From Photos to Places: Generating Personalized Tourist Itineraries with Large Language Models and Image Analysis

Kenny Andres Bermejo-Xiao[1], Ixent Galpin[1,*]

[1]Universidad de Bogotá-Jorge Tadeo Lozano, Bogotá, Colombia

Abstract

This paper presents the development of an application aimed at the automatic generation of personalized travel itineraries through the integration of two emerging technologies in the field of artificial intelligence: large language models (LLMs) and image recognition systems. The proposal captures user preferences through two mechanisms: (i) a structured form for the explicit collection of interests, and (ii) the analysis of an image provided by the user, from which preferences are implicitly inferred.

The system architecture is based on the combination of multiple artificial intelligence services through APIs, including image interpretation services, location-based services for identifying relevant tourist sites, and language models for the contextual generation of descriptions and suggestions. This integration enables the construction of travel recommendations that consider both logistical constraints (such as flight duration) and the user's specific interests, ultimately presenting a visually enriched and personalized itinerary.

The application was evaluated through a questionnaire adapted from the System Usability Scale (SUS), obtaining an average score of 77.59 out of 100. This result indicates a high level of user acceptance and usability, supporting the effectiveness of the system as an intelligent travel recommendation tool.

Keywords

Personalized travel itineraries, Large Language Models, Image recognition, Recommender systems, Usability evaluation (SUS)

1. Introduction

Nowadays, travel planning has been transformed by the use of intelligent technologies that allow users to access personalized recommendations quickly and efficiently [1]. The integration of artificial intelligence (AI) technologies, such as large language models (LLMs) [2] and image classification systems [3], has opened new possibilities for creating tourism experiences tailored to individual interests and needs. In this context, a *travel itinerary* is considered one of the essential components of the product, with significant commercial relevance depending on its popularity among travelers. It refers to a geographically defined travel route agreed upon for a specific territory and certain points of interest, described in detail when offered to travelers. An itinerary may also refer to a pre-planned route over a set period of time, intended to provide the consumer with services included in the program [4].

AI is a technology that promises to change the world, comparable to the impact of electricity or the Internet. Its influence is expected to be so profound that all perspectives must be represented. AI has the potential to be extremely powerful, with applications in banking, law, and many other fields [5].

ChatGPT is a well-known LLM developed by OpenAI [6]. It operates using natural language processing techniques and is trained on large amounts of text to generate responses to user-provided questions or suggestions. ChatGPT is built on the GPT family, considered one of the most popular language models to date. LLMs like ChatGPT can perform a wide range of tasks, including summarization, information expansion, text translation, prediction, script writing, code generation, and speech synthesis with high accuracy [7].

ICAIW 2025: Workshops at the 8th International Conference on Applied Informatics 2025, October 8–11, 2025, Ben Guerir, Morocco

*Corresponding author.

kennya.bermejox@utadeo.edu.co (K. A. Bermejo-Xiao); ixent@utadeo.edu.co (I. Galpin)

0009-0002-3362-236X (K. A. Bermejo-Xiao); 0000-0001-7020-6328 (I. Galpin)

On the other hand, convolutional neural networks (CNNs) have achieved significant progress in tasks related to image classification and recognition [8]. Recognition is a fundamental, primary, and complex function of computer vision. Through this function, a system is capable of learning to recognize shapes and later classify them correctly, while object tracking makes it possible to analyze and monitor the behavior or trajectories of the studied objects [9].

This work integrates the aforementioned AI technologies to create an innovative project aimed at facilitating the generation of travel itineraries using either images or data provided by the user. This article describes the design, development, and implementation of the application, as well as the tools used, system architecture, and the results obtained. Additionally, a usability evaluation is included through a questionnaire based on the System Usability Scale (SUS), with the goal of validating user experience and the effectiveness of the proposed system.

Our contribution lies in presenting a novel application that integrates existing AI technologies, viz., image analysis and large language models, into a unified framework that supports the creation of personalized tourist itineraries. The emphasis is therefore on the innovative integration and use of these technologies rather than on the creation of new image analysis or language models. We propose an abstract architecture, described in Figure 1, and subsequently show an instantiation of it using Google Vision, ChatGPT, and Google Places

The structure of this article is as follows. Section 2 presents the related work. The design of the application is explained in Section 3. Then, the implementation is described in Section 4. The application is evaluated through usability tests, and the results are discussed in Section 5. The discussion is presented in Section 6, and finally, conclusions are drawn in Section 7.

2. Related Work

Table 1 presents related works that generate personalized travel itineraries, categorized according to whether they use image recognition techniques to capture user preferences, complement the itinerary with image searches for points of interest, and employ LLMs for itinerary generation.

It can be observed that all the projects used AI to generate travel itineraries. Only Zhang *et al.* [20] employs image recognition, using ChatGPT 4.0, which enhances travel planning through multimodal interactions. This approach allows users to compare destinations, plan activities, and choose transportation more efficiently. Other works utilize different LLMs. For example, the design system in Roamify consists of four main stages: data collection, NLP processing, information summarization, and generation of personalized itineraries based on advanced models such as LLaMA and T5 [31]. Similarly, the integration of AI technologies such as Gemini AI improves the efficiency of travel planning by providing real-time, data-driven, and user-preference-based recommendations. Travel planning, once fragmented and tedious, is now transformed by AI into a seamless and effortless experience through automated recommendations and real-time coordination [32].

According to Gupta *et al.* [16], traditional itinerary planning methods are inflexible and time-consuming, often generating generic recommendations that do not account for travelers' individual preferences. Ilieva *et al.* [13] argue that ChatGPT revolutionizes travel decision-making by enabling tourists to access relevant information in real time through a question-and-answer model. Li *et al.* [29] consider that the communication quality of ChatGPT—especially its accuracy, timeliness, and clarity—is fundamental to fostering users' cognitive trust in the context of travel recommendations. Štilić *et al.* [28] point out that most experts consider AI-generated itineraries acceptable, highlighting their organization and suggested activities, although they also stress the need for greater personalization and cultural experiences.

Our work differs from the other proposals described, as it is the only one that combines image recognition, LLMs for itinerary generation, and image search for the points of interest in the itineraries. By unifying these complementary technologies into a single framework, our proposal demonstrates how AI-driven systems can move beyond fragmented solutions toward fully integrated platforms for tourism. This integration has the potential to transform how travelers plan their journeys, reducing the

Citation	Year	Article Name	Image recognition	Image search	Uses LLM
[10]	2023	EverywhereGPT: An AI Travel Planning Assistant Based on ChatGPT			X
[11]	2023	Autonomous travel decision-making: An early glimpse into ChatGPT and generative AI			X
[12]	2024	ChatGPT as a travel itinerary planner.		X	
[13]	2024	Effects of Generative AI in Tourism Industry.		X	X
[14]	2024	DestinAi: Your personalized travel itinerary planner and chat catalyst using generative AI			X
[15]	2024	Revolutionizing tourism: The power of generative ai			X
[16]	2024	Travel With Generator AI: A Novel Approach to Itinerary Creation			X
[17]	2024	Enhanced travel experience using artificial intelligence: a data-driven approach			X
[18]	2024	Transforming Personalized Travel Recommendations: Integrating Generative AI with Personality Models			X
[19]	2024	Using generative artificial intelligence (ChatGPT) for travel purposes: parasocial interaction and tourists' continuance intention			X
[20]	2024	Enhancing Travel Planning and Experiences with Multimodal ChatGPT 4.0	X		X
[21]	2024	Personalized Travel Itinerary Generator System			X
[22]	2024	Integrating generative AI and IoT for sustainable smart tourism destinations.		X	X
[23]	2024	Automated Travel Planning Via Multi-Agent Systems and Real-Time Intelligence		X	X
[24]	2025	Travel Planning with Artificial Intelligence			X
[25]	2025	Enhancing Tailored Travel by Integrating Generative AI with Insights Driven by Personality			X
[26]	2025	Exploring the role of large language model in collaborative travel planning task			X
[27]	2025	Generative artificial intelligence in hospitality and tourism: future capabilities, AI prompts and real-world applications			X
[28]	2025	The Role of Artificial Intelligence in Shaping The Future of Travel Industry: An Expert Analysis of Artificial Intelligence-Generated Travel Itineraries			X
[29]	2025	Navigating the generative AI travel landscape: The influence of ChatGPT on the evolution from new users to loyal adopters			X
[30]	2025	Travel itinerary recommendation using interaction-based augmented data			X
[31]	2025	Roamify: Designing and Evaluating an LLM Based Google Chrome Extension for Personalised Itinerary Planning		X	X

Table 1
Research related to this study, in the areas of travel itinerary generation.

time and effort traditionally required and enabling experiences that are more personalized, visually enriched, and aligned with user preferences.

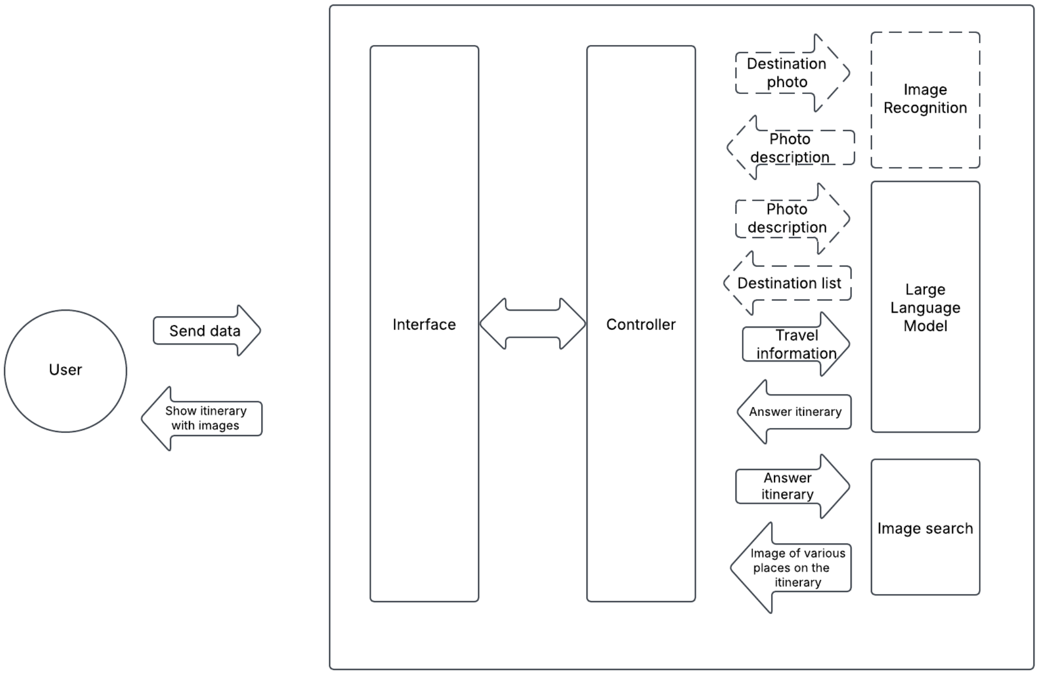


Figure 1: Abstract application architecture.

3. Design

To facilitate personalized travel planning, we present a web application that guides the user through different stages of destination and itinerary selection. The core functions of this web application rely on interaction with external APIs that enable image recognition and text generation. This section describes the proposed architecture. Figure 1 shows the general architecture of the web application, which consists of the following components:

- User: Represents the entity interacting with the web application.
- Interface: Contains the logic and the aesthetic layer of the application, guiding the user through the different stages.
- Controller: Connects and organizes information between the user, internal logic, and external services such as large language models (LLMs), image recognition, and visual content search.
- Image Recognition: Used to analyze photos and extract relevant concepts.
- Large Language Model: Utilizes an AI language model to suggest destinations based on an image or generate personalized itineraries from user preferences.
- Image Search: Used to find images related to some of the locations included in the generated itinerary.

Figure 2 presents the control flow. Initially, the user accesses the main screen and is asked whether they have a destination in mind or prefer to upload an image as a reference. If the user chooses to type in a destination, they are directed to *Screen 2a*, where they are asked to manually enter the name of the destination. Alternatively, if they choose to upload an image, they are directed to *Screen 2b*, where they must upload the image and indicate their city of origin and the maximum number of flight hours they are willing to travel.

In the latter case, when the user uploads an image, it is processed by an image recognition API, which returns descriptions of the content. This information, along with the origin and flight time parameters, is then sent to the language model API, which generates a list of tourist destinations aligned with the image's theme and the allowed flight duration.

Once a destination is selected or entered, the user proceeds to *Screen 3*, where they are prompted to provide additional information: number of travel days, available budget, and personal preferences (such as nature, culture, adventure, etc.). Note that if the user came from *Screen 2a*, they are also asked for their city of origin. With this information, the user can click the "Generate Itinerary" button, which sends the entered data to the language model API to construct a day-by-day travel itinerary.

The generated itinerary is displayed on *Screen 4*, divided by days and accompanied by images obtained through the image search API, related to the tourist locations included in the travel plan.

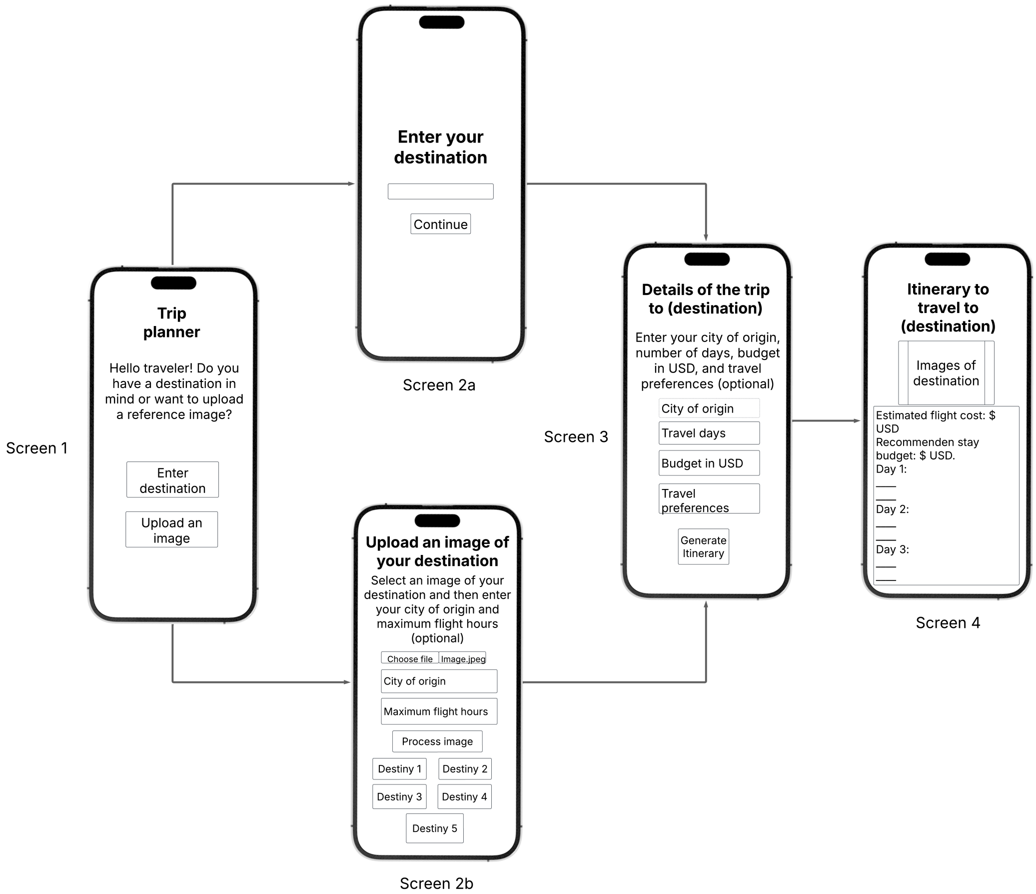


Figure 2: Control flow of our application.

4. Implementation

4.1. Selected Technologies

We instantiate the abstract architecture described in the previous section as a web application, using the technologies described in this section.

Framework and Development Tools: For this application, we selected technologies that are widely used in the modern software industry. On the frontend, we use React[1], a JavaScript library developed by Meta that enables the construction of reactive and modular user interfaces through components. React was chosen for its ease of integration, strong community support, and its ability to create highly interactive, dynamic interfaces with excellent performance due to its optimization capabilities.

On the backend, we use Node.js[2], a server-side JavaScript runtime environment. This technology

[1]https://react.dev
[2]https://nodejs.org/

facilitates efficient communication between the client and external services through an event-driven architecture, which is essential for managing interactions with the third-party APIs involved in itinerary generation.

The development environment used is Visual Studio Code, as it integrates well with modern technologies, has a vast extension ecosystem, and supports collaborative development through integration with Git and other tools.

Image Recognition: To integrate the image analysis feature, we explored different options that allow users to upload an image and generate possible travel destinations based on it. The integrated technology is Google Vision AI[3], a model that analyzes user-provided images (e.g., photos of landscapes or reference locations) and returns descriptions and associated tags. The generated information is essential for inferring potential tourist destinations related to the image.

Generation of Itineraries and Destinations Based on Images: We use a large language model (LLM) from OpenAI to generate a list of tourist destinations related to the uploaded image, as well as to build a personalized itinerary based on the user's preferences. GPT-4o Mini[4] is selected due to its cost effective advanced natural language interpretation, contextual content generation, and its simple integration through JSON-based requests.

Retrieving Images of Tourist Locations Included in the Itinerary: To enrich the itinerary with more informative and visually appealing content, we integrated real images of some of the suggested tourist locations. For this purpose, we used Google Places[5], an API that provides detailed information about tourist spots, including images, names, and descriptions of points of interest.

The combination of these tools enables the development of a smooth and modular application, capable of processing both visual and textual inputs, generating personalized itineraries, and displaying the results clearly and attractively.

4.2. Screen 1: Main Menu

The application's home screen presents the user with two main options: either to enter a tourist destination directly or to upload an image related to a place of interest. This allows the system to accommodate both users who already have a destination in mind and those seeking inspiration. Upon selecting either option, the user is redirected to the corresponding screen (2a or 2b).

4.3. Screen 2a: Direct Destination Input

On this screen, the user manually enters the name of the tourist destination they wish to visit. This information is essential for subsequent steps, where a personalized itinerary will be generated based on the selected location. Once the destination is entered, the user proceeds to the travel configuration screen (Screen 3).

4.4. Screen 2b: Image Analysis and Destination Recommendation

If the user chooses to upload an image, they are directed to Screen 2b. Here, they can upload a photo of a reference location (such as a beach, city, monument, etc.), and they are also asked to provide their city of origin and the maximum number of flight hours they are willing to consider. Once this information is submitted, the image is processed by the Google Vision API, which returns a textual description of the visual content.

This description is then sent as a prompt to the ChatGPT API (OpenAI), along with the user's flight time and origin constraints. ChatGPT analyzes this data and returns a list of possible destinations that match the visual content and user limitations. The user can then select one of the suggested destinations to continue.

[3]https://cloud.google.com/vision
[4]https://openai.com/index/gpt-4o-mini-advancing-cost-efficient-intelligence/
[5]https://developers.google.com/maps/documentation/places/web-service/overview

4.4.1. Extracting Visual Elements with Google Vision

Once the user uploads an image on Screen 2b, it is sent to the server via an HTTP POST request. The image is read in binary format and processed by the Google Vision client, previously configured with a service key. Initially, the `landmarkDetection()` function is used to detect whether the image contains recognizable landmarks or places. If reference points are identified, their names are extracted (e.g., "Eiffel Tower," "Chichen Itzá").

If no landmarks are detected, the `labelDetection()` function is executed, which identifies objects or visual features present in the image, such as "beach," "mountains," "snowy," or "city."

```
import vision from "@google-cloud/vision";

const visionClient = new vision.ImageAnnotatorClient({
  keyFilename: "./clave-google.json",
});
const [landmarkResult] =
  await visionClient.landmarkDetection({ image: { content: imageBuffer } });
// If landmarkAnnotations is empty, fallback to labelDetection
```

4.4.2. Image Visual Element Analysis with ChatGPT

The ChatGPT API, provided by OpenAI, enables programmatic interaction with language models such as gpt-4o-mini, which is used in this project. This API receives as input a set of messages (in JSON format) that simulate a conversation between a system and a user. Each message has a role (system, user, or assistant) and textual content. The model's generated response is also returned as a message, which can be easily processed.

In this web application, the API is used to transform the visual elements detected in the uploaded image by Google Vision (such as "mountains," "beach," "sunset") into a list of possible real-world tourist destinations. To achieve this, a message with the role "user" is constructed, which includes the visual elements identified by the Google Vision API and asks the model to suggest three to five related tourist destinations.

Additionally, if the user entered their city of origin and a maximum flight duration, these constraints are added to the prompt so the model can geographically filter the suggestions. An example of a prompt sent to the model might be the following:

```
let prompt = ‘Based on these visual elements: beach, palm trees, sunset...
what tourist destinations match these characteristics?...
Only consider destinations reachable by plane from Bogotá in a maximum of 6 hours.‘
```

The model's response is a plain-text list of destinations, which is then split into individual lines to convert it into an array that can be visually rendered in the interface.

4.5. Screen 3: Travel Configuration

In the next step, the user provides additional trip details: number of days, available budget, and personal preferences (such as interest in beaches, nature, urban life, etc.). If the user came from Screen 2b, their city of origin has already been captured and is not requested again. Once the data is entered, the itinerary generation is triggered via an action button.

4.6. Screen 4: Itinerary Generation

In this final screen, the personalized travel itinerary is presented, generated through a query to the ChatGPT API. The itinerary is displayed day by day, with suggested activities tailored to the selected destination, budget, duration, and user preferences. Additionally, relevant images of tourist spots are integrated using the Google Places API, enhancing the user's visual experience and immersion in the trip planning process.

Below is the prompt sent to ChatGPT for itinerary generation:

```
let prompt = ‘Generate a detailed 3-day itinerary in Cartagena
based on the following
preferences: "beach and history"

It must include:
- An estimated flight cost from Bogotá.
- Accommodation, food, and daily activities aligned with the user’s preferences.
- A clear breakdown of daily costs.
- Comparison with the budget of 1500 USD.

**Important**: Return the response in valid JSON format, with no additional text,
no comments, and no Markdown formatting.

Expected format:
{
  "flightCost": "Estimated flight cost in USD",
  "totalBudgetNeeded": "Recommended total budget",
  "lowBudgetWarning": true/false,
  "itinerary": [
    {
      "day": "Day 1",
      "activities": [
        "Description of activity 1",
        "Description of activity 2",
        "Dinner at a recommended restaurant",
        "Total cost of the day: X USD"
      ]
    },
    ...
  ]
}‘
```

Once the itinerary is received, at least four key places mentioned in the generated text are extracted. For each of them, a query is sent to the Google Places API to retrieve representative images. These images visually enrich the presentation of the itinerary to the user.

The Google Places API provides access to a large database of geographic locations, including tourist attractions, restaurants, museums, parks, and other points of interest.

The system automatically detects the names of key places mentioned in the text (e.g., "Parc Güell," "Louvre Museum," "Central Park"). Once those names are extracted, the backend sends a query to the Google Places API, specifically using the `findPlaceFromText` or `textSearch` methods, to search for the place in Google's database.

Then, using the `placeid` returned by that search, a second request is made to the API to retrieve place details, including available photos. These photos are accessed via the 'photos' field, which contains an array of image references. To display an image, a special URL provided by Google is used: https://maps.googleapis.com/maps/api/place/photo.

With all this, the images are retrieved and then displayed above the generated itinerary text.

4.7. Example Execution

Figure 3 shows screenshots using the manual destination input function. In Figure 3a, the initial screen is presented, where the user can choose between two options: enter a destination manually or upload an image. Next, Figure 3b shows *Screen 2a*, intended for the user to input the destination. Meanwhile, Figure 3c corresponds to *Screen 3*, where the user must fill in the trip details: city of origin, number of days to stay at the destination, budget, and optionally, personal preferences. Finally, Figures 3d and 3e show the generated itinerary screen, which includes illustrative images of the recommended places obtained through the Google Places API.

Figure 4 shows screenshots using the image upload function. Figure 4a displays screen 2b, where the user can upload an image, indicate their city of origin, and optionally set the maximum flight hours they are willing to take. Then, Figure 4b shows screen 3, where the user must enter the number of days they want to stay at the destination, budget, and optionally, personal preferences. Lastly, Figure 4c presents the generated itinerary screens, which include illustrative images of the recommended places obtained through the Google Places API.

5. Evaluation

Usability testing plays a fundamental role in the development and refinement of products and systems to ensure that user needs and expectations are met. By conducting systematic evaluations of the user experience, it is possible to identify interaction obstacles, detect design flaws, and gather valuable insights into how users perceive and use the application. These findings enable iterative adjustments that improve the system's effectiveness, satisfaction, and accessibility, contributing to a more intuitive, functional, and user-centered solution.

5.1. Usability Testing

On May 6, 2025, a usability test of the application was conducted with a group of 27 students from Universidad Jorge Tadeo Lozano,during a Data Structures class. The test began with a brief explanation of the application's functionality so that the group could understand its purpose and features. After this introduction, each participant was asked to download an image from the web related to the travel theme (landscape, city, monument) in order to test the image recognition function.

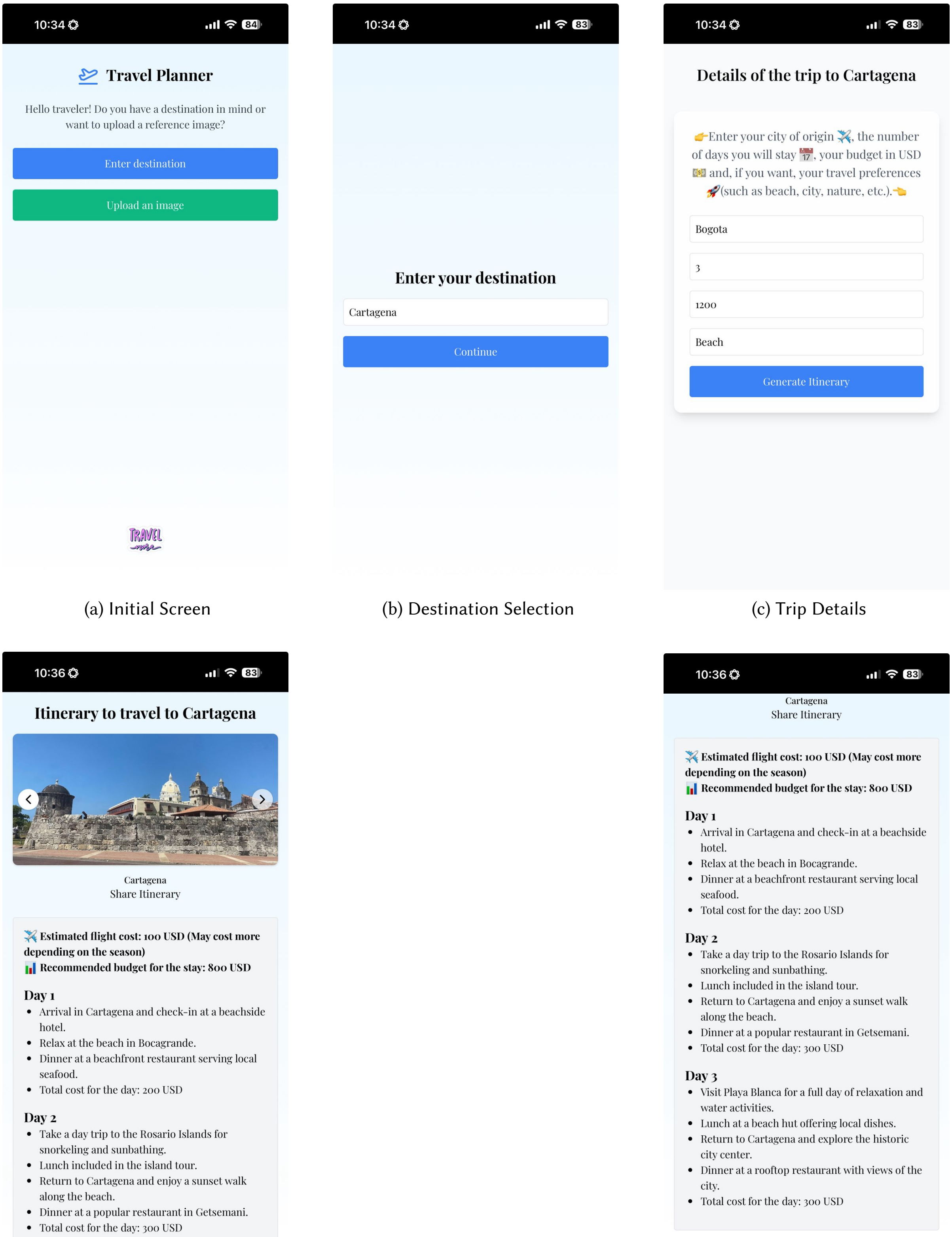
After each student tested the image recognition feature, they were asked to complete a questionnaire evaluating usability and user experience. This questionnaire included statements from the System Usability Scale (SUS), a method that helps assess the usability of any system. The statements were rated from 1 (Strongly Disagree) to 5 (Strongly Agree). The responses were collected using the Microsoft Forms platform. Additionally, users were asked to upload the image they used with the image recognition function and describe its content. Finally, they were invited to write a comment or suggestion about the application.

Table 2 shows the questionnaire adapted from the System Usability Scale along with the average scores obtained for each item.

Statement	**Average Score**
I would consider this application a useful tool for planning future trips.	4.33
I felt the application has unnecessary functions or that complicate its use.	2.15
Navigating the application was intuitive.	4.33
I thought it would be difficult to understand how to use the application without prior assistance.	2.56
The different sections of the application work well together.	4.48
During use, I noticed errors or failures that affected my experience.	2.11
I believe anyone could quickly become familiar with the application.	4.37
The structure or design of the application confused me.	2.00
Using this application made me feel confident and comfortable.	4.22
I had to spend time understanding the application's logic before using it.	2.07

Table 2
System Usability Scale (SUS) questionnaire with average scores obtained.

Results: Figures 5 and 6 present graphical representations of the responses to statements 1 through

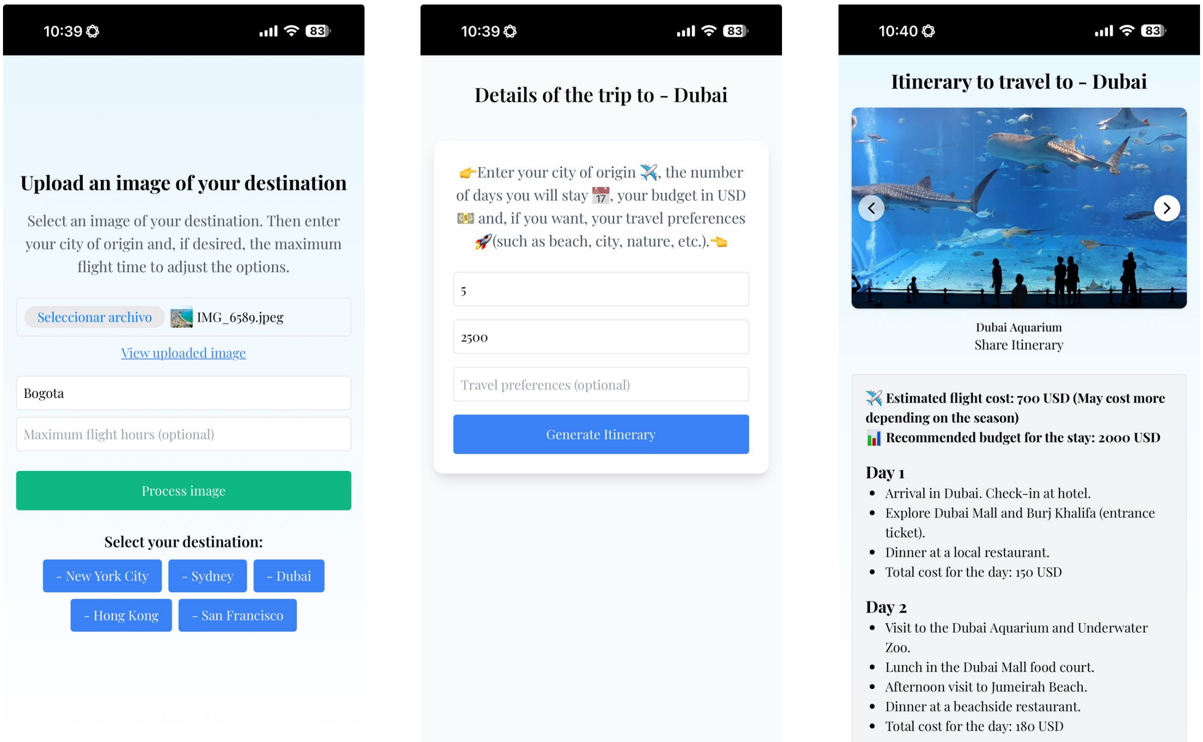


(a) Initial Screen (b) Destination Selection (c) Trip Details

(d) Destination Images (e) Generated Itinerary

Figure 3: Screenshots using the manual destination input function.

6 and 7 through 10, respectively. Figure 5a shows the results for the statement "I would consider this application a useful tool for planning future trips." Most participants responded with 4 (Agree) or 5 (Strongly Agree), with only three participants selecting 3 (Neither Agree nor Disagree). The



(a) Image Upload (b) Selected Destination Details (c) Selected Destination Images

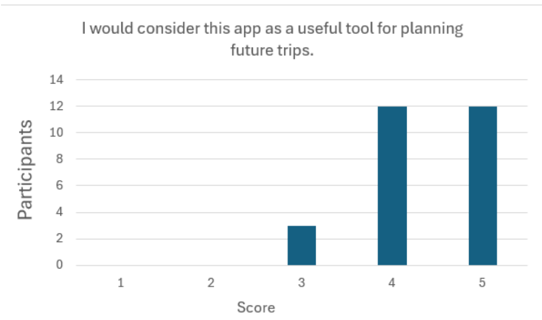
Figure 4: Screenshots using the image upload function.

average score for this statement was 4.33, indicating that the majority consider this application a useful travel planning tool. Figure 5b displays results for "I felt the application has unnecessary functions or that complicate its use." The majority answered 2 (Disagree), three participants answered 1 (Strongly Disagree), two answered 3 (Neither Agree nor Disagree), and one participant each answered 4 (Agree) and 5 (Strongly Agree). The average score was 2.15, suggesting most participants do not find the application unnecessarily complicated. Figure 5c presents results for "Navigating the application was intuitive." Thirteen participants rated 4 (Agree), twelve rated 5 (Strongly Agree), one rated 2 (Disagree), and one rated 3 (Neither Agree nor Disagree). The average score was 4.33, showing the application is considered intuitive to navigate.

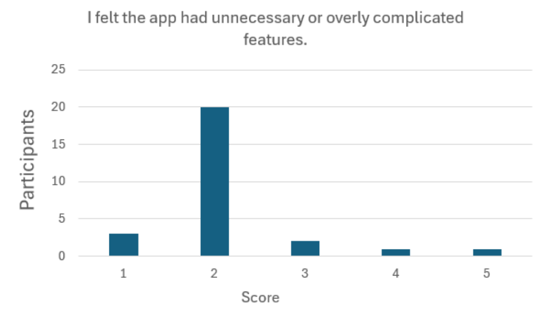
Figure 5d shows the results for "I thought it would be difficult to understand how to use the application without prior assistance." Ten participants responded 2 (Disagree), nine 3 (Neither Agree nor Disagree), four 1 (Strongly Disagree), two 4 (Agree), and two 5 (Strongly Agree). The average score was 2.56, indicating that most participants would understand how to use the application without help. Figure 5e presents results for "The different sections of the application work well together." Fourteen participants answered 4 (Agree), thirteen 5 (Strongly Agree). The average score was 4.48, meaning nearly all participants agreed with this statement. Figure 5f shows results for "During use, I noticed errors or failures that affected my experience." Ten participants answered 1 (Strongly Disagree), nine 2 (Disagree), four 3 (Neither Agree nor Disagree), and one 5 (Strongly Agree). The average score was 2.11, indicating that participants rarely experienced errors affecting their use.

Figure 6a shows results for "I believe anyone could quickly become familiar with the application." Thirteen participants answered 4 (Agree), twelve 5 (Strongly Agree), and two 3 (Neither Agree nor Disagree). The average score was 4.37, suggesting most believe the application is easy to learn. Figure 6b shows results for "The structure or design of the application confused me." Fourteen participants answered 2 (Disagree), seven 1 (Strongly Disagree), five 3 (Neither Agree nor Disagree), and one 4 (Agree). The average score was 2.00, meaning the application structure is generally easy to understand. Figure 6c presents results for "Using this application made me feel confident and comfortable." Fifteen participants answered 4 (Agree), ten 5 (Strongly Agree), one 3 (Neither Agree nor Disagree), and one 1 (Strongly Disagree). The average score was 4.22, showing that most users felt comfortable and confident using the app. Finally, Figure 6d shows results for “I had to spend time understanding the application’s logic before using it.” Thirteen participants answered 2 (Disagree), eight 1 (Strongly Disagree), two 3 (Neither Agree nor Disagree), and four 4 (Agree). The average score was 2.07, indicating that the application is easy to understand and use.

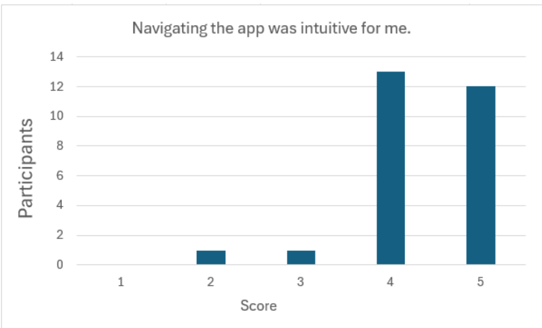
After analyzing the statistical data from the 27 participants’ questionnaires, the overall System Usability Scale (SUS) score was calculated as 77.59, a score considered excellent in terms of usability.



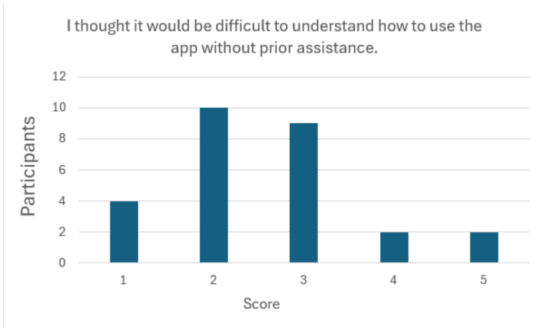
(a) "I would consider this application a useful tool for planning future trips"



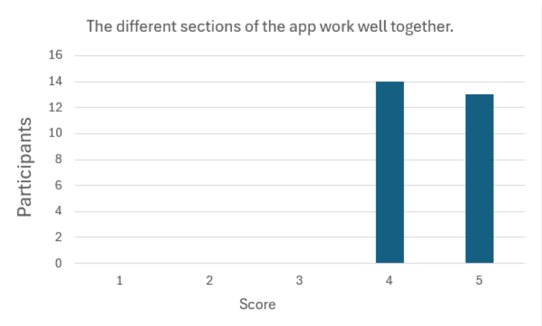
(b) "I felt that the application has unnecessary functions or that complicate its use."



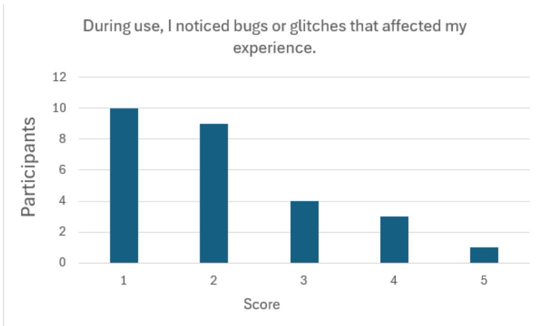
(c) "Navigating the application was intuitive for me."



(d) "I thought it would be difficult to understand how to use the application without prior assistance."

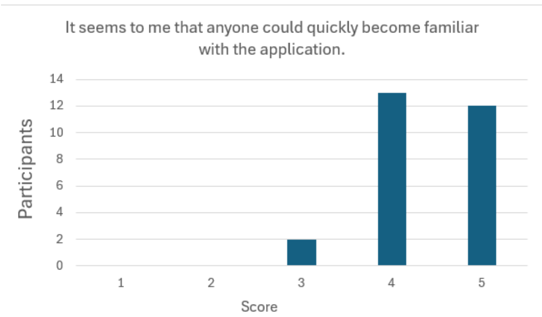


(e) "The different sections of the application work well together."

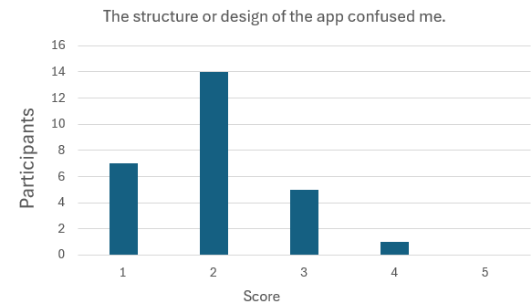


(f) "During use, I noticed errors or glitches that affected my experience."

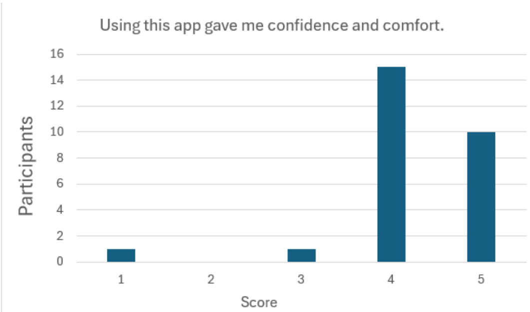
Figure 5: Response distribution for statements 1 to 6 of the usability questionnaire.



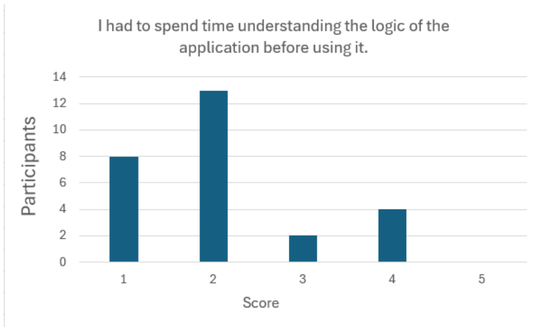
(a) "It seems to me that anyone could quickly become familiar with the application"



(b) "The structure or design of the application confused me."



(c) "Using this application gave me confidence and comfort."



(d) "I had to spend time understanding the logic of the application before using it."

Figure 6: Response distribution for statements 7 to 10 of the usability questionnaire.

6. Discussion

When reviewing the results, the opinions and suggestions that participants wrote in the form were also taken into account. The vast majority of these were positive comments, but some participants offered suggestions, which were:

- *My only suggestion would be to add transportation costs between the different points recommended by the app, or if these are already included, to clarify that. Otherwise, I found it very useful.*
- *Make it more visually appealing.*
- *It was pleasant, it has good recommendations and provides results according to your request. It would be nice if it showed ratings for these recommendations, such as stars or suggestions.*
- *It would be very interesting if it not only located similar cities, but also, with the uploaded image, gave you the itinerary of the city shown in the image. Other than that, everything was excellent.*
- *I would like it to provide more details about the recommended places, like their address or, if it is a restaurant, the menu with prices.*

From these suggestions, several areas for improvement can be identified, such as making the application interface more visually attractive, or generating itineraries with much more detailed information like transportation costs or the address of each suggested place. Incorporating these ideas could considerably improve the user experience. Nevertheless, the application received excellent feedback from participants.

The ChatGPT API effectively fulfilled its role of generating travel itineraries and also suggesting destinations thanks to the image recognition by Google Vision. However, it often suggested the same destinations because many participants uploaded similar images, such as beaches, cities, and monuments. Despite this, no participant reported any errors using this function, which indicates that this component may not require immediate adjustments.

7. Conclusions

The development of this artificial intelligence-based application for planning tourist itineraries represents a significant advancement in how users interact with technology to address real travel needs. Throughout the project, various APIs such as Google Vision, Google Places, and ChatGPT were effectively integrated, enabling the system to interpret images, detect landmark locations, consider constraints like flight time and user preferences, and ultimately generate personalized and visually enriched travel proposals. This approach not only improves the user experience but also reduces the time and effort traditionally required to organize a trip.

From a technical perspective, the combination of image analysis through Google Vision and content generation with language models like ChatGPT allowed the system to provide destination suggestions even when the place could not be directly identified. This aspect is key, as it gives the system flexibility to operate under visual uncertainty, using tags detected in the image to infer possible destinations. Likewise, the use of Google Places to enrich the interface with real images of suggested places not only enhances the presentation but also creates a more immediate emotional connection between the user and the suggested itinerary.

The system also proved to be adaptable, allowing users to input relevant data such as origin city, budget, number of days, and personal preferences. This customization capability, combined with the generation of contextual texts and descriptions, turns the tool into a powerful solution for travelers seeking options tailored to their actual conditions.

During usability testing, the usability score of 77.59/100 verified that the application was well received by participants, who highlighted its ease of use, clear interface, and the added value of having a detailed itinerary in just a few steps. The use of the SUS questionnaire provided quantitative data that supports the effectiveness of the design from the end-user perspective.

Finally, this project lays the foundation for future improvements and integrations. Future versions could add features such as real-time hotel recommendations, integration with airlines to check flight availability according to user constraints, or even a gamification system to encourage frequent use. In summary, the developed application not only fulfills its functional purpose but also demonstrates the transformative potential of artificial intelligence in the field of smart and personalized tourism. Furthermore, the abstract architecture we propose can be instantiated with different AI services, ensuring adaptability to future technological advances and resilience to the evolution of available APIs. In this way, the contribution of this article extends beyond the case study presented, offering a flexible and impactful model that can guide future developments in smart and personalized tourism applications.

Declaration on Generative AI

During the preparation of this work, the authors used GPT-4o to perform grammar and spelling checks. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] P. Fakfare, N. Manosuthi, J.-S. Lee, H. Han, M. Jin, Customer word-of-mouth for generative ai: Innovation and adoption in hospitality and tourism, International Journal of Hospitality Management 126 (2025) 104070.

[2] Y. Chang, X. Wang, J. Wang, Y. Wu, L. Yang, K. Zhu, H. Chen, X. Yi, C. Wang, Y. Wang, et al., A survey on evaluation of large language models, ACM transactions on intelligent systems and technology 15 (2024) 1–45.

[3] J. Maurício, I. Domingues, J. Bernardino, Comparing vision transformers and convolutional neural networks for image classification: A literature review, Applied Sciences 13 (2023) 5521.

[4] D. Alieva, Redes de organizaciones turísticas, itinerarios de viaje y experiencias interculturales de los turistas ruso-hablantes en Andalucía, Ph.D. thesis, Universidad de Sevilla, 2018.
[5] J. C. Niebles, Inteligencia artificial en todo y para todos, Revista Digital Universitaria 21 (2020) 5.
[6] T. Brown, B. Mann, N. Ryder, M. Subbiah, J. D. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, et al., Language models are few-shot learners, Advances in neural information processing systems 33 (2020) 1877–1901.
[7] M. Morales-Chan, Explorando el potencial de Chat GPT: Una clasificación de Prompts efectivos para la enseñanza, Technical Report, GES Department, Galileo University, Guatemala, Guatemala, 2023.
[8] A. Krizhevsky, I. Sutskever, G. E. Hinton, Imagenet classification with deep convolutional neural networks, Advances in neural information processing systems 25 (2012).
[9] A. D. Cantero Alonso, E. A. Martínez Jara, Visión por computadora: identificación, clasificación y seguimiento de objetos, Technical Report, Facultad Politécnica, Universidad Nacional del Este, Ciudad del Este, Paraguay, 2014. `1davidcantero.99@gmail.com`, `2eamartinez@fpune.edu.py`.
[10] B. Li, Everywheregpt: An ai travel planning assistant based on chatgpt, in: Proceedings of the 4th International Conference on Artificial Intelligence and Computer Engineering, 2023, pp. 995–1003.
[11] I. A. Wong, Q. L. Lian, D. Sun, Autonomous travel decision-making: An early glimpse into chatgpt and generative ai, Journal of Hospitality and Tourism Management 56 (2023) 253–263.
[12] K. Volchek, S. Ivanov, Chatgpt as a travel itinerary planner, in: ENTER e-Tourism Conference, Springer, 2024, pp. 365–370.
[13] G. Ilieva, T. Yankova, S. Klisarova-Belcheva, Effects of generative ai in tourism industry, Information 15 (2024) 671.
[14] A. Kanhed, P. Bhagwat, K. Palande, N. Palav, S. Balpande, K. Deshpande, U. Patil, Destinai: Your personalized travel itinerary planner and chat catalyst using generative ai, in: International Conference on Information and Communication Technology for Intelligent Systems, Springer, 2024, pp. 47–55.
[15] R. George, J. Atluri, Revolutionizing tourism: The power of generative ai, in: Generative AI for Transformational Management, IGI Global, 2024, pp. 271–302.
[16] M. Gupta, R. Dias, N. Jain, R. R. Dhamija, R. V. Bidwe, S. Mishra, G. Deshmukh, Travel with generator ai: A novel approach to itinerary creation, in: 2024 8th International Conference on Computing, Communication, Control and Automation (ICCUBEA), IEEE, 2024, pp. 1–6.
[17] K. Londhe, N. Dharmadhikari, P. Zaveri, U. Sakoglu, Enhanced travel experience using artificial intelligence: a data-driven approach, Procedia Computer Science 235 (2024) 1920–1928.
[18] E. Aribas, E. Daglarli, Transforming personalized travel recommendations: Integrating generative ai with personality models, Electronics 13 (2024) 4751.
[19] C. D. Duong, T. H. Nguyen, T. V. N. Ngo, T. T. P. Pham, A. T. Vu, N. S. Dang, Using generative artificial intelligence (chatgpt) for travel purposes: parasocial interaction and tourists' continuance intention, Tourism Review (2024).
[20] J. Zhang, D. J. Mills, H.-W. Huang, Enhancing travel planning and experiences with multimodal chatgpt 4.0, in: Proceedings of the 2024 International Conference on Innovation in Artificial Intelligence, 2024, pp. 12–19.
[21] G. Kalaiarasi, M. Selvi, et al., Personalized travel itinerary generator system, in: 2024 International Conference on Intelligent Systems for Cybersecurity (ISCS), IEEE, 2024, pp. 1–6.
[22] P. Suanpang, P. Pothipassa, Integrating generative ai and iot for sustainable smart tourism destinations, Sustainability 16 (2024) 7435.
[23] A. Singh, R. Madhogaria, A. Misra, E. Elakiya, Automated travel planning via multi-agent systems and real-time intelligence, Available at SSRN 5089025 (2024).
[24] M. Balawardhana, Travel planning with artificial intelligence, The Role of Artificial Intelligence in the Tourism and Hospitality Sector (2025).
[25] A. Garg, D. Verma, L. Pandey, K. S. Kumar, R. Singh, A. Sharma, Enhancing tailored travel by integrating generative ai with insights driven by personality, in: 2025 International Conference on Intelligent Control, Computing and Communications (IC3), IEEE, 2025, pp. 404–409.

[26] H. Kumamoto, H. Joho, Exploring the role of large language model in collaborative travel planning task, Information Research an international electronic journal 30 (2025) 1123–1130.
[27] M. I. Saleh, Generative artificial intelligence in hospitality and tourism: future capabilities, ai prompts and real-world applications, Journal of Hospitality Marketing & Management (2025) 1–32.
[28] A. Štilić, A. Puška, M. Nicić, et al., The role of artificial intelligence in shaping the future of travel industry: An expert analysis of artificial intelligence-generated travel itineraries, DETUROPE-The Central European Journal of Regional Development and Tourism 16 (2025) 57–79.
[29] Y. Li, S. O. Lee, Navigating the generative ai travel landscape: The influence of chatgpt on the evolution from new users to loyal adopters, International Journal of Contemporary Hospitality Management 37 (2025) 1421–1447.
[30] K. Otaki, Y. Baba, Travel itinerary recommendation using interaction-based augmented data, Expert Systems with Applications 269 (2025) 126294. URL: https://www.sciencedirect.com/science/article/pii/S0957417424031610. doi:https://doi.org/10.1016/j.eswa.2024.126294.
[31] V. Udandarao, N. A. Tiju, M. Vairamuthu, H. Mistry, D. Kumar, Roamify: Designing and evaluating an llm based google chrome extension for personalised itinerary planning, arXiv preprint arXiv:2504.10489 (2025).
[32] S. Mohith S, H. H, K. .K, N. N, S. . R, A Review Paper On AI-Driven Travel Planning, Technical Report, Department of Computer Science and Engineering, BMS Institute of Technology and Management, Bengaluru, Karnataka., 2025. doi:10.5281/zenodo.14591573.