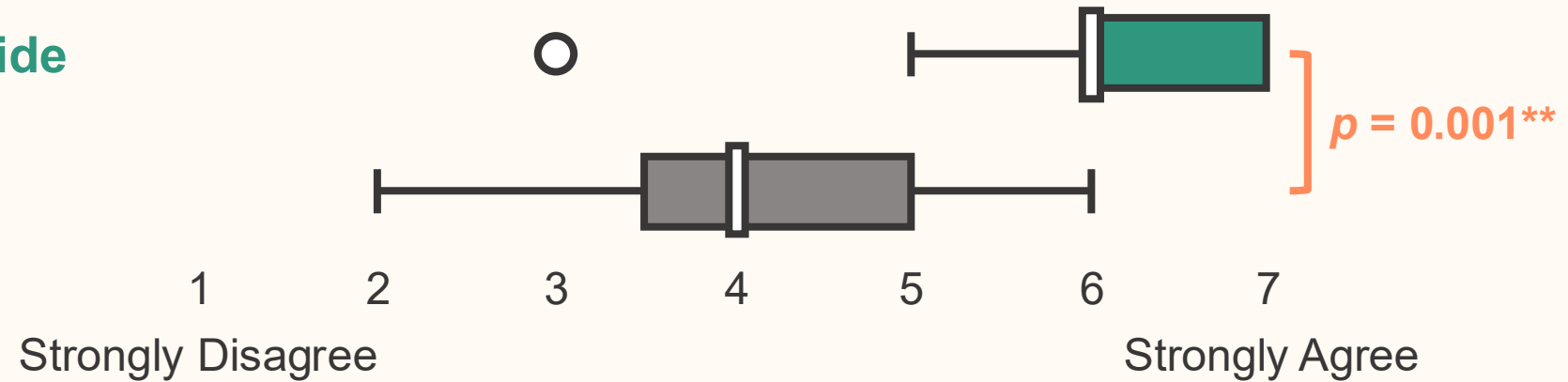


Exploration Experience when Using ChitChatGuide

I **enjoyed** exploring the facility with the system.

ChitChatGuide

Baseline



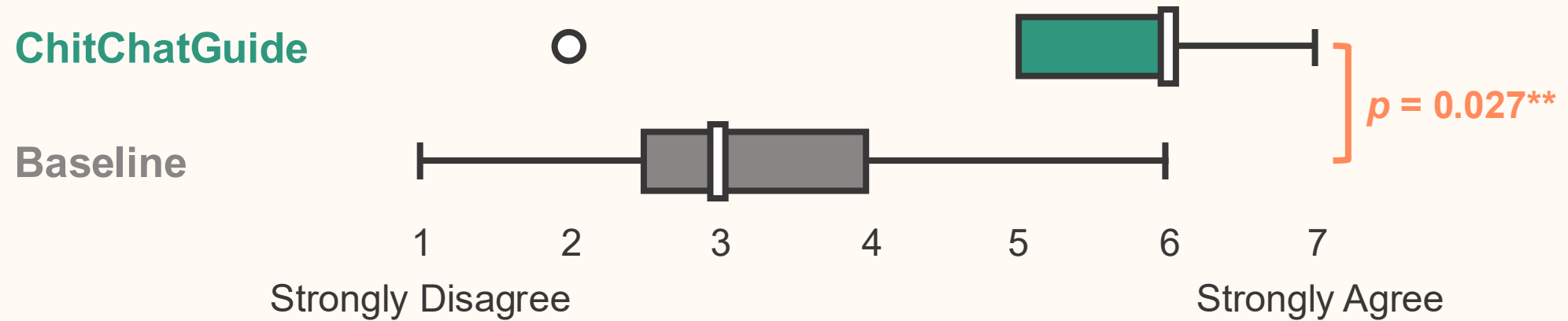
When I go around a certain floor by selecting a predefined route, I can enjoy and see what kind of stores are on the floor, like window-shopping.

Three Criteria for Integrating LLMs into Navigation Systems

- 1) The Balance of Attractiveness and Length of Descriptions**
- 2) Trust in Responses of Conversation Systems**
- 3) The Requirement for the Depth of Information**

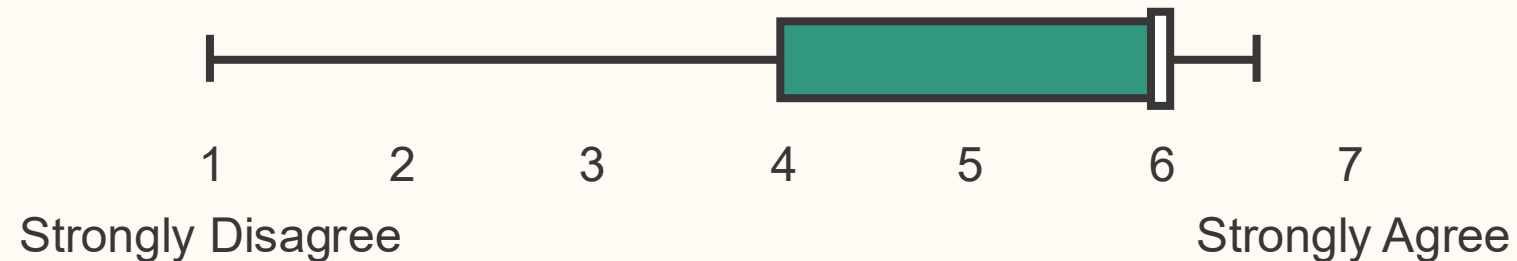
1) The Balance of Attractiveness and Length of Descriptions

The length of descriptions was **appropriate**.



1) The Balance of Attractiveness and Length of Descriptions

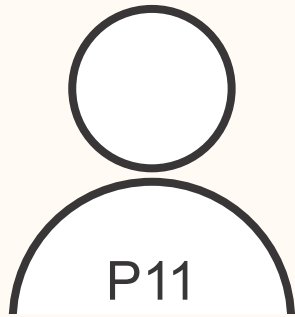
It was useful for exploring the facility that the **length of POI** descriptions was **changed based on transit time**.



Design Implication:

Provide a short description first, details upon user action

2) Trust in Responses of Conversation Systems



*I felt like it was **doubtful** that no store matched my question.
I thought **the responses were not reliable**.*



LLM

False information was occasionally observed

- Misidentifying the **category of stores**
- Offering incorrect **floor information**
- Hallucinations

2) Trust in Responses of Conversation Systems

Design Implication:

Reduce mistakes by using **retrieval-augmented generation (RAG)** method

System



External Database



LLM

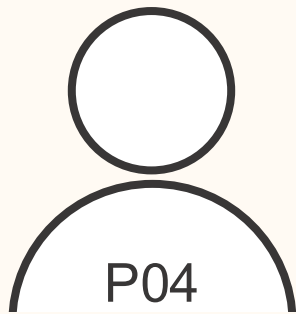


Request to search

Information relevant to user input

3) The Requirement for the Depth of Information

There were cases when the information requested by the participants was not included in the database



Usually, I have to enter a store and ask clerks for details.

It (the Q&A functionality) would help me get information beforehand and decide whether to enter.

Details like recommended items, cuisine, and prices would be very helpful.

Design Implication:

Create a collaborative framework between facility managers and developers

Conclusion

- We developed **ChitChatGuide**, a system that **assists PVI in exploring a shopping mall through conversational interaction** by integrating an **LLM** with a navigation system.
- Our study at a public shopping mall revealed that visually impaired participants were able to **engage in exploration with increased enjoyment**.
- The study also suggested **three criteria for Integrating LLMs into navigation systems**, and we aim to improve the system based on these criteria in future work.

Supplementary Materials

Process Overview

Destination Setting

Summarization

Questions
&
Answering

Request for
POI Description

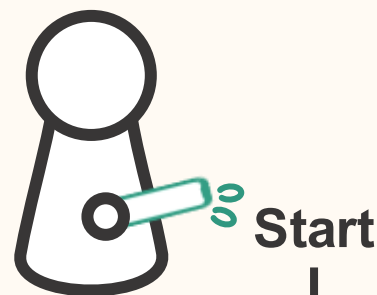
Grasping POIs

POI Description
Generation

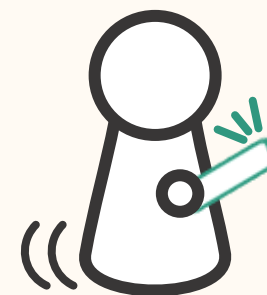
Further
Q&A



Before Navigating



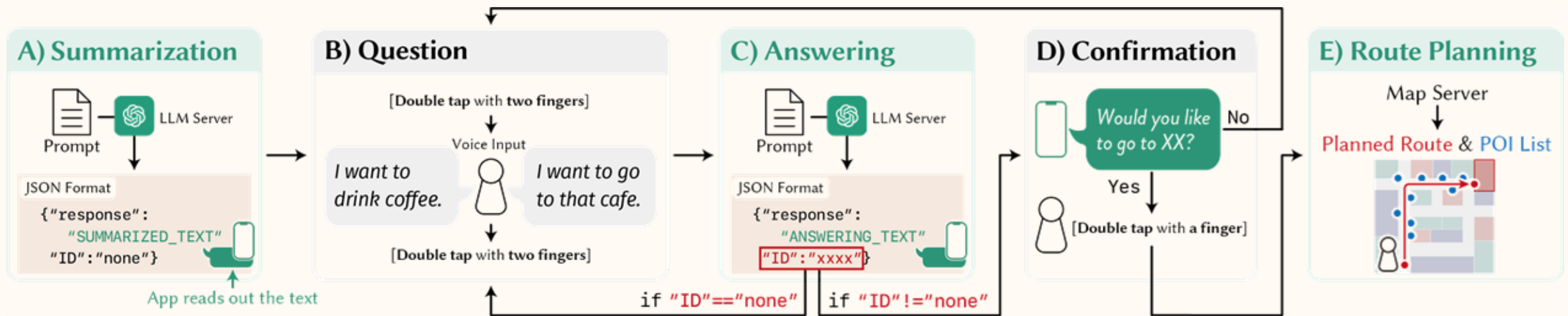
Start



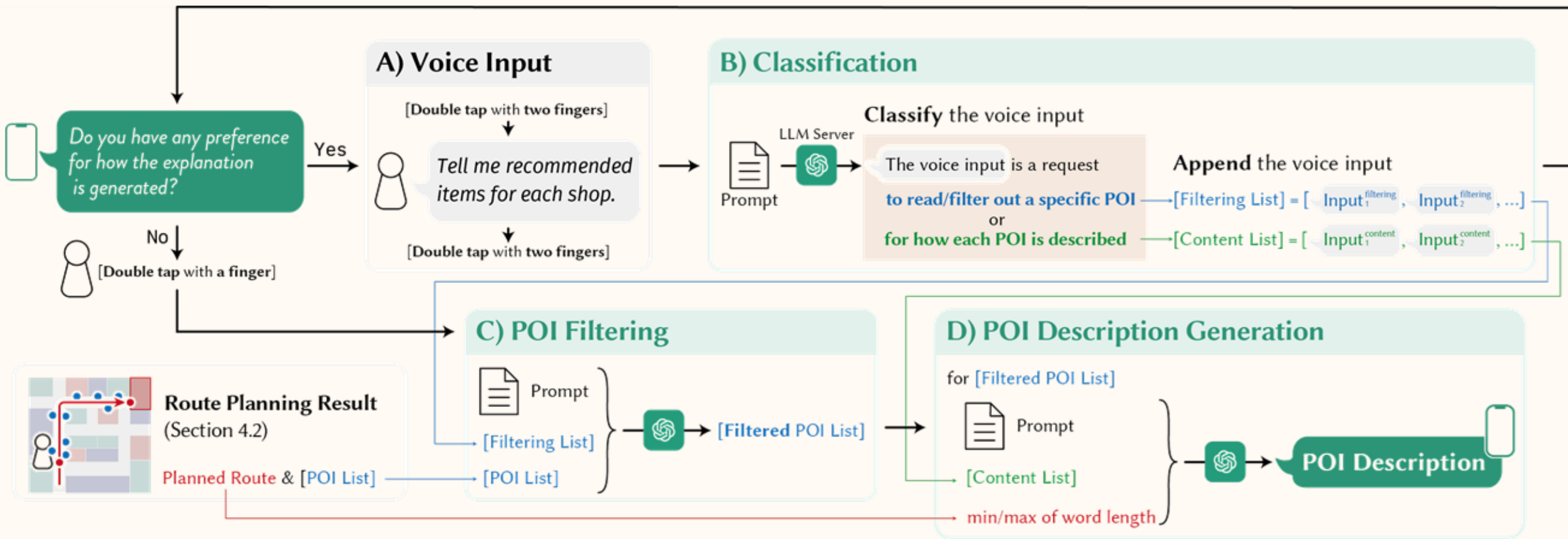
While Navigating



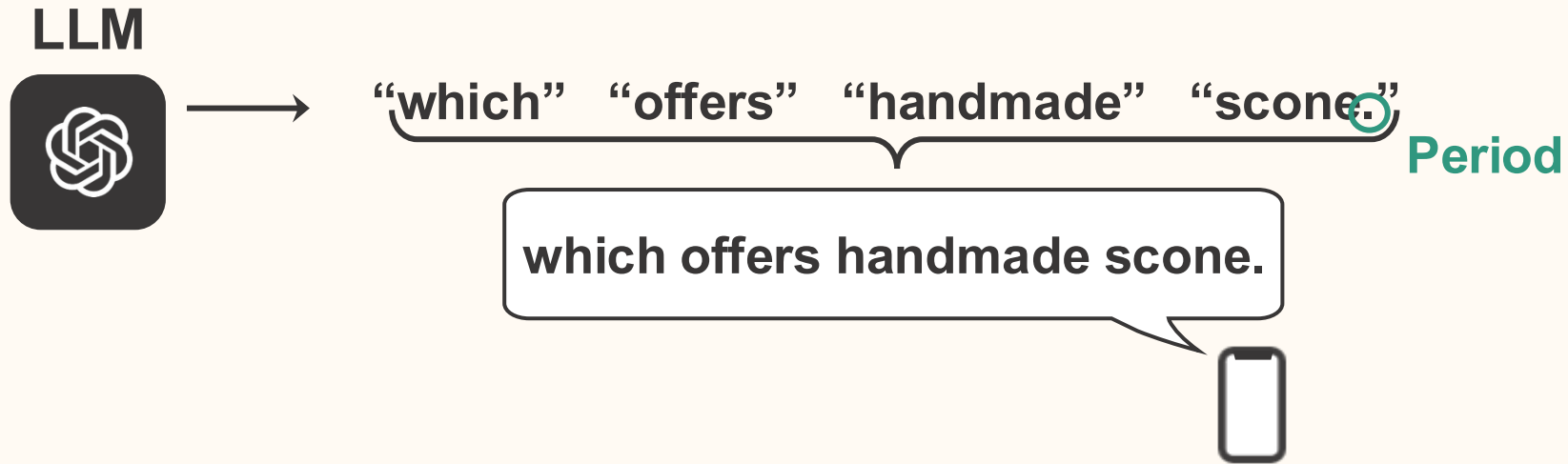
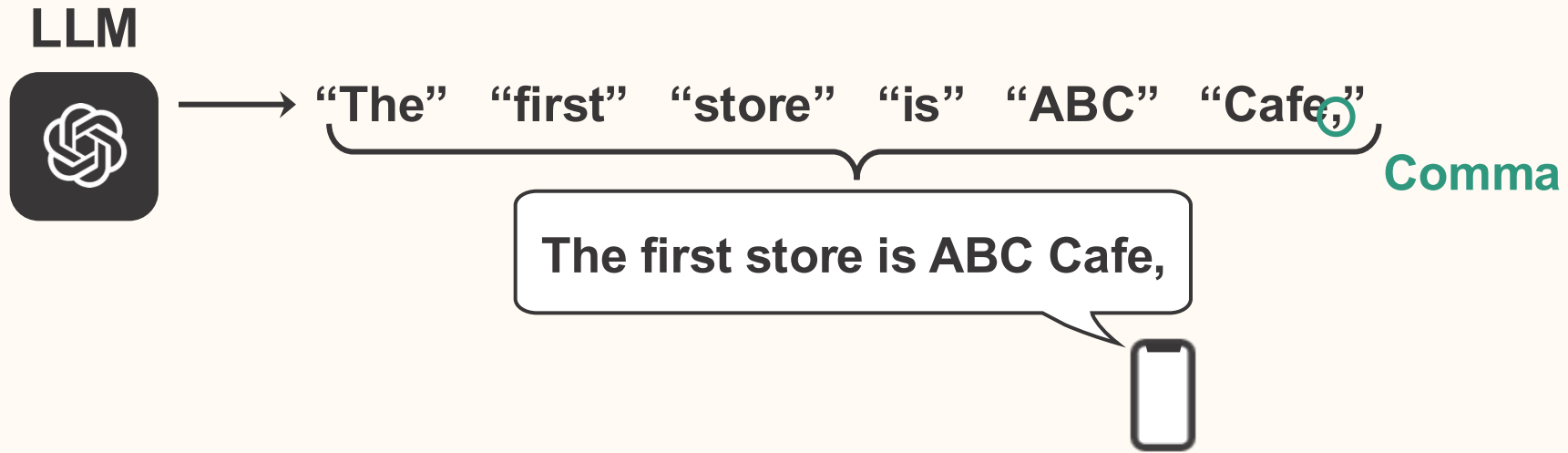
Planning a Tour



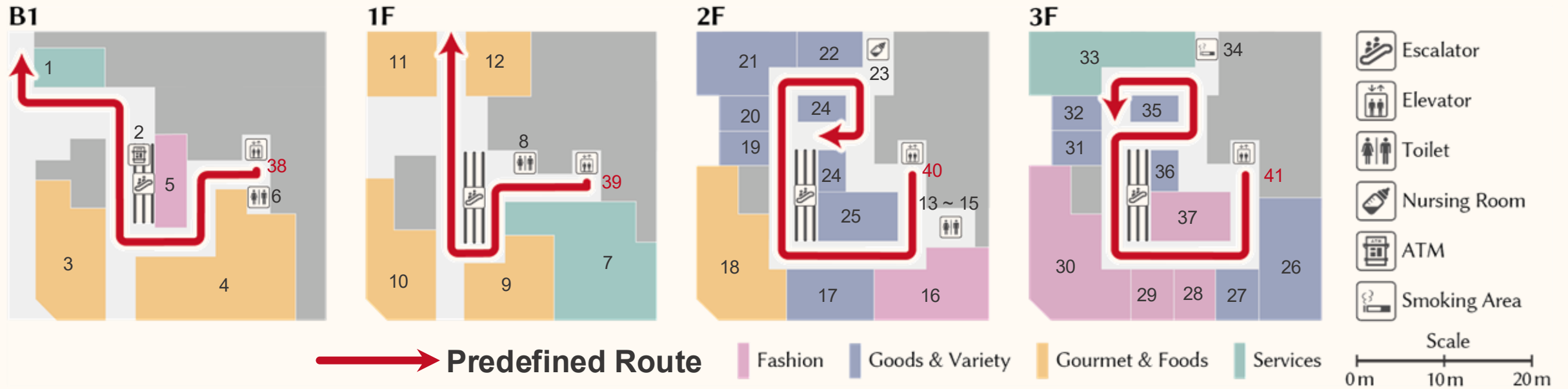
Generating POI Descriptions - Process Overview



Real-Time Reading Out



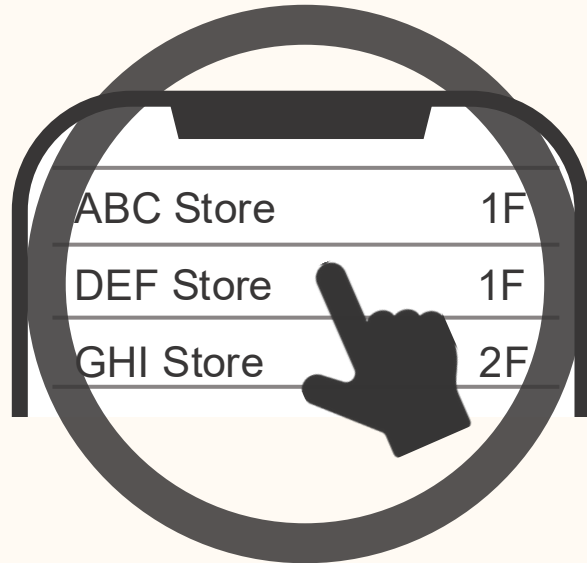
Shopping Mall Information



Data Handling Constraints

Functionality of selecting a destination with voice input was **not** able to use

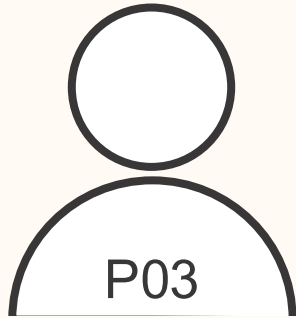
List of store names



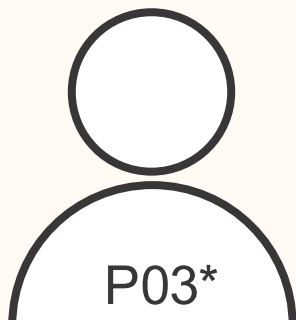
Simple voice input



Another User Comment for the First Criteria



I would prefer all descriptions to be the same length and short. I want to ask questions if the short description interests me.



The system produces lengthy descriptions for stores, regardless of their size or product range. I prefer concise summaries with the option to request more details. I also recommend shorter narrations while users are walking, with automatic expansion when they stop out of interest.


Another Design Implication for the Second Criteria

Provide **relevant recommendations** on **similar or alternative stores**

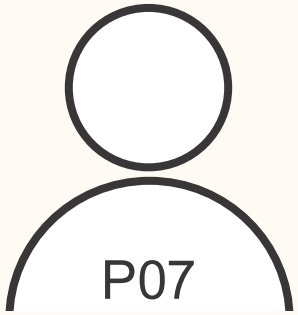
I'm sorry, but there is no coffee shop.



There is no coffee shop at the facility, but the bakery offers drinks as a set menu so that coffee may be available.



Another User Comment for the Third Criteria



*For bargain and sale items, getting a shop assistant to show you those is difficult. And sometimes it's hard to ask for cheaper options. **The system naturally, and fairly, provides that information, which is probably something we often miss out on.***