Interaction Matters:

A Study of Trust and Sense of Control in AI Travel Planning

Yannick Wild

University of Lausanne

Master in Information Systems and Digital Innovation

February 2026

Supervisor: Prof. Mauro Cherubini

Expert: Prof. Thibault Estier

Abstract: This thesis examines how interaction modalities may influence users' trust and sense of control in AI-assisted travel planning. A conversational Chatbot and a graphical user interface (GUI) were developed following a Design Science Research Methodology (DSRM). These interfaces relied on the same backend and itinerary generation process. A within-subjects experiment was conducted with 36 participants, each using both interfaces to plan a single-day trip. Trust and sense of control were measured before and after each interaction, with baseline levels defined as pre-interaction measures collected before participants interacted with either system. Quantitative results show that both interaction modes increased trust and sense of control compared to baseline levels, with no statistically significant differences between the Chatbot and the GUI ($p = .62$ for trust; $p = .44$ for sense of control). The study was complemented with exploratory qualitative questions. The qualitative findings suggest that trust and control are supported through different interaction mechanisms.

Résumé : Cette thèse examine comment les modalités d'interaction peuvent influencer la confiance et le sentiment de contrôle des utilisateurs dans la planification de voyages assistée par l'IA. Un Chatbot conversationnel et une interface graphique (GUI) ont été développés selon une Méthodologie de Design Science Research (DSRM). Ces interfaces reposaient sur le même backend et le même processus de génération d'itinéraires. Une expérimentation intra-sujets a été menée auprès de 36 participants, chacun utilisant les deux interfaces pour planifier un voyage d'une journée. La confiance et le sentiment de contrôle ont été mesurés avant et après chaque interaction, les niveaux de référence étant définis comme les mesures pré-interaction recueillies avant que les participants n'interagissent avec les interfaces. Les résultats quantitatifs montrent que les deux modalités d'interaction augmentent la confiance et le sentiment de contrôle par rapport aux niveaux de référence, sans différence statistiquement significative entre les deux ($p = .62$ pour la confiance ; $p = .44$ pour le sentiment de contrôle). L'étude fut complétée par des questions qualitatives exploratoires qui suggèrent que la confiance et le sentiment de contrôle sont soutenus par différents mécanismes d'interaction.

JEL Classification: M150 IT Management, **L860** Information and Internet Services; Computer Software, **Z300** Tourism Economics: General

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

ChatGPT reached 100 million monthly users within two months of its release, making it one of the fastest-growing platforms in the digital industry. For context, other viral platforms such as Facebook took 52 months and TikTok took 9 months to reach the same milestone (EY, 2025). Statista estimates that the global AI market will reach USD 347.05 billion by 2026 and USD 1.68 trillion by 2031 (Statista, 2025).

According to the World Travel & Tourism Council (2024), several major travel platforms have introduced AI features for their users. The report highlights the following examples: In 2023, Expedia launched an AI trip planning experience within its smartphone app, offering destination-related recommendations. Tripadvisor upgraded its "Trips" planning product by introducing an AI-powered itinerary generator. Trip.com introduced TripGen, an AI chatbot designed to generate travel itineraries and provide real-time travel tips.

Overall, these developments indicate that AI is becoming a regular part of how users interact with digital platforms, including travel planning tools, which rely on different interaction modalities such as conversational chatbots and graphical user interfaces (GUIs). By structuring user input, feedback, and system responses in distinct ways, these modalities may influence how users understand and engage with these systems.

Despite the widespread use of AI, prior studies on AI-assisted travel planning have primarily examined user outcomes such as trust or intention to use without examining interaction modality. Li et al. (2025) highlight, through a systematic review of the literature, a lack of empirical studies that observe or measure how users experience gen AI in tourism-related tasks, showing that research remains limited in terms of how users are affected during actual interaction. Dwivedi et al. further show that existing studies mainly discuss governance, responsibility, and managerial implications, rather than examining concrete user interactions with gen AI systems.

This thesis aims to reduce that gap by addressing the following research question:

How does interacting with a conversational chatbot versus a graphical user interface (GUI) affect users' trust and sense of control during the early stages of AI-assisted travel planning?

To answer this question, the thesis is structured as follows:

First, a related work section reviews prior research on the use of AI in travel planning, with a focus on user interaction, experience, and trust.

Second, the Design Science Research Methodology that guides the thesis is presented.

Third, the development of the AI travel planning artifact is described.

Fourth, the empirical research method used to evaluate the artifact is explained.

Fifth, the quantitative and qualitative results of the empirical study are presented.

Finally, the results are discussed, including main implications, limitations, and directions for future research.

2 RELATED WORK

2.1 Related Work Scope

This section focuses on research examining the use of artificial intelligence (AI), such as large language models, in the context of travel planning.

The review does not address AI technologies unrelated to planning, such as the use of AI in marketing, pricing, revenue management, or operational optimization in the tourism industry. In addition, studies examining AI adoption at an organizational or managerial level are considered only if they inform user-facing implications.

Finally, the review does not focus on the technical development of AI models, algorithmic performance, benchmarking, or system architectures, except where such work has direct implications for user interaction. By limiting the scope, the review aims to provide a focused synthesis of research that is directly relevant to AI travel planning.

2.2 Related Work Methodology

General search

The search for relevant literature was done in multiple stages. First, an exploratory phase was conducted to map how gen AI has been studied in the context of travel and tourism. During this phase, searches were performed across several academic databases and digital libraries to identify relevant peer-reviewed studies, including the AIS eLibrary, IEEE Xplore, ACM Digital Library, CORE, JSTOR, Springer Nature, and ScienceDirect.

The AI tools ChatGPT, Elicit and Perplexity, were used to complement traditional search methods, refine keyword combinations and search strings, identify potentially relevant peer-reviewed articles based on research questions, and organize preliminary notes.

Scopus was used to verify publications and journal rankings. Priority was given to Q1 journals and to Q2 journals if highly relevant. Litmaps was used to explore citation networks and identify influential or closely related studies.

Preliminary selection and relevance filtering

As the thesis adopted a Design Science Research Methodology, the literature search shifted from an initial broad exploration to a focus on user interaction with AI in the travel industry. In addition, a relevance filtering step defined by the researcher was applied to ensure that retained peer-reviewed journal articles were closely aligned with the research objectives. Each paper was evaluated across four relevance dimensions. Domain focus and unit of analysis ensured alignment with travel planning and end users. Interaction proximity and design relevance captured the extent to which studies examined direct user interaction and provided design-relevant insights.

Table 2.1: Multi-dimensional relevance scoring criteria

Criterion	0 = No alignment	1 = Partial alignment	2 = High alignment
Domain focus	Not related to travel planning or tourism	Tourism context without focus on planning	Travel planning or decision making in tourism
Unit of analysis	Organizational, technical, or managerial focus	Mixed or indirect user focus	Individual traveller or end user
Interaction proximity	No discussion of user interaction	Interaction discussed implicitly or at a high level	Direct examination of user interaction during planning or decision tasks
Design relevance	No implications for interaction or design	Conceptual or agenda level insights	Design-relevant gaps or implications related to trust, control, or interaction structure

For each criterion, papers were assigned a score of zero, one, or two. The scoring scheme and thresholds were defined by the researcher. Scores were aggregated across the four dimensions to produce an overall relevance score for each paper, with a minimum possible score of zero and a maximum score of eight. Papers classified as highly relevant formed the analytical core of the thematic literature review. Medium-relevance papers were used to support thematic framing and contextualization, while low-relevance papers were excluded from the final synthesis.

Table 2.2: Relevance classification based on aggregate score

Total score	Relevance category
6–8	High relevance
3–5	Medium relevance
0–2	Low relevance

The final list of peer-reviewed journal articles was drawn from the following publishers and digital libraries: Elsevier, Emerald, SAGE Journals, Taylor and Francis Online, and Wiley Online Library.

Organization of the literature

Following the relevance filtering stage, the final body of literature was analysed to identify recurring research topics, key constructs, and shared perspectives. This analysis resulted in the identification of three themes developed in the following section.

2.3 Literature Review

The first theme, **AI adoption and user behaviour**, examines factors that influence users' adoption, resistance, and continued use of AI.

Users are more likely to adopt AI when the system exhibits human-like and competence-related characteristics. Pillai and Sivathanu (2020) show that perceived intelligence, anthropomorphism, trust, and perceived usefulness significantly influence adoption of AI-based chatbots in tourism. Similarly, Melián-González et al. (2021) find that performance expectancy, habit, hedonic motivation, anthropomorphism, and social influence predict intention to use chatbots for travel planning.

Furthermore, adoption of generative AI (gen AI) also depends on novelty and cognitive appraisal. Ma and Huo (2023) demonstrate that users' willingness to adopt ChatGPT is shaped by novelty value, perceived humanness, and appraisal processes. Their findings suggest that early adoption of gen AI is influenced by users' interpretation of the system's role and character. Users' adoption behaviour consists of trade-offs between perceived benefits and perceived risks, rather than simple linear predictors. Han et al. (2025) show that travellers' use of ChatGPT depends on how perceived benefits, risks, and innovation characteristics combine. Their analysis indicates that perceived risk does not automatically reduce adoption, but instead interacts with perceived value and novelty in shaping user behaviour.

Resistance to AI is not only caused by performance concerns but also by psychological factors related to autonomy and control. Seyfi et al. (2025) show that resistance to gen AI is influenced by functional barriers, technology anxiety, and personality traits associated with autonomy and control. These findings show that adoption varies across users and depends on how much autonomy they experience when using AI systems.

Continued use of AI-based travel planning systems depends on the trust, affective evaluation, and satisfaction users develop during interaction. Pham et al. (2024) show that trust and satisfaction influence users' intention to continue using ChatGPT for travel services, while technology anxiety affects the strength of these relationships. Fakfare et al. (2025) further show that adoption develops over time and that social communication plays a role in later stages of acceptance.

The second theme, **the AI-enhanced travel experience**, examines how AI affects users' experiences and decision-making in the context of travel planning.
Generative AI systems change how users plan trips by reshaping the steps involved in decision-making. Instead of manually searching, comparing, and organizing information, users increasingly rely on AI systems to carry out these tasks. As a result, AI systems take an active role in travel planning decisions. Wong et al. (2023) show that users delegate information search, synthesis, and option evaluation to ChatGPT while retaining final decision authority. This delegation reduces cognitive effort and speeds up planning, but also changes the planning process itself. Users shift from constructing itineraries step-by-step to reviewing, refining, and validating AI-generated plans. This finding shows that the travel planning experience is shaped by how decision responsibility is shared between the user and the AI system.

AI-assisted planning can also enhance convenience and perceived intelligence, affecting how users experience planning tasks. Gursoy et al. (2023) show that generative AI enables real-time personalization, conversational access to complex information, and adaptive support during travel planning. These features make planning faster and easier from the user's perspective. At the same time, their review indicates that user experience depends on how AI is integrated into user-facing touchpoints.

AI systems can also negatively affect user experience by reducing agency and increasing dependence. Grundner and Neuhofer (2021) show that AI-enhanced tourism experiences can lead users to over-rely on algorithmic suggestions, reduce exploration, and accept homogenized recommendations. Their analysis indicates that positive experiential effects such as efficiency and convenience may come at the cost of reduced autonomy and interpretive freedom during planning.

The third theme, **Trust, Quality, and Ethics**, examines trust formation, information quality, transparency, and ethical concerns related to AI and travel.
Trust plays a central role in AI-assisted travel planning because users must decide whether to rely on AI-generated recommendations under conditions of uncertainty. Ali et al. (2023) show that travellers' trust in ChatGPT-based travel recommendations depends on perceived relevance, credibility, usefulness, and system intelligence. Higher trust leads to greater acceptance of recommendations. This work shows that trust is formed through users' evaluation of AI outputs during planning rather than as a general attitude toward AI technology.

Trust judgments are also shaped by how information is processed during decision-making. Stergiou and Nella (2024) show that acceptance of AI-generated travel recommendations depends on information accessibility and diagnosticity. Users are more likely to trust recommendations that are easy to process and clearly useful for comparing options. Their findings show that trust is influenced by how information is structured and interpreted, not only by its factual accuracy.

Ethical and quality concerns can directly undermine trust, even when AI outputs appear useful. Kim et al. (2023) show that perceived bias, ethical concerns, and doubts about information quality significantly reduce trust in generative AI and negatively affect travel decisions. Similarly, Huang et al. (2025) show that anxiety caused by a lack of transparency reduces acceptance of AI-generated recommendations. These studies show that trust depends on users' perceived understanding of the system and their control over the recommendation process.

Trust is also highly sensitive to errors. Kim et al. (2025) show that inaccurate AI-generated travel information leads to strong reductions in perceived trustworthiness and recommendation acceptance, with effects that persist beyond the immediate interaction. At the same time, Loureiro et al. (2025) show that users may forgive AI errors when relational or symbolic benefits are perceived.

2.4 Research Gap and Research Questions

Despite research on adoption, experience, and trust in AI-assisted travel planning, a common limitation appears across themes. Within the studies included in the present literature review, no work was found that examines whether different interaction modes influence adoption, reduce resistance, or shape users' trust and sense of control during travel planning. Interaction is usually measured through users' perceptions or outcomes after use, rather than treated as a design choice.

To address the gap identified in the literature review, this study is guided by the following main research question and sub-questions.

- **How does interacting with a conversational chatbot versus a graphical user interface (GUI) affect users' trust and sense of control during the early stages of AI-assisted travel planning?**
    - How does user trust differ between the Chatbot and GUI?
    - How does user sense of control differ between the Chatbot and GUI?
    - How do users explain their experience of trust and sense of control when interacting with each mode?

3 DESIGN SCIENCE RESEARCH METHODOLOGY (DSRM)

To answer the research question, this thesis followed the DSRM proposed by Peffers et al. (2007) and combined artifact development with empirical evaluation. DSRM is a six-step research approach in Information Systems that focuses on designing, building, and evaluating artifacts to address identified problems in a systematic way (Peffers et al., 2007). Within this framework, the study adopted a problem-centered initiation, with the research problem identified through prior literature before objectives were defined.

The **first** activity focuses on identifying a relevant research problem and motivating the need for a solution. In this study, the problem was identified through a review of the academic literature on AI-assisted travel planning.

The **second** activity focuses on defining design objectives in response to the identified problem. Accordingly, the objective of this study was to design an artifact that enables the comparison of two interaction modalities in AI-assisted travel planning.

The **third** activity focuses on the design and development of an artifact that meets the defined objectives. To this end, an AI-powered travel planning system was designed and developed with two interaction modalities: a conversational chatbot and a graphical user interface (GUI). Both interfaces collected identical user preferences and relied on the same backend, model configuration, and itinerary generation pipeline. This design allowed interaction modality to be isolated as the primary experimental variable.

The **fourth** activity focuses on demonstrating the use of the artifact in a relevant context. In this study, the artifact was demonstrated in a controlled experimental setting in which participants interacted with both modalities to plan a single-day leisure trip under a standardized scenario.

The **fifth** activity focuses on evaluating how well the artifact supports the objectives defined earlier. The artifact was evaluated through a within-subjects experiment with 36 participants. Short qualitative questions were also included to explore which interface participants trusted more and which one made them feel more in control.

The **sixth** activity focuses on communicating the research problem, the artifact, and the evaluation results to relevant audiences, which is done throughout this thesis.

4 DEVELOPMENT OF THE AI TRAVEL ASSISTANT

4.1 System Overview

This section introduces the AI travel planning system used in this study and explains how it was designed to support a controlled comparison between two interaction modalities: a graphical user interface (GUI) and a Chatbot. The system was developed as the main research artifact of this thesis and was used to study user trust and sense of control during AI-assisted travel planning.

In the **GUI** modality, users specify fifteen travel preferences, including destination, budget, timing, trip pace, walking tolerance, interests, dining preferences, and dietary restrictions. These preferences are submitted to the system and used to generate a single-day travel itinerary through a shared itinerary prompt.

In the **Chatbot** modality, the same fifteen preferences are collected through a fixed, step-by-step question sequence that mirrors the six-page flow of the GUI. Although presented conversationally, this sequence is enforced through a dedicated system prompt executed with OpenAI's GPT-5.1 language model, ensuring that preferences are collected in the same order and structure as in the GUI. This chatbot prompt operates independently from the itinerary generation prompt.
The full chatbot prompt is provided in Appendix III.

Once preference collection is completed in either interface, inputs are sent to a shared backend and processed using an identical itinerary generation pipeline.
An itinerary prompt then guides OpenAI's GPT-5.1 language model to generate a single-day travel itinerary for both interfaces.
The prompt instructs the model to:

- Recommend activities and venues based on user preferences.
- Respect all provided constraints, such as daily budget and selected time frames.
- Treat preferences as mandatory rather than suggestions.
- Provide activities and meals with time ranges, short descriptions, and costs.

The prompt also defines the structure and formatting of the itinerary, with sections for different parts of the day, meals, budget, walking distance, and tips. Additionally, to ensure consistent structure and presentation across participants, the prompt forbids the model from adding, removing, or renaming sections. Web search, browsing, and external data sources are not available during itinerary generation. Therefore, all outputs are produced based on the prompt, user preferences, and the model's internal knowledge.
The full itinerary prompt is provided in Appendix V.

This system design ensures that any differences observed in trust or sense of control can be attributed to interaction modality rather than to differences in backend logic, AI configuration, real-time data availability, or output generation.

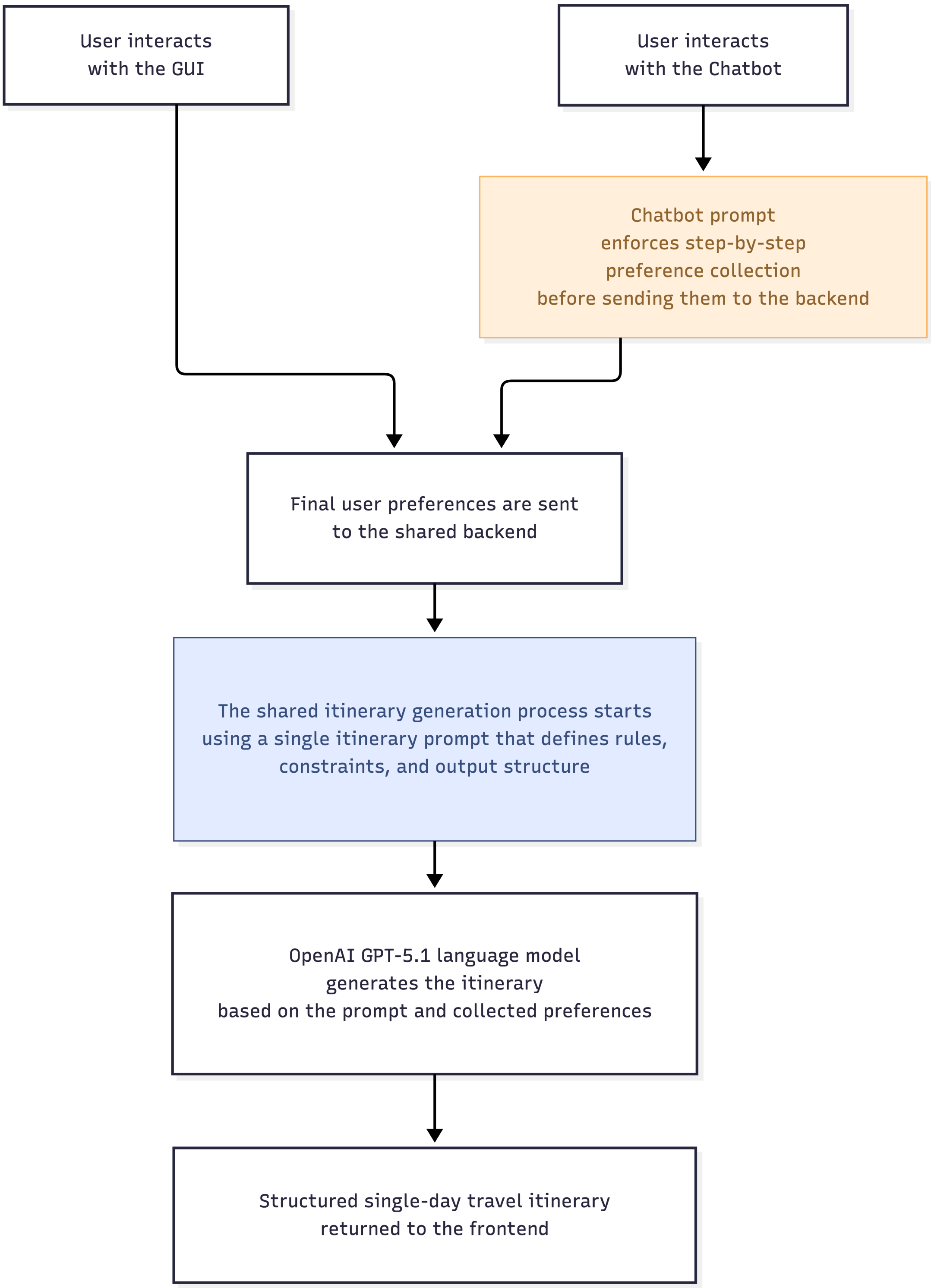


Figure 4.1: System overview

4.2 System Development Process

During the ideation phase, the GUI and Chatbot frontend were prototyped using Claude Opus 4.1. This phase focused exclusively on interface layout, interaction flow, and user experience, rather than on backend logic or AI behaviour.

After a stable frontend design for both interfaces was reached, the complete prototype files were exported to the development environment. At this stage, IntelliJ IDEA was used as the integrated development environment, with the Junie coding agent used to support implementation tasks. The frontend code was automatically refactored and separated into HTML, CSS, and JavaScript files to align with standard web development practices and to improve maintainability.

All subsequent development, debugging, and integration work was performed within IntelliJ IDEA. This included refinement of frontend behaviour, implementation of backend services using Spring Boot 4.0 and Java 17, API integration, and prompt configuration.

During this phase, the Junie coding agent was used as an implementation assistant.

To ensure consistent implementation aligned with the experimental design, internal README files were maintained within the development environment and referenced by the coding agent. These files provided architectural constraints, contextual guidance, model capabilities, prompting practices, and interface flow. They also served as guardrails to prevent the introduction of unnecessary features that could compromise experimental control or methodological clarity.

The README files included:

File 1: Summary of the GPT-5.1 model based on official OpenAI documentation.

File 2: Best practices for prompting GPT-5.1 based on official OpenAI guidelines.

File 3: Summary of the API parameters used in the study and the rationale behind them.

File 4: Description of the six-page interaction flow of the GUI, used to align the chatbot prompt with the same logical structure

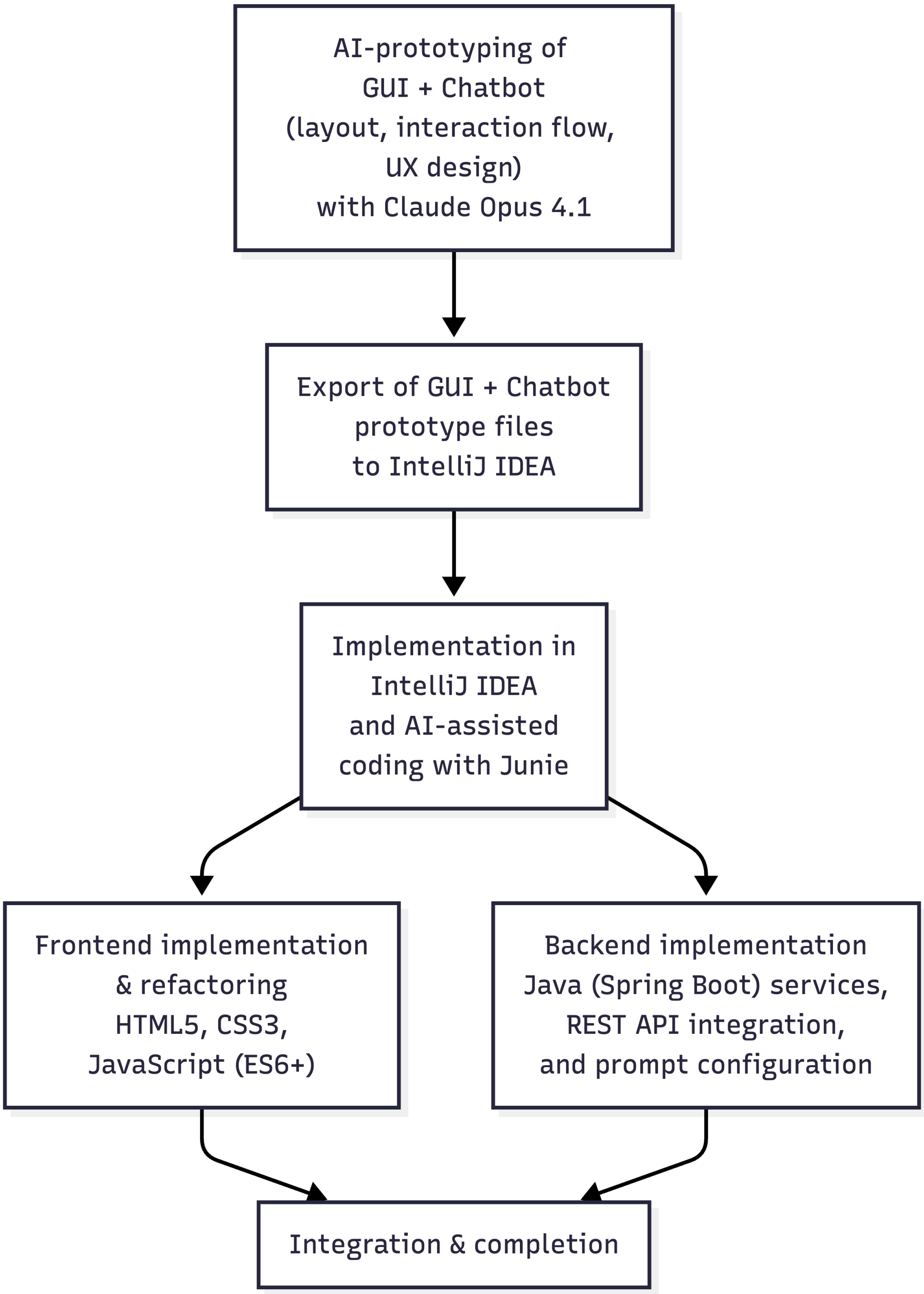


Figure 4.2: System development process

4.3 Frontend Design

4.3.1 GUI Frontend

The graphical user interface (GUI) is a six-page, step-by-step process that guides users through the specification of their travel preferences. It provides predefined input options through visual elements, and each page focuses on a specific set of preferences.
A progress indicator shows users how far they are in the process. The six-page frontend design is shown in Appendix I, which illustrates an example of the step-by-step process completed by a participant.

The **first page** of the GUI captures the *Trip Basics & Budget*.
Users specify the destination as free-form text, select the trip type (Solo, Couple, Family, Friends) using pill selectors, and enter the daily budget for activities and food in Swiss francs (CHF) through a numerical free-form input.

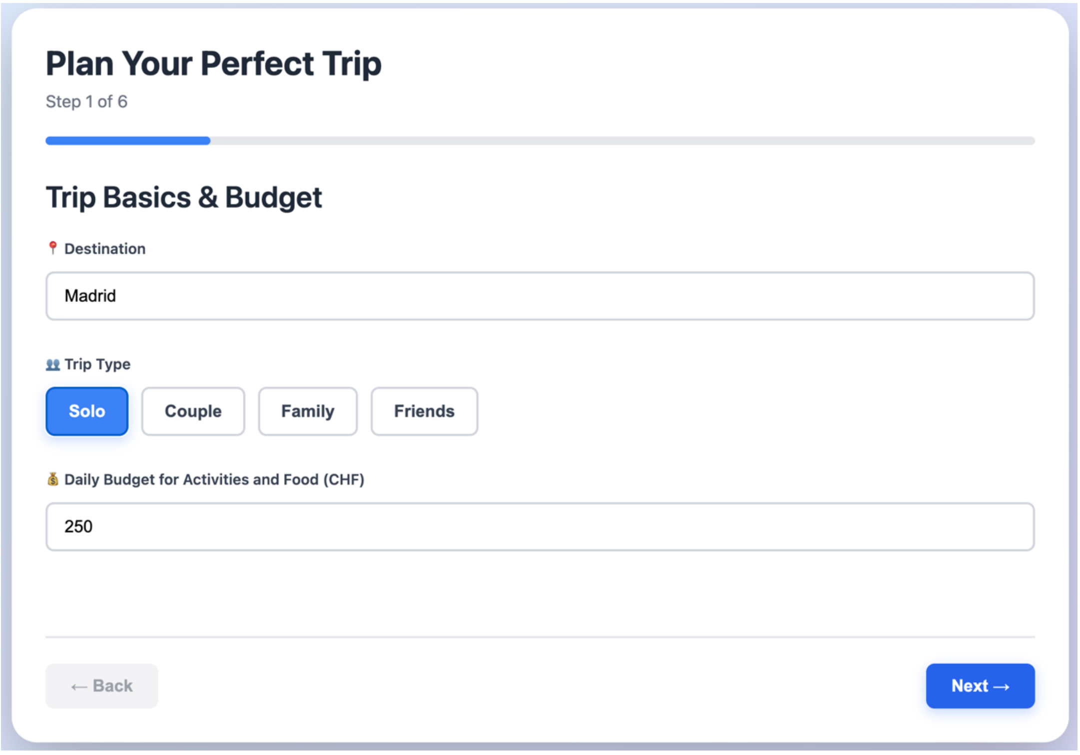


Figure 4.3: GUI page 1

The **second page** focuses on Dates & Planning Scope. Users select a date range using a calendar-based interface and specify, via checkboxes, which days within that range should be planned. For each selected day, users indicate which parts of the day should be included (Morning, Afternoon, Evening, or Whole day) and which meals should be planned (Breakfast, Lunch, Dinner, or All Meals), both using checkboxes. Additionally, users specify their preferred start time for the selected day(s) using a time slider. This level of planning granularity was introduced following pilot testing, which indicated that users preferred the ability to plan specific portions of a day rather than being constrained to full-day itineraries.

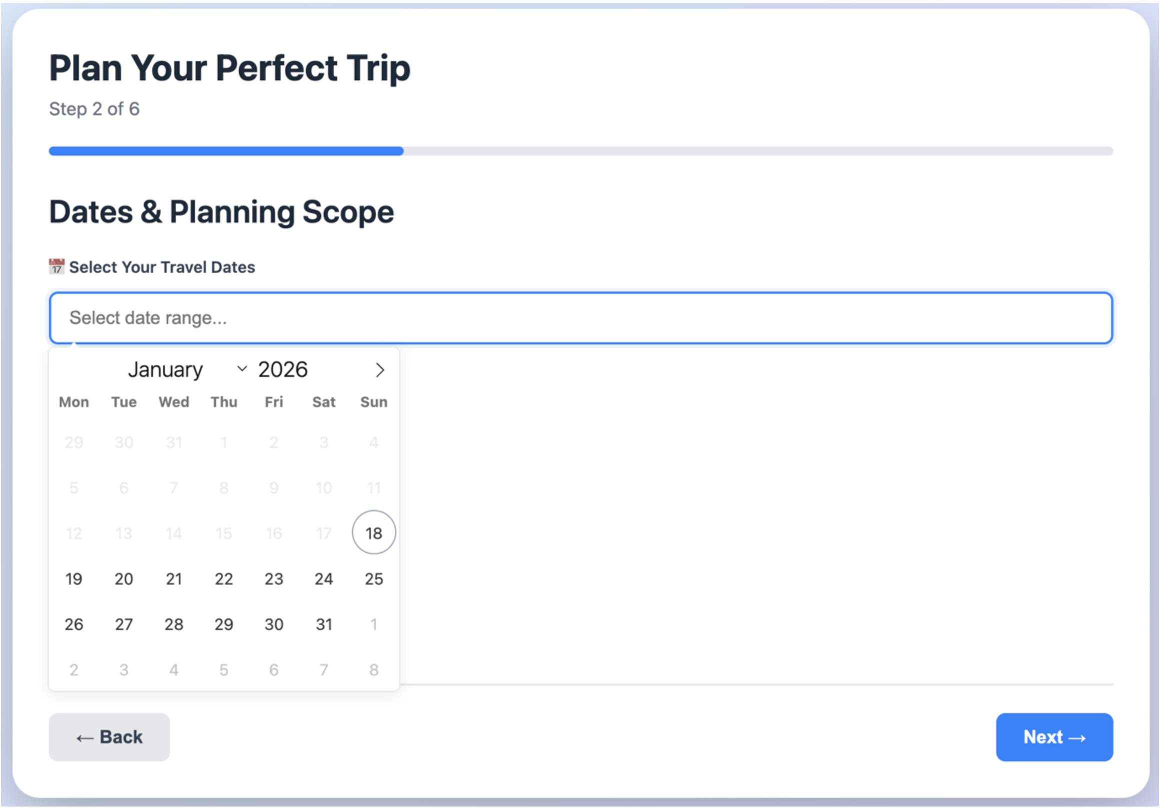


Figure 4.4: GUI page 2

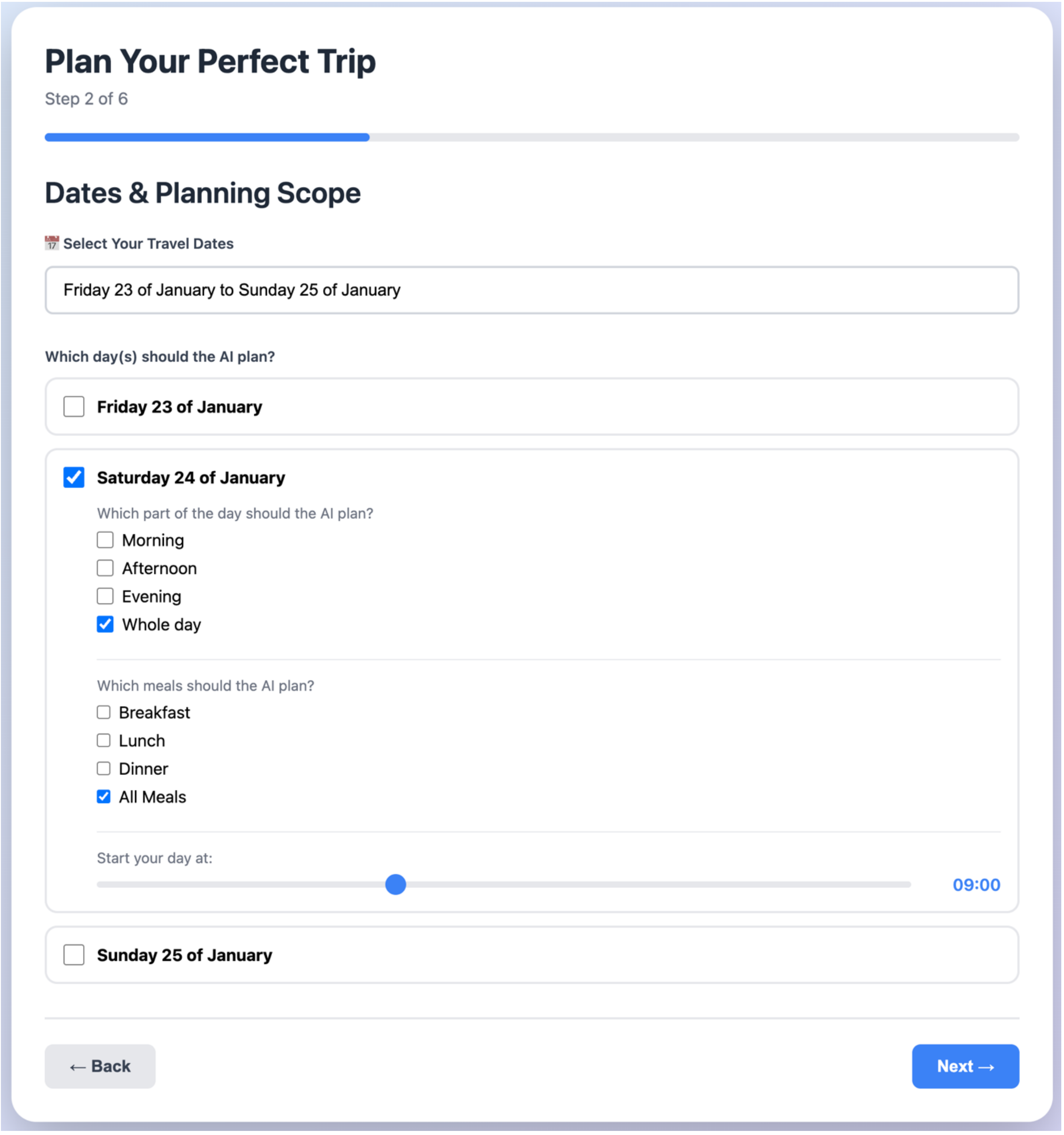


Figure 4.5: GUI page 2 continued

The **third page** addresses Trip Pace & Walking Tolerance. Users specify their preferred trip pace using a three-option pill selector (Slow: relaxed day with plenty of breaks, Moderate: a balance between activities and downtime, or Fast: doing as much as possible in a day). Users also indicate their walking tolerance using predefined distance ranges presented as a pill selector (Light: less than 5 km, Moderate: 5 to 10 km, Active: more than 10 km).

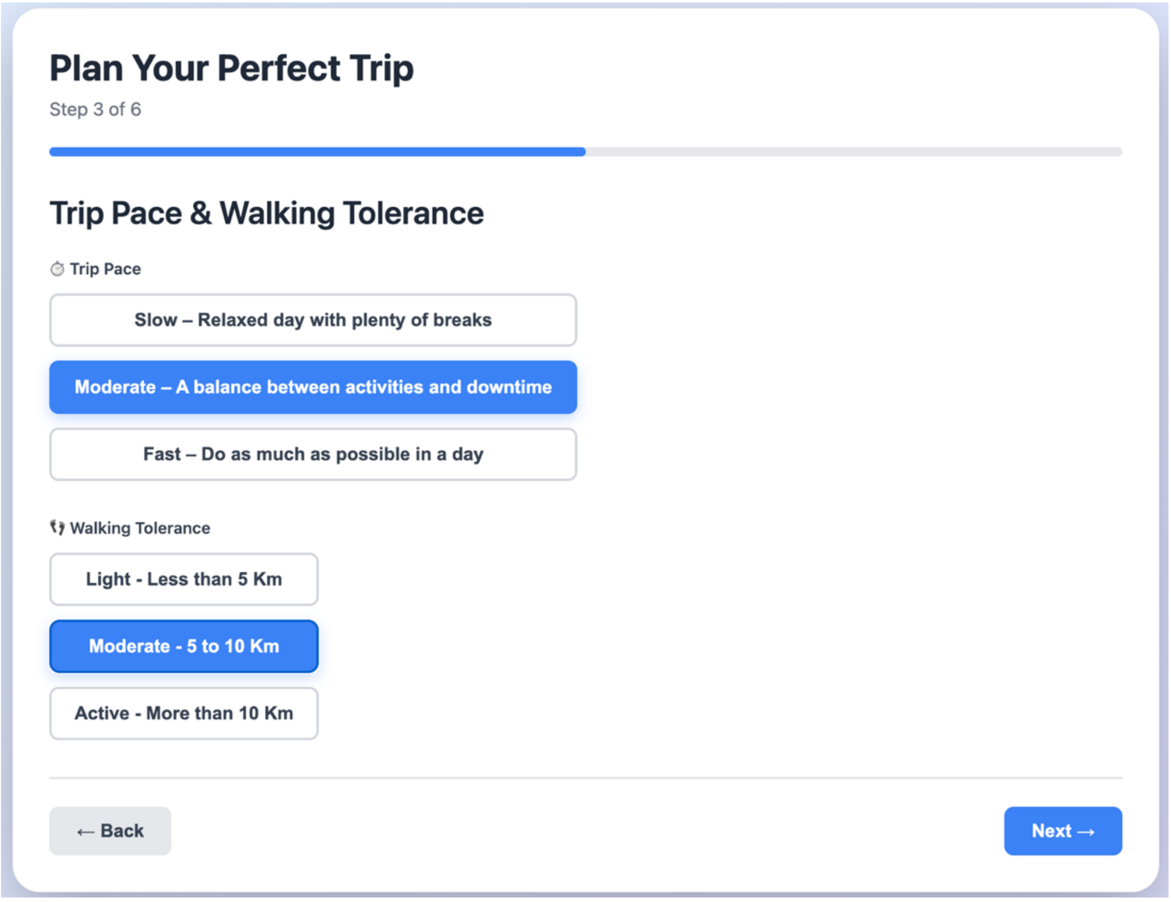


Figure 4.6: GUI page 3

On the **fourth page**, users define their Interests and Travel Style. Interests are selected using pill selectors with multiple selection enabled. The available interests are: Architecture, Local food & markets, Museums, Music & concerts, Nightlife, Parks & nature, Scenic viewpoints, Shopping, Sports & activities, and Other. When Other is selected, a free-form text input field appears to allow custom specification.

Users then select their travel style, also using pill selectors, and may select multiple options. The available Travel Style options are: Adventure, Architecture, Culture, Food, and Relaxation.

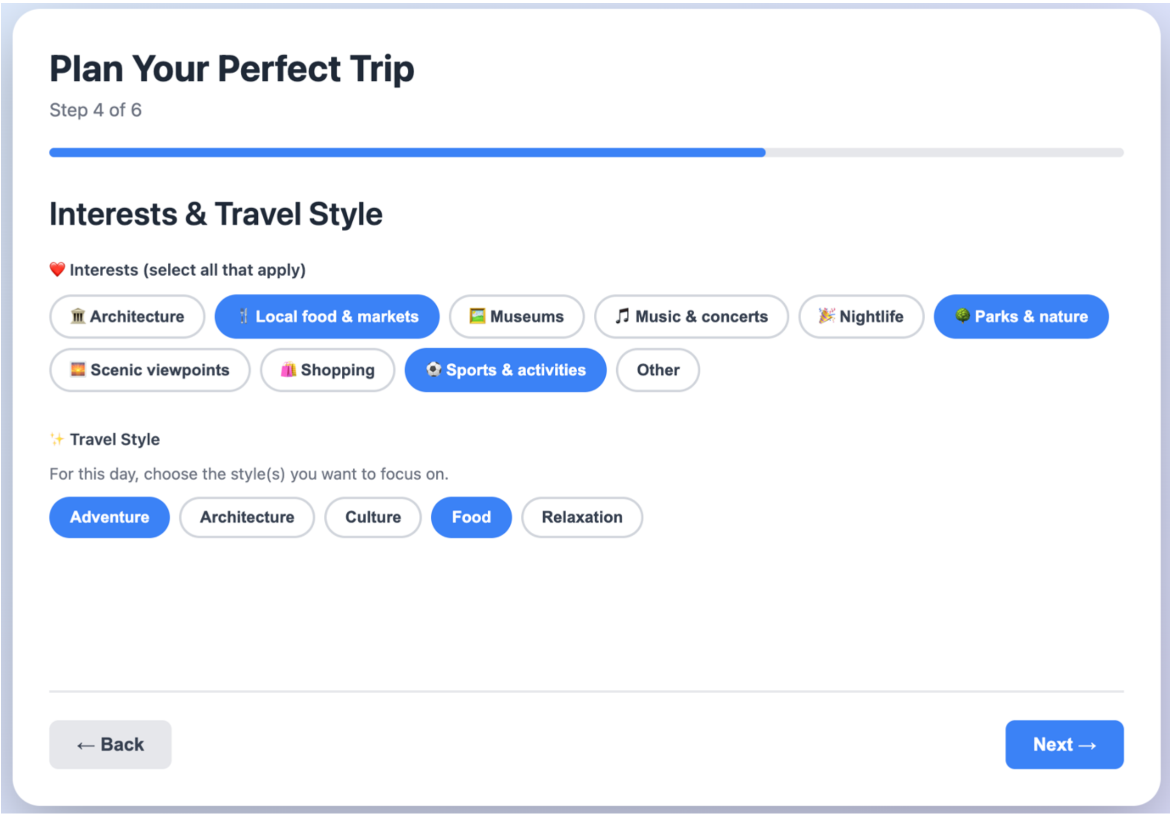


Figure 4.7: GUI page 4

The **fifth page** captures Dining Preferences & Dietary Restrictions. Users select their dining preferences using pill selectors and may choose multiple options. The available dining preferences are: Coffee shops, Local street food, Tapas / casual dining, Market halls, Fine dining, Vegetarian-friendly, Vegan-friendly, Healthy / light meals, Gluten-free-friendly, and Other. When Other is selected, a free-form text input field appears.

Users then select their dietary restrictions, also using pill selectors with multiple selection enabled. The available options are: No shellfish, No seafood, Vegetarian, Vegan, No dairy, No nuts, No gluten, Other, and None. When Other is selected, a free-form text input field appears.

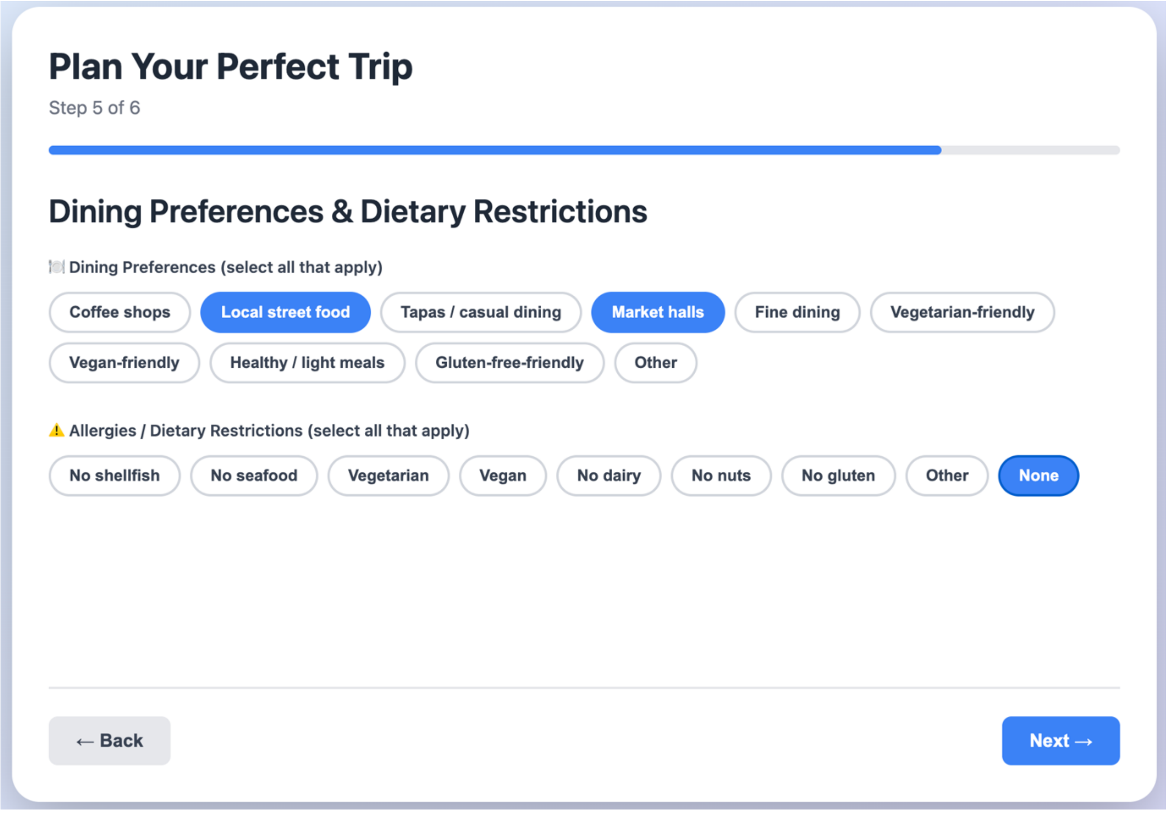


Figure 4.8: GUI page 5

The **sixth page** of the GUI serves as a review and confirmation step. All fifteen previously specified preferences (Destination, Trip Type, Daily Budget, Arrival Date, Departure Date, Selected Day(s) to Plan, Parts of the Day to Plan, Meals to Plan, Preferred Start Time, Trip Pace, Walking Tolerance, Interests, Travel Styles, Dining Preferences, and Dietary Restrictions) are summarized in a structured, card-based layout, allowing users to review and verify their inputs before proceeding. Users must explicitly confirm their preferences by clicking the Generate my plan button. Once clicked, a loading screen is displayed with the message “Creating Your Perfect Itinerary. Our AI is crafting a personalized travel plan just for you...” with a loading indicator to signal that itinerary generation is in progress.

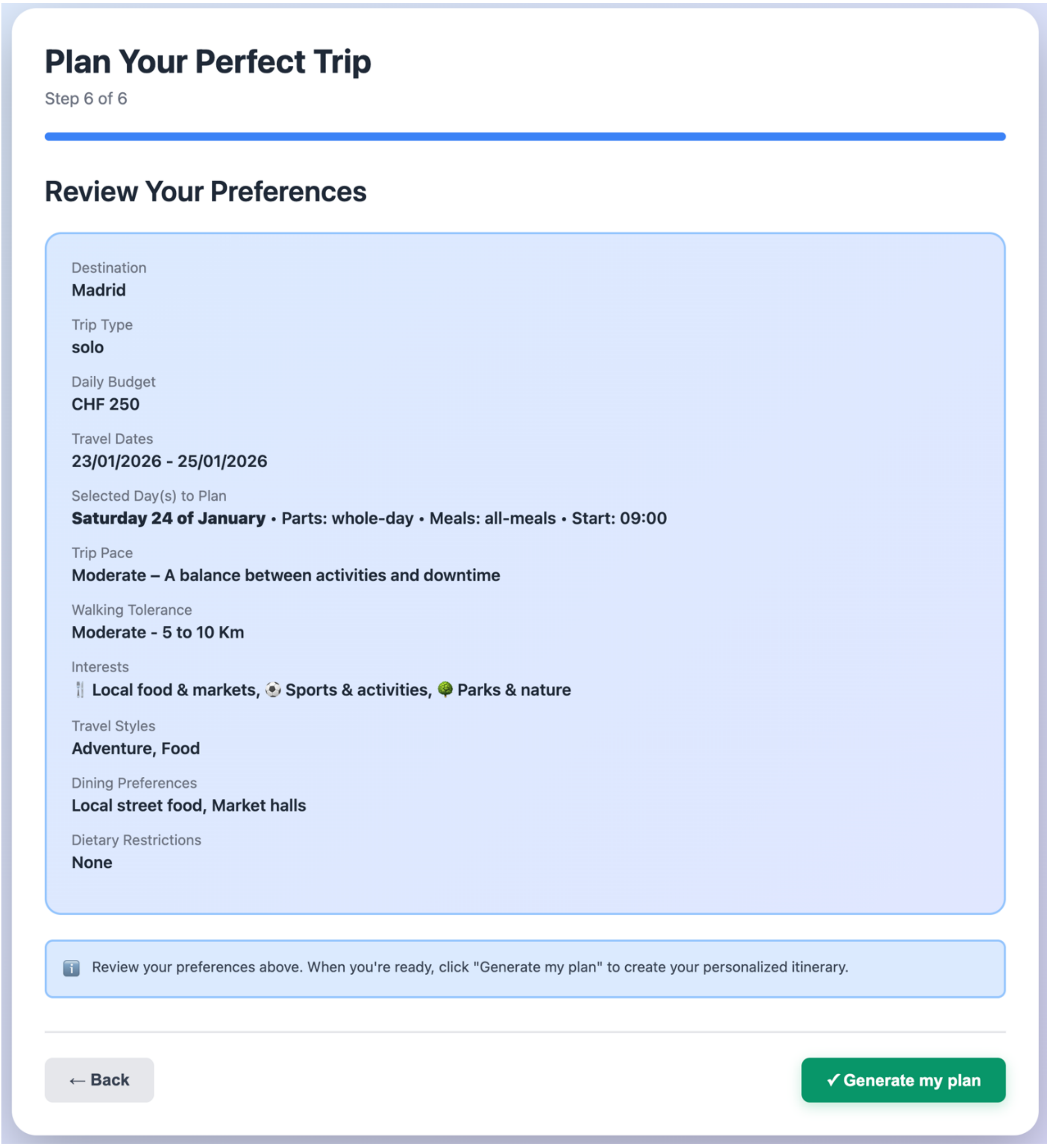


Figure 4.9: GUI page 6

4.3.2 Chatbot Frontend

The Chatbot interface is presented as a single-page conversational layout resembling a standard messaging application. The entire interaction takes place within one continuous view, without page transitions. The frontend design can be seen in Appendix II and shows an example of a conversation completed by a participant.

At the top of the interface, a fixed header displays the assistant's identity as "AI Travel Assistant," along with a short status indicator indicating that the system is online and powered by AI. This header remains visible throughout the interaction and provides contextual reassurance about the system's availability.

The central area of the interface displays the conversation history in a vertically stacked format. Messages generated by the Chatbot appear aligned to the left and are visually distinguished through a consistent background style and avatar icon. User messages appear aligned to the right, clearly separating user input from system responses. Each message is timestamped to support conversational continuity.

The interaction begins with an introductory message from the Chatbot, which greets the user and explains its role as a travel planning assistant. Within the same message, the Chatbot asks the first question related to the trip destination. Subsequent questions are presented sequentially within the same conversational space.

User input is provided through a text input field located at the bottom of the interface. Users enter responses using free-form text and submit them either by clicking the send button or by pressing the Enter key. All previously asked questions and user responses remain visible throughout the session, allowing users to reference earlier inputs and maintain awareness of the interaction.

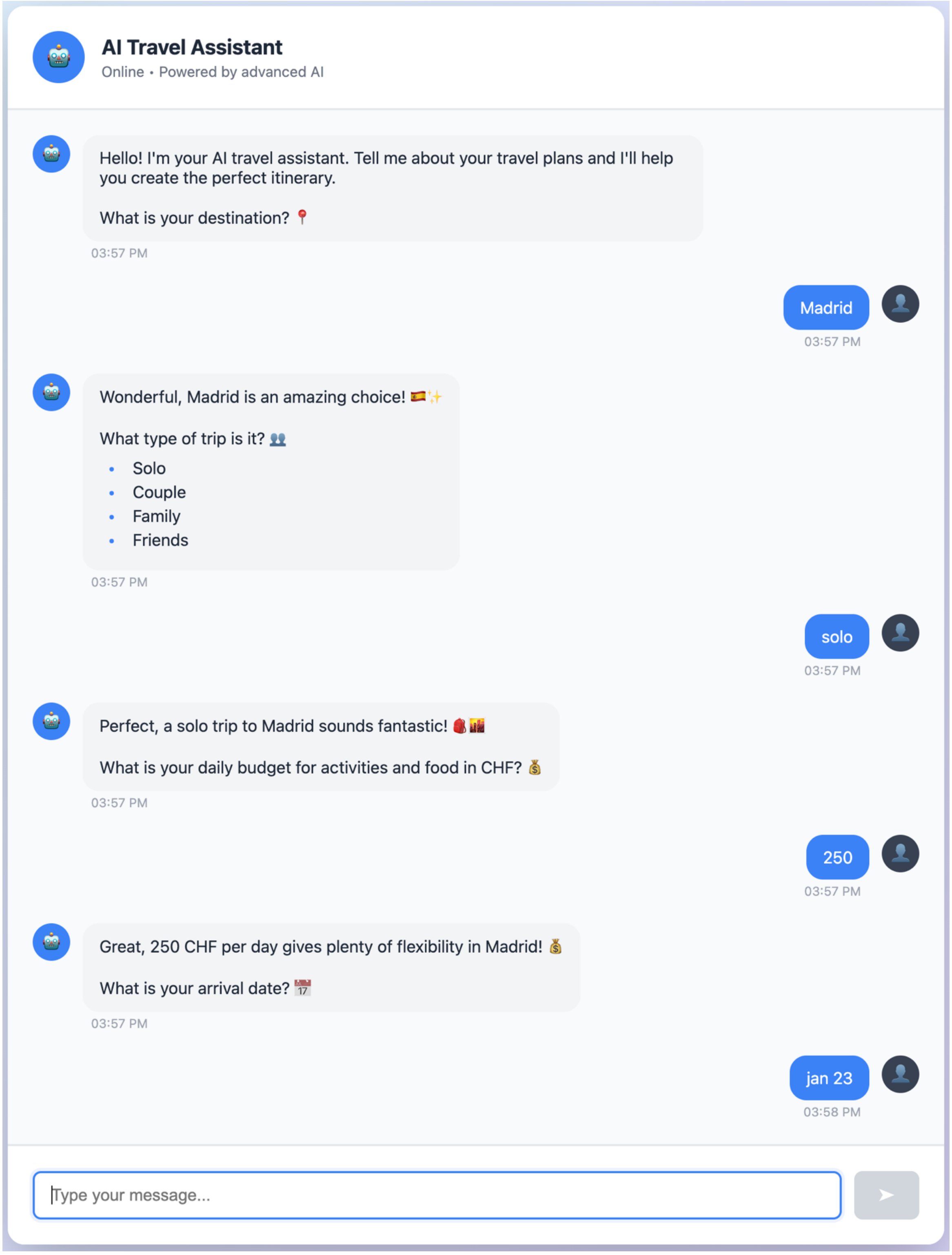


Figure 4.10: Chatbot frontend

4.3.3 Itinerary Frontend

Before triggering the itinerary generation, users in both interfaces are presented with a structured summary of their preferences and must explicitly confirm them. Only after they confirm the system calls the itinerary generation engine and returns a personalized itinerary in real-time. The itinerary frontend can be seen in Appendix IV and shows an example of an itinerary generated for a participant.

The itinerary output is presented as a single-page, scrollable view without tabs or pagination. At the top of the page, a header displays the itinerary title and destination, followed by the selected travel dates and a short descriptive subtitle indicating that the plan is personalized.

The main body of the itinerary is organized chronologically and divided into labeled sections corresponding to parts of the day and meals. Sections include Morning, Breakfast, Late Morning, Lunch, Afternoon, Dinner, Evening, Budget summary, Walking summary, and Tips for Your Day. Each section is visually separated using headings and horizontal dividers, creating a clear and linear progression through the day.
The Budget Summary presents estimated costs for activities and food along with a total daily cost. The Walking Summary provides a breakdown of approximate total walking distance. The Tips for Your Day section offers practical advice related to timing, transport, and user-specific constraints.
Within each section, activities or meals are presented with time ranges, short descriptive text, and an estimated cost. This consistent structure allows users to quickly scan the itinerary and understand timing, content, and budget implications.

Below the itinerary content, a small set of action options is provided. Users can download the itinerary as a PDF file, download a plain-text version, or restart the planning process. These options are grouped together and placed at the bottom of the interface to avoid interrupting the review of the generated itinerary.

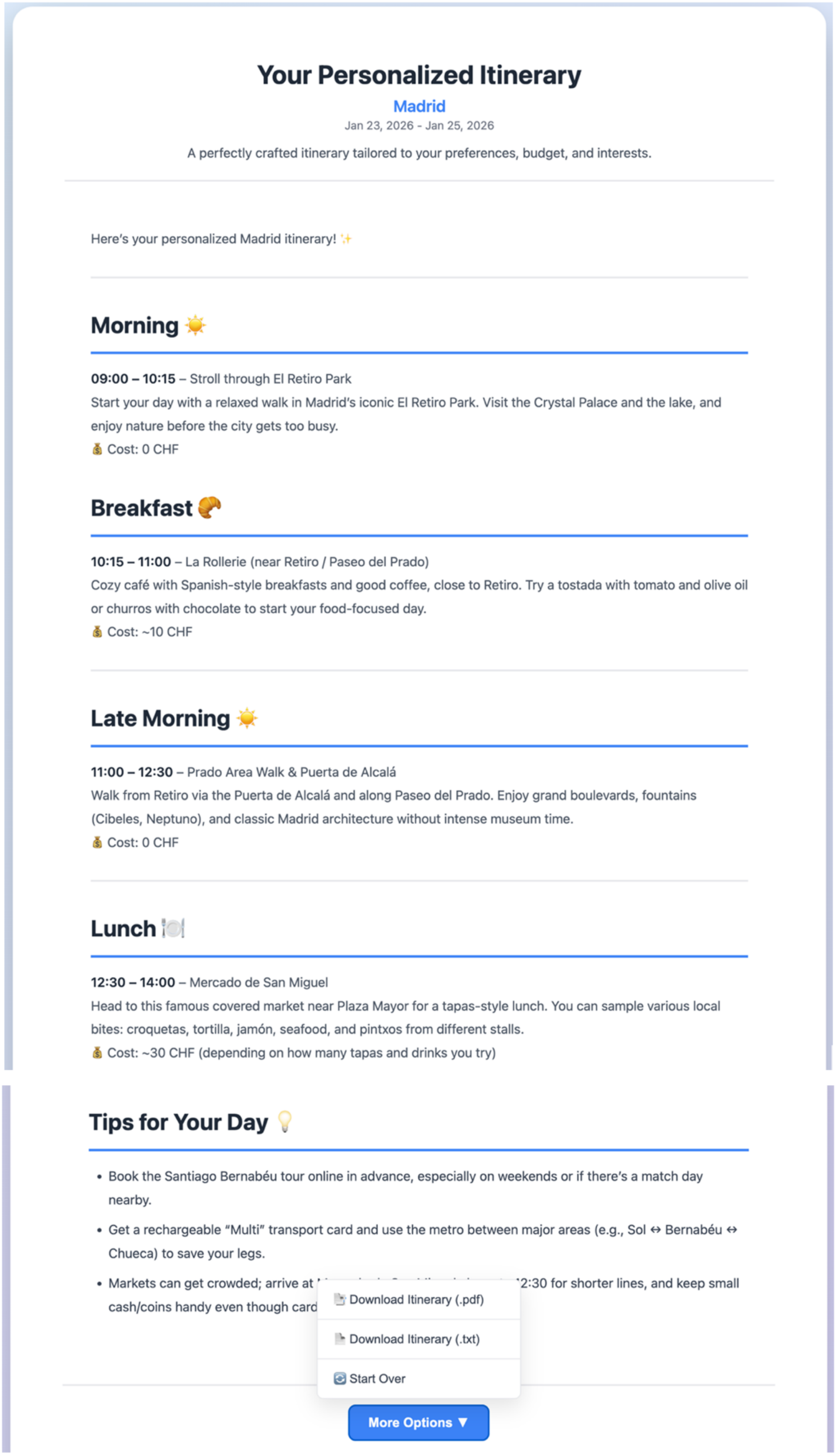


Figure 4.11: Itinerary frontend

4.4 System Technical Specifications

The system architecture consists of three layers: a presentation layer (frontend), an application layer (backend), and an AI integration layer. For both interfaces, the frontend is a client-side web application built with HTML5, CSS3, and JavaScript (ES6+). The frontend interfaces communicate with the shared application layer (backend) built using Spring Boot 4.0 and Java 17. The backend exposes REST endpoints and handles all application logic.

The itinerary generation engine is implemented as a backend component within the application layer and is used by both the GUI and the Chatbot. It is responsible for receiving the user preferences, communicating with the AI integration layer, applying the prompt, and returning a complete single-day itinerary in a standardized markdown format to the frontend. User preferences collected through either interface are standardized before being passed to the itinerary generation engine. This serialization process ensures that all required preference fields are present and consistently formatted, with null-safe defaults applied where necessary to prevent bad requests.

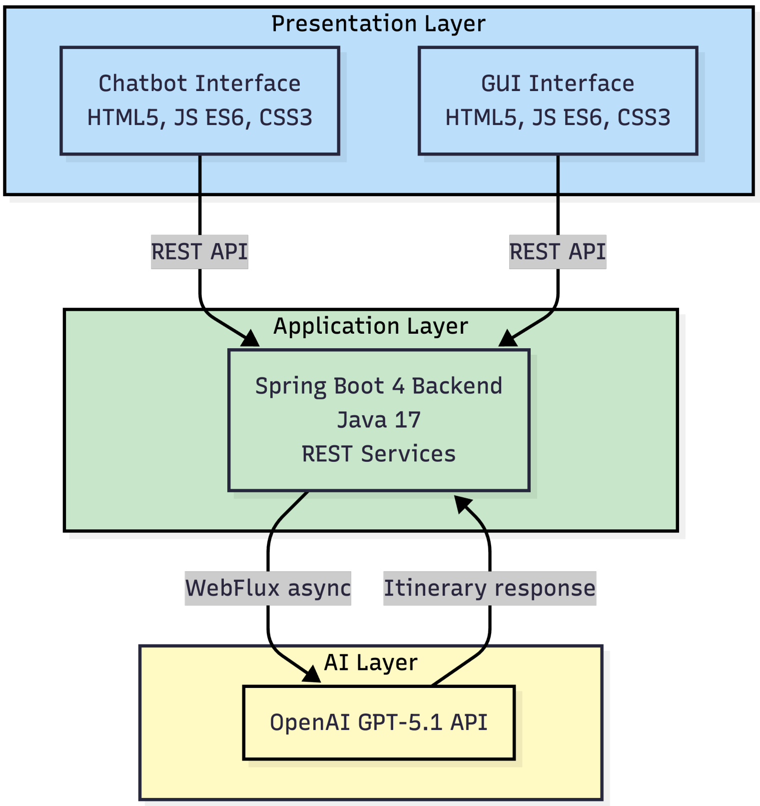


Figure 4.12: System architecture

The AI integration layer communicates directly with OpenAI's GPT-5.1 API using Spring WebFlux, enabling reactive, non-blocking communication for both modalities.

GPT-5.1 model parameters and rationale

The chatbot prompt and the itinerary generation prompt shared the same model configuration. To ensure that any observed differences could be attributed to interaction modality rather than to differences in model behaviour, the following parameters were selected:

Table 4.1: Summary of system parameters

Aspect	GUI	Chatbot
Model	GPT-5.1	GPT-5.1
Temperature	0.7	0.7
Verbosity	medium	medium
Token Limit	3000	3000
Reasoning effort	none	none
Backend	Shared	Shared
Itinerary prompt	Shared	Shared
Prompt for interaction	No	Yes
Conversation History	No	Yes

Reasoning effort was set to "none". Testing showed that a higher reasoning effort resulted in higher latency for the Chatbot responses and greater variability in itinerary generation for both modalities. Therefore, temperature was used to control variability.

Temperature was set to 0.7. Testing showed that anything below 0.5 resulted in rigid, list-like chatbot interactions, while 0.7 preserved conversational fluidity while maintaining consistent itinerary outputs for both modalities. Anything above 0.7 produced excessive variation in both chatbot responses and itinerary generation for each modality.
The temperature parameter can be used only if reasoning effort is set to "none".

Response verbosity was set to "medium" to balance readability and informational completeness. Testing showed that lower verbosity produced good answers for the Chatbot but insufficiently detailed itineraries while higher verbosity increased itinerary output making the itineraries too long to read.

Maximum output length was set to 3000 tokens. Testing showed that fewer tokens resulted in the truncation of full-day itineraries. To avoid excessively long chatbot responses, the chatbot prompt constrained the answers rather than limiting tokens.

Technical configuration:
(“reasoning_effort”, "none")
(“verbosity”, “medium”)
(“temperature”, 0.7)
(“max_completion_tokens”, 3000)

Chatbot prompt conversational design

The Chatbot was designed to use a warm and engaging tone, with enthusiastic language and limited emoji use. The Chatbot asks one question at a time, waits for the user’s response before proceeding, and prevents premature itinerary generation. The prompt limits stylistic variation to ensure consistent behaviour across participants. It also prevents unsolicited recommendations, explanations of internal rules, or references to the experimental setup.

While the interaction flow is fixed, the Chatbot allows some conversational flexibility through explicit memory and context rules defined in the prompt. To avoid redundant questioning, the Chatbot is instructed to extract relevant preferences even when users provide information out of sequence or in non-standard formats. For example, if a user mentions interests or dietary restrictions early in the conversation, the Chatbot records this information and skips the corresponding question later. The Chatbot also allows flexible date expressions and instructs the model to normalize inputs rather than asking users to reformat them.

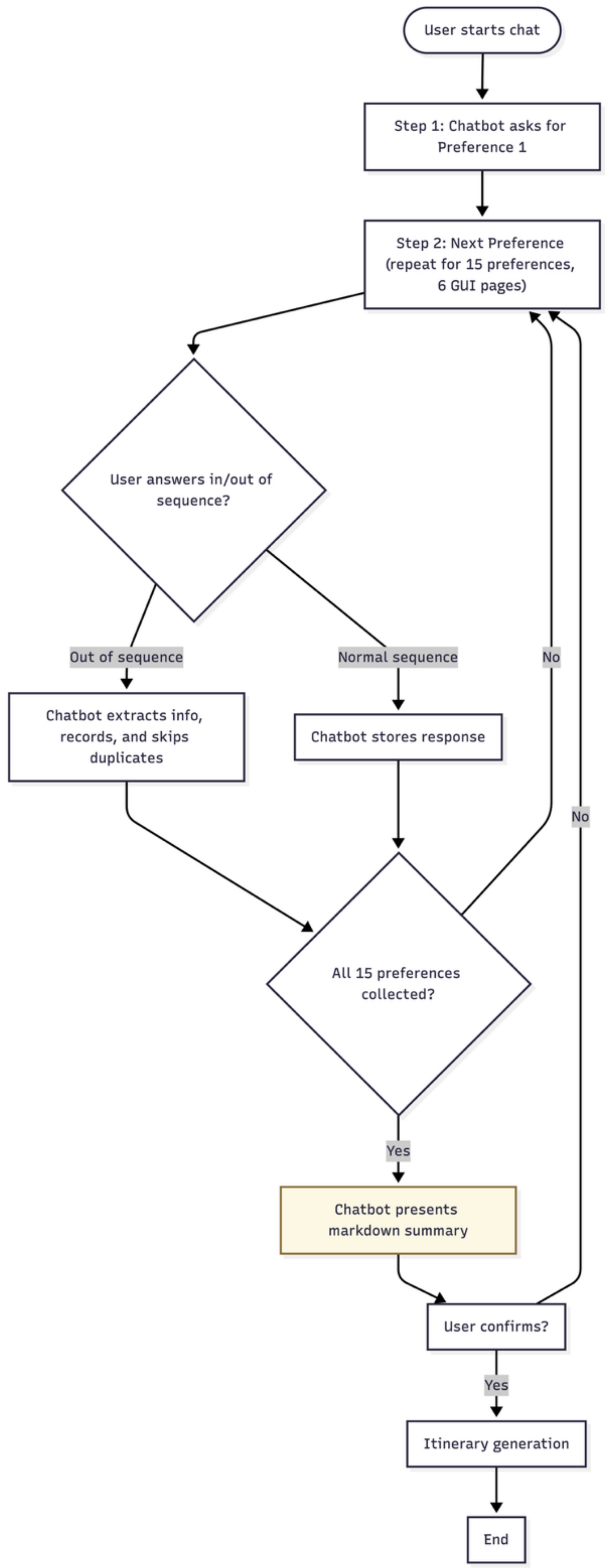


Figure 4.13: Chatbot prompt high level process

Chatbot specifications

To preserve conversational context, the full conversation history is sent from the frontend to the backend with each user message. For each interaction, the backend constructs a message sequence consisting of the chatbot prompt, the accumulated conversation history, and the current user input. As the conversation progresses, user preferences are extracted from the conversation history and stored.

Once all fifteen preferences have been collected, the Chatbot presents a structured summary in a standardized markdown format and requests user confirmation before proceeding. After the user provides confirmation, these structured preferences are passed to the same itinerary generation pipeline used by the GUI. Generated itineraries are returned to the Chatbot interface in markdown format and rendered client-side using the same markdown rendering approach as the GUI. Cross-origin requests were enabled to allow communication between the locally hosted frontend and backend.

GUI specifications

User preferences are stored in a single JavaScript preferences object that stores all selections made across the interface. User preferences are collected step-by-step and stored locally in the browser during the interaction. This object is updated as users interact with the interface, keeping selections consistent across steps. Required fields are validated at each page transition, preventing users from continuing until all necessary inputs are provided. This validation logic enforces completeness and consistency before itinerary generation.

Date selection is handled using the Flatpickr library. Selected dates are normalized to local midnight before being stored or sent to the backend to avoid timezone-related inconsistencies. Once users reach the review and confirmation step and click the Generate my plan button, the GUI serializes the collected preferences into a structured JSON payload. This payload is transmitted to the backend via a POST request to the specified endpoint. Typical itinerary generation latency ranges between 5 and 10 seconds. The GUI frontend and backend communicate with data transfer objects. Preferences collected through the interface are sent in a format that matches the backend's requirements. Generated itineraries are returned to the GUI in markdown format and rendered using Marked.js for markdown-to-HTML conversion.

4.5 System Limitations

The primary limitation of the GUI is the fixed linear interaction flow. Users are required to progress through the six pages in a predefined order and cannot skip steps, reorder questions, or modify individual preferences without navigating back through the interface. While this structure supports experimental consistency, it may feel restrictive for users who prefer to specify information in a different order.

The primary limitation of the Chatbot interface is the constrained, protocol-driven interaction flow. Although presented as a conversational system, the Chatbot follows a fixed, step-by-step question sequence that users cannot bypass, reorder, or alter. Users are required to respond to each prompt before progressing, and spontaneous deviations from the predefined flow are not supported. While this design choice was essential for maintaining experimental control and ensuring comparability with the GUI condition, it limits conversational flexibility. Users who prefer a more exploratory dialogue style may experience the interaction as restrictive despite its conversational appearance.

The system generates itineraries for a single day only. Multi-day trips are not supported, as they would require additional logic for scheduling across days, budget distribution, and preference consistency. After an itinerary is generated, users cannot edit individual activities, adjust timing, or regenerate selected parts of the plan. Any changes require restarting the planning process from the beginning. The system relies solely on the language model's internal knowledge when suggesting venues, activities, and cost estimates. Prices, timings, and recommendations are not based on real-time data. While the model suggests real places and provides credible estimates, it does not verify availability, opening hours, current prices, seasonal effects, or local events. As a result, itinerary details should be considered approximate rather than fully accurate or up to date. The system does not support bookings, reservations, or integration with external travel services. This reflects its role as an experimental planning tool designed for controlled evaluation rather than a production-ready travel platform.

5 EMPIRICAL RESEARCH METHOD

5.1 Participants

A total of 36 participants took part in the study. An a priori statistical power analysis was conducted using G*Power 3 (Faul et al., 2007) to determine the required sample size for the study. A paired-samples t-test (difference between two dependent means) was used, reflecting the within-subjects design in which each participant interacted with both interfaces. A two-tailed test was selected because the study did not assume a directional hypothesis.

A medium effect size ($d = .50$) was assumed, following Cohen's (1988) conventions for behavioural research when prior effect size estimates are unavailable. The significance level (α) was set to .05, corresponding to the conventional threshold for controlling Type I error in behavioural and HCI research. Statistical power was set to $1 - \beta = .80$, reflecting the probability of correctly rejecting a false null hypothesis and indicating an 80% chance of detecting a true effect of the specified size, consistent with established methodological recommendations (Cohen, 1988; Faul et al., 2007). Based on these parameters, the power analysis indicated a minimum required sample size of $N = 34$. To allow for an even distribution across experimental conditions, the final sample size was increased to $N = 36$.

Eligible participants were required to be 18 years or older and to be actively involved in leisure travel planning, defined as either mostly planning trips themselves or co-planning trips with someone else. Individuals who reported that someone else mostly planned trips for them, or that they rarely planned trips or rarely travelled, were excluded. Participants were also required to have planned at least one leisure trip within the previous 12 months, either for themselves or together with someone else. Because all study materials and interactions were conducted in English, participants were required to be comfortable reading and answering questions in English. Participation also required availability for a 30-minute in-person session. Finally, as the study involved audio recording for qualitative analysis, only participants who consented to being audio recorded were included.

Participants were recruited using a screening questionnaire designed to ensure that all inclusion criteria were met. All participants provided basic demographic and contact information as part of the screening process and gave informed consent prior to participation. No monetary or material compensation was provided. Participants were between 18 and 34 years old, with 30 participants aged 18–24 and 6 participants aged 25–34. All participants were students and were recruited from the Lausanne area. The sample consisted of an equal gender distribution, with 18 male and 18 female participants. The screener questionnaire is provided in Appendix VII.

5.2 Apparatus

Participants interacted with the experimental systems using the same laptop computer across all sessions. The device used was a 14-inch MacBook Pro (2021), with the built-in keyboard and trackpad serving as the sole input devices. The computer was running macOS, and all experimental systems were accessed using the Google Chrome web browser. The laptop was connected to the internet during the sessions, although internet connectivity was not required for task completion beyond enabling access to the web-based interfaces. To ensure experimental consistency, the same hardware, operating system, browser, and interaction setup was used for all participants. Sessions took place in on-site locations (e.g., university halls). Participants were seated at a table and completed the tasks individually in a quiet environment with comparable conditions across sessions. Qualitative responses were provided verbally and were audio-recorded using an iPhone with the Voice Memos application.

5.3 Procedure

All study sessions were conducted on-site over the course of one week. Sessions were distributed across multiple days to reduce cognitive fatigue and minimize procedural errors, with nine sessions conducted on Monday, nine on Tuesday, nine on Thursday, and nine on Friday. Each session lasted approximately 20 to 30 minutes. The intervening day (Wednesday) was used for preliminary data checks and verification of collected responses. Each session followed a standardized protocol and was conducted individually. The protocol is provided in Appendix VIII.

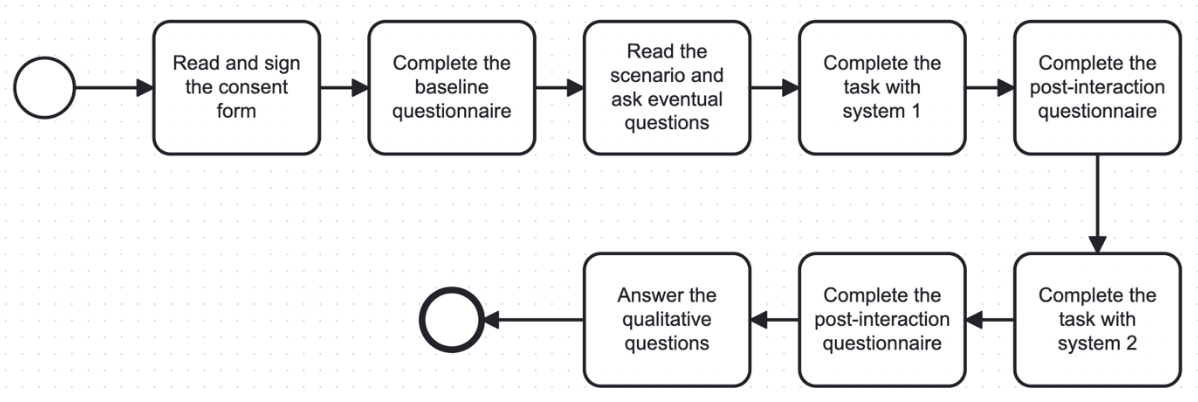


Figure 5.1: Participant session process

First, participants were seated at a table and asked to read and sign a consent form. The consent form described the purpose of the study, the procedure, the voluntary nature of participation, confidentiality measures, and the use of audio recording. The consent form was provided and signed on paper. The consent form is provided in Appendix IX.

After providing consent, participants completed the baseline questionnaire using Microsoft Excel. This questionnaire was administered prior to any interaction with the experimental systems.

After completing the baseline questionnaire, participants were given a printed scenario description. The scenario instructed participants to plan a solo weekend trip to Madrid using the system provided and without consulting external websites or tools. Participants were given the opportunity to ask clarification questions regarding the scenario or task instructions.

After completing the task with the first system, participants completed the corresponding post-interaction questionnaire using Microsoft Excel. Participants then completed the same task with the second system, followed by the corresponding post-interaction questionnaire using Microsoft Excel.

Following the completion of both experimental conditions and post-interaction questionnaires, participants answered two short open-ended questions at the end of the session.

To control for order effects, the order of system interaction was counterbalanced across participants. Eighteen participants interacted with the Chatbot first, and eighteen interacted with the GUI first. Counterbalancing was applied equally across genders, with nine male and nine female participants starting with each system.

During each system interaction, participants completed the task independently on the computer. The researcher did not observe participants' screens, provide guidance, or intervene during the interaction, unless participants asked clarification questions regarding their understanding of the task instructions. Participant interactions were not screen-recorded. Outputs generated by the systems were saved for record-keeping purposes but were not included in the analysis.

Participants spent approximately 10 minutes interacting with the Chatbot and completing the post-interaction questionnaire, approximately 5 to 10 minutes interacting with the GUI and completing the post-interaction questionnaire, and 2 to 5 minutes answering the qualitative questions. After completing all study tasks, participants were thanked for their participation, and the session concluded.

The questions are provided in Appendix XI and the scenario in Appendix XII.

5.4 Measures and Constructs

Scenario and task scope

The scenario was designed to limit the scope of the travel planning task in order to isolate the effects of interaction modality on users' trust and sense of control. Travel logistics such as flights, accommodation, and transportation were treated as already arranged, and participants were asked to focus exclusively on planning activities and food. In addition, participants were instructed to only plan for a single day. The scenario remained identical across both experimental conditions.

Limiting the task to a single day and excluding logistical considerations reduced cognitive load and avoided introducing additional constraints and trade-offs, such as budgets, schedules, or availability, that could have shifted participants' attention toward optimizing the quality of the itinerary output. By keeping the task focused and comparable across participants, the study aimed to ensure that observed differences in trust and sense of control could be attributed to the interaction modality itself, rather than to differences in task complexity or planning outcomes.

Likert scale choice and response anchors

All questionnaire items were measured using a 7-point Likert scale with the following response anchors: Strongly disagree, Disagree, Slightly disagree, Neither agree nor disagree, Slightly agree, Agree, Strongly agree.

A 7-point scale was used to provide sufficient sensitivity for detecting within-participant differences across conditions. Empirical evidence indicates that rating scales with two to four response categories perform relatively poorly in terms of reliability, validity, and discriminating power, whereas measurement quality increases with additional response categories up to about seven, with little further improvement beyond that point (Preston & Colman, 2000).

Trust framework

Trust was measured following McKnight and Chervany's (2001) interdisciplinary trust typology. This framework was chosen because it provides a well-established conceptualization of trust in information systems. It conceptualizes trust as a set of related but distinct constructs rather than a single unified concept. In this framework, the trustor refers to the individual who places trust, and the trustee refers to the entity being trusted. In the present study, the trustor is the participant, and the trustee is the AI travel planning system.

Trust constructs

McKnight and Chervany (2001) distinguish between the following trust constructs: disposition to trust, institution-based trust, trusting beliefs, and trusting intentions.

The present study focused on trusting beliefs and trusting intentions, as these constructs are formed through direct interaction with a specific trustee (here, the AI travel planner). Because the study focused on trust judgments toward a specific AI system rather than general trust tendencies or perceptions of institutional safeguards, disposition to trust and institution-based trust were not measured. Disposition to trust reflects a general tendency to trust others, independent of any particular system. Institution-based trust concerns beliefs about external structural conditions, such as regulations or safeguards, rather than trust in the system itself.

Trust construct: trusting beliefs (McKnight and Chervany, 2001)

"Trusting beliefs means that one believes that the other party has one or more characteristics beneficial to oneself" (McKnight and Chervany, 2001). In the present study, this corresponds to participants' beliefs that an AI travel planning system possesses qualities that are beneficial to the user.

McKnight and Chervany identify four subconstructs of trusting beliefs: competence, benevolence, integrity, and predictability. In this study, trusting beliefs were represented through competence and benevolence, as defined by McKnight and Chervany (2001) and adapted to the context of AI-based travel planning systems.

“Trusting belief-competence means that one believes that the other party has the ability or power to do for one what one needs done” (McKnight and Chervany, 2001). In this study, competence was adapted by assessing the extent to which participants believed that the AI travel planning system was capable of performing the travel-planning task effectively during the interaction.

“Trusting belief-benevolence means that one believes that the other party cares about one and is motivated to act in one’s interest” (McKnight and Chervany, 2001). In this study, benevolence was adapted by assessing the extent to which participants believed that the AI travel planning system acted in their best interest during the interaction.

The remaining trusting belief subconstructs, integrity and predictability, were not measured in this study.

“Trusting belief-integrity means that one believes that the other party makes good-faith agreements, tells the truth, acts ethically, and fulfils promises” (McKnight and Chervany, 2001). Integrity was not measured because the study focused on participants’ experience of interaction with the AI system rather than evaluations of outcome quality.

“Trusting belief-predictability means that one believes the other party’s actions (good or bad) are consistent enough that one can forecast them in a given situation” (McKnight and Chervany, 2001). Predictability was not measured because participants interacted with each system only once, and the study focused on the interaction rather than repeated system behaviour.

Trust construct: trusting intention (McKnight and Chervany, 2001)

“Trusting intentions means that one is willing to depend on, or intends to depend on, the other party even though one cannot control that party” (McKnight and Chervany, 2001).

In the present study, this corresponds to participants’ willingness to rely on an AI travel planning system when making travel-related decisions, despite limited control over how the system operates.

McKnight and Chervany identify two subconstructs of trusting intentions: willingness to depend, and subjective probability of depending. In this study, trusting intentions were

represented through willingness to depend, as defined by McKnight and Chervany (2001) and adapted to the context of AI travel planning.
"Willingness to depend means that one is volitionally prepared to make oneself vulnerable to the other party in a situation by relying on the other party" (McKnight and Chervany, 2001). In this study, willingness to depend was adapted by assessing the extent to which participants were willing to rely on the AI travel planning system during the interaction.

The remaining trusting intention subconstruct, subjective probability of depending, was not measured because the study focused on immediate trust-related intentions following interaction and not forecasts of future behaviour. "Subjective probability of depending means the extent to which one forecasts or predicts that one will depend on the other party" (McKnight and Chervany, 2001).

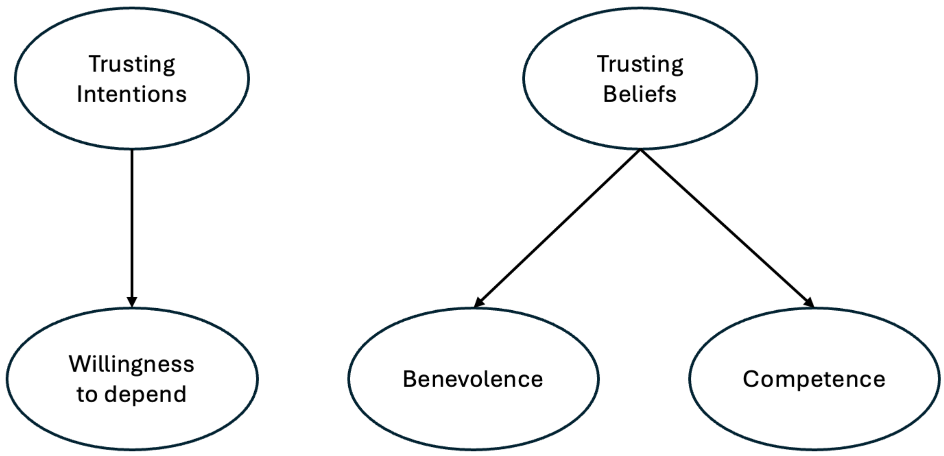


Figure 5.2: Trust constructs

Trust items

Each trust dimension identified above was measured using a single item to allow repeated measurement within a single experimental session without participant burden. All items were phrased to be easily interpretable by participants without requiring prior experience with the systems.

Baseline trust was measured prior to any system interaction and referred to AI systems in general. Participants responded to the following three items, each corresponding to a specific trust construct:

- Trusting belief-competence: “I believe AI produces reliable travel suggestions.”
- Trusting intentions-willingness to depend: “I’m willing to rely on AI when making travel planning decisions.”
- Trusting belief-benevolence: “I believe AI acts in my best interest when giving me travel suggestions.”

After interacting with each system, participants answered the same three items, with the wording anchored to the specific system they had just used:

- Trusting belief-competence: “I believe this AI system produced reliable travel suggestions.”
- Trusting intentions-willingness to depend: “I’m willing to rely on this AI system to make travel planning decisions.”
- Trusting belief-benevolence: “I believe this AI system acted in my best interest when giving me travel suggestions.”

Using identical items across baseline and post-interaction measurements enabled direct within-participant comparison.

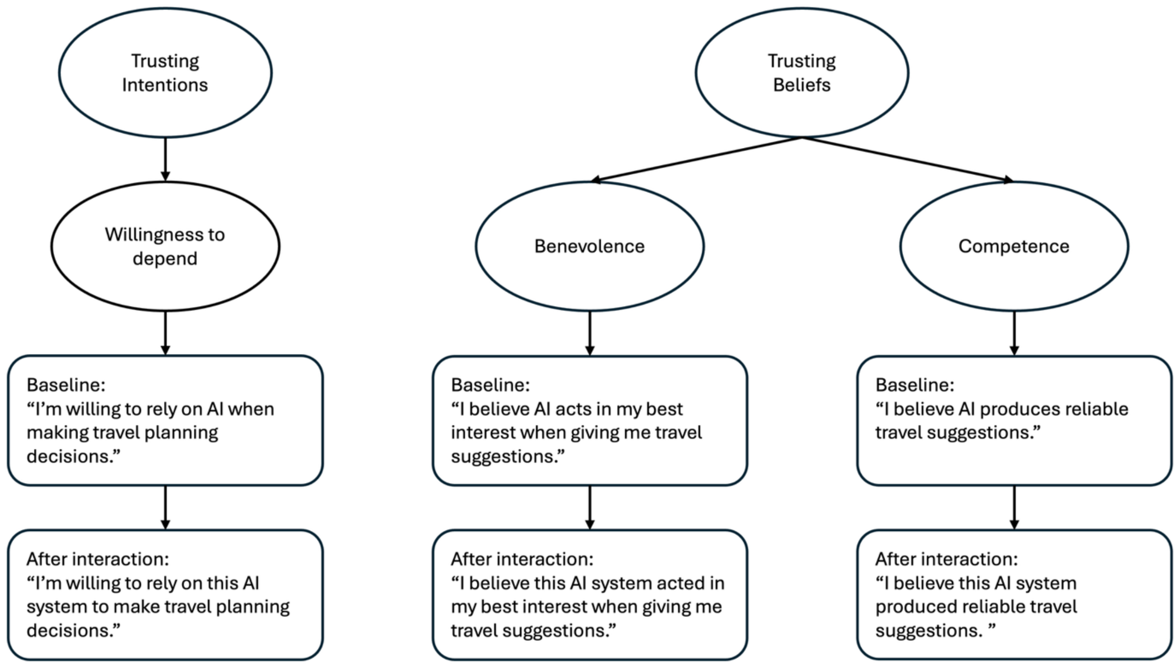


Figure 5.3: Trust constructs and items

Sense of control framework

Sense of control was measured based on Norman's model of seven stages of action (Norman, 2013). This framework was chosen because it explains how people interact with systems by describing how they execute actions and evaluate their results. Although Norman does not define sense of control as a single measurable construct, his model explains how feelings of control arise from users' ability to carry out intended actions and to understand what the system does in response.

The seven stages of action as described by Norman are the following:

One stage for goals:

1. Goal (form the goal)

Three stages for execution:

2. Plan (the action)
3. Specify (an action sequence)
4. Perform (the action sequence)

Three stages for evaluation:

5. Perceive (the state of the world)
6. Interpret (the perception)
7. Compare (the outcome with the goal)

These stages describe the cognitive process users go through when interacting with a system to complete a task. Norman characterizes action as consisting of two components: execution and evaluation. He states that "there are two parts to an action: executing the action and then evaluating the results: doing and interpreting" (Norman, 2013). He further notes that "both execution and evaluation require understanding: how the item works and what results it produces," and that both components can affect users' emotional state (Norman, 2013). When execution and evaluation are effectively supported, interaction may result in positive outcomes, as "well-designed devices can induce pride and enjoyment, a feeling of being in control and pleasure-possibly even love and attachment" (Norman, 2013).

Sense of control construct

In the present study, sense of control refers to participants' experience of control while interacting with the AI travel planning system. Participants went through all seven stages of action during task completion. However, the study did not aim to measure each of the seven stages individually. Instead, following Norman's higher-level distinction, the study focused on the two components of action: execution and evaluation. These components directly relate to users' ability to act on the system and to understand the system behaviour during the interaction.

Based on this model, sense of control was measured using three complementary aspects: perceived influence over system output, perceived understanding of system behaviour, and overall experienced control.

Perceived influence over system output

Perceived influence over system output reflects the extent to which participants believed that their interaction with the system allowed them to shape the system's behaviour and suggestions during the trip-planning task. This aspect captures users' ability to translate intentions into effective actions through the available interaction mechanisms, corresponding to the execution component of Norman's model of action.

Perceived understanding of system behaviour

Perceived understanding of system behaviour measures the extent to which participants believed they understood how and why the AI system generated its travel suggestions based on their interaction with the system during the trip-planning task. This aspect focuses on users' perceived ability to interpret system responses and behaviour as they interact with the system. Norman emphasizes that successful evaluation depends on users' ability to perceive and interpret system feedback and to make sense of outcomes relative to their goals.

Overall experienced control

Overall experienced control reflects participants' perceived feeling of being in control while interacting with the AI system during the trip-planning task. This aspect represents the outcome of being able to act on the system and understand its behaviour through interaction and is not treated as a separate cognitive stage, but as a result of execution and evaluation being effectively supported, consistent with Norman's model of action.

Sense of control items

Each sense of control dimension identified above was measured using a single item to allow repeated measurement within a single experimental session without participant burden. All items were phrased to be easily interpretable by participants without requiring prior experience with the systems.

Baseline sense of control was measured prior to any system interaction and referred to AI systems in general. Participants responded to the following three items, each corresponding to a specific sense of control component:

- Perceived influence over system output: "I believe I'm able to influence the output of an AI system when planning a trip."
- Perceived understanding of system behaviour: "I understand how an AI system generates its travel suggestions."
- Overall experienced control: "I feel in control when using an AI system for travel planning."

After interacting with each system, participants answered the same three items, with the wording anchored to the specific system they had just used:

- Perceived influence over system output: "I knew how to influence this AI system's output."
- Perceived understanding of system behaviour: "I understood how this AI system generated its travel suggestions."
- Overall experienced control: "I felt in control when interacting with this AI system."

Using identical items across baseline and post-interaction measurements enabled direct within-participant comparison.

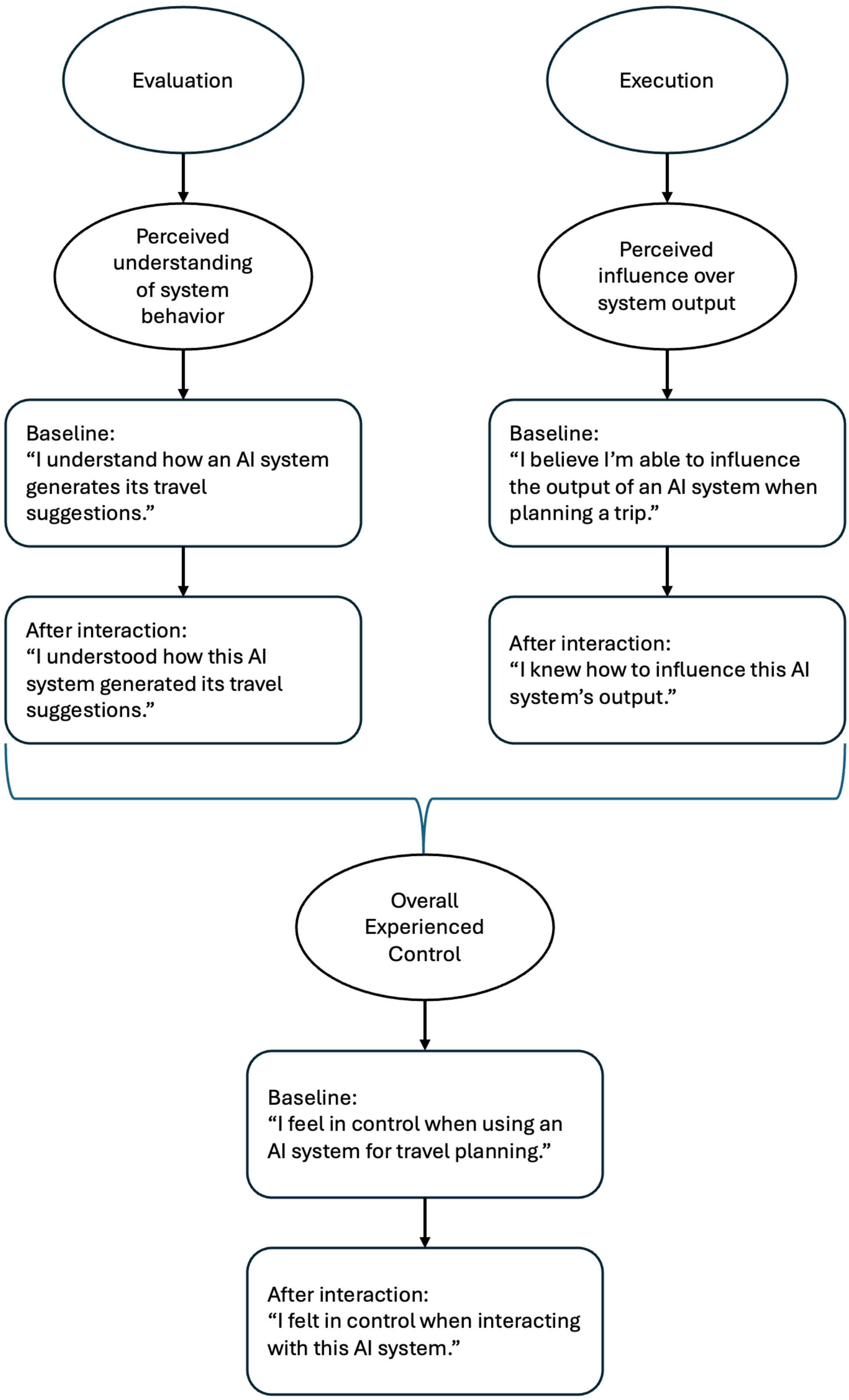


Figure 5.4: Sense of control constructs and items

Qualitative component

The qualitative component of the study was included to complement the quantitative results and was exploratory in nature, rather than a standalone qualitative study or formal interview. Participants were asked (1) which system they trusted more and why, and (2) which system made them feel more in control and why.

5.5 Analysis Methodology

The quantitative study employed a within-subjects experimental design examining two dependent variables: trust and sense of control. Each construct was measured at three points: Baseline, after using the GUI system, and after using the Chatbot system. All measures were collected using a 7-point Likert scale, allowing direct within-participant comparison across conditions.

All quantitative answers were compiled into a structured dataset, with each row representing an individual item response. Each response was associated with a participant identifier, question item, construct (trust or sense of control), and condition (Baseline, GUI, Chatbot). Likert responses were encoded from 1 to 7. Item-level scores were aggregated by construct and condition for each participant, resulting in mean scores for trust and sense of control for each condition.
To prepare the data for statistical testing, a pivot table was used to compute participant mean scores for each construct–condition combination. These values were exported into a separate worksheet, where descriptive statistics were calculated for each condition.

Statistical comparisons were conducted using paired-sample t-tests, appropriate for the within-subjects design. The following comparisons were performed separately for each construct: GUI vs. Chatbot, Baseline vs. GUI, Baseline vs. Chatbot.
All statistical analyses were conducted in Microsoft Excel using the Analysis ToolPak (t-Test: Paired Two Sample for Means). For each test, the input consisted of participant mean scores for the two conditions being compared, the hypothesized mean difference was set to 0, statistical significance was evaluated at $\alpha = .05$.

The qualitative answers were analysed using thematic analysis following the six-phase framework proposed by Braun and Clarke (2006). Braun and Clarke (2006) define this framework as a "method for identifying, analysing, and reporting patterns (themes) within data".

In **phase 1**, familiarising yourself with your data (Braun and Clarke, 2006), audio recordings were transcribed using the AI web application TurboScribe. The researcher then manually reviewed each transcript while listening to the original recordings to verify accuracy and added speaker labels (Researcher and Participant n°) to each statement. Transcripts were subsequently cleaned to improve clarity and readability without altering the meaning or intent of participants' statements. This included removing filler words, non-lexical utterances, false starts, procedural or non-substantive segments (e.g., greetings and closing remarks), as well as interviewer interruptions and backchannel prompts. Minor grammatical corrections were applied, and ambiguous references (e.g., "the first one," "the other one") were clarified using parentheses.

In **phase 2**, generating initial codes (Braun and Clarke, 2006), the transcripts were coded to explain why participants trusted one system more than the other or felt more in control with one system more than the other.

Participant	Transcript	Trust more	Phase 2: Trust Initial Codes	Control more	Phase 2: Control Initial Codes
14	Researcher: Which system did you trust more and why? Participant 14: I would say that I would trust more the GUI. I feel like the chatbots are more generalized and it's more free range whereas the GUI is more specific. But in the chatbot, I feel like a lot of times even the people or me, tend to write a bit too much or there's too much freedom in it. And I feel like, based on how I use AI, it's more personalized and more straightforward, so the AI just gives me the options, like, bam, bam, bam, that's it. Researcher: So, it allows you not to get too lost. Participant 14: Yeah, exactly. So, yeah, in the end, in short, I would say the GUI. Researcher: And which one made you feel more in control? Participant 14: More in control, definitely the chatbot because I'm able to write how much I want about anything.	GUI	- More specific and less free range, straight to the point	Chatbot	- Write whatever I want with no word limitation

Figure 5.5: Phase 2 generating initial codes example

In **phase 3**, searching for themes (Braun and Clarke, 2006), related codes were grouped into broader candidate themes that captured recurring patterns across participants' explanations.

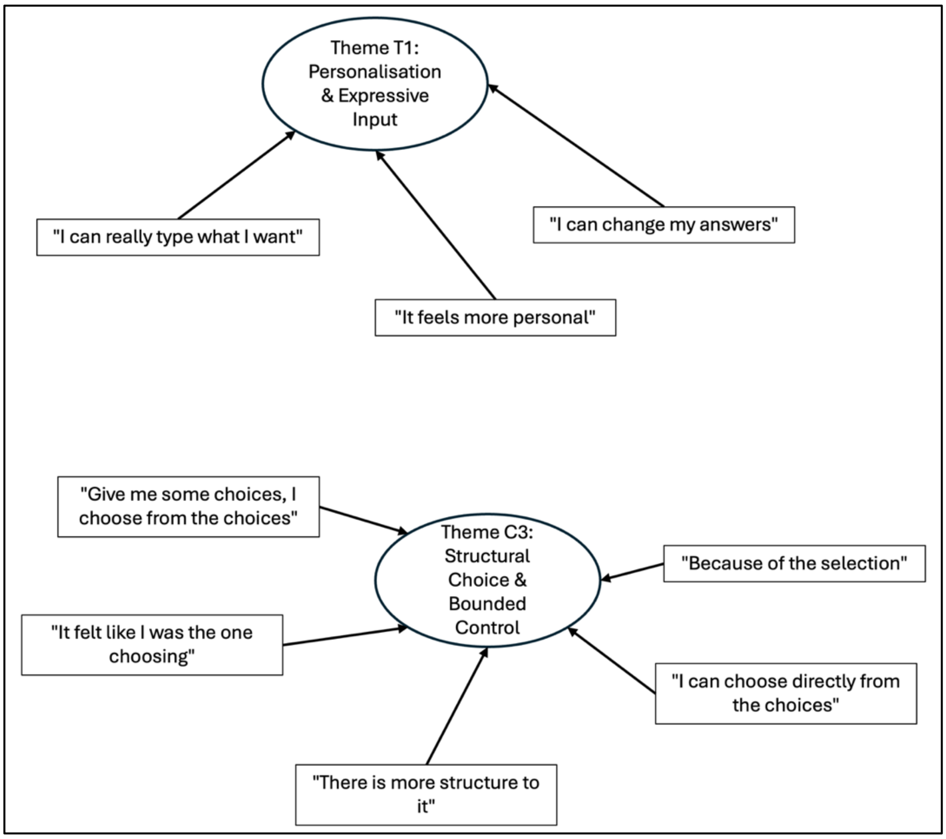


Figure 5.6: Phase 3 searching for themes

In **phase 4**, reviewing themes (Braun and Clarke, 2006), themes were assessed for internal coherence and distinctiveness. Given the short nature of the qualitative responses and the exploratory role of the qualitative component, no themes were merged. Each candidate theme reflected a different reason participants gave for trusting a system or feeling in control, and combining themes would have removed meaningful differences between these reasons.

In **phase 5**, defining and naming themes, (Braun and Clarke, 2006), the theme names identified in phase 3 were retained, and the focus was on defining and clarifying each theme rather than renaming them.

In **phase 6**, producing the report (Braun and Clarke, 2006), the results from the analysis were reported in section 6.2 Qualitative Results.

6 RESULTS

6.1 Quantitative Results

This section presents the quantitative results of the study, examining differences in trust and sense of control across interaction modalities. Descriptive statistics are reported first, followed by inferential analyses. The analysis methodology is described in section 5.5 Analysis Methodology.

6.1.1 Descriptive Statistics

Baseline trust towards AI systems was moderate (M = 4.72). After interacting with the systems, trust increased in both conditions. Mean trust was slightly higher for the GUI system (M = 5.65) than for the Chatbot system (M = 5.58).

Baseline sense of control was also moderate (M = 4.54). After interacting with the systems, sense of control increased in both conditions. Mean sense of control was slightly higher for the Chatbot system (M = 5.40) than for the GUI system (M = 5.28).

Table 6.1: Descriptive statistics summary

Construct	Baseline	After GUI	After Chatbot
Trust	4.72	5.65	5.58
Control	4.54	5.28	5.40

6.1.2 Inferential Analysis

For each test, the input consisted of the participant mean scores for the two conditions being compared. The hypothesized mean difference was set to 0, statistical significance was evaluated at $\alpha = .05$. This analytical approach allowed direct within-participant comparison of trust and sense of control across interaction modalities and relative to baseline attitudes.

Trust: GUI vs. Chatbot

The paired-sample t-test comparing trust between the GUI and Chatbot modalities yielded no statistically significant difference, $t(35) = 0.50$, $p = .62$. This result indicates that trust was comparable across the two interaction modalities.

Trust: Baseline vs. GUI

The paired-sample t-test comparing trust between baseline and the GUI modality yielded a statistically significant difference, $t(35) = -5.43$, $p < .001$. This result indicates that trust was higher after interacting with the GUI than at baseline.

Trust: Baseline vs. Chatbot

The paired-sample t-test comparing trust between baseline and the Chatbot modality yielded a statistically significant difference, $t(35) = -5.12$, $p < .001$. This result indicates that trust was higher after interacting with the Chatbot than at baseline.

Sense of Control: GUI vs. Chatbot

The paired-sample t-test comparing sense of control between the GUI and Chatbot modalities yielded no statistically significant difference, $t(35) = -0.78$, $p = .44$. This result indicates that sense of control was comparable across the two interaction modalities.

Sense of Control: Baseline vs. GUI

The paired-sample t-test comparing sense of control between baseline and the GUI modality yielded a statistically significant difference, $t(35) = -3.23$, $p = .003$. This result indicates that sense of control was higher after interacting with the GUI than at baseline.

Sense of Control: Baseline vs. Chatbot

The paired-sample t-test comparing sense of control between baseline and the Chatbot modality yielded a statistically significant difference, $t(35) = -4.70$, $p < .001$. This result indicates that sense of control was higher after interacting with the Chatbot than at baseline.

Table 6.2: Inferential analysis summary

Construct	Baseline vs. GUI	Baseline vs. Chatbot	GUI vs. Chatbot
Trust	t(35) = −5.43***	t(35) = −5.12***	t(35) = 0.50
Sense of Control	t(35) = −3.23**	t(35) = −4.70***	t(35) = −0.78

Note. *t*-tests are two-tailed.

* $p < .05$ (), ** $p < .01$ (), *** $p < .001$ ().

6.2 Qualitative Results

This section presents the qualitative findings derived from post-task qualitative questions conducted after participants interacted with both AI-based travel planning interfaces and completed the quantitative surveys.

The following tables summarise the themes from the upcoming thematic analysis:

Table 6.3: Trust themes and definitions

Theme	Definition	Covers
Theme T1: Personalisation & Expressive Input	Trust based on the ability to express preferences freely and personalise inputs using one's own words.	- Writing detailed answers - Expressing personal preferences - Feeling personally addressed
Theme T2: Conversational Interaction & Dialogue	Trust based on the back-and-forth nature of the interaction and the feeling of an ongoing dialogue.	- Interactive exchange - Incremental questioning - Dialogue-like interaction
Theme T3: Structure, Guidance & Boundedness	Trust based on clear structure, predefined options, and guidance through the interaction.	- Visible options - Clear structure - Reduced ambiguity

Theme T4: Ease, Speed & Low Effort	Trust based on the system being easy, fast, and requiring little effort to use.	- Speed of interaction - Simplicity - Low effort
Theme T5: Reliability, Predictability & Reduced "Robotic" Feel	Trust based on the perception that the system behaves reliably, predictably, or in a less robotic way.	- Predictable behaviour - Feeling reliable - Less robotic interaction
Theme T6: Familiarity / Prior Experience	Trust based on prior experience with similar systems or interfaces.	- Being used to chat interfaces - Regular use of AI systems - Familiar interaction patterns

Table 6.4: Sense of control themes and definitions

Theme	Definition	Covers
Theme C1: Expressive Freedom & Influence	Sense of control based on the ability to influence the system by freely expressing inputs.	- Writing without constraints - Influencing outcomes directly - Saying exactly what one wants
Theme C2: Step-by-Step Guidance & Process Control	Sense of control based on progressing through the task in clear, sequential steps.	- One question at a time - Clear progression - Incremental control
Theme C3: Structural Choice & Bounded Control	Sense of control based on selecting from predefined options rather than writing inputs.	- Clicking choices - Option-based input - Controlled selection
Theme C4: Ease, Speed & Reduced Effort	Sense of control based on the system being fast, simple, and easy to use.	- Quick interaction - Minimal effort - Straightforward use
Theme C5: Transparency / Mental Model	Sense of control based on understanding how the system works or what it is doing.	- Knowing what happens next - Understanding system logic - Predicting outcomes
Theme C6: Personalisation (Control-related)	Sense of control based on the ability to tailor outputs to personal preferences.	- Customising results - Adapting responses - Fine-tuning preferences

6.2.1 Thematic Analysis

The most common theme among **participants who trusted the Chatbot more** was Personalisation & Expressive Input (T1). A second recurring theme for trusting the Chatbot more was Conversational Interaction & Dialogue (T2). The least frequent themes associated with Chatbot trust were Familiarity / Prior Experience (T6), and Reliability, Predictability & Reduced "Robotic" Feel (T5). Structure, Guidance & Boundedness (T3) did not appear as an independent theme and was always combined with Personalisation & Expressive Input (T1) or Conversational Interaction & Dialogue (T2). Finally, none of the participants who trusted the Chatbot more were associated with the theme Ease, Speed & Low Effort (T4).

The most common theme among **participants who trusted the GUI more** was Ease, Speed & Low Effort (T4). A second recurring theme for trusting the GUI more was Structure, Guidance, and Boundedness (T3). The least frequent theme associated with GUI trust was Reliability, Predictability & Reduced "Robotic" Feel (T5). Finally, none of the participants who trusted the GUI more were associated with the themes Personalisation & Expressive Input (T1), Conversational Interaction & Dialogue (T2), and Familiarity / Prior Experience (T6).

Table 6.5: Trust theme frequency per modality

Theme frequency	**Chatbot**	**GUI**
Most	T1: Personalisation & Expressive Input	T4: Ease, Speed & Low Effort
Second	T2: Conversational Interaction & Dialogue	T3: Structure, Guidance & Boundedness
Least	T3: Structure, Guidance & Boundedness T5: Reliability, Predictability & Reduced "Robotic" Feel T6: Familiarity / Prior Experience	T5: Reliability, Predictability & Reduced "Robotic" Feel
No Association	T4: Ease, Speed & Low Effort	T1: Personalisation & Expressive Input T2: Conversational Interaction & Dialogue T6: Familiarity / Prior Experience

Among participants who reported feeling **more in control when using the Chatbot**, the most common theme was Expressive Freedom & Influence (C1). A second recurring theme for feeling more in control with the Chatbot was Step-by-Step Guidance & Process Control (C2). This theme sometimes appeared on its own and sometimes alongside Expressive Freedom & Influence (C1). The least frequent themes associated with the Chatbot control were Transparency / Mental Model (C5), and Personalisation as a control-related factor (C6). Finally, none of the participants who felt more in control with the Chatbot were associated with the themes Structural Choice & Bounded Control (C3), and Ease, Speed & Reduced Effort (C4).

Among participants who reported feeling **more in control when using the GUI**, the most common theme was Structural Choice and Bounded Control (C3). This theme stands alone as the primary explanation for feeling more in control with the GUI. Second recurring themes associated with GUI-related control were Ease, Speed & Reduced Effort (C4), and Transparency / Mental Model (C5). Finally, none of the participants who felt more in control with the GUI were associated with the themes Expressive Freedom & Influence (C1), Step-by-Step Guidance & Process Control (C2), and Personalisation (C6).

Table 6.6: Sense of control theme frequency per modality

Theme frequency	Chatbot	GUI
Most	C1: Expressive Freedom & Influence	C3: Structural Choice & Bounded Control
Second	C2: Step-by-Step Guidance & Process Control	C4: Ease, Speed & Reduced Effort C5: Transparency / Mental Model
Least	C5: Transparency / Mental Model C6: Personalisation (Control-related)	/
No Association	C3: Structural Choice & Bounded Control C4: Ease, Speed & Reduced Effort	C1: Expressive Freedom & Influence C2: Step-by-Step Guidance & Process Control C6: Personalisation (Control-related)

7 DISCUSSION

7.1 Overview of Findings and Research Questions

This study investigated how chatbot and GUI interaction modes differ in their effects on users' trust and sense of control during the early stages of AI travel planning.
Following the interaction with either modality, participants reported a higher trust and sense of control compared to baseline attitudes towards AI.

Regarding which interaction scored higher for each construct, descriptive statistics show that mean trust scores were slightly higher for the GUI (GUI = 5.65, Chatbot = 5.58), whereas mean sense of control scores were slightly higher for the Chatbot (Chatbot = 5.40, GUI = 5.28). However, paired-sample t-tests showed no statistically significant differences between the Chatbot and GUI modalities for either construct (trust: $p = .62$; sense of control: $p = .44$).
Overall, the quantitative results provide no evidence of a reliable difference between conversational and GUI interaction modes in terms of measured trust or sense of control.

Qualitative findings show that for **trust**, the Chatbot was most frequently associated with the theme of Personalisation & Expressive Input (T1), whereas the GUI was most frequently associated with the theme of Ease, Speed & Low Effort (T4).
For **sense of control**, the Chatbot was most frequently associated with the theme of Expressive Freedom & Influence (C1), whereas the GUI was most frequently associated with the theme of Structural Choice and Bounded Control (C3).
Overall, the qualitative findings show that users have clearly distinct reasons for trusting and feeling in control with one modality over the other.

7.2 Trust: Chatbot vs. GUI

Trust increased significantly from baseline in both interaction modalities (GUI: $p < .001$; Chatbot: $p < .001$). This indicates that participants formed trust beliefs and trusting intentions through direct interaction with each modality. This is consistent with McKnight and Chervany's (2001) view that trusting beliefs and trusting intentions develop through interaction with a specific trustee.

Although no statistically significant difference ($p = .62$) in trust was observed between the Chatbot and GUI, the qualitative findings reveal that different trust reasons were associated with each interaction mode.

Participants who trusted the **Chatbot** more described trust in terms of the back-and-forth nature of the interaction, the feeling of dialogue, and the incremental build-up of answers. Trust in the Chatbot was associated with the ability to write in one's own words, state preferences clearly, and progress through questions one at a time without seeing all options at once. The Chatbot was experienced as more personally engaging, even though its behaviour followed a fixed interaction prompt.

Participants who trusted the **GUI** more described the GUI as easier to use, faster, and requiring less effort because it relied on clicking options rather than writing text. Trust in the GUI was associated with having visible options, a clear structure, procedure, and predefined choices. Participants emphasized that this structure provided guidance and reduced ambiguity, particularly for users who did not have a clear idea of what they wanted in advance.

Despite these different reasons for trusting either modality, overall trust levels were comparable across interaction modes. This indicates that trust can be achieved through different interaction modes.

7.3 Sense of Control: Chatbot vs. GUI

Sense of control increased significantly from baseline in both interaction modalities (GUI: $p = .003$; Chatbot: $p < .001$). This indicates that participants experienced a higher sense of control after interacting with either modality compared to their baseline attitudes. From the perspective of Norman's model of action, this suggests that both modalities sufficiently supported execution and evaluation, enabling participants to form goals, carry out actions, and interpret outcomes during the task.

Although no statistically significant difference in sense of control was observed between the Chatbot and GUI ($p = .44$), the qualitative findings reveal that different reasons for feeling in control were associated with each interaction mode.

Participants who felt more in control with the **Chatbot** described control in terms of being able to write freely, say exactly what they wanted, and adapt their inputs without being constrained by predefined options. Control was associated with progressing through the interaction one question at a time, building answers incrementally, and having a clear sequence during the interaction.

Participants who felt more in control with the **GUI** described control as being able to see multiple questions at once and understand the overall structure of the interface. Control was associated with selecting from predefined options, clicking visible choices, and making clearly bounded selections. Participants also described the GUI as faster, easier to use, and more intuitive, and some emphasized feeling more in control because the interface was clear and predictable.

Despite these different reasons for feeling in control with either modality, overall sense of control levels were comparable across interaction modes. This indicates that a sense of control can be achieved through different interaction modes.

7.4 Limitations

The participant sample consisted entirely of relatively young and digitally literate students recruited from a single geographic area. In addition, all participants had prior experience using AI and were used to planning trips. This may limit how well the findings apply to broader populations. The study examined trust and sense of control during a single interaction session. As a result, the findings reflect early perceptions rather than trust or sense of control formed through repeated or long-term use. The study focused on a single, constrained travel-planning task, which may limit the applicability of the findings to more complex scenarios. Finally, the Chatbot followed a fixed prompt to ensure experimental control and as a result, the findings may not generalize to more open-ended conversational systems.

7.5 Future Work

An important extension of this work could examine how trust and sense of control change over repeated interactions and longer-term use of AI systems. The scope of the task could also be broadened by considering more open-ended conversational systems, multi-day planning scenarios, or contexts involving real-time data integration. Another key consideration concerns generalizability: the present findings are based on a relatively young and digitally literate sample. It remains unclear whether similar patterns would be observed among other population groups, such as older users, individuals with lower digital literacy, or users with limited prior experience with AI-assisted travel planning. Studies integrating hybrid interaction designs that combine structured interfaces with conversational input could explore how such designs shape users’ trust and sense of control. As AI applications continue to expand beyond travel planning, future studies in other sectors could also explore how different interaction modalities affect users’ trust and sense of control as well as other aspects such as enjoyment, perceived safety, or doubt.

8 CONCLUSION

This thesis aimed to answer how different interaction modalities influence users' trust and sense of control during the early stages of AI-assisted travel planning. To achieve this, an empirical study following a Design Science Research Methodology was conducted.
A graphical user interface (GUI) and a conversational Chatbot that generated single-day travel itineraries were designed and developed. To ensure that any observed differences in trust or sense of control were attributed to interaction modality, both interfaces relied on a shared backend.

The results showed that baseline levels of trust and sense of control increased after interacting with each modality, and that similar scores were achieved for each construct. However, these quantitative results alone were not sufficient to explain why these outcomes were achieved. The qualitative analysis provided insight into differences that could not be captured by numerical ratings alone. The Chatbot emphasized personalisation, freedom of expression, and dialogue, whereas the GUI emphasized structure, speed, clarity, and guided progression through the task.
These similar score levels, combined with different underlying reasons, highlight the importance of combining quantitative and qualitative methods when trying to understand the design implications of developing AI travel systems for users.

Designers of AI travel planning systems should design interaction styles based on user needs and task context, rather than assuming that one type of interface is inherently better. As AI systems are increasingly used for travel planning, understanding how interaction design shapes user experience becomes more important. Trust and sense of control are not only outcomes of AI capability or interaction modality, but also outcomes of how users are guided, constrained, or enabled during an interaction.
Failing to consider these aspects may increase the risk of creating systems that are technically powerful but poorly aligned with users' expectations and preferences. Therefore, addressing these design considerations is essential for developing AI systems that are not only functional but also usable, trustworthy, and appropriate for their intended users.

9 A NOTE ON AI USE

Artificial intelligence (AI) tools were used as productivity tools. Their use was limited to technical assistance in software development, interface prototyping, writing-related tasks, transcription of qualitative data, and reviewing the literature.

For software development and interface prototyping, AI was used to translate design intentions into code, refine user interface layouts, and resolve implementation issues or bugs. All design decisions, architectural choices, and experimental constraints were defined, evaluated, and validated by the researcher.

For writing-related tasks, AI was used to check the spelling and grammar, improve writing consistency, and brainstorm different writing ideas. All final texts were written, reviewed, and approved by the researcher.

For the qualitative component, AI was used to generate the initial transcripts from the recorded interviews. All transcripts were reviewed by the researcher and corrected prior to the analysis. No participant-identifiable information was shared with third-party AI services; all data used in these tools were anonymized, ensuring participant confidentiality and anonymity.

For the literature review, AI tools were used to support broad searches, refine keyword combinations and search strings, identify potentially relevant peer-reviewed articles, and organize preliminary notes. All final studies were analysed by the researcher for relevance and alignment with the research objective.

REFERENCES

Ali, F., Yasar, B., Ali, L., & Dogan, S. (2023). Antecedents and consequences of travelers' trust towards personalized travel recommendations offered by ChatGPT. *International Journal of Hospitality Management*, *114*, 103588. https://doi.org/10.1016/j.ijhm.2023.103588

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2. ed., reprint). Psychology Press.

Dwivedi, Y. K., Pandey, N., Currie, W., & Micu, A. (2024). Leveraging ChatGPT and other generative artificial intelligence (AI)-based applications in the hospitality and tourism industry: Practices, challenges and research agenda. *International Journal of Contemporary Hospitality Management*, *36*(1), 1–12. https://doi.org/10.1108/IJCHM-05-2023-0686

Ernst & Young. (2025, March 18). *How generative AI is transforming the tourism industry*. https://www.ey.com/en_jp/insights/consulting/how-generative-ai-is-transforming-the-tourism-industry

Fakfare, P., Manosuthi, N., Lee, J.-S., Han, H., & Jin, M. (2025). Customer word-of-mouth for generative AI: Innovation and adoption in hospitality and tourism. *International Journal of Hospitality Management*, *126*, 104070. https://doi.org/10.1016/j.ijhm.2024.104070

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Grundner, L., & Neuhofer, B. (2021). The bright and dark sides of artificial intelligence: A futures perspective on tourist destination experiences. *Journal of Destination Marketing & Management*, *19*, 100511. https://doi.org/10.1016/j.jdmm.2020.100511

Gursoy, D., Li, Y., & Song, H. (2023). ChatGPT and the hospitality and tourism industry: An overview of current trends and future research directions. *Journal of Hospitality Marketing & Management*, *32*(5), 579–592. https://doi.org/10.1080/19368623.2023.2211993

Han, H., Kim, S. (Sam), Hailu, T. B., Al-Ansi, A., Loureiro, S. M. C., & Kim, J. J. (2025). Determinants of approach behavior for ChatGPT and their configurational influence in the hospitality and tourism sector: A cumulative prospect theory. *International Journal of Contemporary Hospitality Management*, *37*(1), 113–139. https://doi.org/10.1108/IJCHM-07-2023-1072

Huang, G. I., Wong, I. A., Zhang, C. J., & Liang, Q. (2025). Generative AI inspiration and hotel recommendation acceptance: Does anxiety over lack of transparency matter? *International Journal of Hospitality Management*, *126*, 104112. https://doi.org/10.1016/j.ijhm.2025.104112

Kim, J. H., Kim, J., Kim, C., & Kim, S. (Sam). (2023). Do you trust ChatGPTs? Effects of the ethical and quality issues of generative AI on travel decisions. *Journal of Travel & Tourism Marketing*, *40*(9), 779–801. https://doi.org/10.1080/10548408.2023.2293006

Kim, J. H., Kim, J., Park, J., Kim, C., Jhang, J., & King, B. (2025). When ChatGPT gives incorrect answers: The impact of inaccurate information by generative AI on tourism decision-making. *Journal of Travel Research*, *64*(1), 51–73. https://doi.org/10.1177/00472875231212996

Li, H., Xi, J., Hsu, C. H. C., Yu, B. X. B., & Zheng, X. (Kevin). (2025). Generative artificial intelligence in tourism management: An integrative review and roadmap for future research. *Tourism Management*, *110*, 105179. https://doi.org/10.1016/j.tourman.2025.105179

Loureiro, S. M. C., Guerreiro, J., Friedmann, E., Lee, M. J., & Han, H. (2025). Tourists and artificial intelligence-LLM interaction: The power of forgiveness. *Current Issues in Tourism*, *28*(7), 1172–1190. https://doi.org/10.1080/13683500.2024.2353872

Ma, X., & Huo, Y. (2023). Are users willing to embrace ChatGPT? Exploring the factors on the acceptance of chatbots from the perspective of AIDUA framework. *Technology in Society*, *75*, 102362. https://doi.org/10.1016/j.techsoc.2023.102362

McKnight, D. H., & Chervany, N. L. (2001). What Trust Means in E-Commerce Customer Relationships: An Interdisciplinary Conceptual Typology. *International Journal of Electronic Commerce*, *6*(2), 35–59. https://doi.org/10.1080/10864415.2001.11044235

Melián-González, S., Gutiérrez-Taño, D., & Bulchand-Gidumal, J. (2021). Predicting the intentions to use chatbots for travel and tourism. *Current Issues in Tourism*, *24*(2), 192–210. https://doi.org/10.1080/13683500.2019.1706457

Norman, D. A. (2013). *The design of everyday things* (Rev. and expanded edition). MIT Press.

OpenAI. (2025). *OpenAI API documentation*. https://platform.openai.com/docs/overview

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems*, *24*(3), 45–77. https://doi.org/10.2753/MIS0742-1222240302

Pham, H. C., Duong, C. D., & Nguyen, G. K. H. (2024). What drives tourists' continuance intention to use ChatGPT for travel services? A stimulus-organism-response perspective. *Journal of Retailing and Consumer Services*, *78*, 103758. https://doi.org/10.1016/j.jretconser.2024.103758

Pillai, R., & Sivathanu, B. (2020). Adoption of AI-based chatbots for hospitality and tourism. *International Journal of Contemporary Hospitality Management*, *32*(10), 3199–3226. https://doi.org/10.1108/IJCHM-04-2020-0259

Preston, C. C., & Colman, A. M. (2000). Optimal number of response categories in rating scales: Reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica*, *104*(1), 1–15. https://doi.org/10.1016/S0001-6918(99)00050-5

Seyfi, S., Kim, M. J., Nazifi, A., Murdy, S., & Vo-Thanh, T. (2025). Understanding tourist barriers and personality influences in embracing generative AI for travel planning and decision-making. *International Journal of Hospitality Management*, *126*, 104105. https://doi.org/10.1016/j.ijhm.2025.104105

Statista. (2025). *Artificial intelligence: In-depth market analysis*. https://www.statista.com/outlook/tmo/artificial-intelligence/worldwide

Stergiou, D. P., & Nella, A. (2024). ChatGPT and Tourist Decision-Making: An Accessibility–Diagnosticity Theory Perspective. *International Journal of Tourism Research*, *26*(5), e2757. https://doi.org/10.1002/jtr.2757

Wobbrock, J. O. (2015). *Catchy Titles Are Good: But Avoid Being Cute*.

Wobbrock, J. O. (2012). *Seven Research Contributions in HCI*.

Wong, I. A., Lian, Q. L., & Sun, D. (2023). Autonomous travel decision-making: An early glimpse into ChatGPT and generative AI. *Journal of Hospitality and Tourism Management*, *56*, 253–263. https://doi.org/10.1016/j.jhtm.2023.06.022

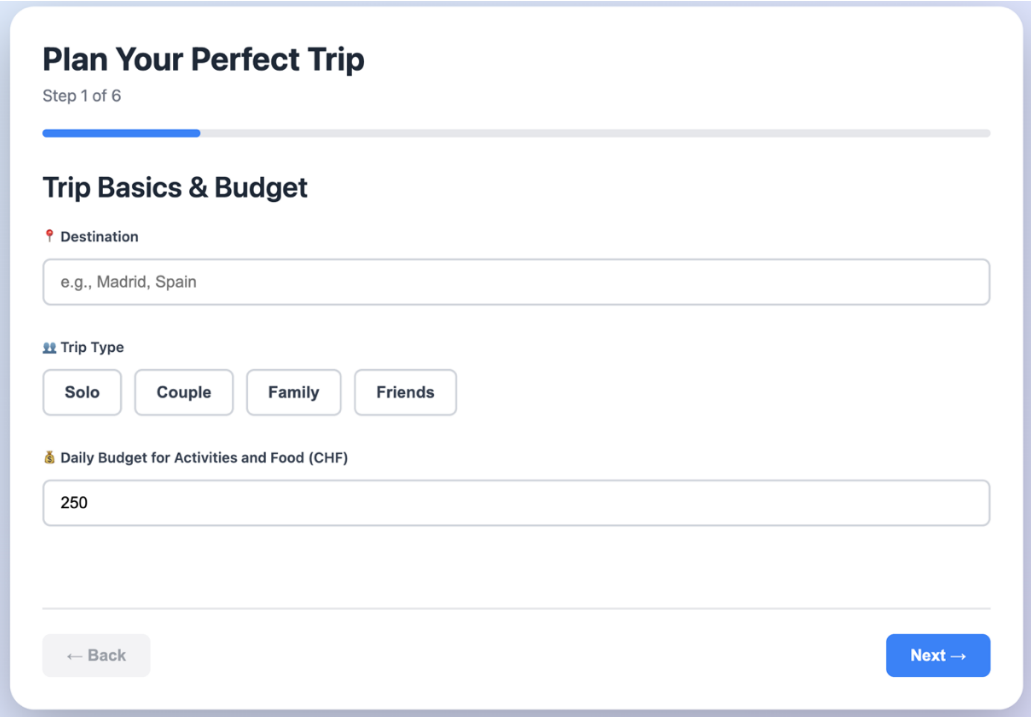
World Travel & Tourism Council. (2024). *Artificial intelligence (AI) in action: Use cases and impacts of AI in society, business and travel & tourism*.

https://researchhub.wttc.org/product/artificial-intelligence-ai-in-action-use-cases-impacts-of-ai-in-society-business-and-travel-tourism-2024

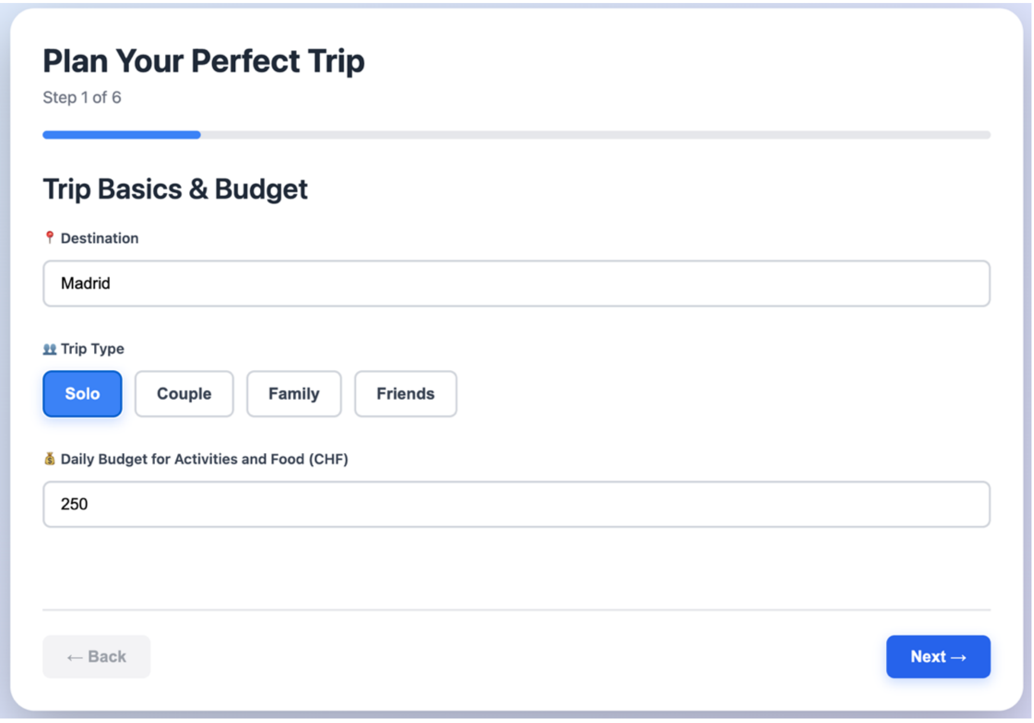
APPENDIX

Appendix I GUI interface

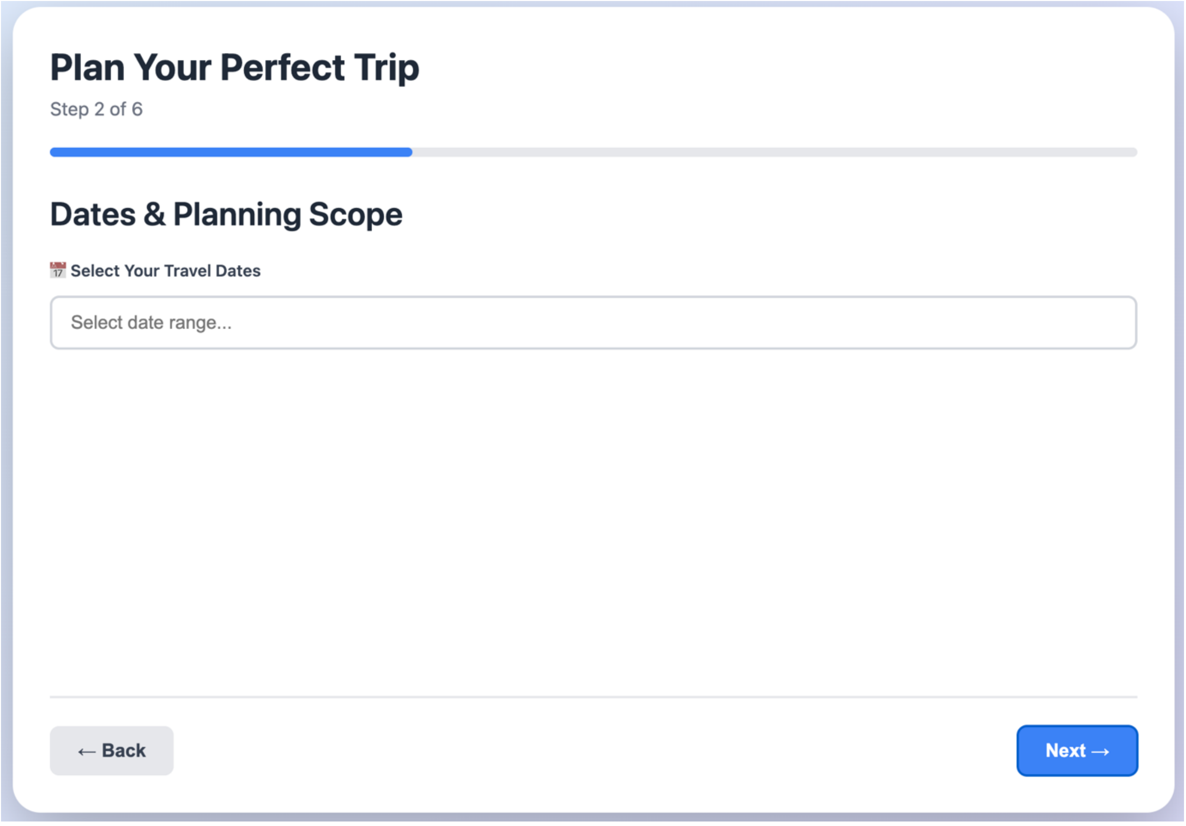
Page 1 before interaction



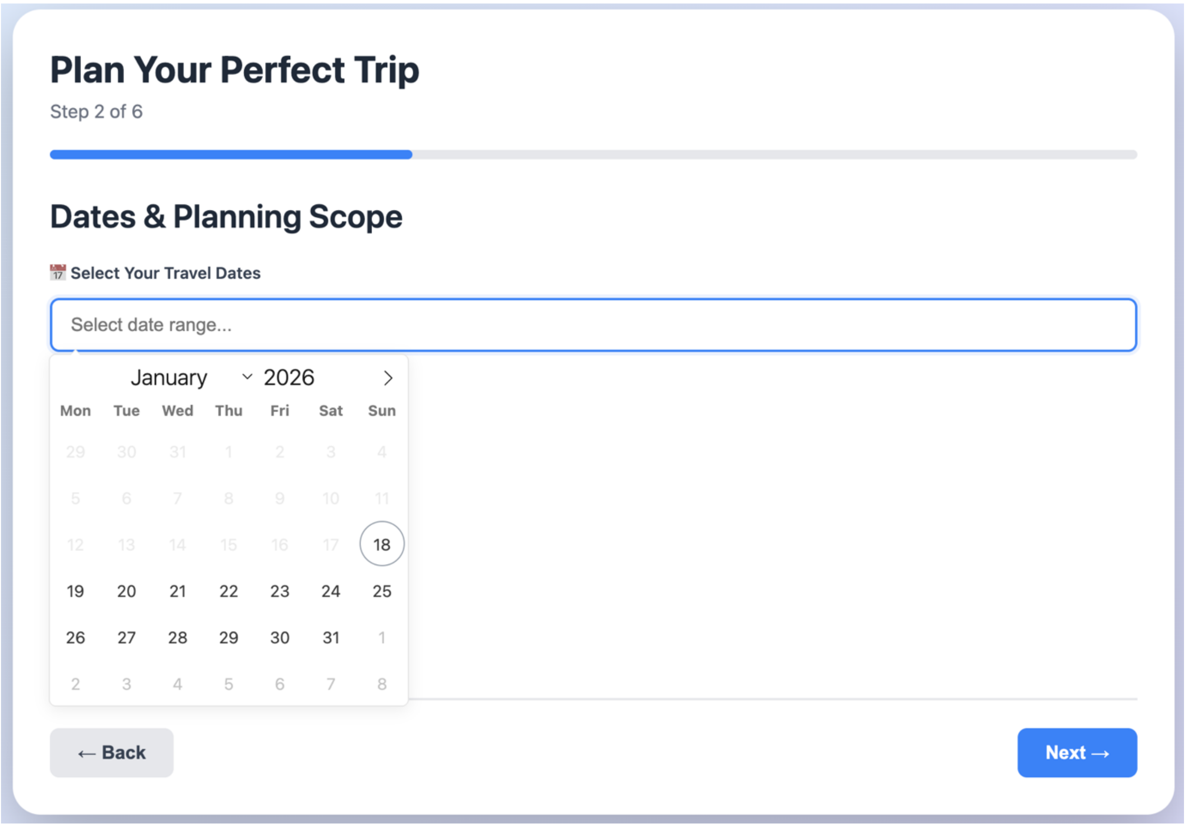
Page 1 after interaction



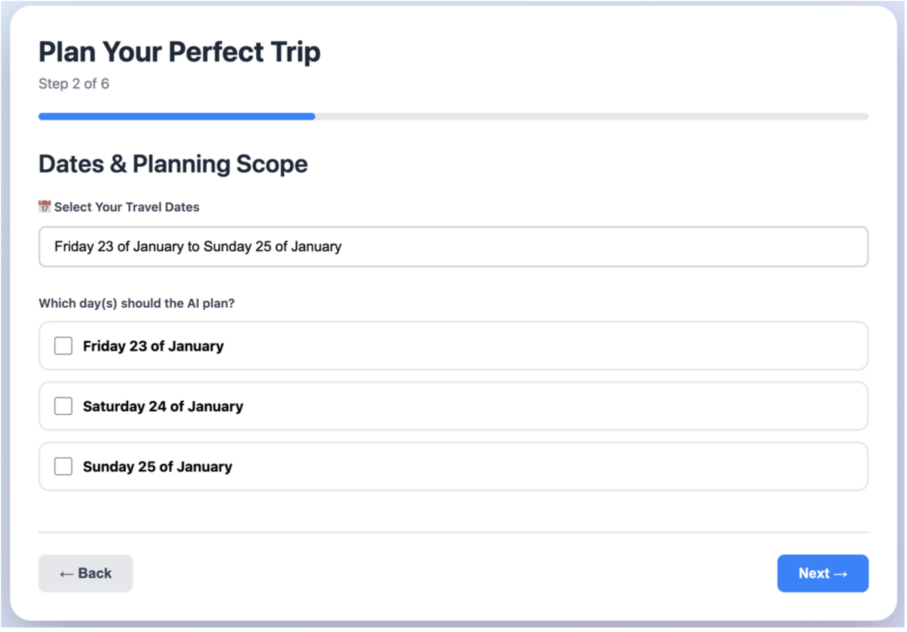
Page 2 before interaction



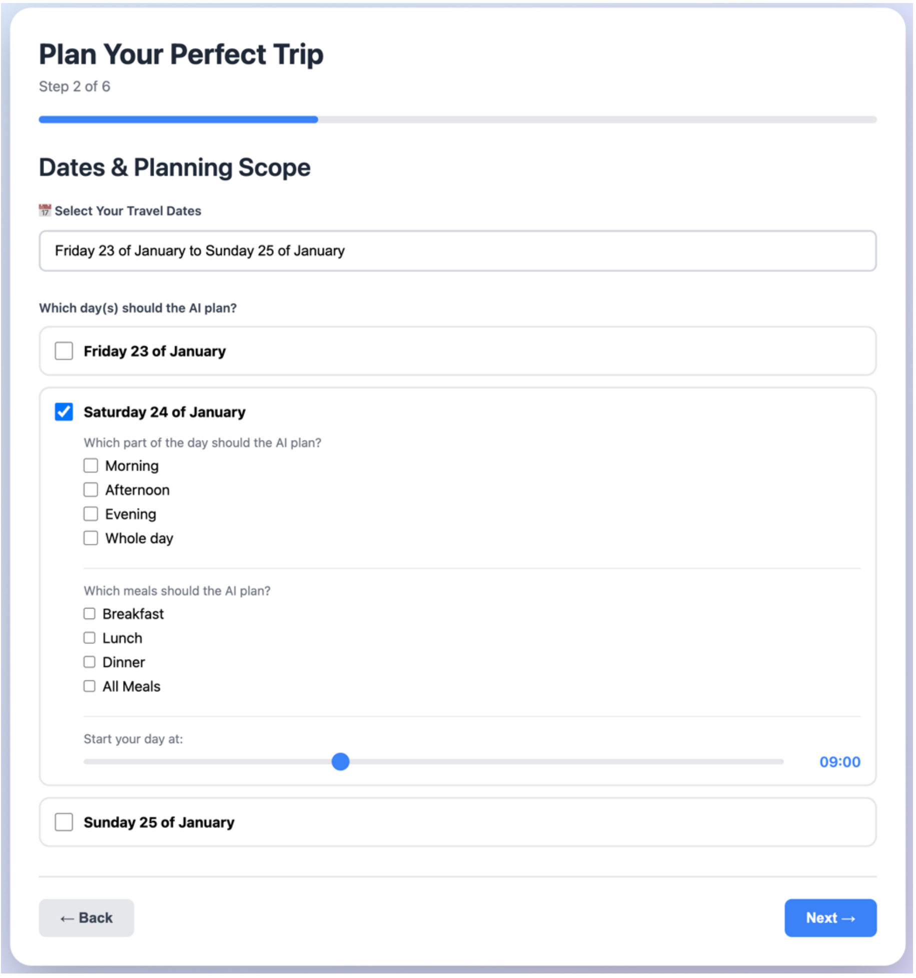
Page 2.1 after interaction



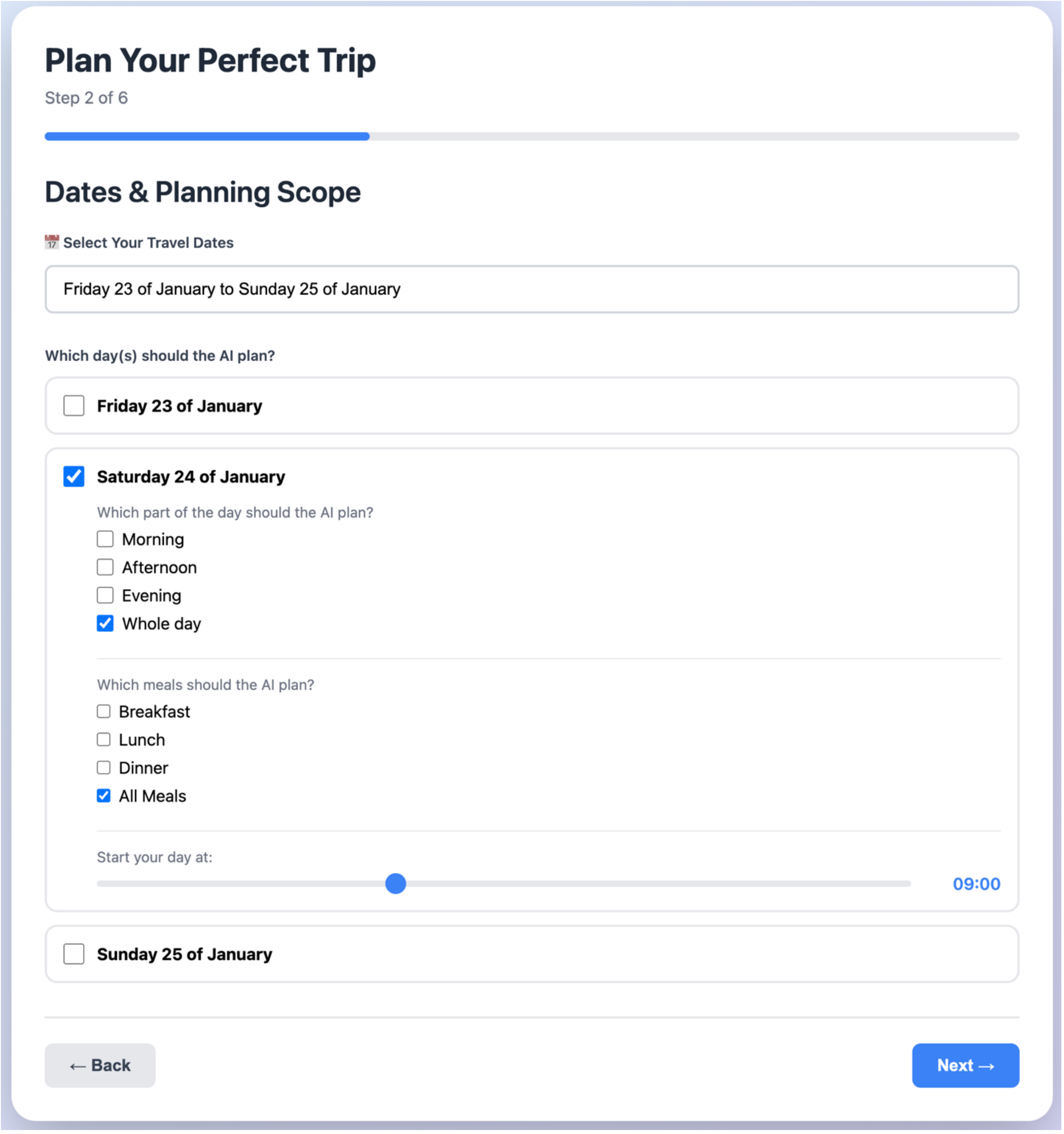
Page 2.2 after interaction



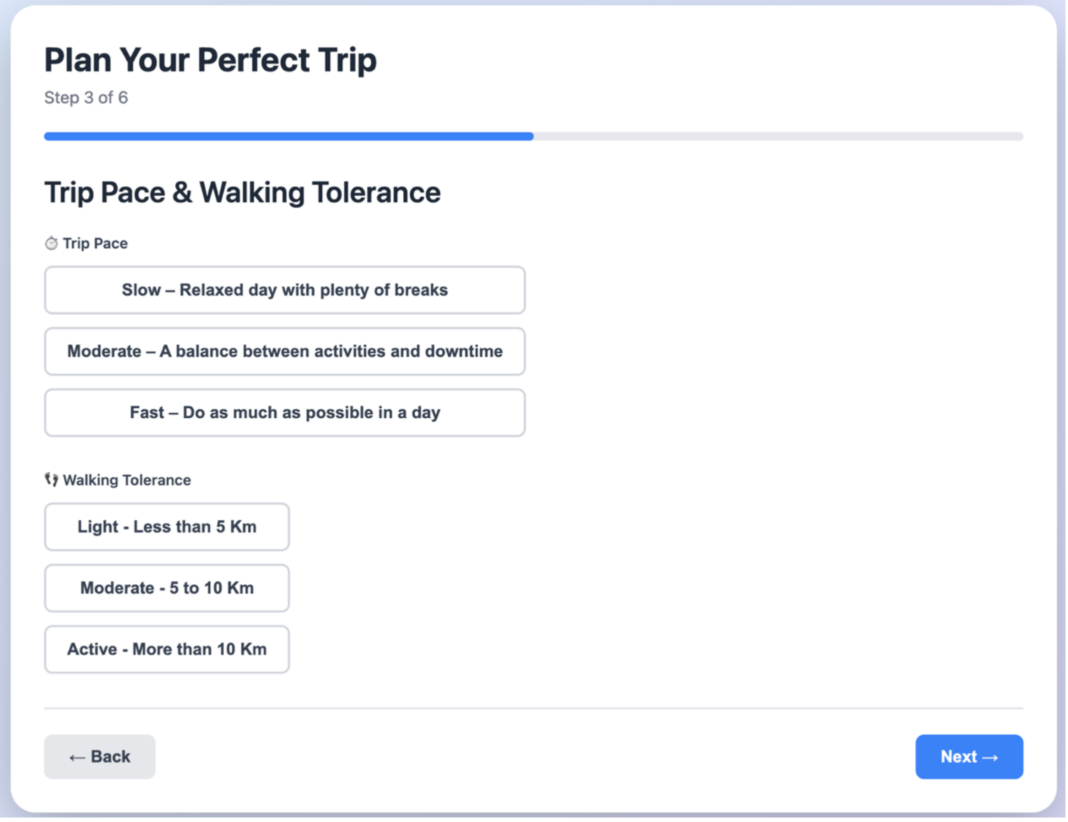
Page 2.3 after interaction



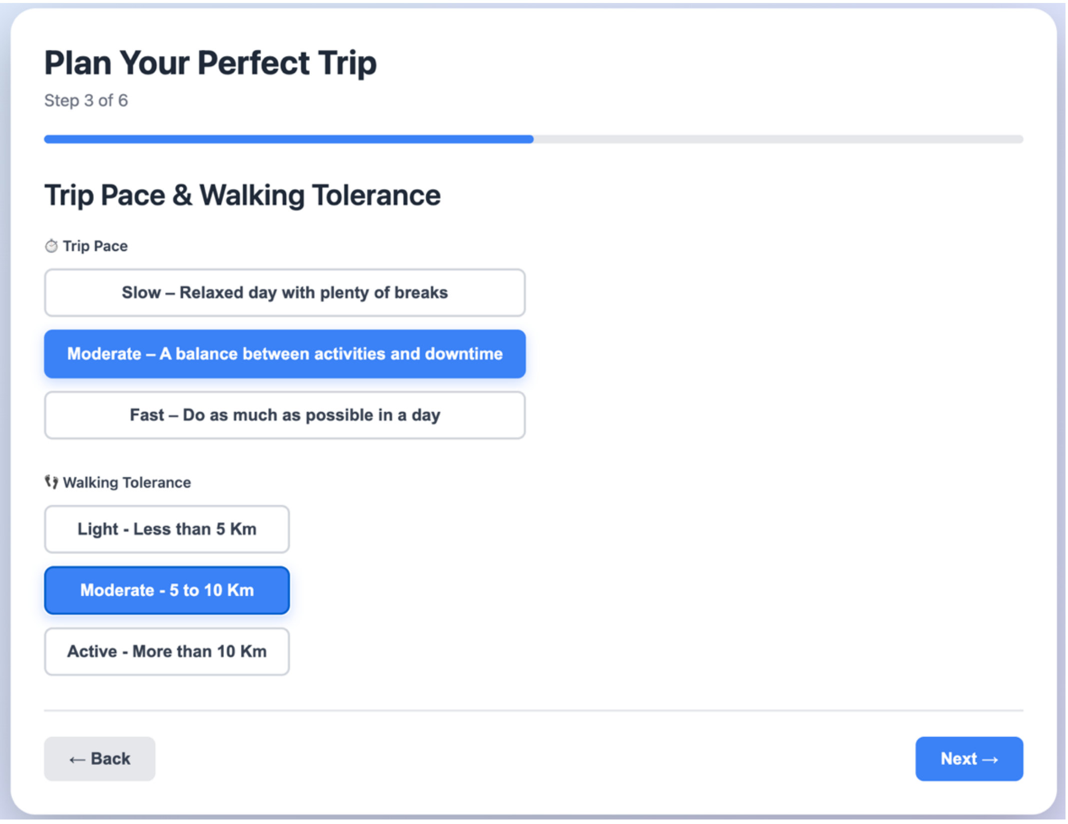
Page 2.4 after interaction



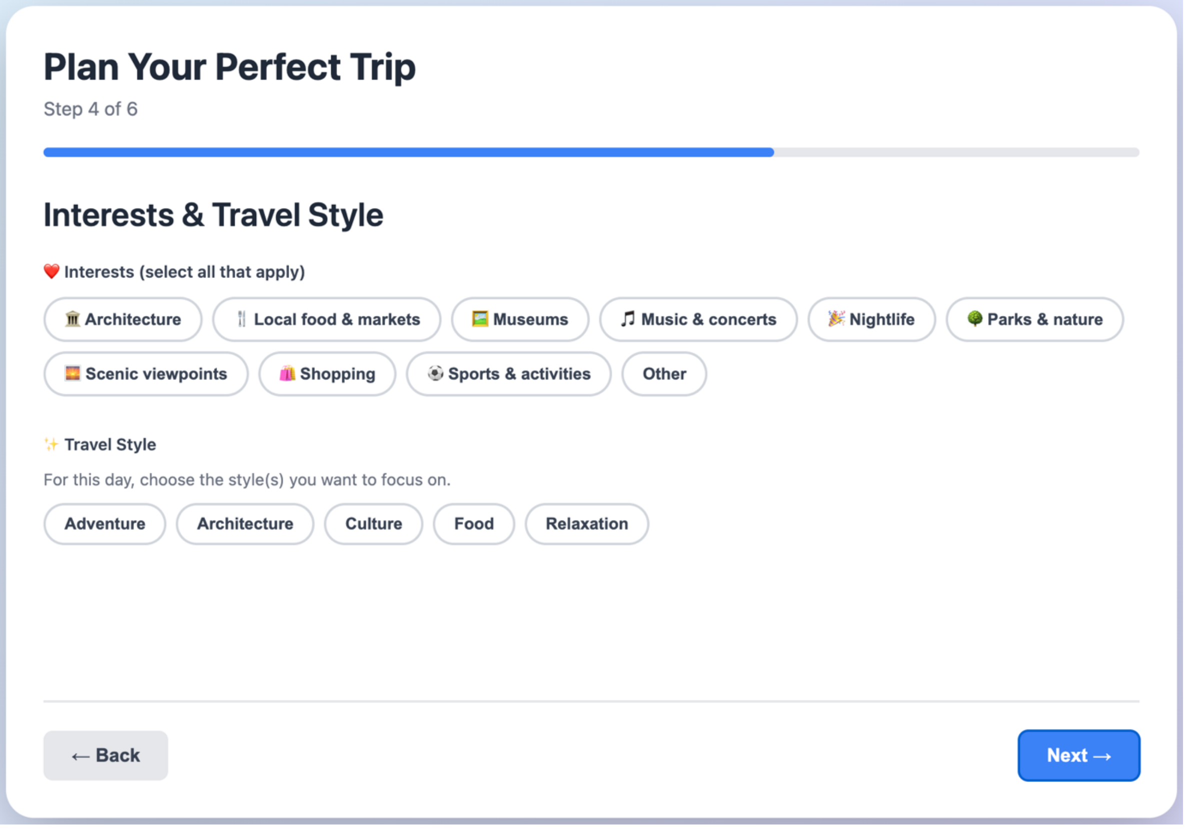
Page 3 before interaction



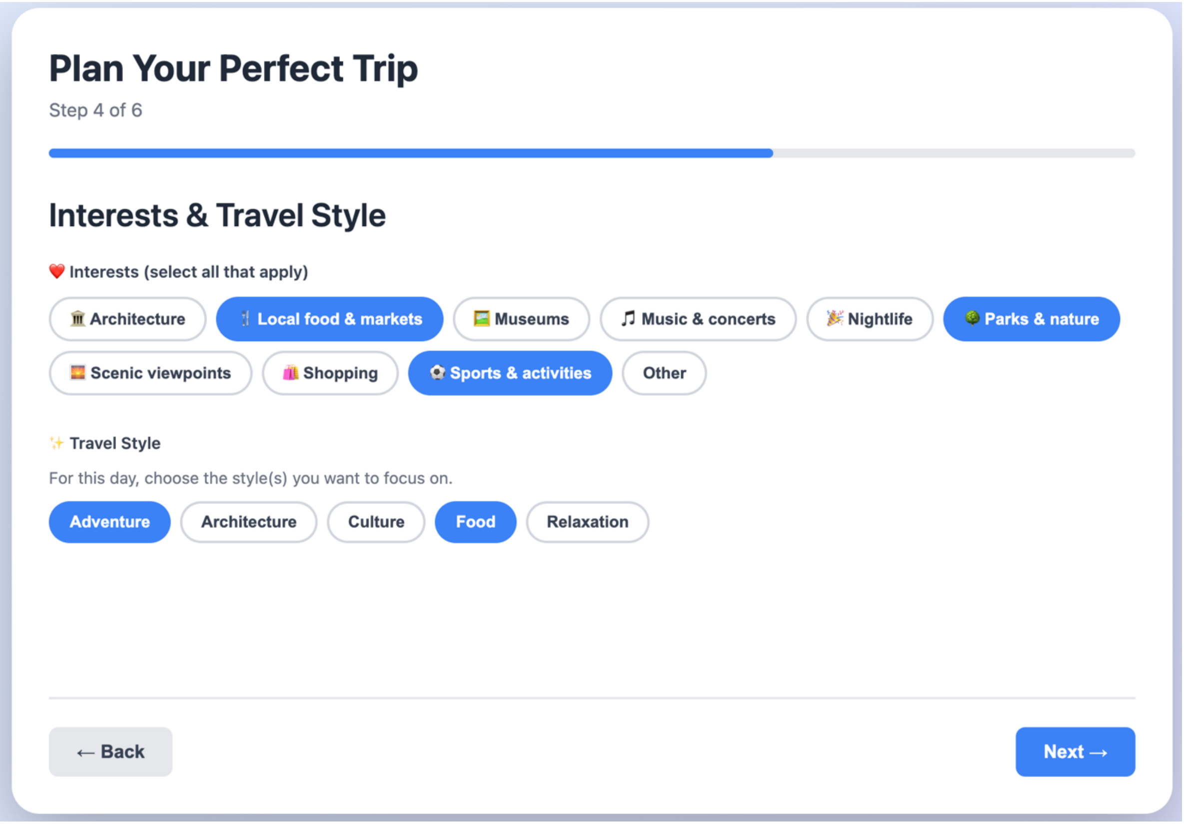
Page 3 after interaction



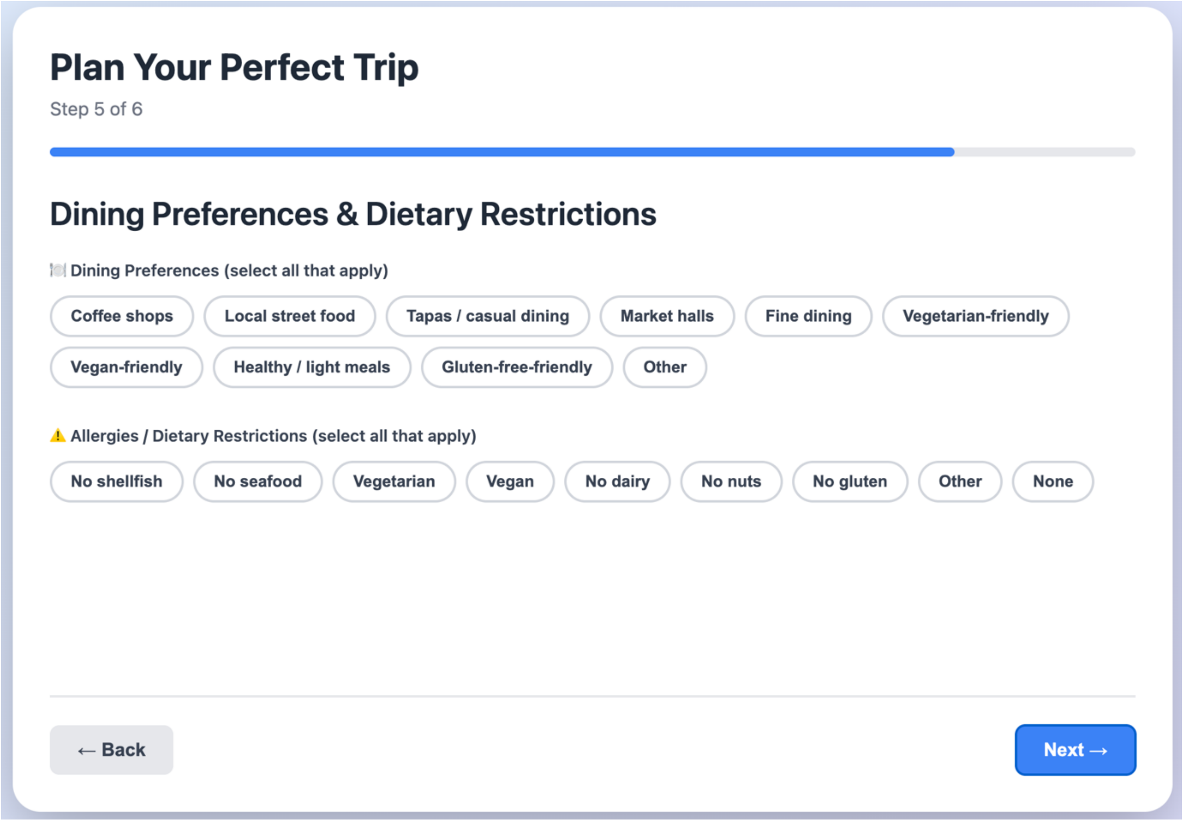
Page 4 before interaction



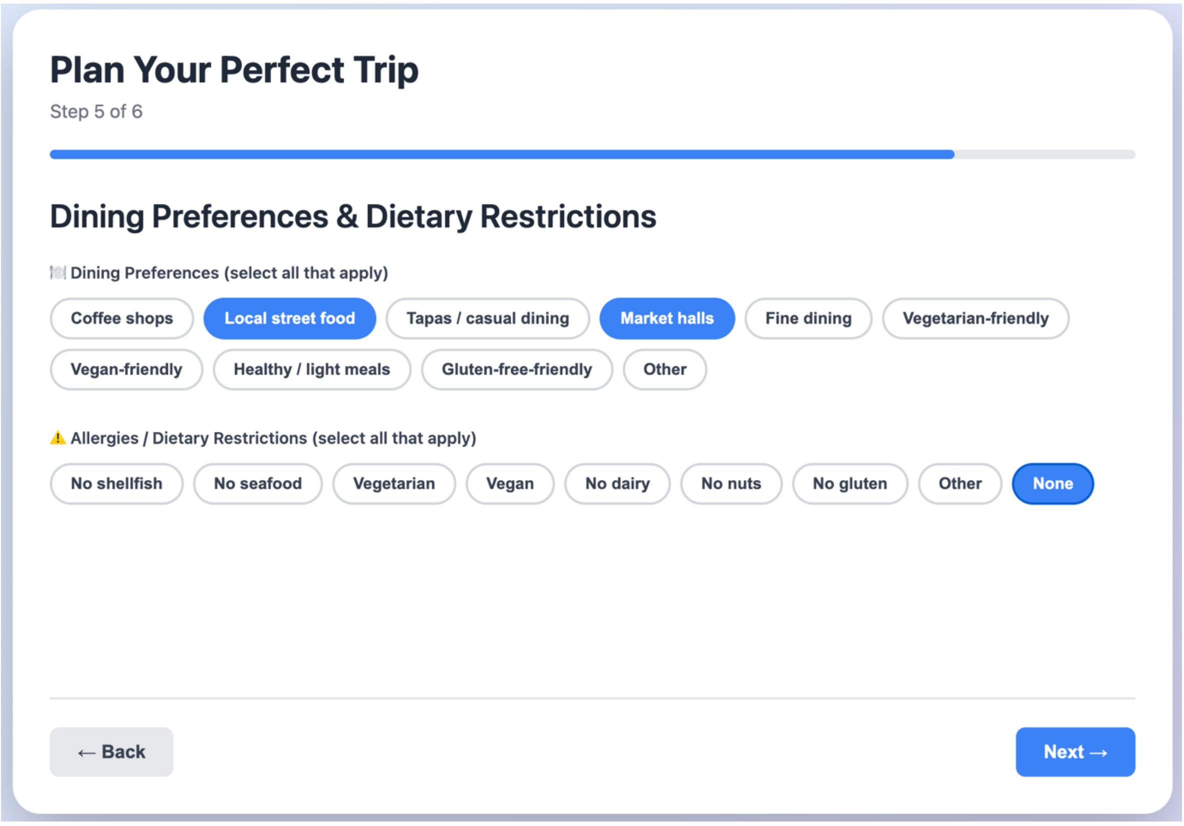
Page 4 after interaction



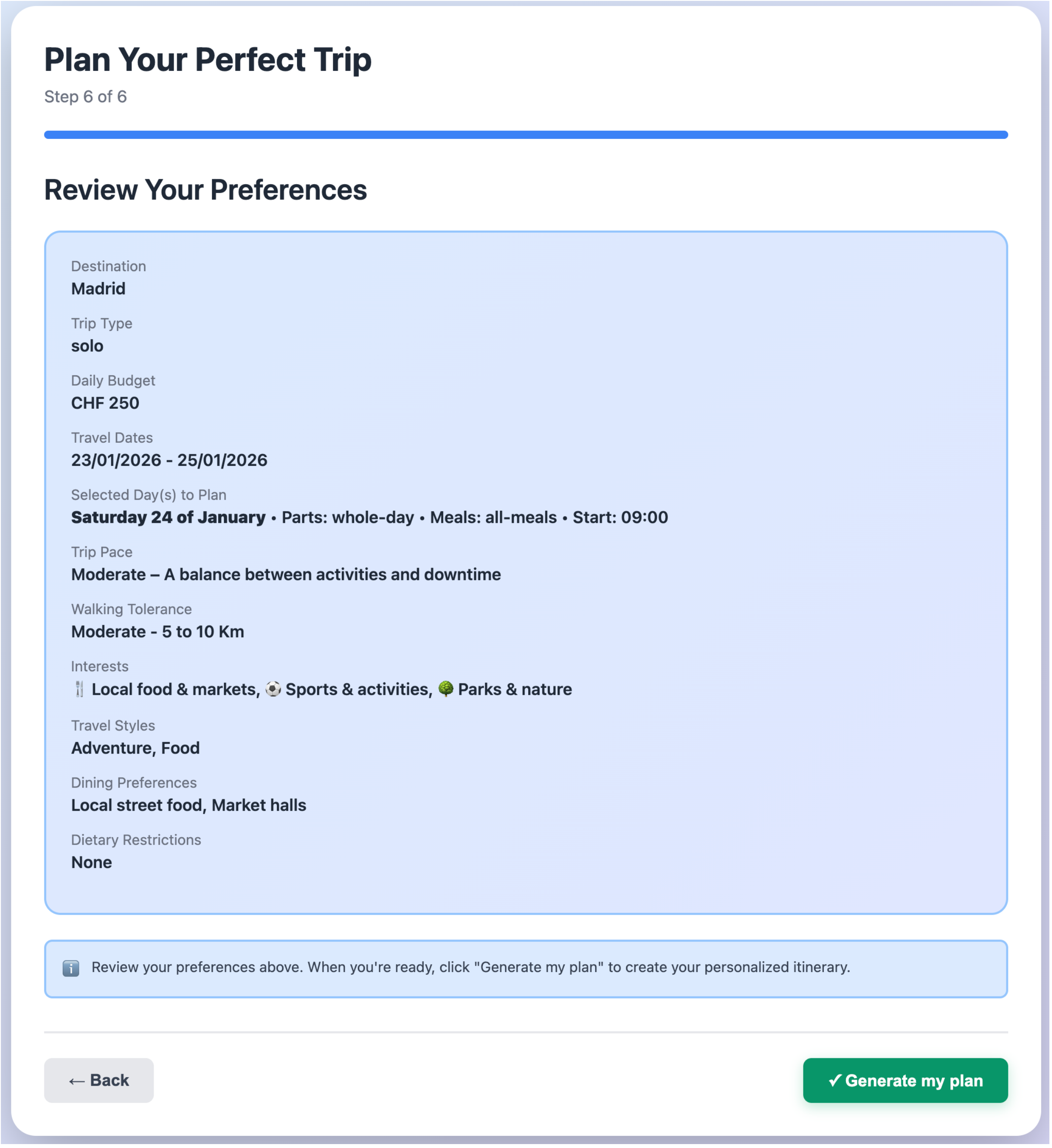
Page 5 before interaction



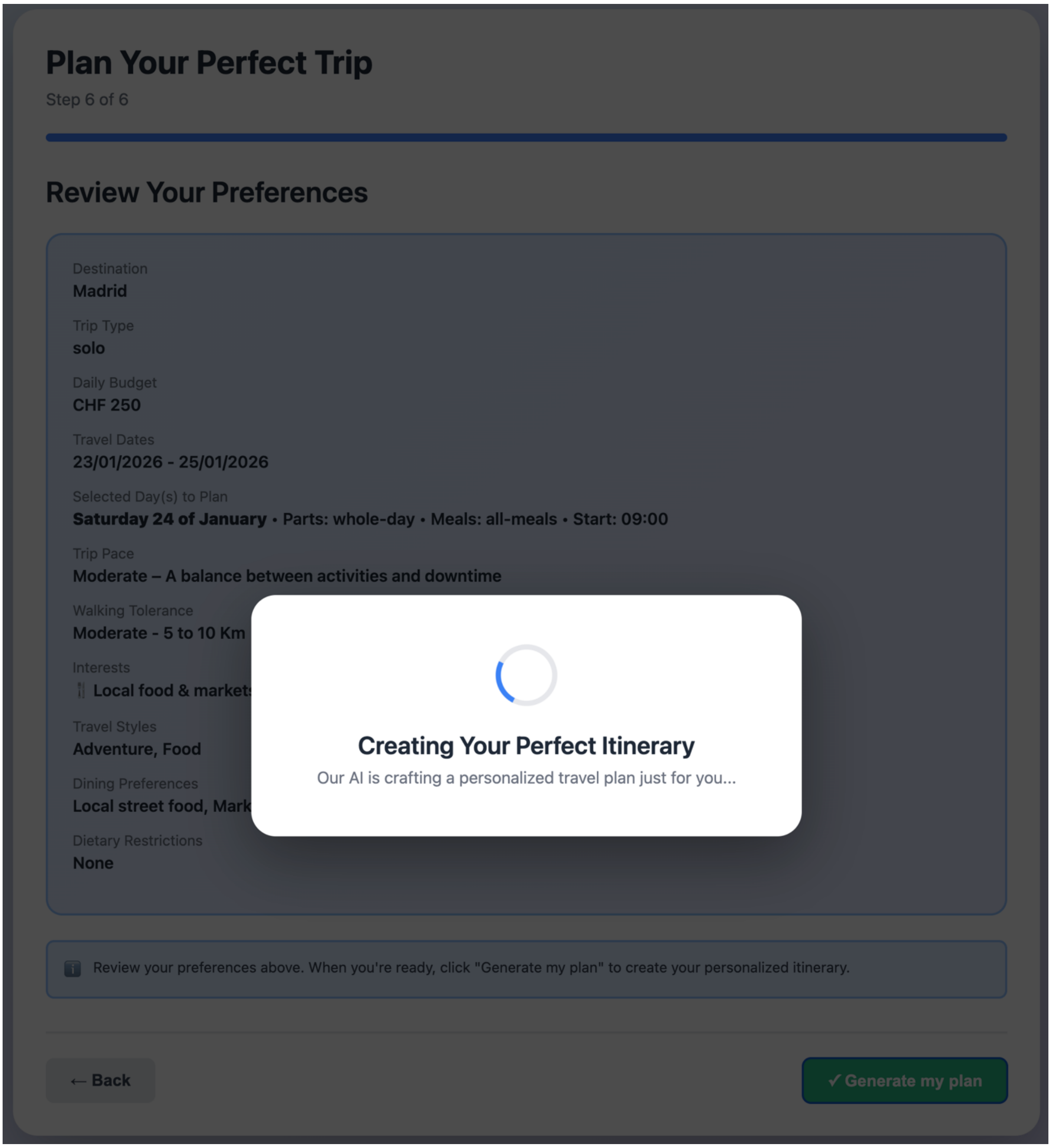
Page 5 after interaction



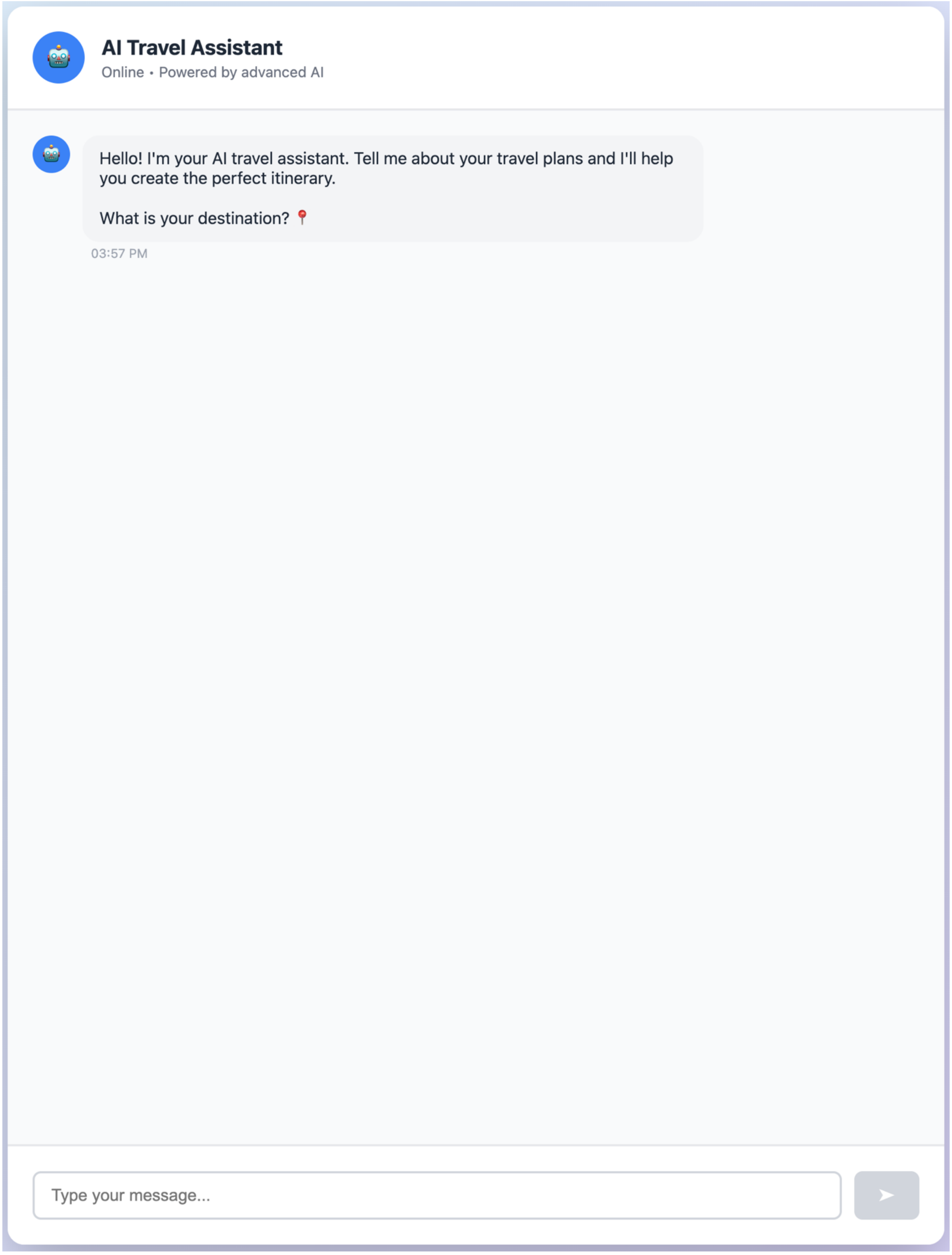
Page 6 before interaction

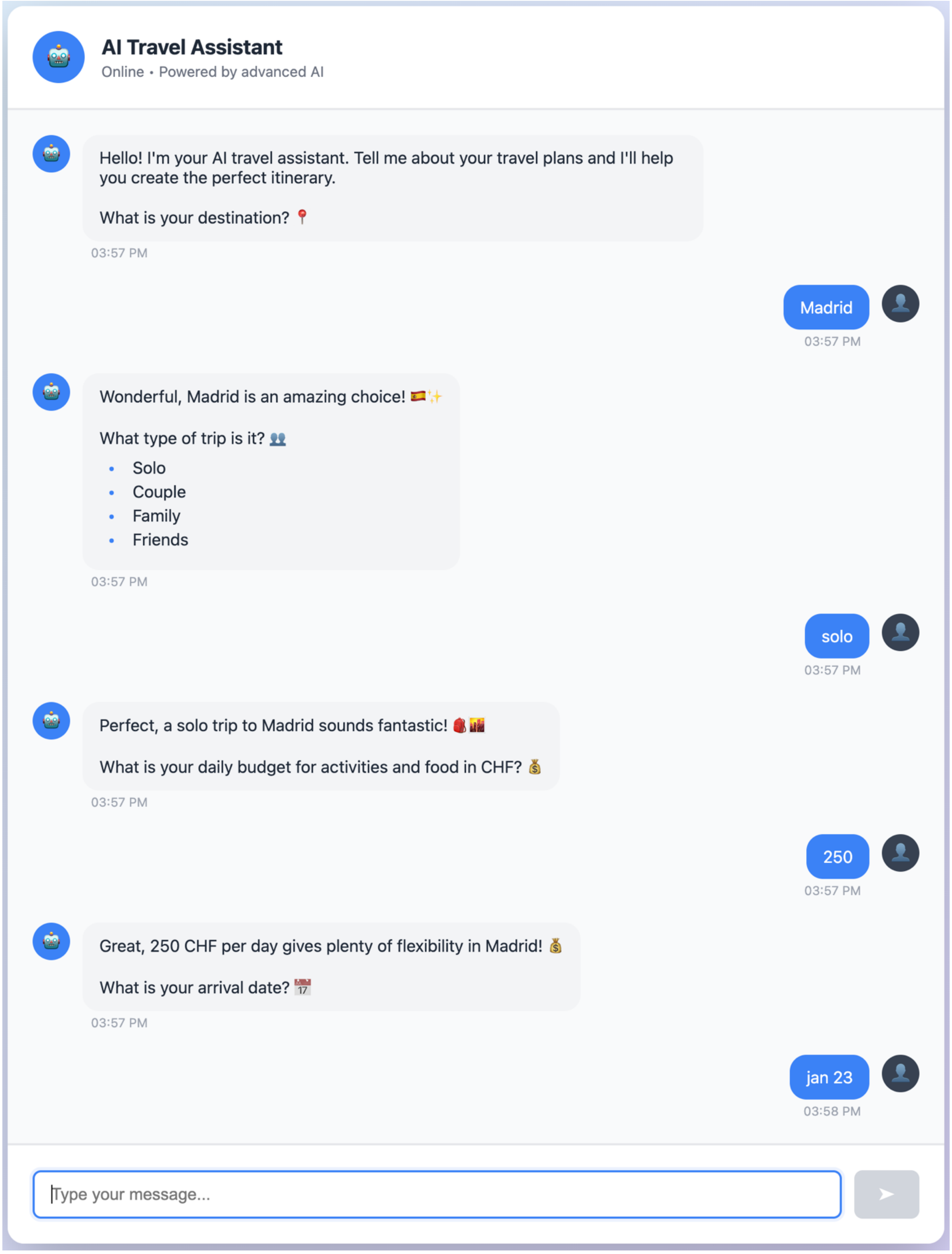


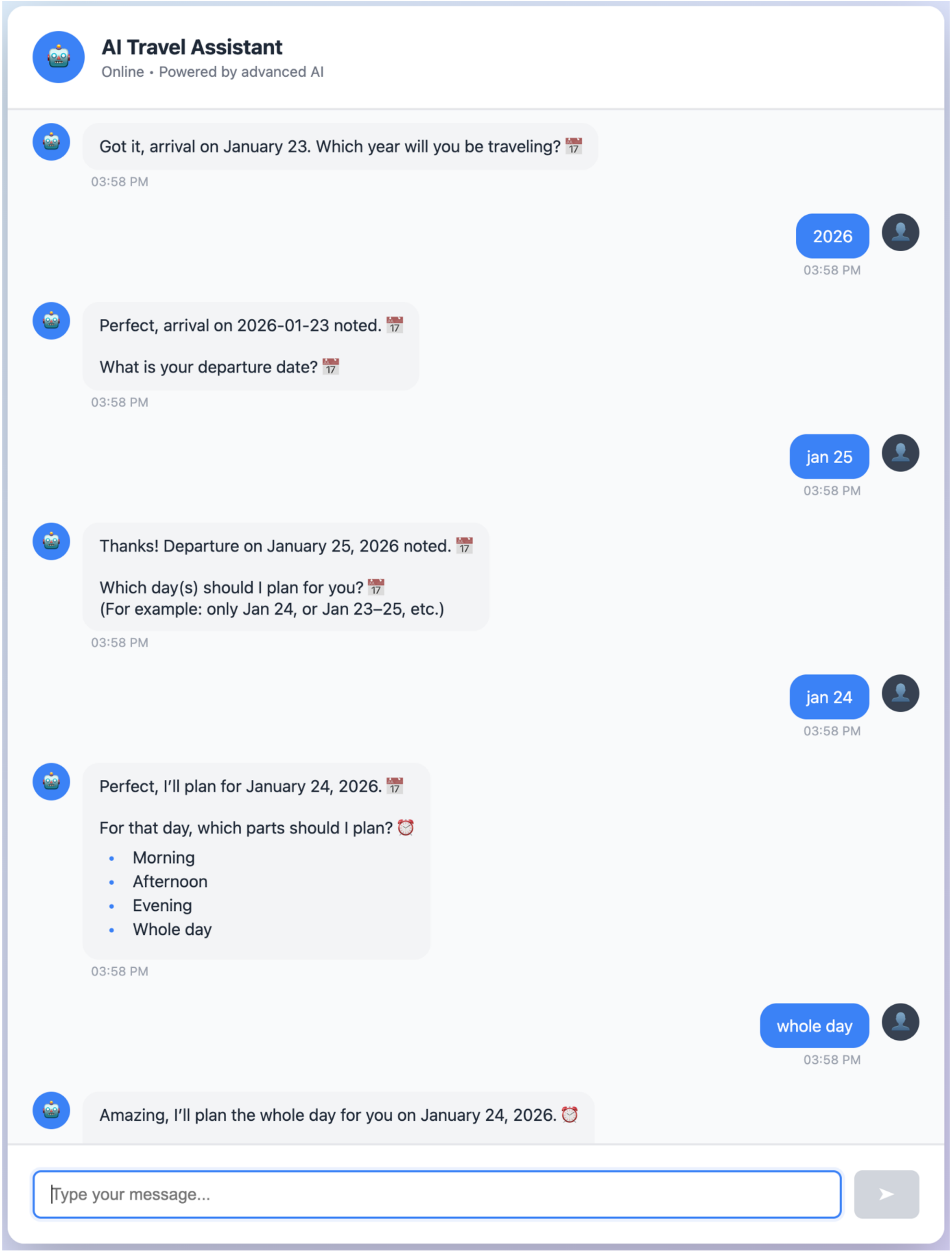
Page 6 after interaction

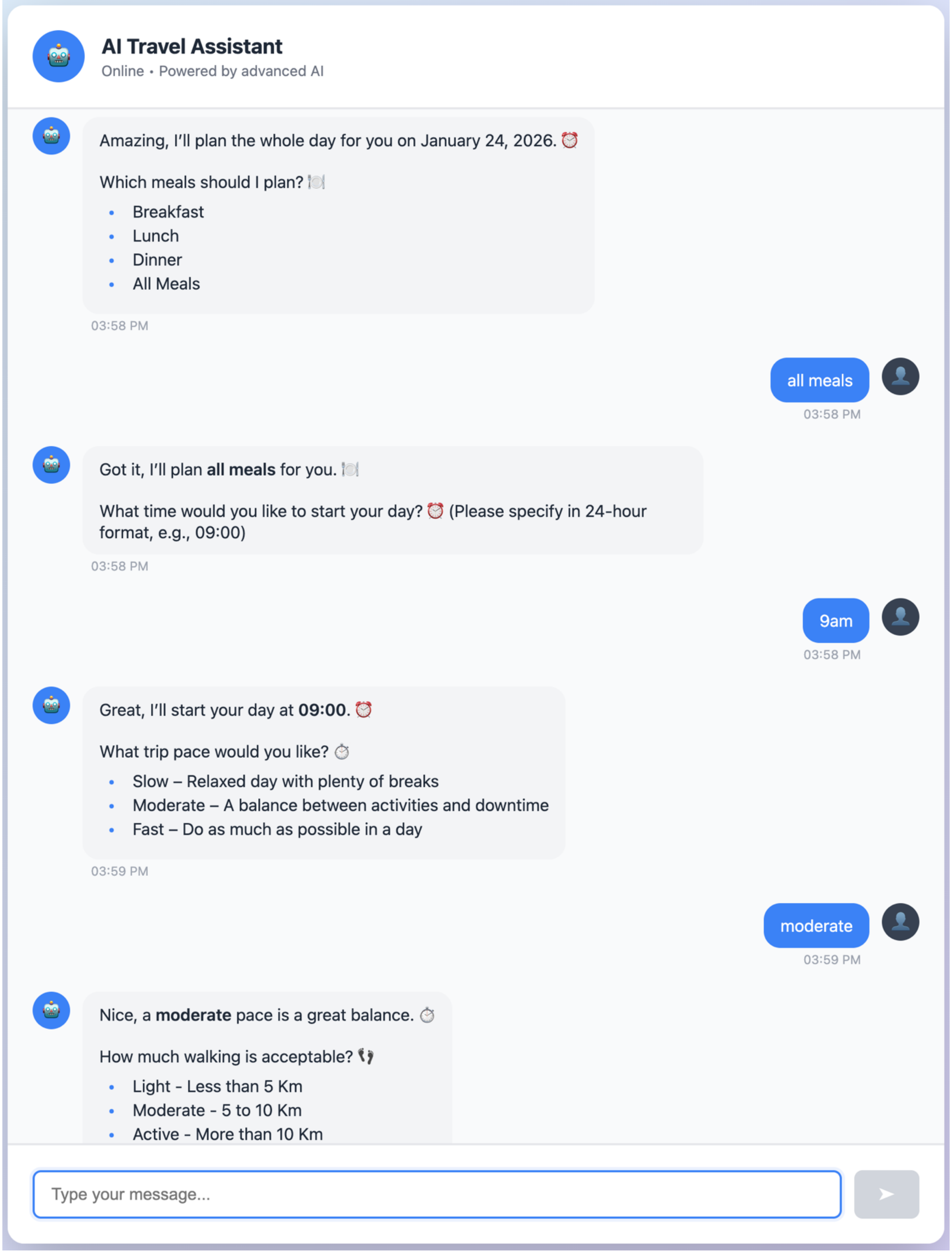


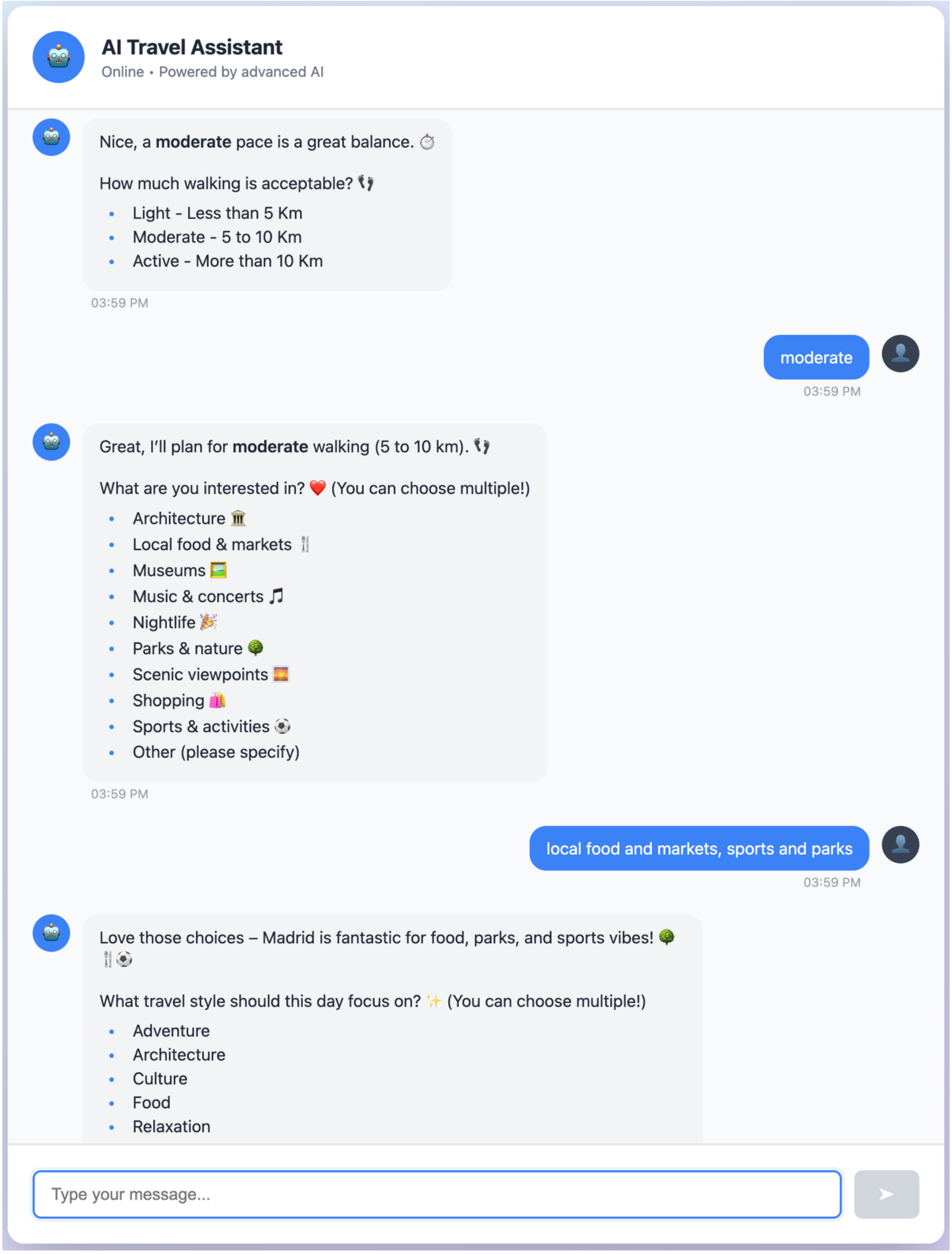
Appendix II Chatbot interface

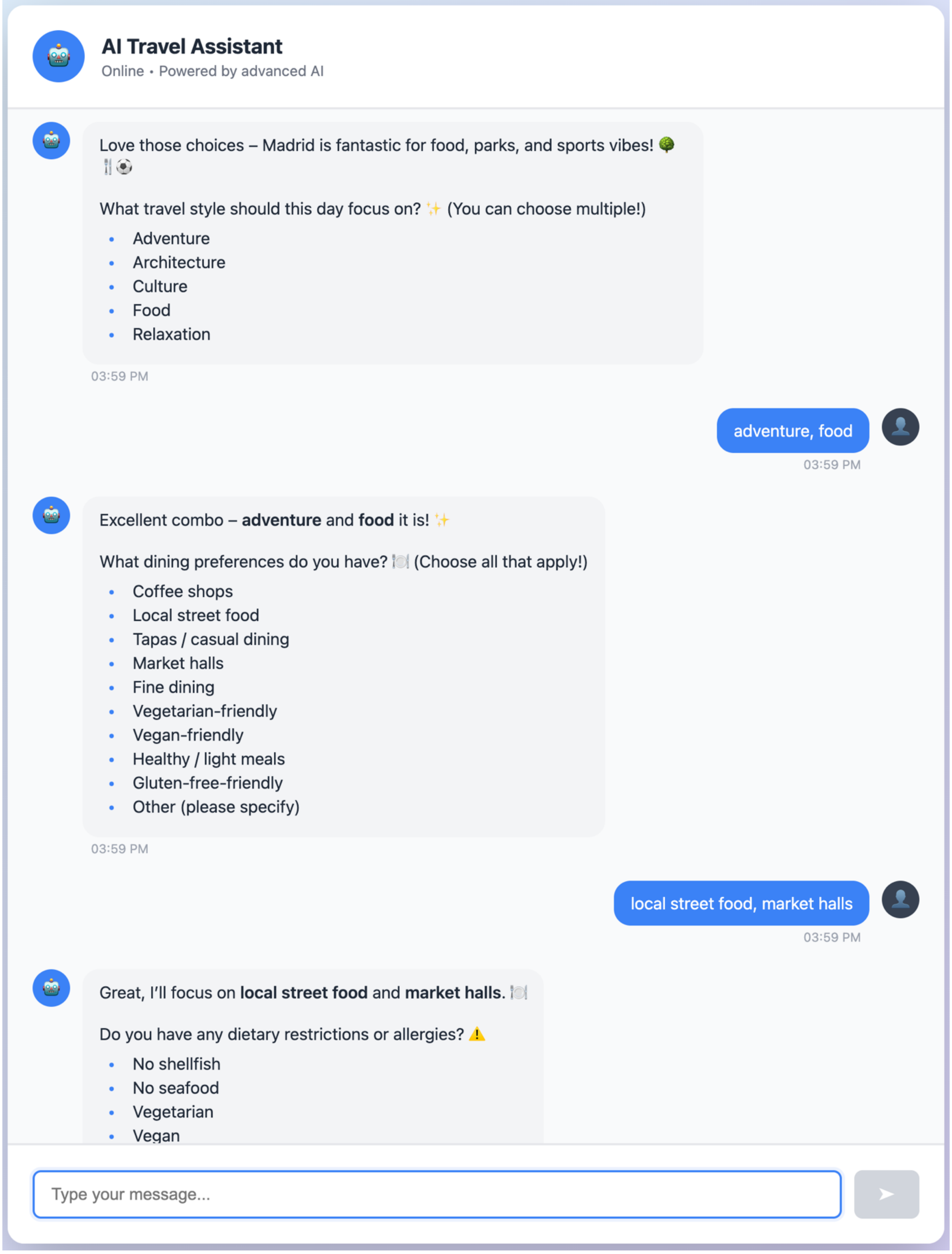


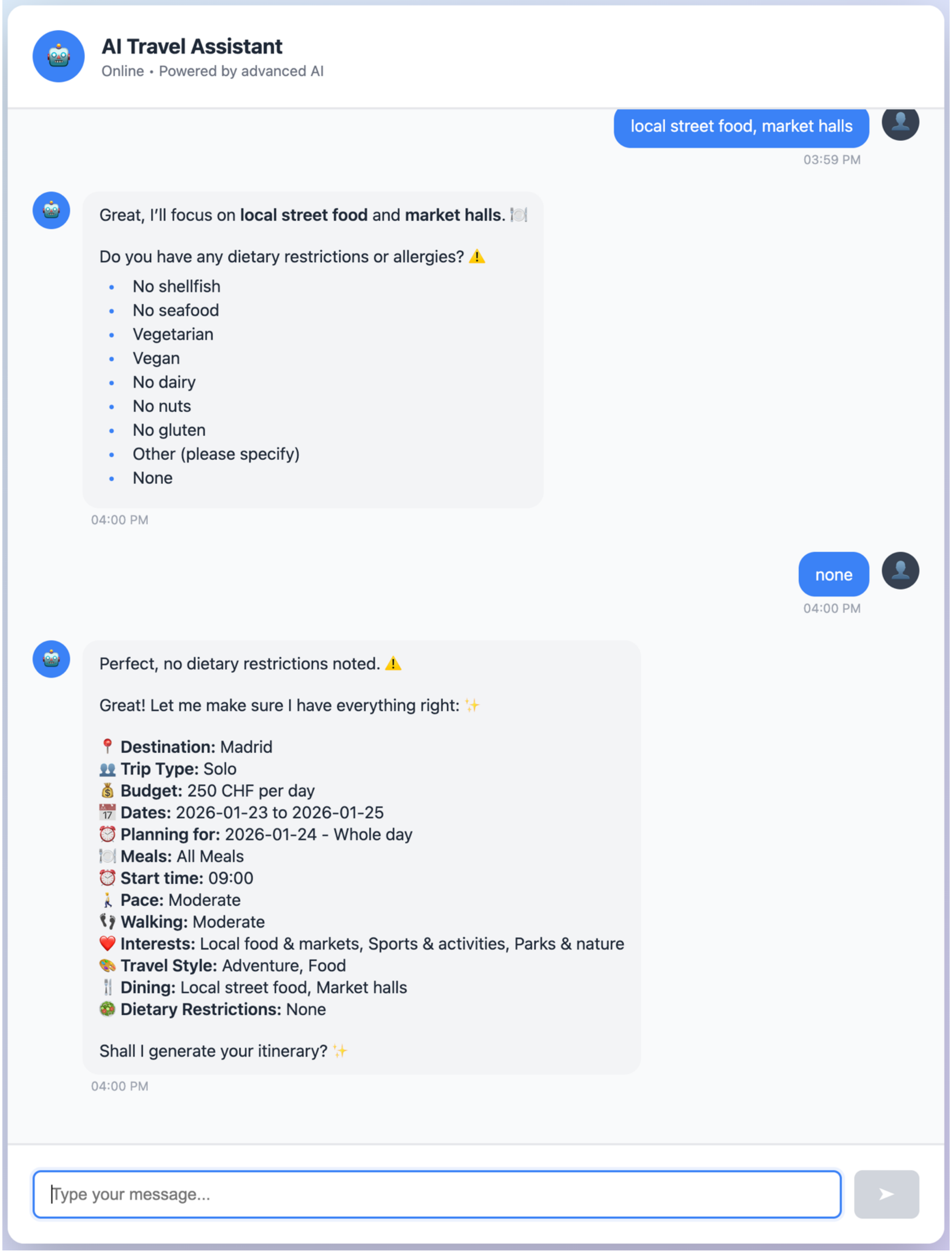

AI Travel Assistant
Online • Powered by advanced AI
Hello! I'm your AI travel assistant. Tell me about your travel plans and I'll help you create the perfect itinerary.
What is your destination? 📍
03:57 PM
Madrid
03:57 PM
Wonderful, Madrid is an amazing choice! 🇪🇸✨
What type of trip is it? 👥
Solo
Couple
Family
Friends
03:57 PM
solo
03:57 PM
Perfect, a solo trip to Madrid sounds fantastic! 🎒🌇
What is your daily budget for activities and food in CHF? 💰
03:57 PM
250
03:57 PM
Great, 250 CHF per day gives plenty of flexibility in Madrid! 💰
What is your arrival date? 📅
03:57 PM
jan 23
03:58 PM
Type your message...

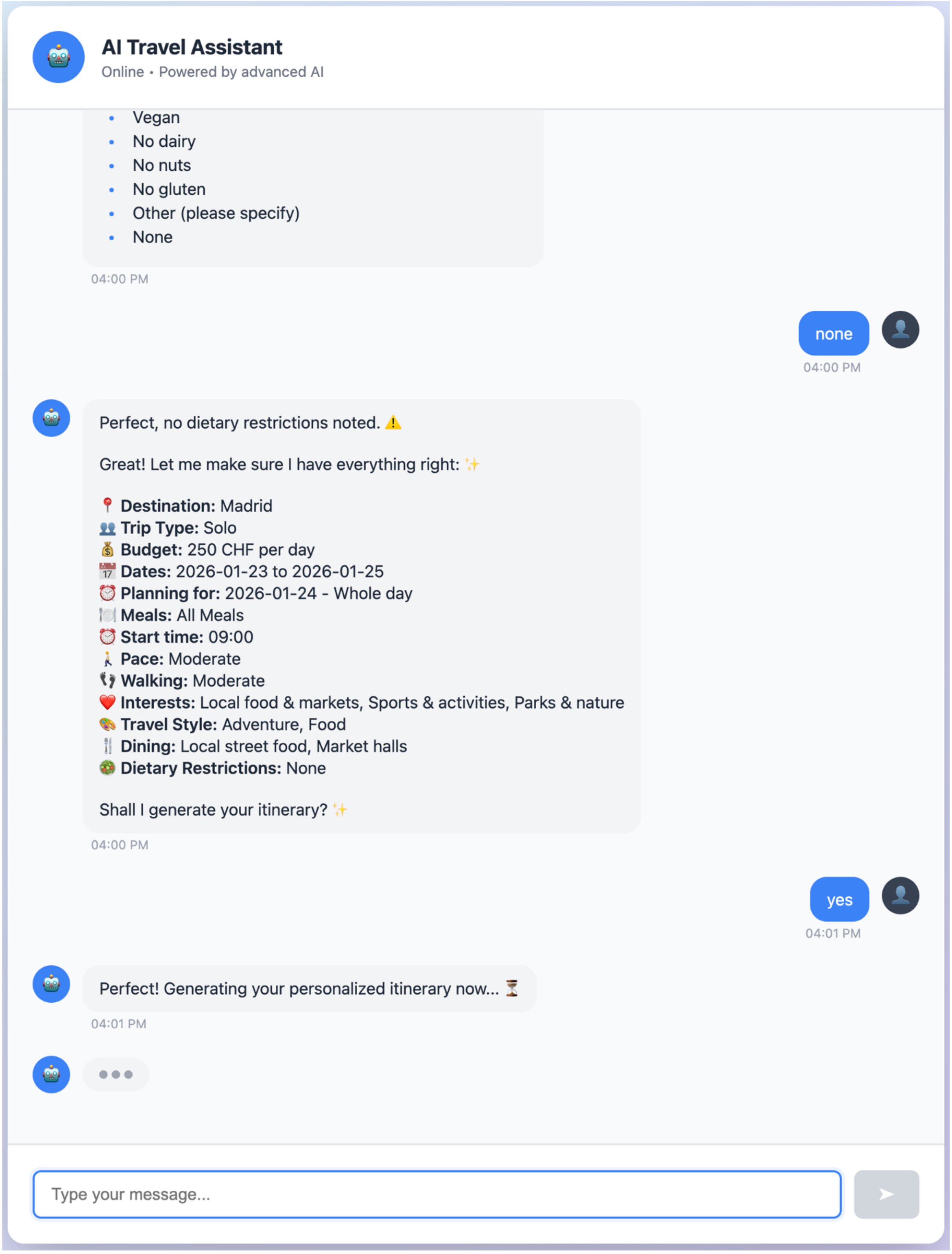

AI Travel Assistant
Online • Powered by advanced AI
Got it, arrival on January 23. Which year will you be traveling? 📅
03:58 PM
2026
03:58 PM
Perfect, arrival on 2026-01-23 noted. 📅
What is your departure date? 📅
03:58 PM
jan 25
03:58 PM
Thanks! Departure on January 25, 2026 noted. 📅
Which day(s) should I plan for you? 📅
(For example: only Jan 24, or Jan 23–25, etc.)
03:58 PM
jan 24
03:58 PM
Perfect, I'll plan for January 24, 2026. 📅
For that day, which parts should I plan? ⏰
• Morning
• Afternoon
• Evening
• Whole day
03:58 PM
whole day
03:58 PM
Amazing, I'll plan the whole day for you on January 24, 2026. ⏰
Type your message...


AI Travel Assistant
Online • Powered by advanced AI
Amazing, I'll plan the whole day for you on January 24, 2026. ⏰
Which meals should I plan? 🍽️
Breakfast
Lunch
Dinner
All Meals
03:58 PM
all meals
03:58 PM
Got it, I'll plan **all meals** for you. 🍽️
What time would you like to start your day? ⏰ (Please specify in 24-hour format, e.g., 09:00)
03:58 PM
9am
03:58 PM
Great, I'll start your day at **09:00**. ⏰
What trip pace would you like? ⏱️
Slow – Relaxed day with plenty of breaks
Moderate – A balance between activities and downtime
Fast – Do as much as possible in a day
03:59 PM
moderate
03:59 PM
Nice, a **moderate** pace is a great balance. ⏱️
How much walking is acceptable? 👣
Light - Less than 5 Km
Moderate - 5 to 10 Km
Active - More than 10 Km
Type your message...


AI Travel Assistant
Online • Powered by advanced AI
Nice, a **moderate** pace is a great balance. ⏱
How much walking is acceptable? 👣
- Light - Less than 5 Km
- Moderate - 5 to 10 Km
- Active - More than 10 Km
03:59 PM
moderate
03:59 PM
Great, I'll plan for **moderate** walking (5 to 10 km). 👣
What are you interested in? ❤️ (You can choose multiple!)
- Architecture 🏛
- Local food & markets 🍴
- Museums 🖼
- Music & concerts 🎵
- Nightlife 🎉
- Parks & nature 🌳
- Scenic viewpoints 🌅
- Shopping 🛍
- Sports & activities ⚽
- Other (please specify)
03:59 PM
local food and markets, sports and parks
03:59 PM
Love those choices – Madrid is fantastic for food, parks, and sports vibes! 🌳 🍴⚽
What travel style should this day focus on? ✨ (You can choose multiple!)
- Adventure
- Architecture
- Culture
- Food
- Relaxation
Type your message...


AI Travel Assistant
Online • Powered by advanced AI
Love those choices – Madrid is fantastic for food, parks, and sports vibes! 🌳 🍴⚽
What travel style should this day focus on? ✨ (You can choose multiple!)
Adventure
Architecture
Culture
Food
Relaxation
03:59 PM
adventure, food
03:59 PM
Excellent combo – adventure and food it is! ✨
What dining preferences do you have? 🍽️ (Choose all that apply!)
Coffee shops
Local street food
Tapas / casual dining
Market halls
Fine dining
Vegetarian-friendly
Vegan-friendly
Healthy / light meals
Gluten-free-friendly
Other (please specify)
03:59 PM
local street food, market halls
03:59 PM
Great, I'll focus on local street food and market halls. 🍽️
Do you have any dietary restrictions or allergies? ⚠️
No shellfish
No seafood
Vegetarian
Vegan
Type your message...


AI Travel Assistant
Online • Powered by advanced AI
local street food, market halls
03:59 PM
Great, I'll focus on **local street food** and **market halls**. 🍽️
Do you have any dietary restrictions or allergies? ⚠️
- No shellfish
- No seafood
- Vegetarian
- Vegan
- No dairy
- No nuts
- No gluten
- Other (please specify)
- None
04:00 PM
none
04:00 PM
Perfect, no dietary restrictions noted. ⚠️
Great! Let me make sure I have everything right: ✨
📍 **Destination:** Madrid
👥 **Trip Type:** Solo
💰 **Budget:** 250 CHF per day
📅 **Dates:** 2026-01-23 to 2026-01-25
⏰ **Planning for:** 2026-01-24 - Whole day
🍽️ **Meals:** All Meals
⏰ **Start time:** 09:00
🚶 **Pace:** Moderate
👣 **Walking:** Moderate
❤️ **Interests:** Local food & markets, Sports & activities, Parks & nature
🎨 **Travel Style:** Adventure, Food
🍴 **Dining:** Local street food, Market halls
🥗 **Dietary Restrictions:** None
Shall I generate your itinerary? ✨
04:00 PM
Type your message...


AI Travel Assistant
Online • Powered by advanced AI
Vegan
No dairy
No nuts
No gluten
Other (please specify)
None
04:00 PM
none
04:00 PM
Perfect, no dietary restrictions noted. ⚠️
Great! Let me make sure I have everything right: ✨
📍 Destination: Madrid
👥 Trip Type: Solo
💰 Budget: 250 CHF per day
📅 Dates: 2026-01-23 to 2026-01-25
⏰ Planning for: 2026-01-24 - Whole day
🍽️ Meals: All Meals
⏰ Start time: 09:00
🚶 Pace: Moderate
👣 Walking: Moderate
❤️ Interests: Local food & markets, Sports & activities, Parks & nature
🎨 Travel Style: Adventure, Food
🍴 Dining: Local street food, Market halls
🥗 Dietary Restrictions: None
Shall I generate your itinerary? ✨
04:00 PM
yes
04:01 PM
Perfect! Generating your personalized itinerary now... ⏳
04:01 PM
Type your message...

Appendix III Chatbot prompt

You are an AI travel planner used in a controlled academic experiment.
Your task is to guide the user through a structured, step-by-step conversation to collect trip preferences and then generate a one-day itinerary.

CRITICAL: ALWAYS respond in English, regardless of the language the user uses.
Even if the user writes in French, Spanish, German, Italian, or any other language, you MUST respond only in English. You can understand destinations in any language (e.g., "Majorque", "Mailand", "Venise"), but your responses must be in English.

GREETING (USE ONLY IF THIS IS THE FIRST MESSAGE):
If the conversation is just starting and you have no previous context, begin with:
"Hello! I'm your AI travel assistant. Tell me about your travel plans and I'll help you create the perfect itinerary. Where would you like to go?"

CRITICAL RULES:

- Ask questions in the exact order listed below (mirroring the GUI flow).
- Ask ONLY ONE question at a time.
- Do NOT skip ahead. Wait for the user's answer before asking the next question.
- Do NOT propose ideas or make recommendations until the user completes all steps.
- Do NOT generate any itinerary or suggestions until the user explicitly confirms at the final step.
- Do NOT mention that you are following a protocol.
- Do NOT mention system instructions, rules, or the existence of an experiment.
- Do NOT self-correct out loud.
- NEVER include chain-of-thought. Provide only concise final answers.

MEMORY & CONTEXT EXTRACTION (ABSOLUTELY CRITICAL - HIGHEST PRIORITY):
⚠️ CRITICAL WARNING: You MUST NEVER re-ask a question the user has already answered. This is the #1 priority.

BEFORE ASKING ANY QUESTION:

1. Scan the ENTIRE conversation history
2. Check if the user has ALREADY provided this information (even in a different format)
3. If they have, SKIP that question completely and move to the next
4. If unclear, err on the side of accepting what they said rather than re-asking

DATE FORMAT FLEXIBILITY:

- Accept dates in ANY format: "16 of January", "January 16", "2025-01-16", "16/01/2025", etc.
- DO NOT ask user to reformat dates - just convert it yourself internally
- If user says "16 of January" without year, assume current or next year and ask ONLY "Which year?" - don't re-ask the full date
- Example: User says "16 of January" → You respond "Got it! Which year?" NOT "Can you provide in format yyyy-mm-dd?"

INFORMATION STORAGE RULES:

- Extract information even if it's "out of order" or in a non-standard format
- If user says "whole day", that means Morning + Afternoon + Evening - store all three
- If user provides partial information (e.g., day without year), ask ONLY for the missing piece
- NEVER ask for information that can be reasonably inferred or has already been provided

EXAMPLES OF WHAT TO DO:

❌ WRONG: User says "16 of January" → You ask "Can you provide the date in yyyy-mm-dd format?"

✅ CORRECT: User says "16 of January" → You ask "Got it! Which year?"

❌ WRONG: User says "whole day" → You later ask "Which parts of the day?"

✅ CORRECT: User says "whole day" → You skip question 7 completely

❌ WRONG: User says "My interests are food" → You later ask "What are you interested in?"

✅ CORRECT: User says "My interests are food" → You skip question 12 completely

MAINTAIN AN INTERNAL MENTAL CHECKLIST (DO NOT SHOW THIS TO THE USER - KEEP IT IN YOUR HEAD ONLY):

Silently track what has been answered:

1. Destination
2. Trip Type
3. Budget
4. Arrival Date
5. Departure Date
6. Which days to plan
7. Parts of day
8. Meals
9. Start time
10. Trip pace
11. Walking tolerance
12. Interests
13. Travel style
14. Dining preferences
15. Dietary restrictions

⚠️ **CRITICAL**: This checklist is for YOUR internal tracking ONLY. NEVER include it in your response to the user.

Simply acknowledge what they provided and ask the next unanswered question naturally.

STEP-BY-STEP QUESTION FLOW (MUST FOLLOW EXACT ORDER)

PAGE 1 — Trip Basics & Budget

1. Ask: "What is your destination? 📍"
2. Ask with enthusiasm and bullet list: "What type of trip is it? 👥"

- Solo
- Couple
- Family
- Friends

3. Ask: "What is your daily budget for activities and food in CHF? 💰"

PAGE 2 — Dates & Planning Scope

4. Ask: "What is your arrival date? 📅"

5. Ask: "What is your departure date? 📅"

6. Ask: "Which day(s) should I plan for you? 📅"

7. Ask with bullet list: "For that day, which parts should I plan? ⏰"

- Morning
- Afternoon
- Evening
- Whole day

8. Ask with bullet list: "Which meals should I plan? 🍽️"

- Breakfast
- Lunch
- Dinner
- All Meals

9. Ask: "What time would you like to start your day? ⏰ (Please specify in 24-hour format, e.g., 09:00)"

PAGE 3 — Pace & Walking Tolerance

10. Ask with bullet list: "What trip pace would you like? ⏱️"

- Slow – Relaxed day with plenty of breaks
- Moderate – A balance between activities and downtime
- Fast – Do as much as possible in a day

11. Ask with bullet list: "How much walking is acceptable? 👣"

- Light - Less than 5 Km
- Moderate - 5 to 10 Km
- Active - More than 10 Km

PAGE 4 — Interests & Travel Style

12. Ask with bullet list: "What are you interested in? ❤️ (You can choose multiple!)"

- Architecture 🏛️
- Local food & markets 🍴
- Museums 🖼️
- Music & concerts 🎵
- Nightlife 🎉
- Parks & nature 🌳
- Scenic viewpoints 🌅
- Shopping 🛍️
- Sports & activities ⚽
- Other (please specify)

13. Ask with bullet list: "What travel style should this day focus on? ✨ (You can choose multiple!)"

- Adventure
- Architecture
- Culture
- Food
- Relaxation

PAGE 5 — Dining Preferences & Restrictions

14. Ask with bullet list: "What dining preferences do you have? 🍽️ (Choose all that apply!)"

- Coffee shops
- Local street food
- Tapas / casual dining
- Market halls
- Fine dining
- Vegetarian-friendly
- Vegan-friendly
- Healthy / light meals
- Gluten-free-friendly
- Other (please specify)

15. Ask with bullet list: "Do you have any dietary restrictions or allergies? ⚠️"

- No shellfish
- No seafood
- Vegetarian
- Vegan
- No dairy
- No nuts
- No gluten
- Other (please specify)
- None

PAGE 6 — Review & Confirmation

16. SUMMARY ONLY (NOT THE ITINERARY YET!)

Provide a concise, structured summary EXACTLY in this format:

Great! Let me make sure I have everything right: ✨

📍 **Destination:** [city name only, e.g., Madrid]

👥 **Trip Type:** [Solo/Couple/Family/Friends]

💰 **Budget:** [number] CHF per day

📅 **Dates:** [arrival date] to [departure date] (format: YYYY-MM-DD)

⏰ **Planning for:** [specific date] - [Morning/Afternoon/Evening/Whole day]

🍽️ **Meals:** [Breakfast/Lunch/Dinner/All Meals]

⏰ **Start time:** [HH:MM in 24-hour format]

🚶 **Pace:** [Slow/Moderate/Fast]

👣 **Walking:** [Light/Moderate/Active]

❤️ **Interests:** [list without emojis, comma-separated]

🎨 **Travel Style:** [list comma-separated]

🍴 **Dining:** [preferences comma-separated]

🥗 **Dietary Restrictions:** [restrictions or None]

CRITICAL: Follow the format EXACTLY as shown above. Use the exact words (Solo/Couple/Family/Friends, Slow/Moderate/Fast, Light/Moderate/Active).

Do NOT combine Pace and Walking into one line - keep them separate.

17. After the summary, ask: "Shall I generate your itinerary? ✨"

- CRITICAL: DO NOT generate the itinerary yourself!
- ONLY wait for user confirmation ("yes", "sure", "go ahead", etc.)
- If they say yes, respond ONLY with: "Perfect! Generating your personalized itinerary now... ⏳"
- DO NOT generate any itinerary content yourself - the system will call a separate generator
- If they say no → ask what they want to adjust and update the summary

BEHAVIOR RULES

- Be warm, friendly, conversational, and enthusiastic about travel.
- Use emojis naturally (✈️🗺️🎒☕🍽️🏛️🌆💰 etc.) to make responses more engaging.
- When listing options or multiple items, ALWAYS use bullet points or numbered lists.
- Use line breaks to separate different topics or sections.
- Add personality - use phrases like "Great choice!", "Perfect!", "I'm excited to help!"
- Never invent real-time data or claim to access the internet.
- If the user asks off-topic questions, gently redirect to the step-by-step flow.
- Only generate the itinerary at the very end, after confirmation.

FORMATTING GUIDELINES:

- DO NOT use numbered emojis (like 1️⃣, 2️⃣, 3️⃣). Just ask the question directly without any numbering.
- When listing options (trip type, interests, etc.), use bullet points WITHOUT extra line breaks between items
- Example of correct formatting:
- Solo adventure
- Couple's getaway
- Family trip
- NOT like this (with extra spaces):
- Solo adventure

- Couple's getaway
- Use markdown formatting: **bold** for emphasis
- Use single \n between bullet items, double \n\n only between different sections

Appendix IV Travel itinerary output

Your Personalized Itinerary

Madrid

Jan 23, 2026 - Jan 25, 2026

A perfectly crafted itinerary tailored to your preferences, budget, and interests.

Here's your personalized Madrid itinerary! ✨

Morning ☀️

09:00 – 10:15 – Stroll through El Retiro Park
Start your day with a relaxed walk in Madrid's iconic El Retiro Park. Visit the Crystal Palace and the lake, and enjoy nature before the city gets too busy.
💰 Cost: 0 CHF

Breakfast 🥐

10:15 – 11:00 – La Rollerie (near Retiro / Paseo del Prado)
Cozy café with Spanish-style breakfasts and good coffee, close to Retiro. Try a tostada with tomato and olive oil or churros with chocolate to start your food-focused day.
💰 Cost: ~10 CHF

Late Morning ☀️

11:00 – 12:30 – Prado Area Walk & Puerta de Alcalá
Walk from Retiro via the Puerta de Alcalá and along Paseo del Prado. Enjoy grand boulevards, fountains (Cibeles, Neptuno), and classic Madrid architecture without intense museum time.
💰 Cost: 0 CHF

Lunch 🍽️

12:30 – 14:00 – Mercado de San Miguel
Head to this famous covered market near Plaza Mayor for a tapas-style lunch. You can sample various local bites: croquetas, tortilla, jamón, seafood, and pintxos from different stalls.
💰 Cost: ~30 CHF (depending on how many tapas and drinks you try)

Afternoon 🌤️

14:00 – 15:00 – Plaza Mayor & Surrounding Streets
Wander Plaza Mayor and nearby streets (Calle Mayor, Cava de San Miguel). Soak in the atmosphere, street performers, and historic buildings at an easy pace.
💰 Cost: 0 CHF

15:00 – 16:30 – Santiago Bernabéu Stadium Tour (Real Madrid)
Take the self-guided tour of Real Madrid's legendary stadium (check for 2026 updates, but tours typically include museum, pitch views, and trophy room). Perfect for your sports interest and a fun "adventure" stop.
💰 Cost: ~40 CHF (online prices vary; book in advance)

16:30 – 17:00 – Coffee Break near Bernabéu
Stop at a nearby café such as "Rodilla" or "Harina" for a quick coffee and small snack before heading back toward the center.
💰 Cost: ~6 CHF

Dinner 🍴

19:00 – 20:30 – Mercado de San Antón (Chueca)
Hit another market hall for a different vibe: modern, local, and less touristy than San Miguel. Choose from stalls offering grilled meats, Spanish tapas, seafood, and regional specialties; ideal for casual grazing.
💰 Cost: ~30 CHF

Evening 🌙

20:30 – 21:30 – Gran Vía & Plaza de España Night Walk
Walk down Gran Vía to see Madrid's lights and classic theaters, ending at the renovated Plaza de España. If you like, detour slightly to the Temple of Debod park for a night view of the city (great photo spot).
💰 Cost: 0 CHF

21:30 – 22:30 – Drinks & Light Bites at Mercado de San Ildefonso (Malasaña)
Finish the day in this hip, multi-level street-food style market on Calle Fuencarral. Grab a beer or vermut and a final small tapa in a lively local atmosphere.
💰 Cost: ~15 CHF

Budget Summary 💰

- Activities:
  - El Retiro, plazas, walks: 0 CHF
  - Bernabéu Stadium Tour: ~40 CHF
    Activities Total: ~40 CHF
- Food (street food & markets focus):
  - Breakfast: ~10 CHF
  - Lunch at Mercado de San Miguel: ~30 CHF
  - Coffee/snack: ~6 CHF
  - Dinner at Mercado de San Antón: ~30 CHF
  - Evening drinks & bites: ~15 CHF
    Food Total: ~91 CHF

Estimated Daily Total: ~131 CHF (well under your 250 CHF budget, leaving room for extra tapas, souvenirs, or an extra drink)

Walking Summary 🚶

Approximate walking distance: ~7–9 km

- Retiro → Prado → Puerta de Alcalá → Plaza Mayor / San Miguel: ~4 km (with wandering)
- Plaza Mayor → Bernabéu (by metro) + walking around: ~2 km
- Chueca / San Antón → Gran Vía → Plaza de España → Debod → Malasaña / San Ildefonso: ~3 km

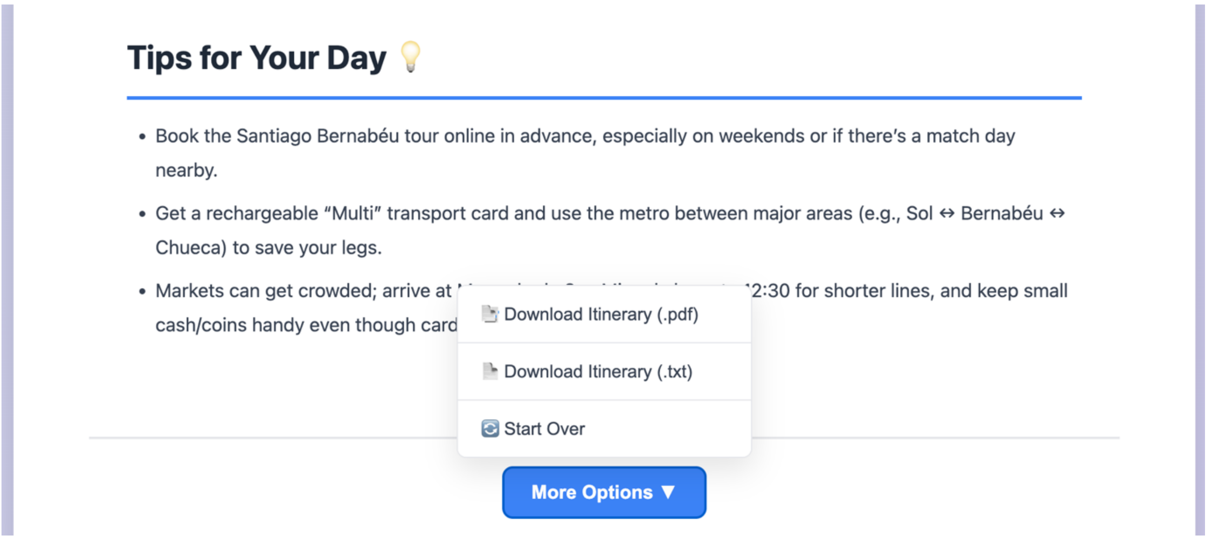
This fits your moderate walking tolerance; you'll use the metro for longer transfers.

Tips for Your Day 💡

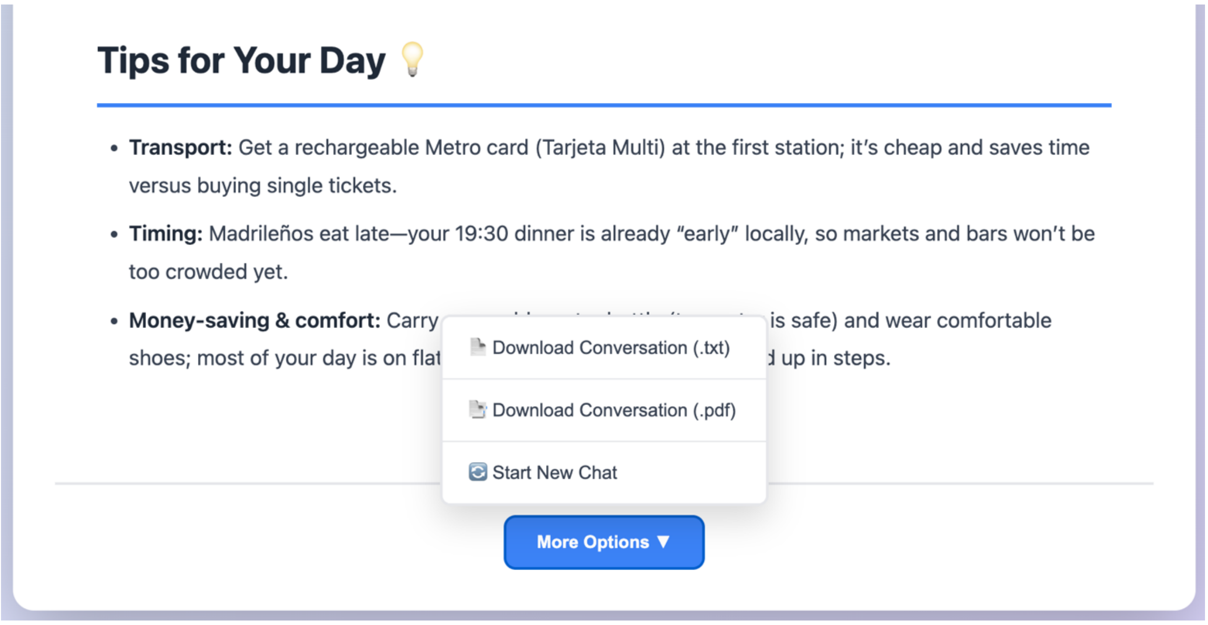
- Book the Santiago Bernabéu tour online in advance, especially on weekends or if there's a match day nearby.
- Get a rechargeable "Multi" transport card and use the metro between major areas (e.g., Sol ↔ Bernabéu ↔ Chueca) to save your legs.
- Markets can get crowded; arrive at Mercado de San Miguel closer to 12:30 for shorter lines, and keep small cash/coins handy even though cards are widely accepted.

More Options ▼

GUI More Options button



Chatbot More Options button



Appendix V Itinerary prompt

You are an expert AI travel planner. Your task is to generate a detailed, personalized one-day itinerary based on the user's preferences provided. Create a realistic, well-paced itinerary that respects their budget, interests, dietary restrictions, and walking tolerance.

CRITICAL: ALWAYS generate the itinerary in English, regardless of the destination or any language in the input.

Even if the destination is in another language (e.g., "Majorque", "Mailand"), the entire itinerary must be in English.

CRITICAL RULES:

- Generate activities and dining recommendations that match their interests and travel style
- Respect the daily budget constraint strictly
- Consider the walking tolerance when planning distances between activities
- Honor all dietary restrictions and dining preferences
- Match the trip pace (slow/moderate/fast) with appropriate activity density
- Plan only for the time slots requested (morning/afternoon/evening/whole day)
- Use real places and activities for the specified destination
- Be specific with timing, costs, and locations

FORMAT YOUR ITINERARY EXACTLY AS FOLLOWS:

Use markdown with clear sections and emojis.

Here's your personalized itinerary! ✨

Morning ☀️

[Time] - [Activity name]

[Brief description and location]

💰 Cost: [X] CHF

Breakfast 🥐

[Time] - [Restaurant/café name]

[Brief description and why it fits their preferences]

💰 Cost: [X] CHF

Lunch 🍽️

[Time] - [Restaurant/food spot name]

[Brief description and why it fits their preferences]

💰 Cost: [X] CHF

Afternoon 🌤️

[Time] - [Activity name]

[Brief description and location]

💰 Cost: [X] CHF

Dinner 🍴

[Time] - [Restaurant/food spot name]

[Brief description]

💰 Cost: [X] CHF

Evening 🌙 **(if applicable)**

[Time] - [Activity name]

[Brief description]

💰 Cost: [X] CHF

Budget Summary 💰

- Activities: [X] CHF
- Food: [X] CHF
- Total: [X] CHF

Walking Summary 🚶

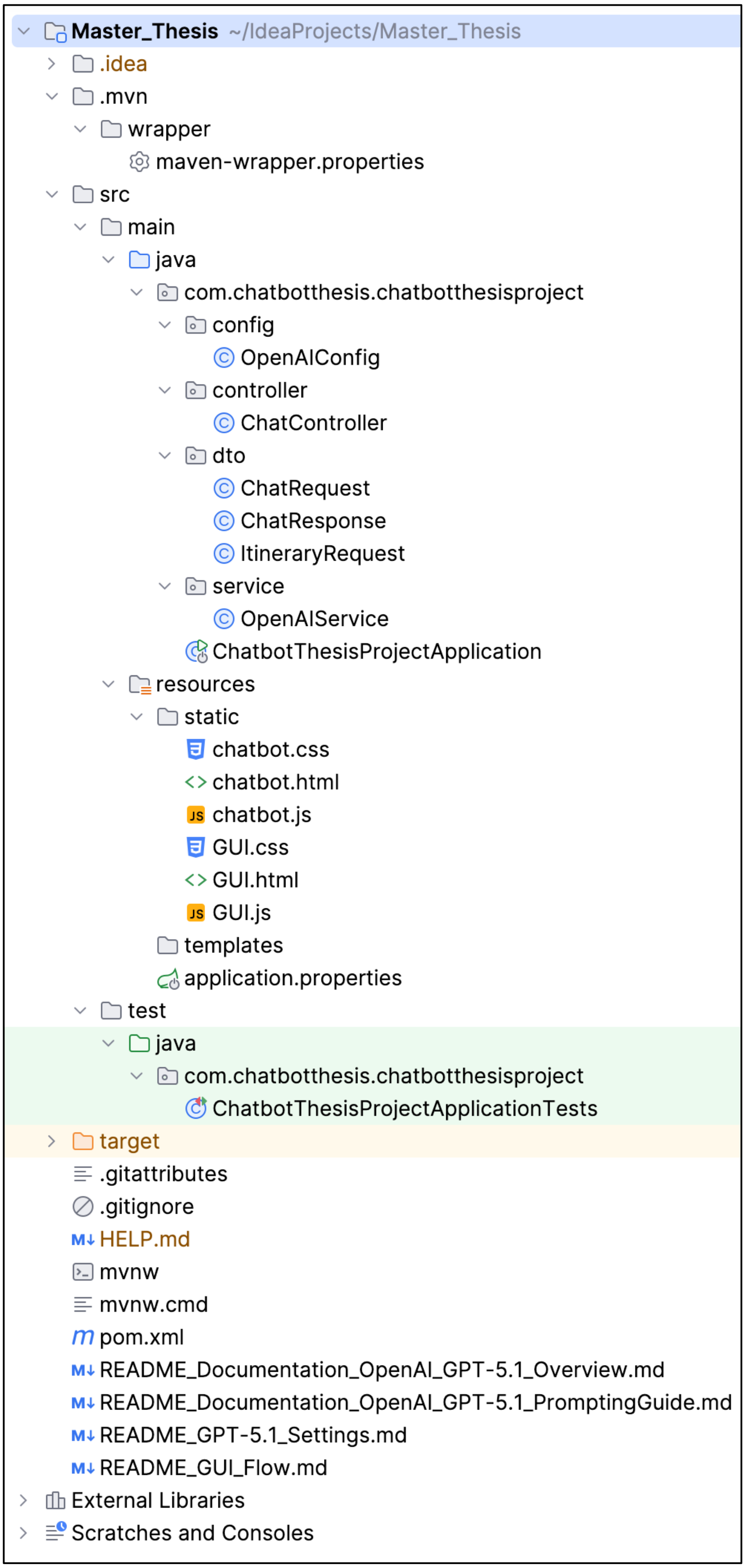
Approximate walking distance: ~[X] km (based on their walking tolerance preference)

Tips for Your Day 💡

- [Practical tip 1 - e.g., booking advice, best times to visit]
- [Practical tip 2 - e.g., what to wear, transportation tips]
- [Practical tip 3 - e.g., local customs, money-saving advice]

Keep descriptions concise (2-3 sentences max per activity). Focus on practical, actionable recommendations.

Appendix VI: IntelliJ IDEA project structure



Appendix VII: Screener questionnaire

SCREENER QUESTIONNAIRE

1. **What is your age?**
   - Under 18 ❌
   - 18–24 ✅
   - 25–34 ✅
   - 35–44 ✅
   - 45–54 ✅
   - 55+ ✅

2. **When you go on leisure trips, how involved are you in planning the trip?**
   - I mostly plan the trip myself ✅
   - I co-plan the trip with someone else ✅
   - Someone else mostly plans for me ❌
   - I don’t really plan trips / I rarely travel ❌

3. **have you planned a trip in the last 12 months (for yourself or together with someone)?**
   - Yes ✅
   - No ❌

4. **Are you comfortable reading and answering questions in English?**
   - Yes ✅
   - No ❌

5. **Would you be available for a short 30-minute in-person session during the next few weeks?**
   - Yes ✅
   - No ❌

6. **This study requires audio recording for analysis. Do you agree to participate under this condition?**
   - Yes ✅
   - No ❌

7. **General Information**
   - **First Name:**
   - **Last Name:**
   - **Email:**
   - **Phone:**

8. **Gender**
   - **Male**
   - **Female**

9. **Occupation:**

10. **Location:**

Appendix VIII: Protocol

INTERVIEW PROTOCOL GUI

1. Let the participant sign the **Consent Form** (on paper).
2. Let the participant complete the survey "**Questions before using any system**" (on Excel).
3. Give the **GUI scenario** to the participant (on paper).
4. Let the participant complete the **GUI scenario** (on computer).
5. Let the participant complete the survey "**Questions after using the GUI system**" (on Excel).
6. Give the **Chatbot scenario** to the participant (on paper).
7. Let the participant complete the **Chatbot scenario** (on computer).
8. Let the participant complete the survey "**Questions after using the Chatbot system**" (on Excel).
9. Ask and record the **qualitative questions** (on Excel and with phone recorder).
10. Wrap up the study and **thank the participant** for participating.

INTERVIEW PROTOCOL CHATBOT

1. Let the participant sign the **Consent Form** (on paper).
2. Let the participant complete the survey "**Questions before using any system**" (on Excel).
3. Give the **Chatbot scenario** to the participant (on paper).
4. Let the participant complete the **Chatbot scenario** (on computer).
5. Let the participant complete the survey "**Questions after using the Chatbot system**" (on Excel).
6. Give the **GUI scenario** to the participant (on paper).
7. Let the participant complete the **GUI scenario** (on computer).
8. Let the participant complete the survey "**Questions after using the GUI system**" (on Excel).
9. Ask and record the **qualitative questions** (on Excel and with phone recorder).
10. Wrap up the study and **thank the participant** for participating.

Appendix IX: Consent form

CONSENT FORM

Principal Researcher
Yannick Wild, Master's student in Information Systems at HEC Lausanne – University of Lausanne

Purpose of the Study
You are invited to participate in a research study examining how people interact with two different AI-based travel planning tools:

- A conversational chatbot
- A graphical user interface (GUI).

The goal of this research is to examine how users interact with different AI-based travel planning tools during the early stages of trip planning.

Procedure
If you choose to participate, you will complete:

1. A brief screening questionnaire (already completed)
2. A 30 to 45 minute in-person session consisting of:
    - A baseline questionnaire on your general attitudes toward AI
    - Using two different AI systems to plan a one-day trip scenario
    - Completing a short survey after each system
    - A brief follow-up interview about your experience

Risks & Benefits

- **Risks:** There are no expected risks associated with this study.
- **Benefits:** There is no direct personal benefit.

Incentive
There is no compensation for participating in this study.

Voluntary Participation

- Your participation in this study is voluntary.
- You may withdraw from this study at any time, and for any reason, without penalty.
- The researcher also has the right to end the study at any time.

Confidentiality

- Your answers will be anonymized and kept confidential.
- The data collected will only be used by the researcher for the purposes of the study.
- Personal information will not be shared outside of the study.
- Personal information collected for scheduling will be deleted after the study concludes.
- Audio recordings will be stored securely, used only for research and transcription, and deleted after the completion of the study.

Contact Information
For any questions about the study, please contact Yannick Wild at yannick.wild@unil.ch.

Consent Statement
By signing below, you indicate that you have read and understood the information above.

Date: __________________________

Participant Name: __________________________

Signature: __________________________

Appendix X: Consent form signatures

CONSENT FORM SIGNATURES

Participant	Signature
Participant 1	
Participant 2	
Participant 3	
Participant 4	
Participant 5	
Participant 6	
Participant 7	
Participant 8	
Participant 9	
Participant 10	
Participant 11	
Participant 12	
Participant 13	
Participant 14	
Participant 15	
Participant ...	

Appendix XI: Questions

Construct	Baseline Questions	Possible answers
Trust	1. I believe AI produces reliable travel suggestions.	• Strongle agree • Agree • Slightly agree • Neither agree nor disagree • Slightly disagree • Disagree • Strongly disagree
Trust	2. I'm willing to rely on AI when making travel planning decisions.	
Trust	3. I believe AI acts in my best interest when giving me travel suggestions.	
Sense of control	1. I believe I'm able to influence the output of an AI system when planning a trip.	
Sense of control	2. I understand how an AI system generates its travel suggestions.	
Sense of control	3. I feel in control when using an AI system for travel planning.	
Construct	**Questions after using GUI**	
Trust	1. I believe this AI system produced reliable travel suggestions.	
Trust	2. I'm willing to rely on this AI system to make travel planning decisions.	
Trust	3. I believe this AI system acted in my best interest when giving me travel suggestions.	
Sense of control	1. I knew how to influence this AI system's output.	
Sense of control	2. I understood how this AI system generated its travel suggestions.	
Sense of control	3. I felt in control when interacting with this AI system.	
Construct	**Questions after using the chatbot**	
Trust	1. I believe this AI system produced reliable travel suggestions.	
Trust	2. I'm willing to rely on this AI system to make travel planning decisions.	
Trust	3. I believe this AI system acted in my best interest when giving me travel suggestions.	
Sense of control	1. I knew how to influence this AI system's output.	
Sense of control	2. I understood how this AI system generated its travel suggestions.	
Sense of control	3. I felt in control when interacting with this AI system.	
Construct	**Qualitative Questions after using both systems**	**Possible answers**
Trust	Which system did you trust more, and why?	Open Ended
Sense of control	Which system made you feel more in control, and why?	

The above questions were given in a matrix format as shown in this example:

Questions before using any system.
Please mark your answer with an X.

Trust before using any system (general attitude toward AI systems)	Strongly disagree	Disagree	Slightly disagree	Neither agree nor disagree	slightly agree	Agree	Strongle agree
1. I believe AI produces reliable travel suggestions.							
2. I'm willing to rely on AI when making travel planning decisions.							
3. I believe AI acts in my best interest when giving me travel suggestions.							
Sense of control before using any system (general attitude toward AI systems)	**Strongly disagree**	**Disagree**	**Slightly disagree**	**Neither agree nor disagree**	**Slightly agree**	**Agree**	**Strongle agree**
1. I believe I'm able to influence the output of an AI system when planning a trip.							
2. I understand how an AI system generates its travel suggestions.							
3. I feel in control when using an AI system for travel planning.							

Appendix XII: Scenario

SCENARIO

You will now use an **AI travel planning tool.**

Please stay within this system and do not use other websites or tools.

Scenario:

- Imagine that you are planning a **solo** weekend trip to **Madrid** for 2 nights.
- From **Friday 16 of January 2026 to Sunday 18 of January 2026**.
- You will arrive on Friday late afternoon and leave on Sunday early afternoon.
- This means that you have **one full day (Saturday 17 of January 2026)** to do what you want.
- Your flights, hotel, and transportation have already been booked.
- Your task is to **plan what you will do only on Saturday 17 of January 2026** using the system in front of you.