CulTour

AI-powered Smart Tourism Application

Group 5: deptrAI

December 16, 2025

Contents

1 Team Information

No.	Full Name	Student ID	Role	Responsibilities
1	Le Dat Phuc An	24125052	Team Lead	Artificial Intelligence
2	Vong Chi Van	24125049	Tech Lead	Artificial Intelligence
3	Tran Ton Minh Ky	24125102	Assistant Lead	Architecture, Backend
4	Pham Tien Dat	24125027	Member	Frontend
5	Nguyen Minh Quan	24125040	Member	Data
6	Pham Nguyen Minh Quan	24125041	Member	Data
7	Quach Thien Lac	24125092	Member	Frontend

2 Idea

The core concept of **CulTour** is to bridge the gap between mass-market tourism and the modern traveler's desire for reality. The idea relies on the increasing global preference for unstaged, authentic experiences - activities that involve discovering local cultures rather than witnessing a curated performance.

CulTour is a **semantic and visual discovery engine**. Instead of treating cultural sites as mere geographical points, we decouple culture into queryable data points (such as Architecture, Building Type, Religion, and Region). This allows us to transition from a **"Staged Authenticity"** model to a **"Utility-First"** discovery model.

2.1 Problem Statement

Cultural tourism in Vietnam often presents a curated performance to meet perceived tourist expectations, while the authentic aspects of daily life remain hidden. Existing solutions offer either **Staged Authenticity** (tourist traps) or **Rigid Context** (generic map markers).

The Authenticity Gap

Travelers face a disconnect between what they want and what is digitally available:

- **High Demand:** 77% of travelers seek authentic experiences representative of local culture[1], and 73% of global travelers now seek to move beyond traditional tourism to find local experiences[2].
- **Low Supply of Information:** Guided tours (e.g., "We Show You Saigon!") and common online resources only show the "front-stage" elements of Vietnamese culture—the popular, crowded sightseeing areas—rather than the full picture of the city.

[1] https://www.booking.com/articles/travel-predictions-2025.html
[2] https://stories.hilton.com/2025-trends-report

The Technical Failure: Rigid Context

Current platforms like Google Maps or Agoda are optimized for logistics, not cultural nuance. This results in a lack of granular search capability:

- **The Pagoda Problem:** A 19th-century pagoda featuring distinct Sino-Vietnamese architecture is treated the exact same way as a newly built pagoda. The platforms lack the metadata to distinguish historical value from utility.
- **Generic Recommendations:** Agoda and similar apps recommend generic places based on commercial ranking, often leading to repetitive experiences.

The Pain Point

- **Authenticity Seekers** cannot find "back-stage" elements or specific cultural filters through generic search engines.
- **Foreigners** find online information about cultural spots overwhelming, generic, and repetitive. Blogs about local cultures are often written in Vietnamese (verbose and inaccessible) or are English articles that lack depth.

2.2 Target Users

- **The Authenticity Seeker:** Travelers who actively avoid tourist traps and are frustrated by the "front-stage" nature of guided tours.
- **The Culturally Curious:** Users who want to understand the difference between specific cultural attributes (e.g., architecture styles, regional history) but lack the local language skills to access native resources.

2.3 System Objectives

To solve the staged authenticity problem, the system aims to achieve the following:

1. **Create a Semantic Discovery Engine:** To decouple cultural locations into extractable data points. The system will allow users to query complex attributes such as *Architecture*, *Building Type*, *Religion*, and *Region*.
2. **Granular Culture Search:** To solve the "Pagoda Problem" by offering deep filtering that distinguishes sites based on historical and cultural context rather than just location.
3. **Auto-Curated Personalization:** To use user preferences to auto-curate events and locations, effectively recommending unstaged experiences that match the user's specific cultural interests.

4. **Utility-First Adoption:** To build a search engine so effective that it attracts a critical mass of users, subsequently incentivizing local organizers to list authentic events on the platform.

3 System Overview

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

4 Sequence Diagram

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu

tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

5 Class Diagram

5.1 Granular Search Service

The search pipeline is built upon a hybrid architecture that combines semantic search and lexical search techniques to optimize both relevance and precision in results. The key components of this pipeline are as follows:

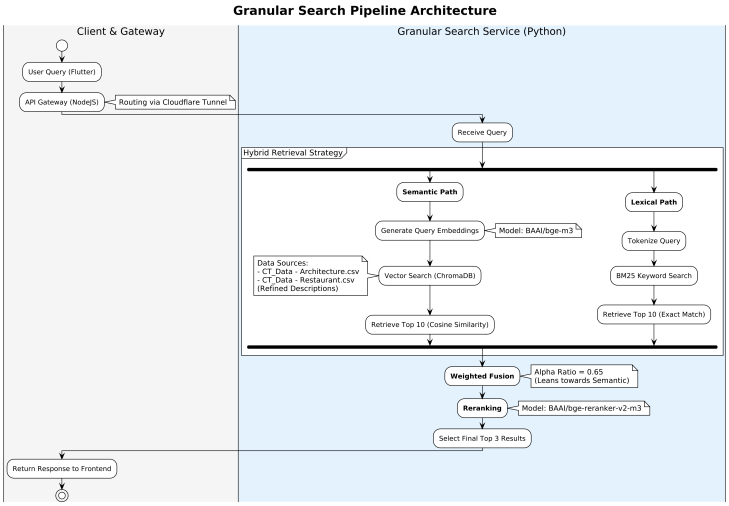
1. **Data Embedding and Indexing:** The system utilizes **ChromaDB** as a vector database. To enhance retrieval accuracy, raw data from architecture and restaurant CSVs is pre-processed into descriptively refined documents before being embedded. We utilize the state-of-the-art `BAAI/bge-m3` sentence transformer model to generate these vector embeddings.

2. **Hybrid Retrieval:** The retrieval phase fuses two distinct methodologies.

   - *Semantic Search:* Calculates **cosine similarity** between the user's query vector and stored document vectors to find the top 10 conceptually relevant results.
   - *Lexical Search:* Utilizes the **BM25** algorithm to identify the top 10 results based on exact keyword matching.

   These results are combined using a weighted fusion mechanism with an alpha ratio (α) of 0.65. This weighting prioritizes semantic search (understanding user intent) while retaining the precision of keyword matching for specific entity names.

3. **Reranking:** The combined pool of candidates undergoes a final filtering stage using the `BAAI/bge-reranker-v2-m3` model. This cross-encoder assesses the relevance of each candidate pair deeply, adding precision and selecting the final top 3 results to be returned to the client.



6 Test Cases and Scenarios

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim.

Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

7 Completed Feature

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

8 Planning

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

9 Technical Solution and Architecture

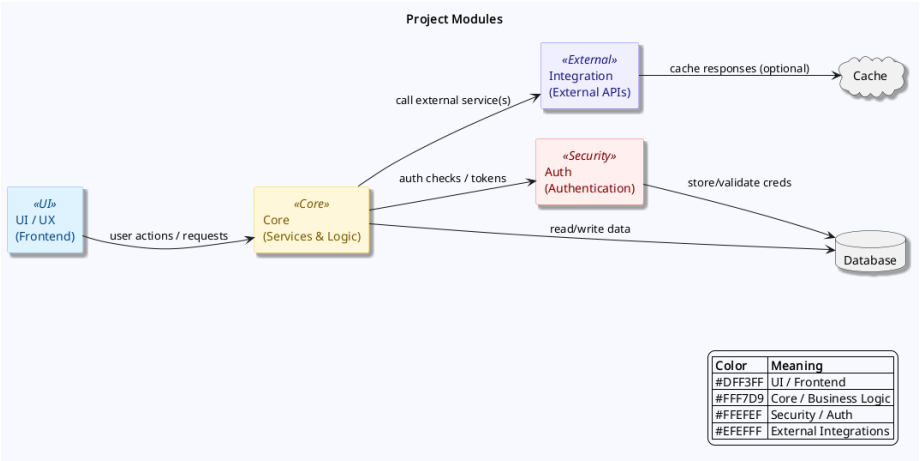
The project, in its core, is a Flutter mobile application that utilizes a NodeJS backend with assistance from Python backend servers that form into a single clean API that the frontend can use. The project's overall architecture can be split into three main architecture styles: the system-level architecture, the client architecture and the backend architecture.

System level architecture. The interaction between the frontend and the backend is a classic three-tier client-server architecture. More specifically, the Flutter client, which we call the *presentation*, utilizes the Node and Python API (the *application*). The backend then utilizes the database and external services to parse upstream data and pass it to the downstream clients. This is the simplest architecture for an MVP achievable within a reasonable time frame.

Client architecture. The frontend is primarily developed vertically via a feature-first architecture. The codebase can be grouped by separate features (including but not limited to login, map screen, trips, saved place, location searching), each has its own logic and UI implementation. We choose this architecture because it is a simple and intuitive architecture primarily derived from the user flow. This ensures ease of development and understanding from the team.

Backend architecture. The backend is a NodeJS express app comprising of multiple endpoints grouped into categories callable from the frontend via dio. In its core, it is an MVC architecture. An endpoint is structured by the core logic's within its service function, express handling and parameters fetching within its controller functions, and routing via an express's Router object. Middlewares for authorization and checking are also included, allowing decoupling and reducing code duplication. Other minor backend servers are also made for training AIs and AI-based searching, and the main backend calls these servers via Cloudflare tunnels. This allows for separation of duties and dependency decoupling.

10 Conclusion and Future Development


Project Modules
«UI»
UI / UX
(Frontend)
user actions / requests
«Core»
Core
(Services & Logic)
call external service(s)
auth checks / tokens
read/write data
«External»
Integration
(External APIs)
cache responses (optional)
Cache
«Security»
Auth
(Authentication)
store/validate creds
Database
Color
Meaning
#DFF3FF
UI / Frontend
#FFF7D9
Core / Business Logic
#FFEFEF
Security / Auth
#EFEFFF
External Integrations