



Smart Tour Guide

OceanCore : Abdulrahim Aljadani - Taif Aladwni

Table of Contents

Introduction

Smart Tour Guide and data

01

02

Tools

image,voice,text

APP

Application structure

03

04

Pros and cons

Future improvements

05

06

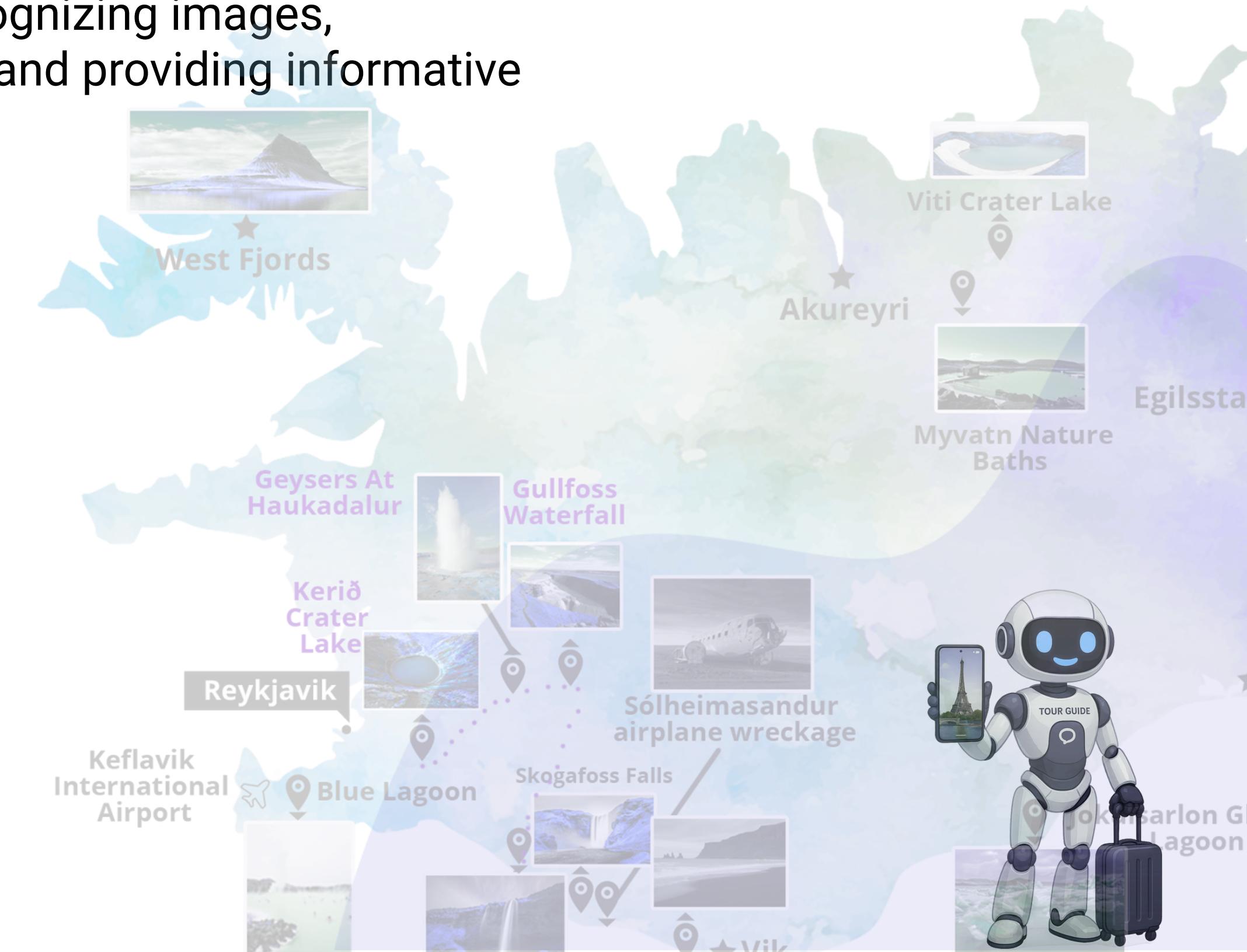
App Demo and Conclusion



◆ Project Introduction

◆ Project Introduction: Smart Tour Guide

The Smart Tour Guide is an AI-powered assistant designed to help users explore tourist landmarks by recognizing images, understanding text or voice questions, and providing informative and interactive answers in real time.



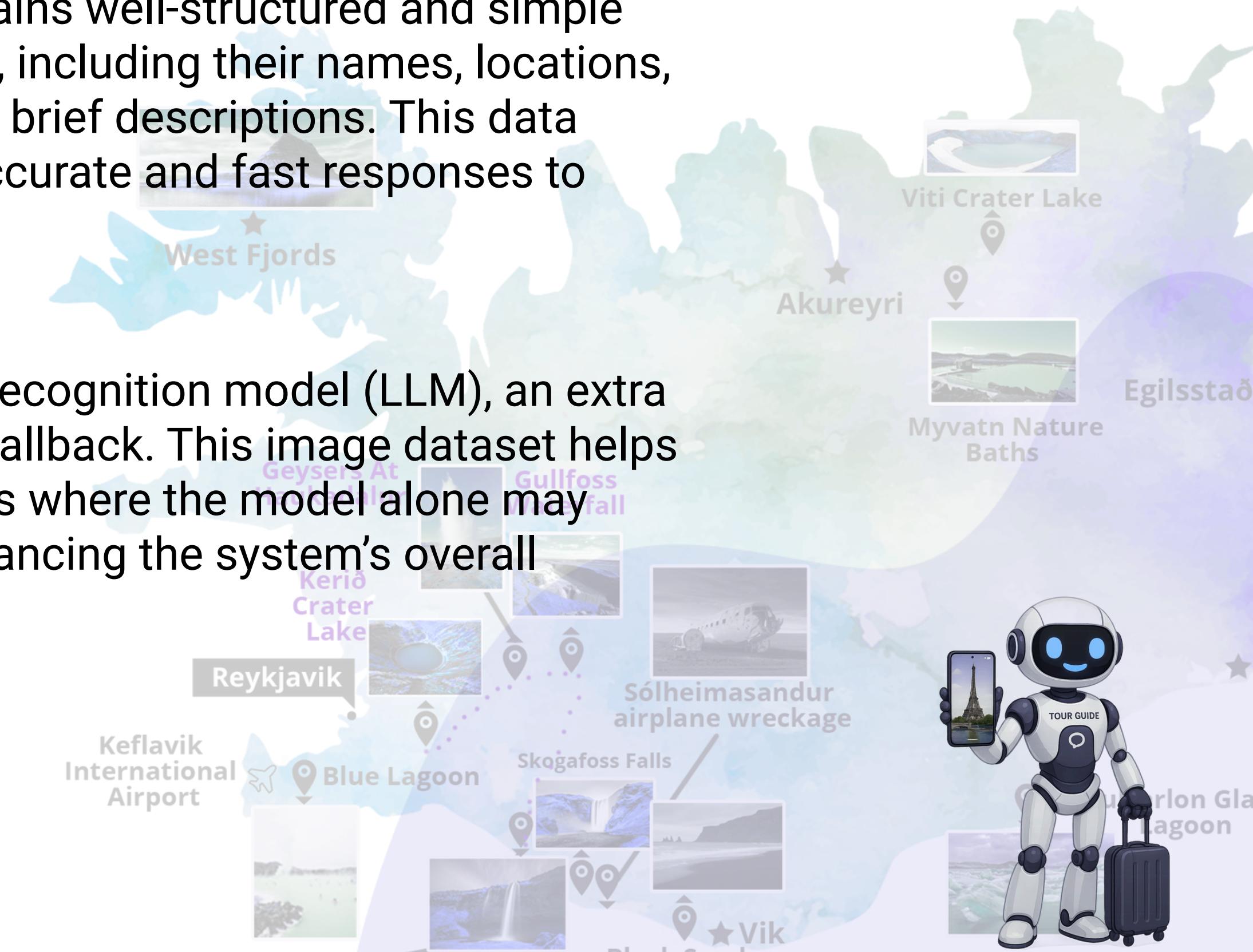
◆ DATASET

◆ Text Dataset

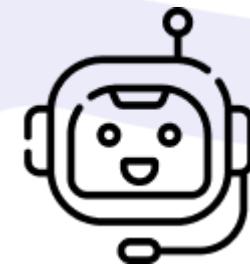
The dataset used in this project contains well-structured and simple information about famous landmarks, including their names, locations, categories (e.g., cultural, natural), and brief descriptions. This data forms the foundation for delivering accurate and fast responses to users.

◆ Image Dataset

In addition to the AI-powered image recognition model (LLM), an extra set of reference images is used as a fallback. This image dataset helps improve recognition accuracy in cases where the model alone may struggle to identify the landmark, enhancing the system's overall reliability.

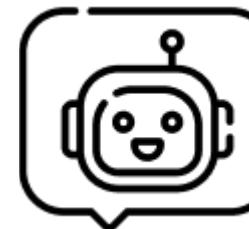


◆ Tools



VOICE

- The user speaks a question using a microphone.
-
- The voice is converted to text using a speech-to-text model (e.g., Whisper).
-
- The converted text is processed just like normal text input.
-



TEXT

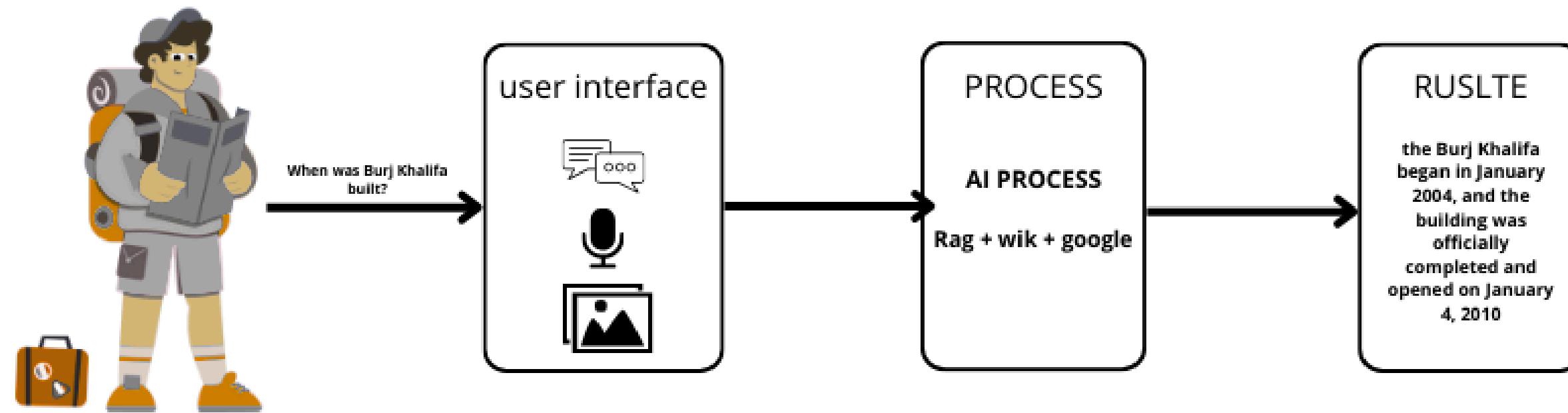
- The user types a question (e.g., "When was the Eiffel Tower built?").
-
- The system uses a language model and the landmark database to generate an accurate answer.
-
- If the answer is not in the database, it may search Wikipedia or use RAG.
-



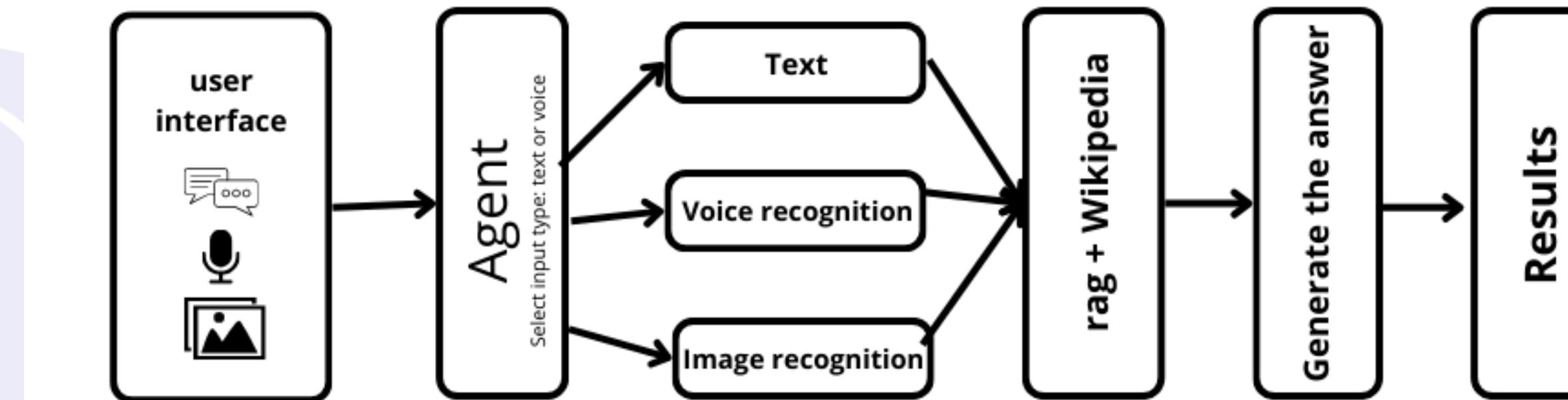
IMAGE

- The user uploads or captures an image of a landmark.
-
- An image recognition model (e.g., CLIP) identifies the landmark.
-
- The identified name is passed to the same system used for text questions to retrieve the answer.
-

Use case



Block Diagram



◆Results

Q1: Where is the Colosseum located?
Prediction: I don't know.
Reference : The Colosseum is located in Rome, Italy.
Score: 0 – Reason: The criterion for this task is the correctness of the submission. The submission should be correct, accurate, and factual.

The input asks for the location of the Colosseum. The reference answer states that the Colosseum is located in Rome, Italy.

The submitted answer is "I don't know." This answer is not correct, accurate, or factual in relation to the question asked. The answer does not provide any specific information about the location of the Colosseum.

Therefore, the submission does not meet the criterion.

N

Q2: What is the name of the famous clock tower in London?
Prediction: The famous clock tower in London is called Big Ben.
Reference : The famous clock tower in London is called Big Ben.
Score: 1 – Reason: The criterion for this task is the correctness of the submitted answer.

The input asks for the name of the famous clock tower in London.

The submitted answer is "Big Ben", which is indeed the name of the famous clock tower in London.

The reference answer also confirms this, as it is identical to the submitted answer.

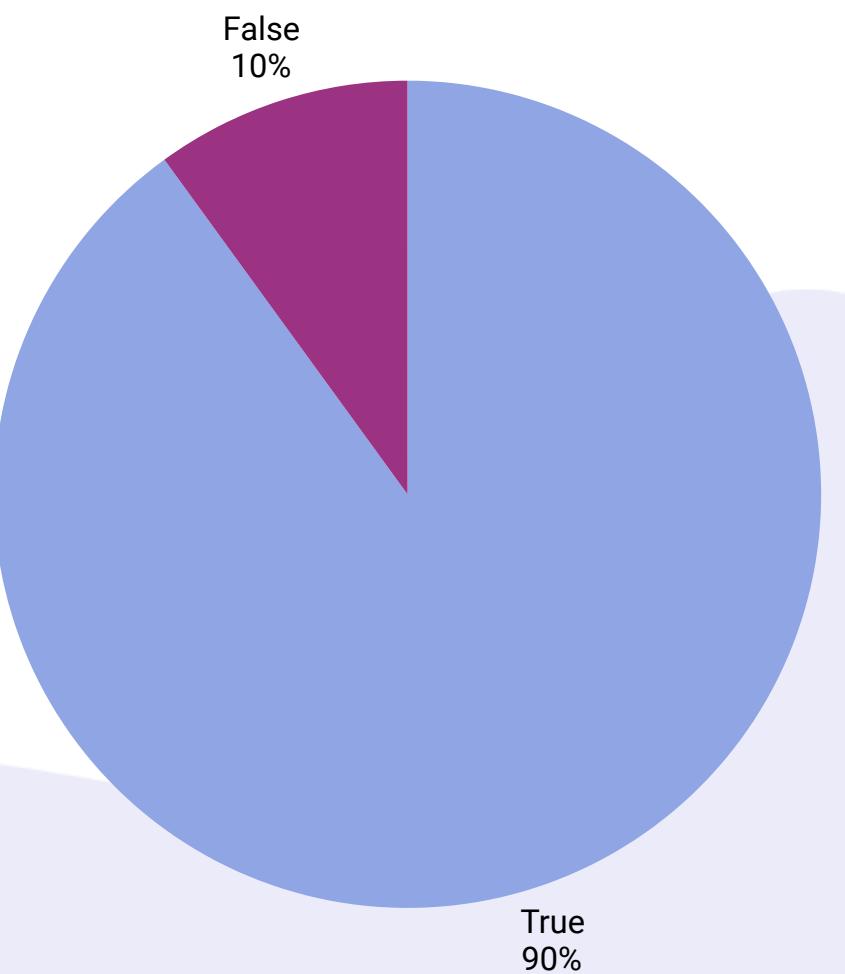
Therefore, the submitted answer meets the criterion of correctness.

...

Y

Agent Accuracy: 90.00% (90/100)

Output is truncated. View as a scrollable element or open in a text editor. Adjust cell output settings.



◆Advantages AND Disadvantages

✓ Advantages

Easy to use with support for text, voice, and image input

Provides instant answers without needing a human guide

Multilingual support (e.g., Arabic and English)

Relies on trusted data sources like Wikipedia or custom databases

Accessible for users with disabilities (visual or hearing)

✗ Disadvantages

Requires constant internet connection for data retrieval

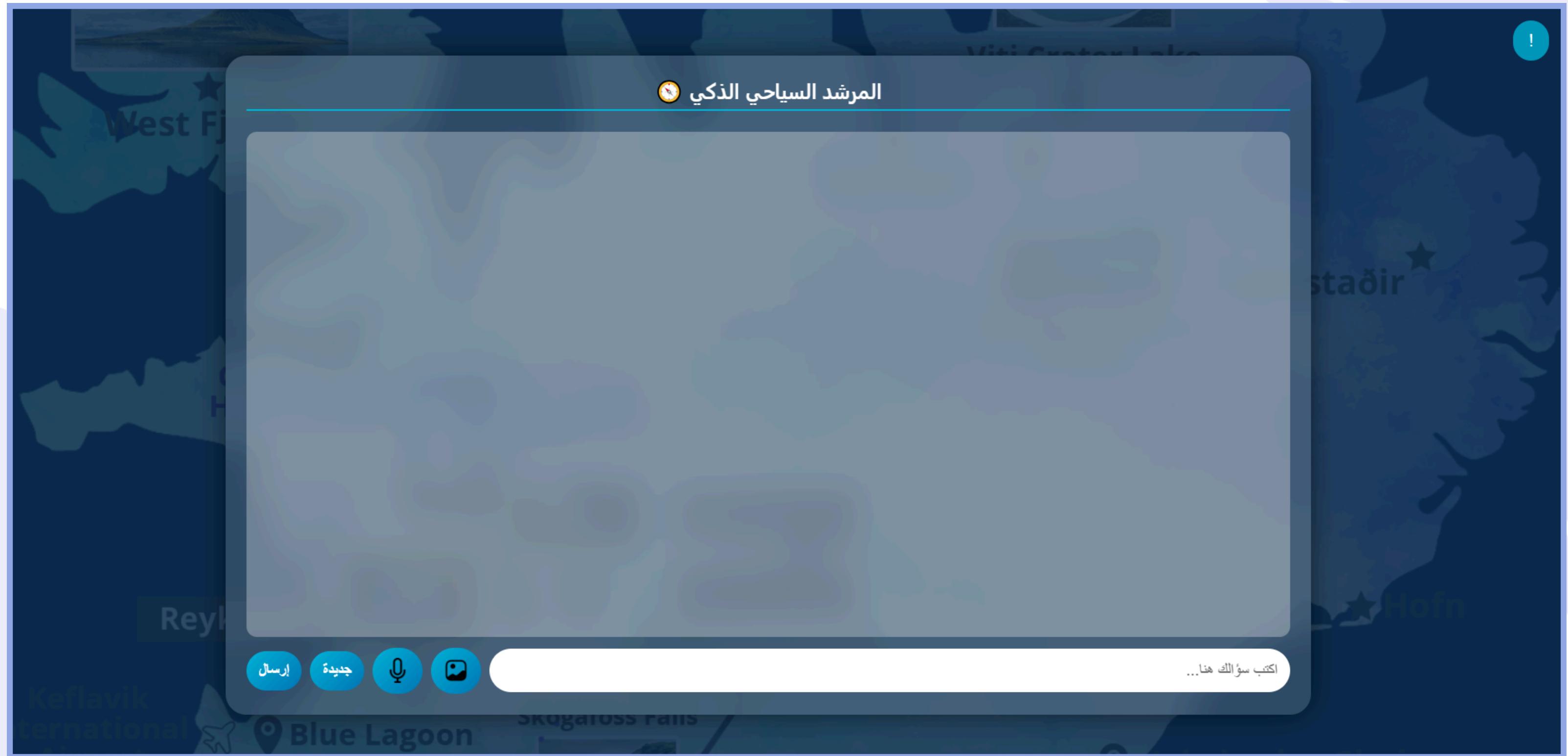
May fail to recognize rare or unclear landmark images

Voice recognition accuracy can be affected by noise or dialects

Limited emotional or cultural interaction compared to human guides

Privacy concerns when handling personal images or location data

◆App Demo



◆Future improvements

- Integrate more advanced AI models for deeper and more natural conversations.
- Add offline mode with a locally stored mini database for remote use.
- Improve image recognition by training on a larger and more diverse image dataset.
- Enhance voice support with better handling of accents, dialects, and noisy environments.
- Use GPS to provide location-based suggestions and nearby landmarks.





Conclusion

Thank you for listening.