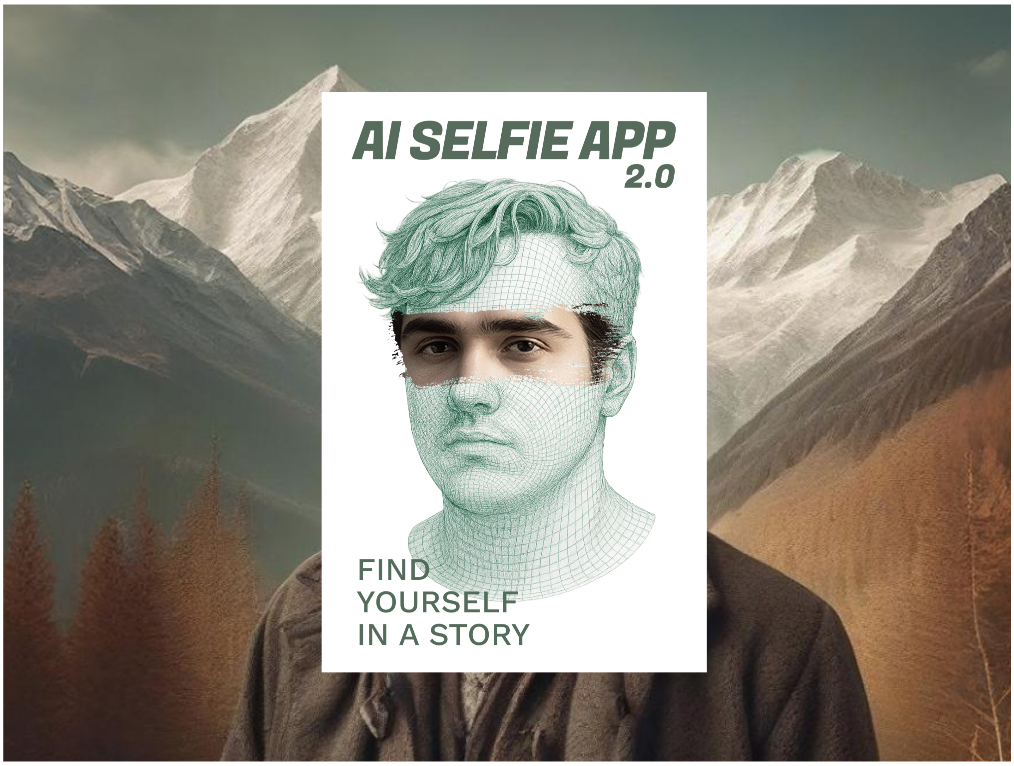
Fachhochschule Nordwestschweiz
Hochschule für Technik

25FS_IIT58: AI Selfie App 2.0: Find yourself in a story

Bachelor's Thesis

Thun, August 2025



Student	Hava Fuga
Expert	Claudia Wipf
Supervisors	Dr. Nitish Patkar, Alain Zanchetta
Client	Prof. Dr. Arzu Çöltekin
Project Number	25FS_IIT58

Fachhochschule Nordwestschweiz, Hochschule für Technik

Abstract

Public exhibitions demand interactions that are fast, safe, and legible. This thesis presents the design and development of the AI Selfie App, a web-based application that lets visitors generate AI portraits of themselves while preserving privacy and making the underlying process understandable. The goals were to ensure smooth, short-session usability, implement an explicit yet lightweight consent flow, and provide a clear, narrative explanation of AI image generation. The project employed a methodology of research and iterative prototyping, supported by two phases of user testing, analysis of task completion times, and the use of established metrics such as the User Engagement Scale (Short Form) for measuring engagement and the System Causability Scale for evaluating explanation clarity. The AI Selfie App demonstrated improved user flows with reduced navigational difficulties, sustained engagement, and explanations perceived as clear and logical. The findings indicate that the developed system successfully integrates explainability and privacy considerations in a manner aligned with the constraints of public exhibition settings.

Keywords: AI Selfie App, Mobile Applications, Artificial Intelligence, Generative AI, User Experience, Explainable AI, Consent Mechanisms, Public Exhibitions, Usability Testing.

Acknowledgements

I would like to express my deepest gratitude to Prof. Dr. Arzu Çöltekin and Mr. Cédric Merz for believing in the project and supporting the idea of turning it into my bachelor thesis.

A special thank you goes to Dr. Nitish Patkar and Mr. Alain Zanchetta as well for their ongoing support and for having the patience and humor to guide me through all the ups and downs. Your insights and feedback were invaluable.

Finally, I want to thank my sister Maide, my greatest role model and source of motivation. Looking up to her has always inspired me to push forward.
To my younger siblings, Hana and Musab, thank you for bringing joy and a bit of fun when I needed it most. I'm proud of you both.
To my friend Lena, thank you for patiently listening to my complaints and always cheering me on.
And to my sweet girl Hurreira, my little angel, I will always keep you in my memory.

This project was a joy to work on, and I often found myself happily working into the early hours simply because it was so meaningful and fun.

Thank you all for being part of this journey.

Contents

List of Figures

List of Tables

1 Introduction

Artificial Intelligence has developed at an extraordinary pace over the past decade, evolving from specialized research projects into technologies that permeate almost every aspect of daily life. This development is evident even within Switzerland, where nearly 80 percent of companies have already established dedicated AI strategies and are preparing substantial investments to integrate these technologies into their core operations (W.I.R.E. and Mindfire, 2022).

Within this rapidly expanding AI ecosystem, generative AI stands out as a particularly dynamic and fast-evolving frontier. Capable of producing novel content such as text, images and even code, this subfield continues to push the boundaries of what machines can create. The pace of progress is so dynamic that, during the course of this thesis itself, Midjourney announced its first video model (Midjourney, 2025), marking a milestone that empowers users to create sophisticated video content from text prompts. Developments like these demonstrate the immense creative potential and growing accessibility of generative AI, which is steadily moving beyond specialized domains into mainstream consumer applications.

However, as the technology improves, many people feel left behind when it comes to understanding how these systems actually work. The KPMG 2023 Global AI Trust Survey found that half of respondents do not know what AI is or how it is applied, while 82% expressed a clear desire to learn more about it (Gillespie et al., 2023).

This contrast between limited understanding and high curiosity highlights the importance of making AI more transparent and accessible to a broad audience. Although there are countless tools and platforms that showcase what AI can create, many of them focus mainly on the fascination of the finished product and rarely provide clear explanations that help users understand the underlying processes in an accessible way (Warren et al., 2022). This results in a gap between technological possibilities and public understanding. Addressing this imbalance between rapid technological progress and limited transparency is therefore both timely and necessary. It raises the question of how interaction and education can be combined to make AI more approachable, understandable, and meaningful for diverse audiences.

In this context, the present work examines how such approaches can be developed and what contribution they can make to strengthening digital literacy and critical engagement with artificial intelligence.

1.1 Initial Situation

This bachelor's project extends the groundwork laid during the IP5 project, which involved the development of a functional application that allows users to upload photographs and generate AI-created images in diverse artistic styles. While the IP5 project successfully showcased the fundamental capabilities of AI-driven image generation and produced a high-fidelity design, the design was not fully integrated into the application. Furthermore, the application was developed without a specific deployment context, focusing primarily on the development and testing of essential AI features. In the current project, it is known that the application will be used in a public exhibition environment. This setting implies the app will be accessible to a selected audience which will be potentially several people simultaneously or just a few at a time. And users will vary in their familiarity with AI technologies. Some will have prior knowledge, while others may have no understanding of how AI-generated images work. These conditions introduce specific requirements for clarity, accessibility, and engagement that were not considered during the original development.

1.2 Problem Statement

The original application, while functionally sound, was not initially conceived for deployment in a public exhibition setting. This introduces new challenges extending beyond purely technical considerations. Key aspects, such as obtaining informed user consent in a clear yet non-intrusive manner, and maintaining user engagement during AI processing time, remain unresolved. These involve complex questions about user behavior, comprehension, and interaction in public, time-limited, and attention-constrained environments. Addressing these challenges requires a deeper understanding of how users interact with AI systems in exhibition settings, how consent can be meaningfully integrated into engaging workflows, and how explanations can be designed to foster trust and understanding without overwhelming the user. These gaps, therefore, necessitate research into user experience design, educational interaction techniques, and responsible AI communication strategies.

1.3 Goal

The objective of this project is to refine and adapt the existing application to meet the specific demands of the public exhibition setting. This entails integrating a user-friendly consent mechanism that is both robust and transparent, ensuring that users fully understand the implications of their consent without negatively impacting their overall experience. To mitigate potential user disengagement during AI training processing times, the project will incorporate informative content designed to explain the AI image generation processes in an accessible manner. Furthermore, the project seeks to fully integrate the frontend with new functionalities, thereby optimizing the user interface and user experience to ensure the application is not only functional and user-friendly but also fully prepared for deployment within a real-world exhibition environment.

1.4 Research Questions

To achieve this goal and to methodically examine the challenges involved, this project formulates and seeks to address the following research questions:

- **RQ1**: What design principles and features, proven to maintain user engagement in mobile applications, can be successfully integrated and validated in the AI Selfie App?
- **RQ2**: How are consent process implementations in mobile applications affected by the constraints of busy and time-limited environments, and which techniques, when implemented, can effectively overcome these challenges?
- **RQ3**: How can the integration of explanation techniques within the AI Selfie App improve user understanding of AI image generation processes?

2 Foundations

This chapter explores the foundations of the project by looking back at the groundwork laid during the IP5 module. It begins by outlining the state of the AI Selfie App at the conclusion of IP5, highlighting its technical components and the design prototype that was created. This retrospective establishes the groundwork for the current project, showing how this thesis builds upon and extends that earlier work.

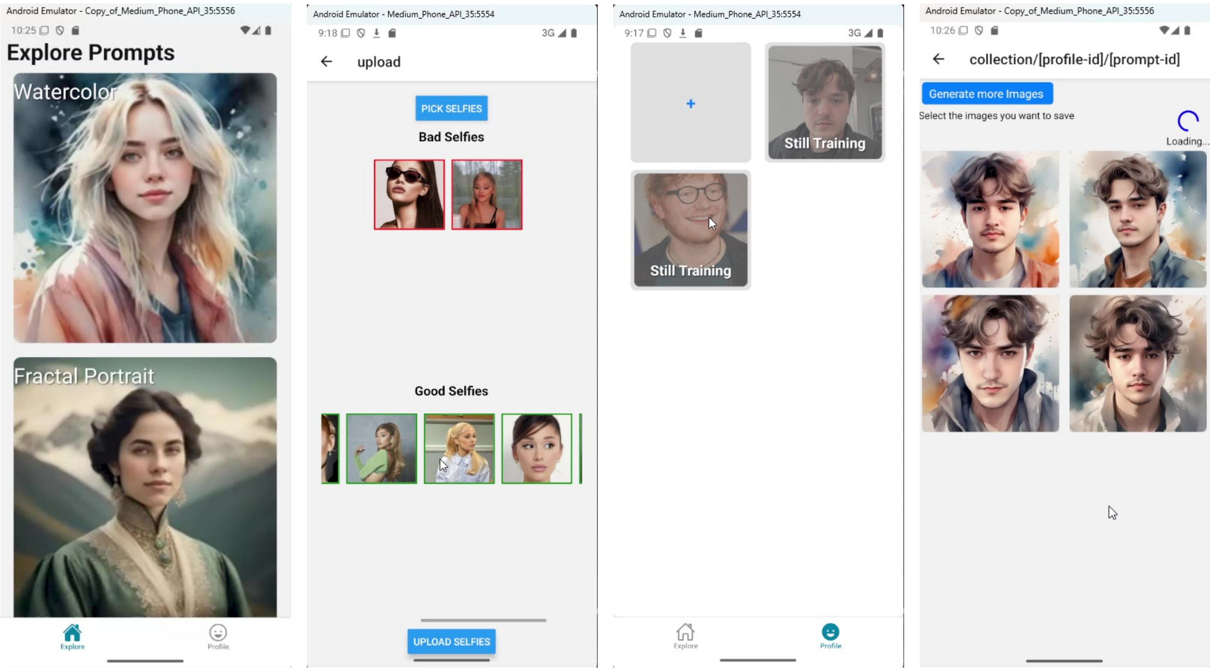
2.1 State of the AI Selfie App at Project Start

The AI Selfie App, initially developed during an IP5 module, aimed to create a mobile counterpart to the "AI Selfie Box" exhibition installation (FHNW, 2024). In the physical exhibition, visitors enter a photobooth where images are captured from multiple angles. After approximately 20 minutes of processing time, visitors can generate AI-based images of themselves. The primary objective of the initial project phase was to ascertain the technical feasibility of replicating this workflow in a mobile application, utilizing the same AI model. Consequently, a functional application was developed, successfully implementing the core functionalities of photo uploading and image generation. In addition to the implemented application, the frontend was prototyped but not yet implemented.

The app was developed using React Native (Meta Platforms, Inc., 2024) in combination with the Expo framework (Expo, n.d.), enabling the creation of a mobile application targeting both iOS and Android platforms. To handle the large volume of uploaded photos and the generated images, a database and storage solution were necessary, for which Supabase (Supabase Inc., n.d.) was selected. Supabase provided both the database backend and storage services for managing and organizing image assets in a scalable and integrated manner. At the conclusion of the IP5 module, the prototype was primarily tested using an Android emulator and could only be deployed on Android devices. More details about this phase can be found in an excerpt from the IP5 paper in Appendix A. While a frontend design was independently developed using Figma Inc., 2025, its full integration was deferred, as the project prioritized the development of a minimal viable product to validate the core workflow's replicability on mobile platforms.

Figure Figure 2.1 presents selected screenshots of this initial version of the application, illustrating the primary screens and providing an overview of the prototype's basic structure and functionality.

- **Prompts Exploring**: Displays a selection of available visual styles that users can browse and choose for generating AI images.
- **Selfies Upload**: Shows selected images automatically classified as "Good" or "Bad" selfies, with visual color coding to guide user selection.
- **Profile Overview**: Shows existing profiles and their training status.
- **Image Generation**: Shows generated image results, with options generate more or save selected outputs.



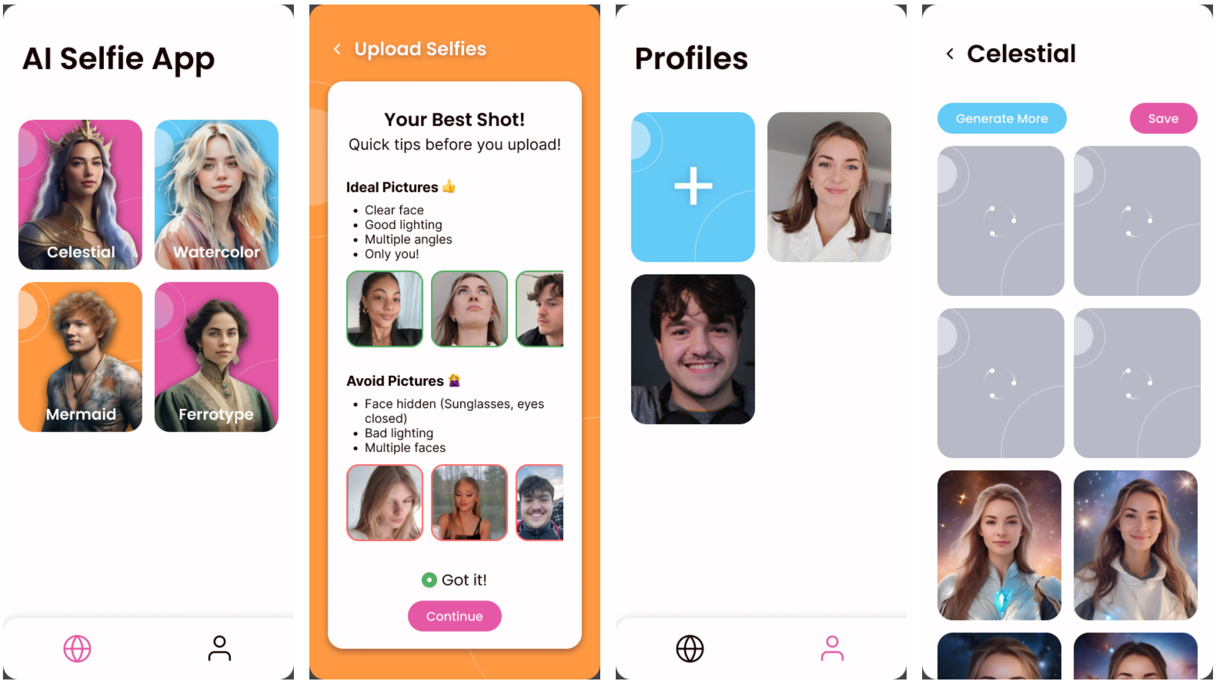
a: Prompts Exploring **b:** Selfies Upload **c:** Profile Overview **d:** Image Generation

Figure 2.1: Screenshots of the initial AI Selfie App developed during IP5

2.1.1 Design Prototype

In conjunction with the technical implementation, a design prototype was developed to illustrate the intended user interface and visual appearance of the application. Due to the absence of a formally specified context of use during the initial IP5 phase, user flows were defined based on general assumptions regarding expected interactions rather than through systematic user research. The prototype represents an initial iteration aimed at visualizing the core navigation and interaction patterns.

Key screens from the prototype are presented in Figure Figure 2.2 to provide an overview of the application's structure and aesthetic concept. The complete design specifications for typography, color schemes and the screen designs are detailed in Appendix B.



a: Prompts Exploring **b:** Upload Guidelines **c:** Profile Overview **d:** Image Generation

Figure 2.2: Screenshots of the AI Selfie App showing the main screens

- **Prompts Exploring**: The explore screen offers a comprehensive overview of the various visual styles available for image generation, encompassing choices like "Celestial," "Watercolor," "Mermaid," and "Ferrotype." Each style is accompanied by a representative example image, facilitating user understanding of the anticipated outcome.
- **Upload Guidelines**: Prior to uploading selfies, the application provides users with upload guidelines that illustrate how to select appropriate images. The interface displays examples of both recommended and discouraged images, with the goal of improving the quality of input data for the AI generation process.
- **Profile Overview**: The profile overview screen enables users to either create new profiles or select from a list of existing profiles to view previously generated images.
- **Image Generation**: This screen shows the image generation process, displaying the images as they are generated using the selected style and profile. Upon navigation to this screen, the image generation is initiated and once completed, four images are presented in a gallery layout. Subsequently, users can select the desired images to save to their profile or opt to generate additional variations.

2.1.2 AI Model

The AI model incorporated into this project was not an original development, but rather a pre-existing resource carried over from the IP5 initiative, consistent with its implementation in the AI Selfie Box installation. Access to the model was provided via a RESTful API, through which the application could submit uploaded images and request generated outputs. The API included dedicated endpoints for uploading image data, initiating the generation process, and retrieving the resulting outputs. Critically, all training, fine-tuning, and core AI processing were handled server-side, placing these aspects outside the scope of the present work.

2.1.3 Supabase

Supabase was chosen as the backend solution due to its capacity for scaling, integrated storage functionalities, and ease of use throughout the development phase. It operates as an open-source Backend-as-a-Service platform, offering a PostgreSQL database, authentication services, and file storage via a consolidated API.

The database schema implemented in this project consists of four main tables, as illustrated in Figure Figure 2.3. Each table is numbered in the diagram for reference:

- **1. Prompt:** Stores information about the predefined AI styles available in the application. Each record defines a visual style that users can select when generating images.
- **2. Image:** Contains metadata for the AI-generated outputs, including references to the selected prompt and the associated user profile.
- **3. Profile:** Represents user-specific data and acts as the central entity linking selfies and generated images to a particular user.
- **4. Selfie:** Stores references to user-uploaded images and links them to the corresponding profile.

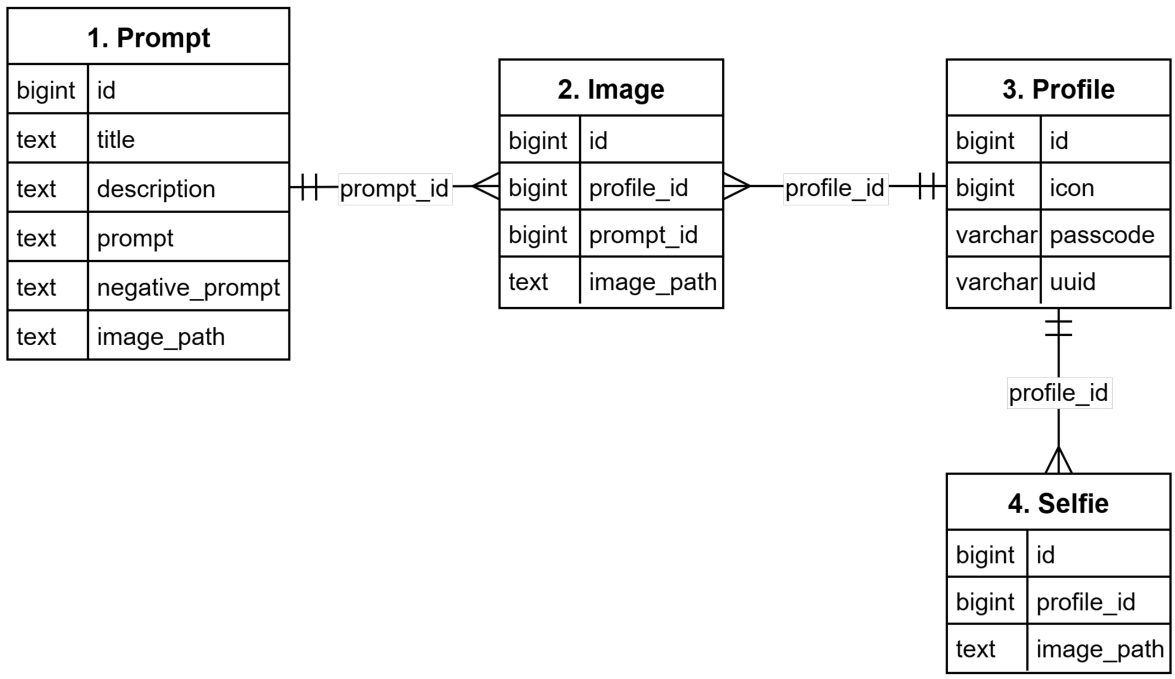


Figure 2.3: IP5 Stand - Database schema diagram

In this paper, the term “selfies” is used to refer to all images uploaded by the user, which may include any type of photograph and are not limited to actual self-portraits. The term “images” refers exclusively to AI-generated outputs. The distinction is maintained throughout to clearly differentiate between uploaded and generated content.

The tight integration between database and storage ensures that each uploaded or generated file is consistently linked to its corresponding record. the Supabase’s storage system was organized hierarchically so that each profile has its own directory containing subfolders for user selfies and generated images. This structure simplifies file management and retrieval.

Connections to both the database and storage are established through the Supabase client, which is configured with the project's unique API credentials. This setup enables the application to securely insert, query, and update data as well as upload and access image files.

The complete schema and relationships between tables are visualized in Figure 2.3.

3 Methods

This chapter details the methodological approach used in this bachelor's thesis. It begins with a detailed presentation of the instruments and procedures employed for data collection, including how this data was accessed and collected, followed by the project management framework. The study was conducted over a period of several weeks, from June 2025 to July 2025, ensuring that all data collection was anonymized and securely handled.

3.1 Data Collection Instruments

This section provides a detailed overview of the various methods and instruments applied for data collection within the scope of this thesis. A targeted literature search was conducted to identify relevant research findings, conceptual frameworks and best practices relevant to the development, implementation, and evaluation of AI-driven applications in public exhibition contexts. Usability testing was performed to gather empirical data on user experience, task completion, and perceived usability under realistic usage conditions. In addition, the standardized User Engagement Scale (Short Form) was used to quantitatively assess subjective engagement across multiple dimensions. To evaluate the effectiveness of explainability features, the System Causality Scale was employed. These instruments collectively informed decisions related to interface design, consent mechanisms, educational content and AI explanation strategies.

3.1.1 Literatre Research

The study's theoretical foundation was informed by two complementary literature search approaches: a targeted structured search and a snowball search. The structured search involved defining key terms and concepts to guide a targeted search of academic databases and search portals, including FHNW swisscovery, Google Scholar, and IEEE Xplore (Solis, 2021). The search queries were structured around the following topics and keywords:

- **UI/UX Design Principles for Public Exhibitions:**
  - "UI UX"
  - "public exhibition"
  - "design exhibition"
  - "mobile app"
- **Consent Mechanisms in High-Interaction Environments:**
  - "user consent"
  - "interactive system"
  - "UI UX"
  - "mobile app"
- **AI Explanation and Visualization Techniques:**
  - "XAI"
  - "explainability methods"

The snowball search technique was used to identify additional relevant sources from the bibliographies of the initially retrieved publications. This iterative process continued until no substantial new literature was found, resulting in a comprehensive theoretical foundation (Krul, 2015).

3.1.2 Usability Testing

To answer the research questions a formative usability test using a think-aloud protocol was chosen to evaluate the effectiveness of the implemented features and identify usability issues early in the development process and to understand users' expectations and experiences in interacting with the application (Charters, 2003).

In the context of the usability testing, the following instruments and procedures ensured a structured approach to data collection:

- **Recruitment email:** Included brief information about the AI Selfie App, study purpose, a Calendly link for scheduling (Calendly, LLC, 2025), session duration (approx. 15 minutes), and technical requirements (internet-enabled device).
- **Pre-session email:** Contained a link to the app, a Google Forms questionnaire (Google LLC, 2025a), and a folder with example selfies from Pexels (Pexels GmbH, 2025), illustrating suitable and unsuitable inputs (see Appendix C).
- **Remote setup:** Sessions were conducted via Google Meet (Google LLC, 2025b) with participant consent. Recordings were used for analysis and deleted afterward; observational notes were also taken.
- **Moderation guide:** A consistent script was used across sessions. As the process became familiar, it was replaced by a simplified checklist of key points.

The evaluation of the explainable AI (XAI) feature was excluded from these usability sessions, as it was assessed separately using the System Causability Scale (see chapter 3.1.4).

3.1.3 User Engagement Scale (Short Form)

The User Engagement Scale Short Form (UES-SF) was used to measure participants' engagement and to evaluate the the second research question (RQ2 in subsection 1.4). The UES-SF is a validated 12-item questionnaire covering four dimensions: Focused Attention, Perceived Usability, Aesthetic Appeal, and Reward (O'Brien et al., 2018). It was chosen because it provides a reliable and efficient way to assess engagement in digital contexts without excessive respondent burden. The scale has shown good validity and reliability in various applications, including web and mobile systems (Wiebe et al., 2014; Holdener et al., 2020). The questionnaire was implemented in Google Forms (Google LLC, 2025a) and evaluated in Google Sheets, which was connected to the Google Forms. Following best practices (O'Brien et al., 2018), items were presented in randomized order to mitigate potential order effects and response bias. In addition to the engagement items, the form included fields for participants' name, age, and an optional comment. However, only age data were used in anonymized form for analysis purposes and the names were collected solely for organizational reasons. To ensure clarity and minimize potential misunderstandings, especially given that most participants were German speakers, minor clarifications were added in brackets to some of the statements. The possible ratings are 1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, and 5 = strongly agree. The full list of all adapted items can be found in Appendix D.

3.1.4 System Causability Scale

To evaluate the third research question (RQ3 in subsection 1.4) the System Causability Scale (SCS) was used to evaluate how comprehensible and clear users perceived the explainable feature in the AI Selfie App. This instrument was chosen because it directly aligns with the research question, which focuses on understandability, and the SCS specifically measures the perceived quality and clarity of explainable AI features (Holzinger et al., 2020). In addition, the scale has been validated in peer-reviewed studies, demonstrating good reliability and construct validity across different contexts (Rinderknecht et al., 2024; Holzinger et al., 2020). The SCS was also developed as a practical tool for rapid assessment, making it suitable for projects with limited time resources available for evaluation procedures (Holzinger et al., 2020).

The SCS consists of 10 items measured on a Likert scale. Consistent with the UES-SF, the response options ranged from 1 (strongly disagree) to 5 (strongly agree). The questionnaire was administered using Google Forms and the connected Google Sheets. One item (originally "I understood the explanations within the context of my work") was adapted to "I understood the explanations within the context of my everyday life" to better suit the general public audience of the AI Selfie App. The exact statements included in the scale are provided in Appendix F.

3.1.5 Participants

A total of 26 volunteers scheduled participation in the study via a Calendly link (Calendly, LLC, 2025) where no compensation was provided. Ultimately, 24 individuals completed all parts of the evaluation. The sample included participants from a diverse range of ages, and professional backgrounds. Age was categorized into broad brackets (18–24, 25–34, 35–44, 45–54, 55–64, and 65+). An overview of the distribution of age groups and participant relationships is provided within Appendix E to ensure the anonymity of individual participants.

3.2 Data Access and Collection

While the previous section focused on the instruments and tools used to gather data, the following section outlines the practical procedures by which these instruments were deployed in the field and how the data was collected. It describes how the usability-sessions were conducted and moderated, and how specific tools such as surveys or usability protocols were administered.

3.2.1 Usability Test

Building on the instruments and communication procedures outlined in chapter 3.1.2, the usability testing sessions were conducted remotely via Google Meet and recorded with the consent of the participants. Each session followed a structured protocol to ensure consistency and comparability across participants.

During the usability session the purpose of the study was clearly communicated to the participants including an introduction into the project. To create a comfortable and realistic testing environment, participants were encouraged to think aloud while using the application, completing two core tasks as they naturally would:

- **T1:** Upload your selfies (or the provided example selfies) to create a profile.
- **T2:** Generate images with a sample profile.

In addition to real-time observation and note-taking during the sessions, the recordings were reviewed afterward to support a more detailed analysis. This included noting the time taken to complete each task, the transition point between the two tasks, and other temporal or behavioral patterns relevant to evaluating the interface flow and user engagement.

The usability evaluation followed a two-phase iterative format. After the initial cohort of participants completed their sessions, the collected data were reviewed to identify recurring usability issues. While time constraints limited the extent of modifications, a series of focused interface improvements were implemented before proceeding with the second round of tests. The rest of the inputs were added to the backlog for future work.

A detailed description of the usability test procedure and the changes implemented between the two phases is provided in chapter 6.1.

3.2.2 Administration and Evaluation of the UES-SF

Upon completing the primary usability tasks, participants were asked to fill out a questionnaire before leaving the Google Meet session. The UES-SF was administered using Google Forms, and responses were automatically recorded in a linked Google Sheets document for further processing. Data analysis followed the guidelines provided by O'Brien et al., 2018:

- The three items associated with the dimension **Perceived Usability** (PU-S.1, PU-S.2, PU-S.3) were reverse-coded to align their scoring direction with the other subscales. dimensions.
- Subscale scores were then computed by averaging the responses within each of the four dimensions.
- An overall engagement score was calculated as the mean of all 12 item responses.

Descriptive analyses were conducted separately for participants who tested the initial version of the application and those who interacted with the revised version. Responses entered in the optional "comment" field of the form were excluded from the quantitative UES-SF analysis. Instead, relevant comments were added to the usability observation notes to enrich qualitative insights.

3.2.3 System Causability Scale

Following the initial usability testing, participants were invited via email to participate in an additional evaluation focused on the explainable AI feature. During this phase, participants were instructed to interact with the XAI feature until they achieved a satisfactory level of understanding. Subsequently, they completed the System Causability Scale questionnaire, which consisted of 10 statements designed to measure their perception of the system's comprehensibility. From all participants who were contacted, five submitted the completed questionnaire.

3.3 Project management

In parallel to the research activities, the software development process was guided by the principles of the Agile Manifesto (Beck et al., 2001). The project adopted a flexible, values-oriented approach that prioritized iterative development, early delivery of working software, continuous feedback, and close collaboration with stakeholders. This way of working was chosen to accommodate evolving requirements and to maintain adaptability throughout the project duration, which is particularly important in research-oriented and time-constrained academic settings (Manamendra et al., 2013).

4 Research Findigs

This section presents the relevant findings from the research methods detailed in chapter 3.1.1, undertaken to address the research questions. The investigation focused on three primary areas:

1. UI/UX design principles tailored for public-access applications characterized by short interaction periods.
2. Consent protocols within the context of high-interaction environments.
3. Explanation approaches for AI image generation.

4.1 UI/UX Design Findings for Applications in Public Exhibitions

This section provides a summary of important research findings concerning the design of user interfaces (UI) and user experiences (UX) for public exhibitions or applications.

4.1.1 User-Centric Design and Environmental Integration

A key principle for applications in public spaces is a user-centered design, which considers the diverse user population and the specific environment.

- **Diverse User Population:** Effective design must accommodate a wide spectrum of users, including casual or first-time interactors and individuals with varying levels of technological proficiency and confidence (Maguire, 1999; Tang et al., 2024).
- **Temporal Constraints:** Applications must be engineered to deliver information or services with exceptional speed and efficiency, recognizing that users in public settings frequently operate under severe time limitations (Maguire, 1999).

4.1.2 Information Presentation and Output Design

The clarity, readability, and effectiveness of information presentation directly impact user comprehension and satisfaction.

- **Instructional Clarity:** On-screen instructions must be concise and context-specific, often complemented by graphical representations of interface elements (Maguire, 1999).
- **Text Readability and Language:** Textual content must be easily readable (minimum 16 point, sans-serif fonts, high contrast) and free of jargon or technical terms unless adequately explained within context (Maguire, 1999).
- **Contextual Help: Concise, context-related help information** should be readily accessible with a single action (e.g., a clearly labeled "help" button) and interruptible at any point (Maguire, 1999).
- **System Feedback:** For system response times exceeding 2-3 seconds, **clear feedback indicating ongoing processing** (e.g., "please wait" messages, progress bars) is necessary to manage user expectation (Maguire, 1999).

4.1.3 Use of Visitors' Own Devices

While providing a dedicated device with the app installed is advantageous, it can be restrictive when multiple visitors are present simultaneously. An alternative approach involves integrating visitors' personal mobile devices into the experience. Data from the Eurobarometer 510 survey by the European Commission indicates that 96% of EU respondents possess mobile phones, highlighting their ubiquitous presence in daily life and likelihood of being carried during public activities (European Commission, 2021). Consequently, numerous museums and researchers are exploring methods to leverage the advanced computational, communication, and display capabilities of these devices to enrich the exhibition experience. (e.g., Barbosa et al., 2015, Rhee and Choi, 2015, Lyons, 2009, O'Hara et al., 2007).

4.2 Consent in High-Interaction Environments

High-interaction environments, characterized by frequent, spontaneous interactions and continuous, passive data collection, pose significant challenges for privacy consent. Traditional consent methods, such as lengthy privacy policies, are unsuitable for these fast-paced settings because the time required to engage with them far exceeds the brief interaction duration (Windl et al., 2025).

Effective consent mechanisms in these environments are guided by two core principles:

- **Conserving User Attention:** Human attention is a finite resource. Overuse, particularly through repetitive warnings, can lead to "habituation," where users dismiss prompts without comprehension (Felt et al., 2012). Therefore, attention should only be demanded when consequences are severe (Felt et al., 2012).
- **Avoiding Interruptions:** Security and privacy decisions often interrupt users from their primary tasks, reducing motivation to engage and encouraging quick dismissals (Felt et al., 2012). Consent mechanisms should integrate seamlessly into workflows (Felt et al., 2012; Wegdam and Plas, 2008; Windl et al., 2025).

4.2.1 Spectrum of Consent Mechanisms

Consent mechanisms range from explicit to implicit, impacting user effort and consent transferability. (Windl et al., 2025). Explicit consent methods afford users granular control and ensure informed decision-making; however, they also necessitate significant user attention, potentially leading to fatigue or frustration (Felt et al., 2012; Wegdam and Plas, 2008; Windl et al., 2025). Interactive consent tries to fix these problems by asking users for consent only when needed and letting them apply their choices to future situations (Felt et al., 2012; Wegdam and Plas, 2008). These methods might use physical controls or digital elements, like menus showing data use or icons indicating data collection. By balancing user comfort and ease of use, they offer a good compromise.

Conversely, implicit consent mechanisms necessitate minimal user involvement and are maximally conducive to the transferability of consent. These systems typically exhibit context sensitivity, discerning user preferences and autonomously modulating settings based on observed behavior (Windl et al., 2025). Contextual consent models within this paradigm empower users to predefine privacy zones or regulations predicated on location or activity. More sophisticated implementations encompass automated assistants capable of interpreting situational behaviors to render privacy decisions on behalf of users. While implicit mechanisms proffer considerable convenience, they presuppose a substantial degree of trust in the system and may precipitate a perceived erosion of autonomy and awareness (Windl et al., 2025). To mitigate this potential drawback, it is advisable for users to periodically review and adjust their preferences to sustain a sense of control.

4.2.2 Contextual Factors and Design Considerations

The suitability of a consent mechanism depends on various contextual factors (Windl et al., 2025):

- **Scenario Dynamics:** Static scenarios with consistent activities may suit explicit mechanisms if data requests are infrequent. Dynamic scenarios benefit from transferable (semi-implicit or implicit) mechanisms to avoid information overload.
- **Frequency of Data Requests:** High-frequency requests, even in static contexts, require mechanisms that reduce constant.
- **Sensitivity of Information:** More sensitive data demands greater user control, often favoring semi-implicit options. Less sensitive data may be suitable for implicit mechanisms.
- **Trust Level of the Environment:** Private spaces may accommodate explicit consent, while public spaces often require discrete explicit or semi-implicit mechanisms to prevent social discomfort or security risks from observable interactions. The presence of unfamiliar people can also shift perceptions of trust.

4.3 Explanation Techniques for AI Image Generation

Many AI systems, particularly generative models, operate as "black boxes," making it difficult for users to grasp how outputs are produced (Ehsan and Riedl, 2020; Haque et al., 2023). Explainable AI (XAI) addresses this challenge by providing human-understandable insights into system behaviour. In this context, human-centred approaches are essential, as explanation needs vary depending on the audience as non-expert users may require simplified, visually supported narratives rather than technical details (Ehsan and Riedl, 2020).

Effective explanations can be delivered through multi-level approaches that combine high-level process overviews with step-by-step visualisations of model operations. For example, visualising how Stable Diffusion refines noise into an image over multiple timesteps can help users build a mental model of the generation process (Lee et al., 2024). Interactive elements, such as parameter adjustments or timestep controls, further support understanding by allowing users to explore cause–effect relationships between inputs and outputs (Kim et al., 2025; Lee et al., 2024).

Explanations may also clarify the influence of user inputs, such as how small changes in a text prompt affect the resulting image (Kim et al., 2025). Techniques like feature attribution or saliency maps can highlight which parts of an input image or prompt most strongly shaped the output, while natural language descriptions can translate these technical insights into more accessible explanations (Hudon et al., 2021).

Beyond clarifying processes, explanations help set realistic expectations, communicate system capabilities and limitations, and foster user trust (Bussone et al., 2015; Cai et al., 2019; Eiband et al., 2019). Transparent, well-designed visualisations can reduce cognitive load and make the interaction more engaging, supporting both learning and creativity. In creative contexts actionable explanations—offering guidance on how to adjust prompts or inputs—can encourage experimentation while also promoting critical reflection on the system's outputs (Ehsan and Riedl, 2020).

5 Implementation

This chapter documents the technical implementation of the AI Selfie App. It outlines how the system was developed and deployed, how its main features such as anonymous authentication, the consent process, the explainable AI feature, and image generation were implemented, and how key design decisions were shaped by the research findings. The structure follows the core components of the application, from frontend development and authentication logic to content management and final user flow.

5.1 Frontend

The frontend interface was developed using Tailwind CSS, with certain adjustments made to the initial design to optimize the user experience and streamline the interaction process, reflecting modifications made during development.

5.2 Usage of the App

Based on the findings discussed in chapter 4.1 regarding temporal and infrastructural constraints, the decision was made to deploy the application as a web application rather than a native mobile app. Given that an active internet connection is essential for the application's core functionality, this strategy negates the necessity for installation on end-user devices and facilitates immediate access via the URL https://ai-selfie-app.expo.app, which can be turned into a QR Code and used on other media. To restrict public access while maintaining ease of deployment, a passcode gate was implemented, as detailed in chapter 5.3. Upon opening the web application, users are prompted to enter a code as seen in Figure 5.1. These codes are stored and managed manually in the database, enabling authorized personnel to add, modify, or deactivate access credentials as needed. This ensures that while the URL remains publicly accessible, the core functionalities are restricted to verified users only.

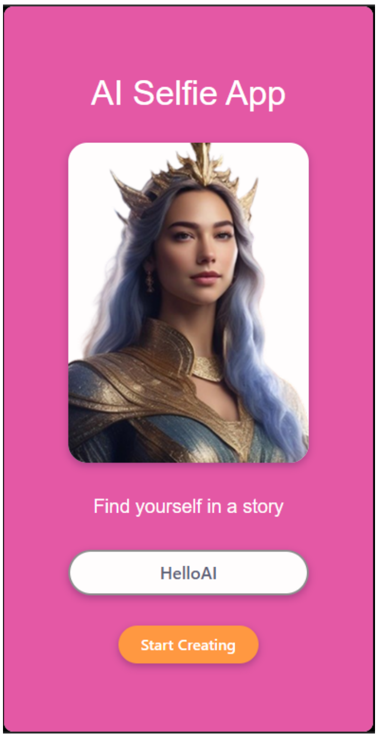


Figure 5.1: Home screen with access-code input.

5.2.1 Deployment via EAS Hosting

The project was deployed using EAS Hosting („EAS Hosting - Expo Documentation“, 2025), a platform provided by Expo for quickly publishing web projects built with React Native and the Expo Router. The deployment process involves logging into the Expo account via the CLI, exporting the application using:

Listing 1: Bash: Exporting the app for the Web platform

```
npx expo export --platform web
```

This generates a production-ready build in the `dist` directory. The deployment can then be triggered with:

Listing 2: Bash: Deploying the app using EAS

```
eas deploy
```

This command uploads the exported project to EAS Hosting, after which a URL is provided. Multiple releases can be deployed and managed simultaneously.

5.2.2 Platform-Specific Code Adjustments

Most of the application logic remained unchanged during the transition to a web-based deployment, thanks to Expo's cross-platform capabilities (Expo, n.d.). However, some platform-specific functionalities had to be adjusted, for example the logic used for uploading files to Amazon S3 which is the saving place for the ai model server.

In the mobile version (`aws-service.ts`), the file upload process relied on React Native's handling of file objects using URIs and the `Platform` module:

Listing 3: Native (mobile): Uploading zip file using file object

```
const fileObject = {
  uri: Platform.OS === 'android' ? `file://${zipFilePath}` : zipFilePath,
  type: 'application/zip',
  name: 'photos.zip',
};
formData.append('file', fileObject as any);
```

This approach is not supported in browsers, as the web platform requires a `Blob` or `File` object for uploads via `fetch` and `FormData`. Therefore, a web-specific implementation was created in a separate file (`aws-service.web.ts`), utilizing Expo's platform-specific file naming convention. The upload function was rewritten as follows:

Listing 4: Web: Uploading zip file using Blob and File object

```
const zipBlob = await fetch(zipBlobUrl).then(r => r.blob());
formData.append('file', new File([zipBlob], 'photos.zip', {
  type: 'application/zip',
}));
```

This ensures compatibility with standard browser APIs and enables successful uploads to the pre-signed S3 URL provided by the backend.

5.3 Anonymous Sign-In

Although the web application is technically accessible via a public URL, access to its functionality is restricted through an authentication mechanism. To achieve this, anonymous sign-in using Supabase Authentication was implemented („Anonymous Sign-ins with Supabase Auth“, 2025). When a user accesses the application, an anonymous sign-in request is triggered automatically. Supabase then creates a temporary user session and assigns a unique `user_id` to that session. This user ID is stored in the database and used to associate the user with a profile entry, without requiring any personally identifiable information (PII). This approach ensures privacy while enabling the system to manage user-specific data such as uploaded files, generated outputs, and session states.

As shown in Figure 5.2, a new attribute was added to the `profile` table to store the `supabase_user_id` generated during the anonymous authentication process. This allows the backend services to retrieve and update the correct profile data throughout the user's interaction with the app.

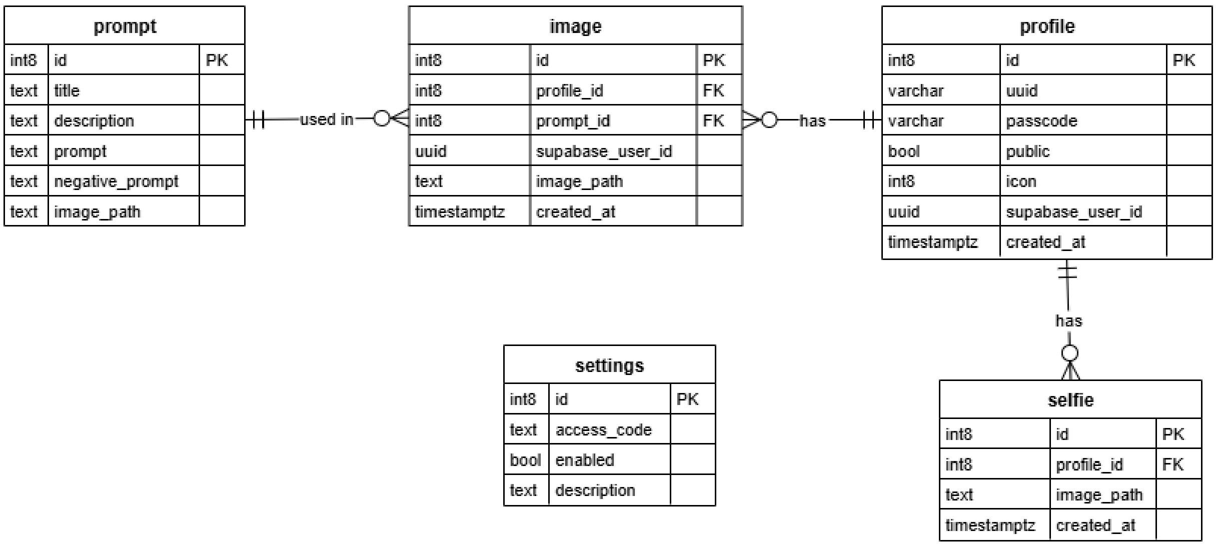


Figure 5.2: Updated database schema and table relations.

5.4 Sample Profile

To support users who prefer not to upload their own images, a publicly available sample profile was introduced. This allows users to experience the application's image generation functionality without providing personal data. To enable this feature, a new attribute named `public` was added to the `profile` table (see Figure 5.2). This boolean flag indicates whether a profile should be accessible to all users. Since the decision to make a profile public should be limited to authorized personnel, the attribute can only be modified through the Supabase interface by an authenticated developer or administrator. When an anonymous user signs in and selects the sample profile, they can generate images as if using a normal profile. However, to ensure privacy and data separation, each image generated using the public profile is linked to the individual anonymous user. For this purpose, a new attribute `supabase_user_id` was added to the `image` table, also visible in Figure 5.2. This structure ensures that although multiple users may interact with the same public profile, the generated images are stored separately and only visible to the user who created them. Thus, the experience remains personalized, and no generated content is shared between users.

5.5 Consent Process

Although the application collects minimal user data due to its anonymous authentication model, an explicit consent mechanism becomes necessary when users choose to upload personal images. In this case, the images may depict identifiable individuals, introducing a level of data sensitivity that warrants user agreement. To address this requirement while maintaining a smooth user experience, a lightweight, interactive consent mechanism was implemented. The design was directly informed by the key research findings in chapter 4.2 regarding privacy and consent in high-interaction environments.

Consent is requested at the moment a user chooses to upload their selfies. This ensures that attention is only demanded at the point where sensitive data is about to be shared. The user is required to confirm that the individual depicted in the images (whether themselves or someone else) has consented to the upload and the processing of their data. The consent dialog explicitly communicates this responsibility to the user, providing clarity while avoiding lengthy or overly technical language.

To integrate the consent step smoothly into the user flow, the mechanism was designed as a modal dialog (see Figure 5.4) that appears when the user selects the option to upload their selfies (see Figure 5.3). The modal contains a brief, clearly worded consent statement accompanied by a question mark icon. Clicking the icon opens a secondary dialog (see Figure 5.5) containing additional information about data usage, deletion policies, and privacy considerations. If the user agrees to the terms, they can proceed to the image upload step. If they do not agree, the user cannot continue with the userflow and is forced to close the modal, where they can instead select are redirected to an alternative option to use the publicly available sample profile instead. This ensures that no image data is collected without informed agreement. The system enforces the consent process programmatically: the upload flow cannot proceed unless the user has confirmed agreement. No images are transmitted to the server unless explicit consent has been given, and a fallback path is always available.

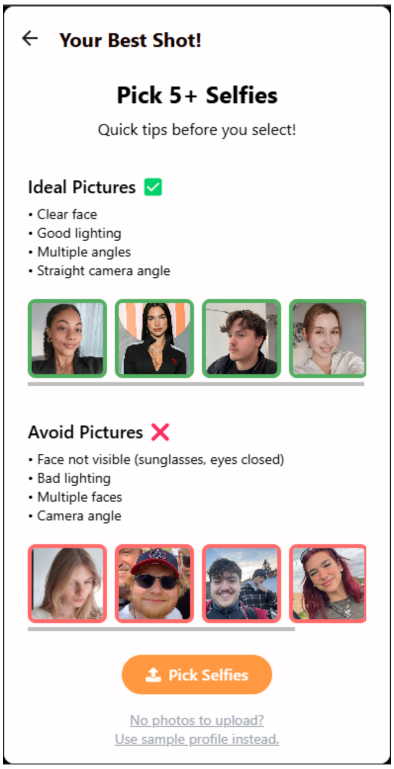


Figure 5.3: Image guidelines with good/bad examples.

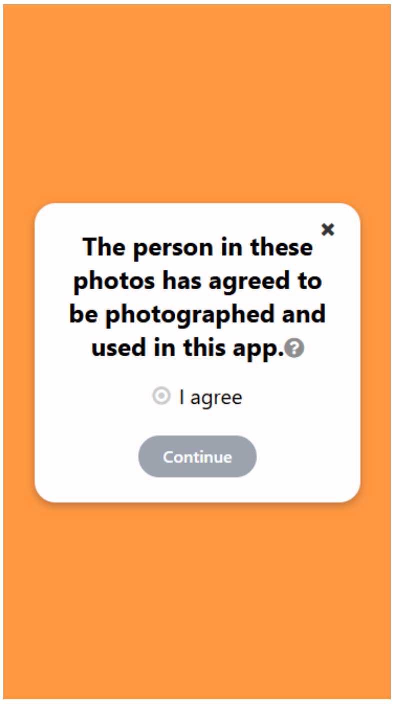


Figure 5.4: Consent dialog shown before uploading personal selfies

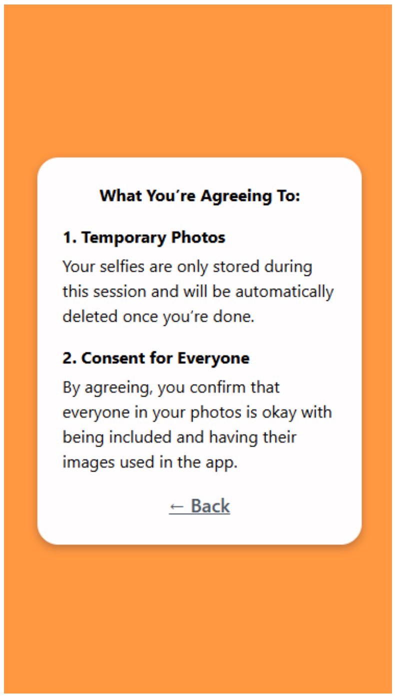


Figure 5.5: Detailed information dialog accessed via the question mark icon

5.6 Daily Cleanup with Supabase Cron

To uphold the promise of data privacy and ensure that all uploaded selfies and generated images are removed after a short period, two automated cleanup jobs were implemented using Supabase Cron („Supabase Cron – Schedule Recurring Jobs in Postgres“, 2025). These cron jobs are scheduled to run daily at 01:00 AM. The first job is responsible for deleting all database entries from the `profile`, `image`, and `selfie` tables that are older than 24 hours. To enable this functionality, a `created_at` timestamp column was added to each of the relevant tables, as shown in Figure 5.2. These timestamps are automatically assigned upon creation of each record, allowing for precise age-based filtering.

Database cleanup alone is insufficient, as actual image files, both uploaded selfies and generated images, are managed separately using Supabase Storage. Due to current limitations of Supabase's SQL access to storage buckets (fenos, 2022), a second seperate cron job was implemented to handle file deletion. This second job invokes a Supabase Edge Function, which is a server-side TypeScript function deployed within Supabase's infrastructure („Supabase Functions Documentation“, 2025). The function iterates over and deletes all the designated files within the storage bucket `profile`, which also includes all the uploaded selfies and generated images. Ensuring that both uploaded and generated image files are permanently removed from the system.

Together, these two cleanup processes ensure that user data is not retained beyond the intended usage period, maintaining user trust and complying with privacy expectations.

5.7 Explainable AI Image Generation Feature: Wizard

The design of the explainable AI feature took inspiration from narrative-driven video games, where user input significantly influences the storyline. A notable example of this approach can be seen in The Walking Dead game, where players make choices that shape the outcome of the narrative as seen in Figure 5.6. Similarly, the explainable AI feature was conceptualized as a multi-step process that guides users through the AI image generation journey.

Figure 5.6: Example of a narrative game where user choices steer the experience.

The wizard begins with a short introduction screen showing a looping animation that transforms noise into an image. The user is prompted with a simple message and a "Start" button to enter the experience, as shown in Figure 5.7. The intention is to make the transition into the explanation lightweight and engaging, in line with findings that recommend low-barrier entry points for time-limited contexts.

Each step in the wizard presents a visual from the Stable Diffusion process alongside a short explanation. These visuals were generated using the open-source Diffusion Explainer tool by the Polo Club at Georgia Tech Lee et al., 2024, which simulates the transformation from pure noise to a finished image. The text for each step is intentionally simple and non-technical, supporting the recommendation to use human-centered, narrative formats for explanation Ehsan and Riedl, 2020.

For users interested in more detail, a "show more" option expands the explanation inline without navigating away, as seen in Figure 5.8. This progressive disclosure structure was chosen to allow users to control the level of information based on their own curiosity or prior knowledge, as suggested in Ehsan and Riedl, 2020; Kim et al., 2025.

To illustrate the effect of user input, one of the steps prompts users to select a sample prompt (*Cityscape*, *Panda*, *Renaissance*), as seen in Figure 5.9. This reinforces the idea that the AI output is shaped by what the user inputs—an important part of the overall mental model.

Later in the wizard, a slider component allows users to manually scrub through intermediate images in the denoising process, as shown in Figure 5.10. This visual interaction helps communicate how the image becomes gradually clearer over time, which reflects research findings that interactive visuals can help users build an understanding of otherwise opaque model behavior Lee et al., 2024.

To manage and store the content of each step, the information, including titles and explanatory text, is temporarily saved in a JSON file. All the steps and corresponding explanations are documented in the attached JSON file in Appendix G, providing a clear, simple and organized structure for the feature's implementation.

Figure 5.7: Wizard entry screen with a transformation animation and Start action.

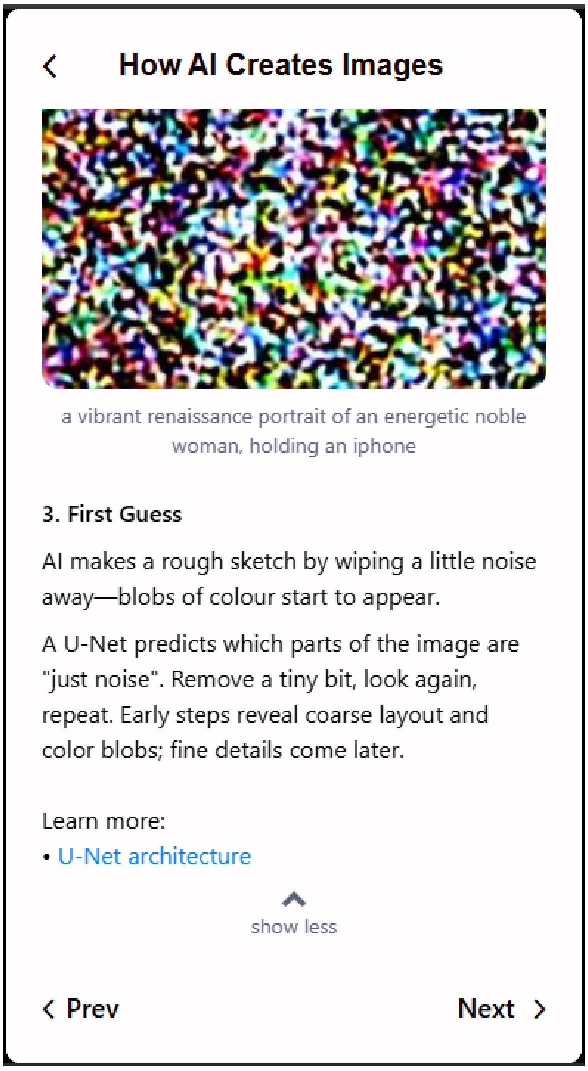


Figure 5.8: Step view with concise text; Expanded detail via show-more.

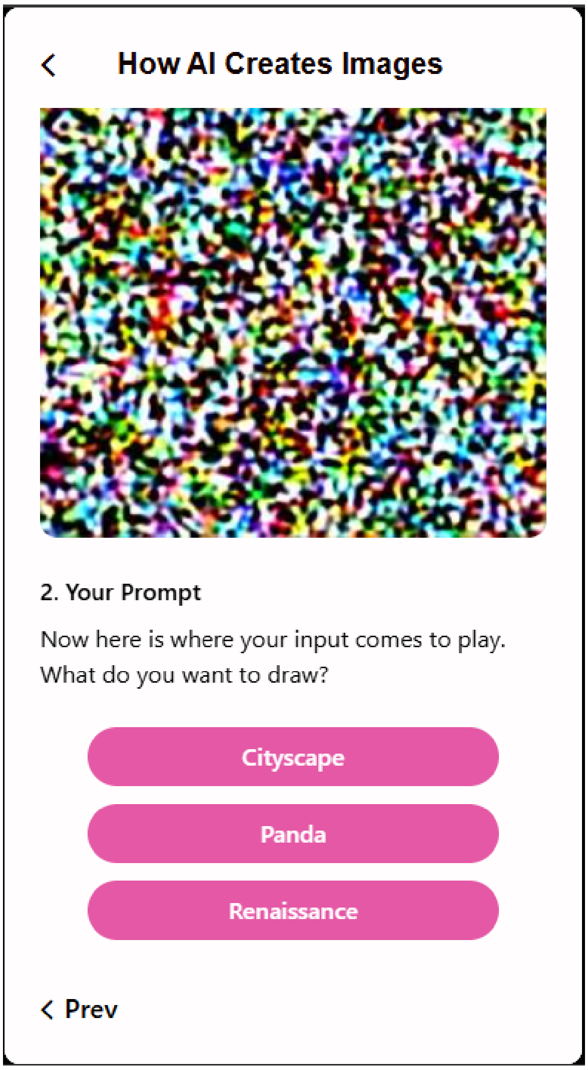


Figure 5.9: Prompt selection illustrates how user input steers generation.

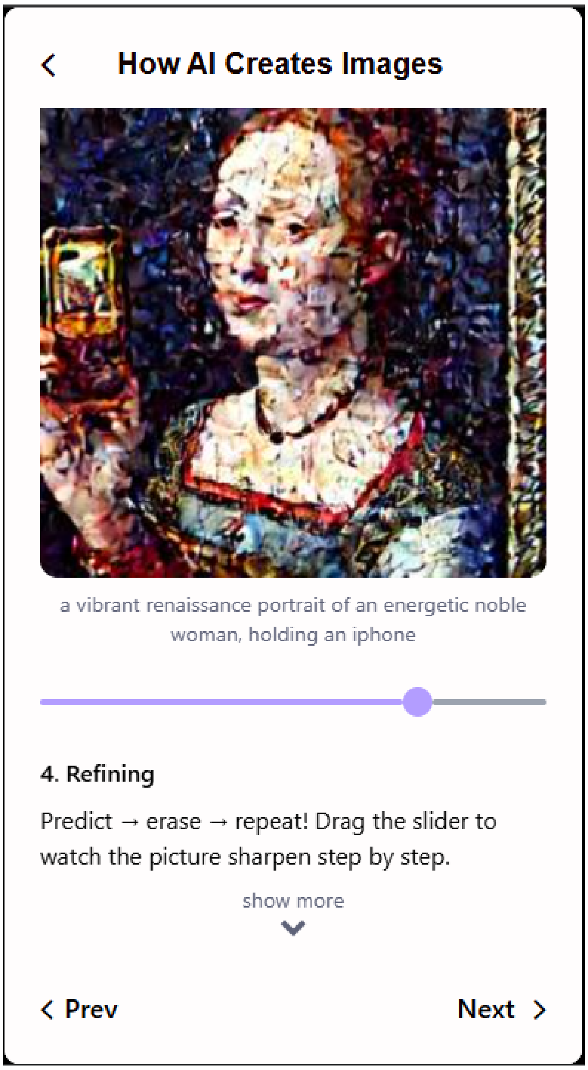


Figure 5.10: Refinement slider revealing iterative denoising across timesteps.

5.8 Final User Flow

The final user flow describes how visitors access the web application, provide consent and images, create a profile, generate results, and optionally proceed using a pre-configured sample profile. Each step references the corresponding screen.

5.8.1 Access and Entry

Visitors arrive via a shared link or by scanning a QR code and land on the home screen with an access-code prompt, as shown in Figure 5.11. After entering a valid code, they are redirected to the image guidelines screen, which explains acceptable and unacceptable selfie examples, as shown in Figure 5.12.

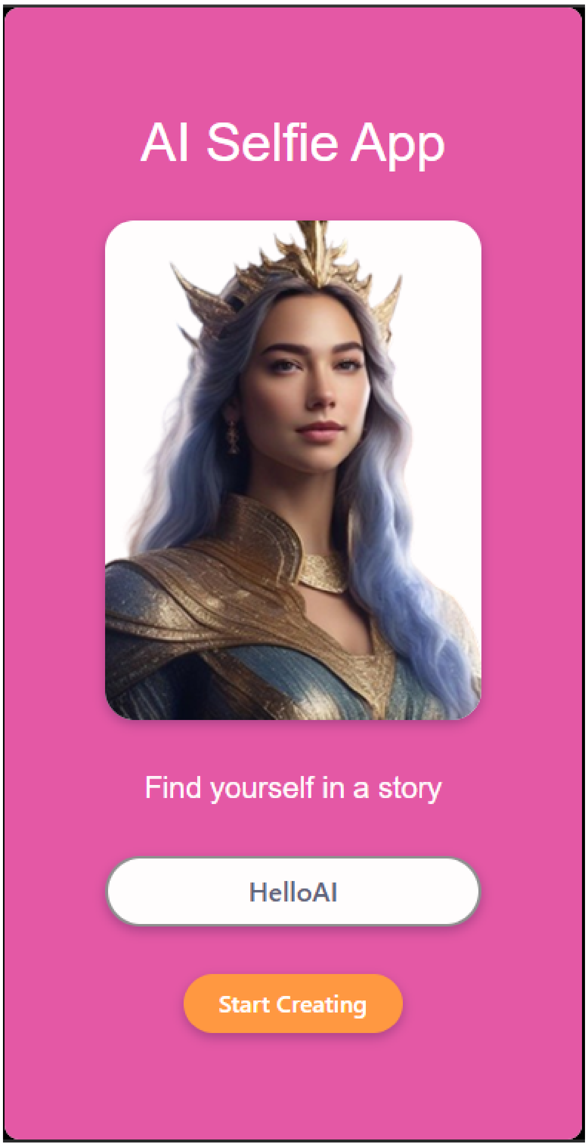


Figure 5.11: Home screen with access-code input.

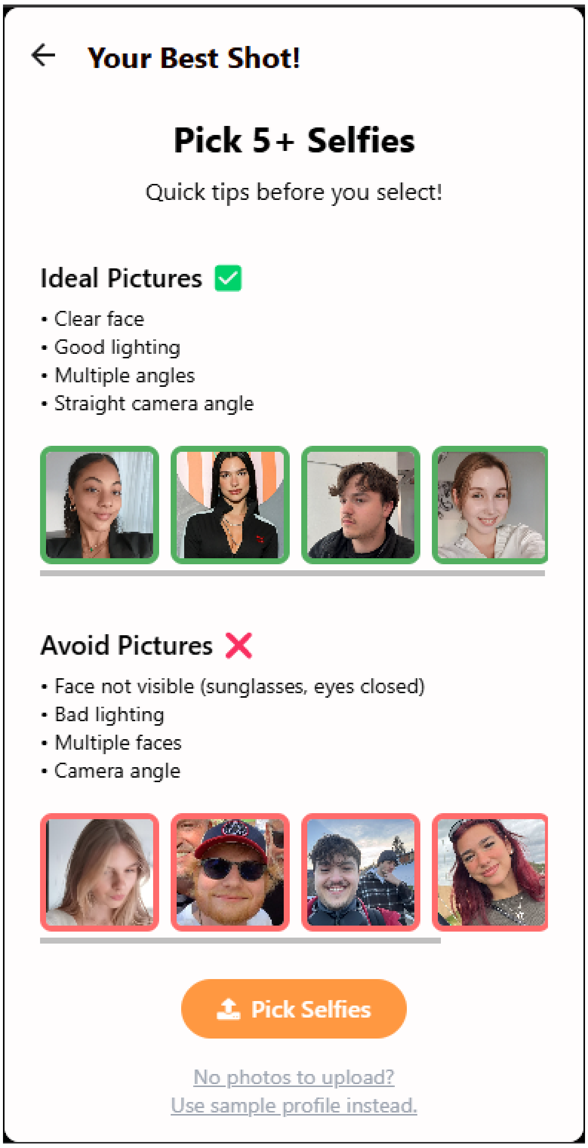


Figure 5.12: Image guidelines with good/bad examples.

5.8.2 Uploading Own Images

If the option to upload personal selfies is chosen, selecting the pick selfies action opens the consent modal, as shown in Figure 5.13. Only after explicit consent is given does the browser invoke the native file picker. Selected images are then analyzed and visually separated into acceptable and not recommended groups to support the minimum-quality requirement, as shown in Figure 5.14. Once at least five acceptable images are confirmed, the upload can be completed. The user is redirected to the profile overview, as shown in Figure 5.15, where a transient confirmation appears indicating successful profile creation, the approximate training duration of twenty minutes, and a brief introduction to the explanation feature. After timeout or manual dismissal, the overview shows all profiles, including the newly created profile in training state and the sample profile.

If the option to upload personal selfies is chosen, selecting the pick selfies action opens the consent modal, as shown in Figure 5.13. Only after explicit consent is given does the browser invoke the native file picker. Selected images are then analyzed and visually separated into acceptable and not recommended groups to support the minimum-quality requirement, as shown in Figure 5.14. Once at least five acceptable images are confirmed, the upload can be completed and the user is redirected to the profile overview. A toast notification confirms successful profile creation and the newly created profile card displays an overlay with the training status and an estimated remaining time as shown in Figure 5.15. On the same screen a contextual prompt beneath the cards provides a link to the explanation feature. The toast auto-dismisses after a short delay or on user dismissal; the overview continues to show all profiles, including the profile in training and the sample profile.

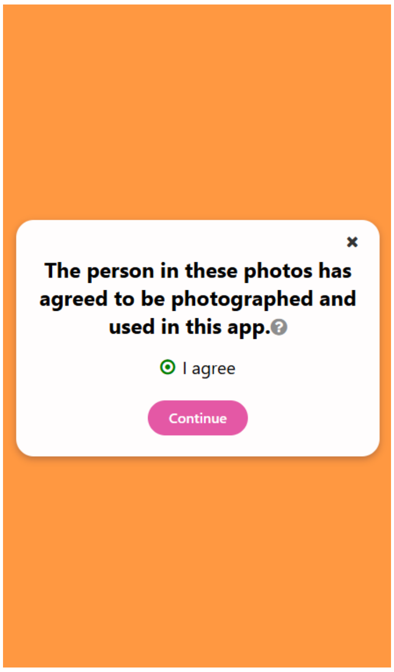


Figure 5.13: Consent modal.

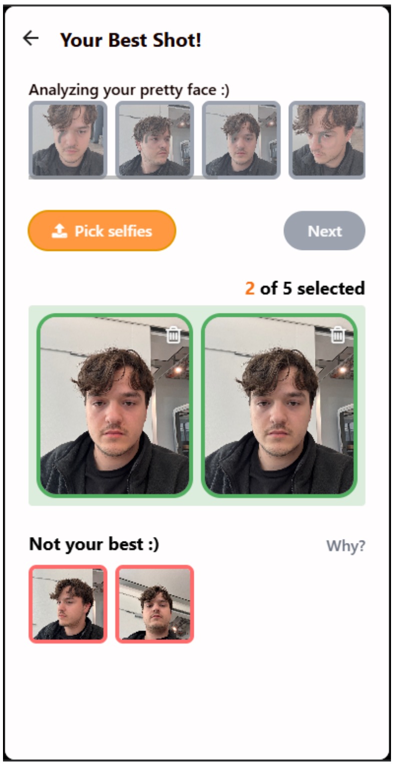


Figure 5.14: Upload review separating acceptable vs. not recommended images.

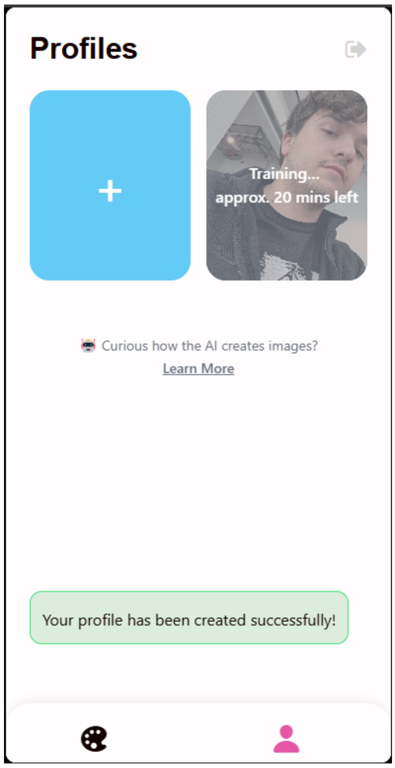


Figure 5.15: Profile overview after creation: success toast, training overlay (ETA), and Learn More link.

5.8.3 Generating Images

Images can be generated either from the profile detail view via a direct generate action, as shown in Figure 5.16, or from the explore view listing available prompt styles, as shown in Figure 5.17. After choosing a prompt style, a trained profile is selected and optional attributes such as gender are set, as shown in Figure 5.18. Generation proceeds in batches of four with a short processing time and the results are presented in a gallery with the option to zoom for closer inspection, as shown in Figure 5.19 and Figure 5.20. Users may generate additional variations or select preferred images and continue. The flow returns to the profile detail view, where generated images are organized by prompt and session for later review, as shown in Figure 5.21.

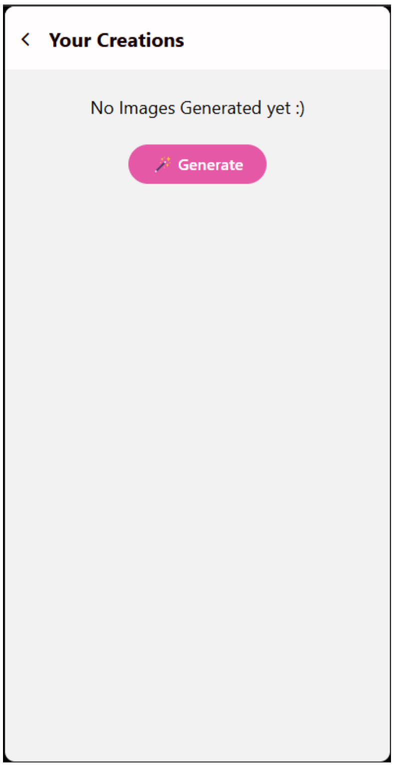


Figure 5.16: Profile detail view with direct generate action.

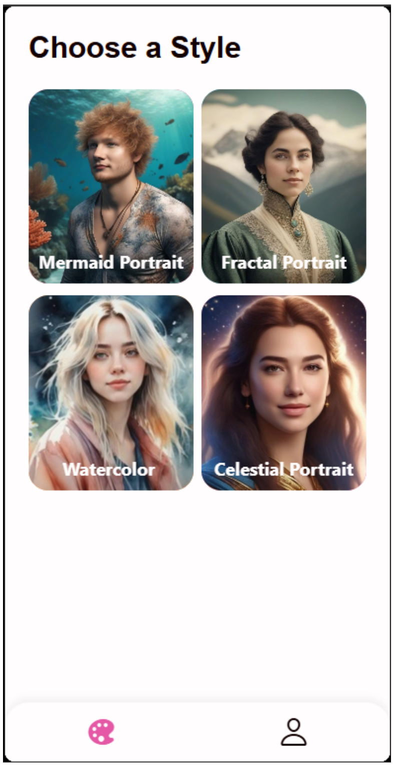


Figure 5.17: Explore view with prompt styles.

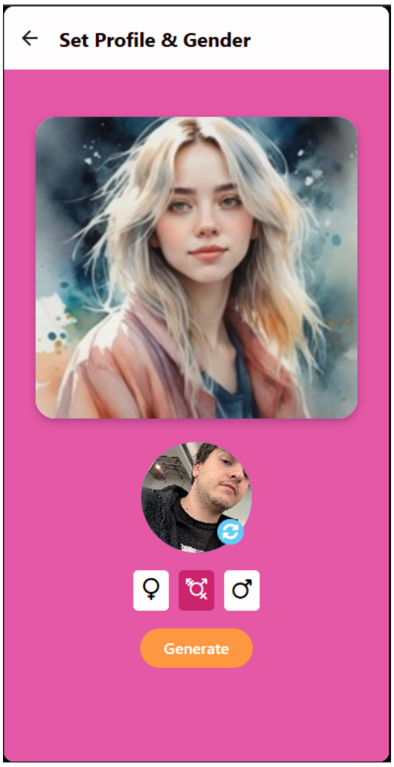


Figure 5.18: Prompt selection and profile chooser.

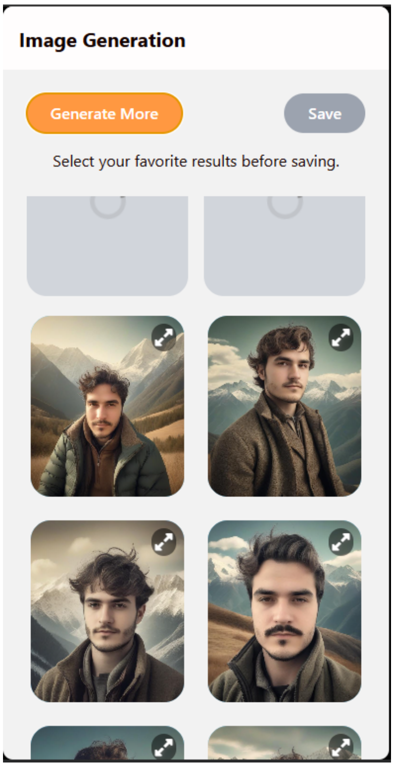


Figure 5.19: Generated results gallery.

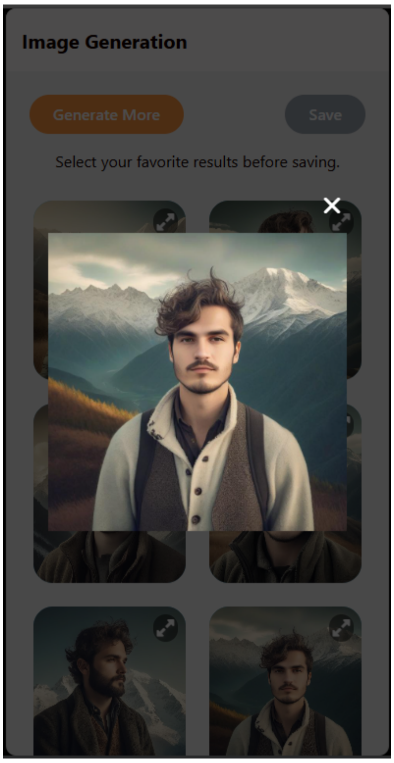


Figure 5.20: Zoomed image view.

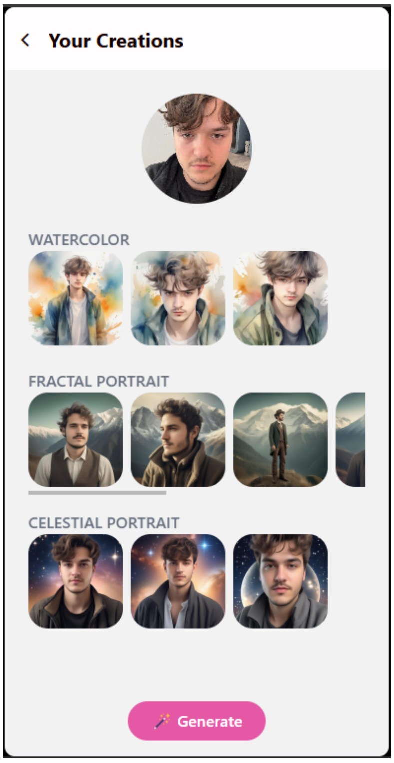


Figure 5.21: Results organized by prompt in profile detail.

5.8.4 Onboarding via Sample Profile

If the sample profile is chosen instead of uploading selfies, the flow jumps directly to the explore view of prompt styles, as shown in Figure 5.17. From there, the process matches the generation sequence described in subsubsection 5.8.3.

5.8.5 Source Code Repository

The complete project source code is hosted on GitLab:
https://gitlab.fhnw.ch/iit-projektschiene/hs24/24hs_iit41-creative-ai-selfie-app.
The repository serves as the canonical record of the implementation and includes the web application, configuration, and deployment assets. Setup and usage instructions are provided in the repository's README. Sensitive credentials are excluded and must be supplied via environment variables.

6 Evaluation and Results

This chapter reports the findings from three complementary evaluation methods: usability testing, the User Engagement Scale (Short Form), and the System Causability Scale. Usability testing compared the initial (Version A) and refined (Version B) designs, capturing observations, participant feedback, and task completion times. The UES-SF measured user engagement across attention, usability, aesthetics, and reward, while the SCS assessed the clarity and effectiveness of the application's explainable AI features.

6.1 Usability Testing

The usability testing was conducted in two phases, with a total of 24 participants: 13 participants tested the initial version (Version A), and 11 participants tested the refined version (Version B).

At the start of each session, participants were welcomed and reminded of the study's objective. To simulate a realistic and comfortable interaction context, participants were encouraged to use the application as they naturally would, completing two central tasks:

- **T1:** Upload your selfies (or the provided example selfies) to create a profile.
- **T2:** Generate images with a sample profile.

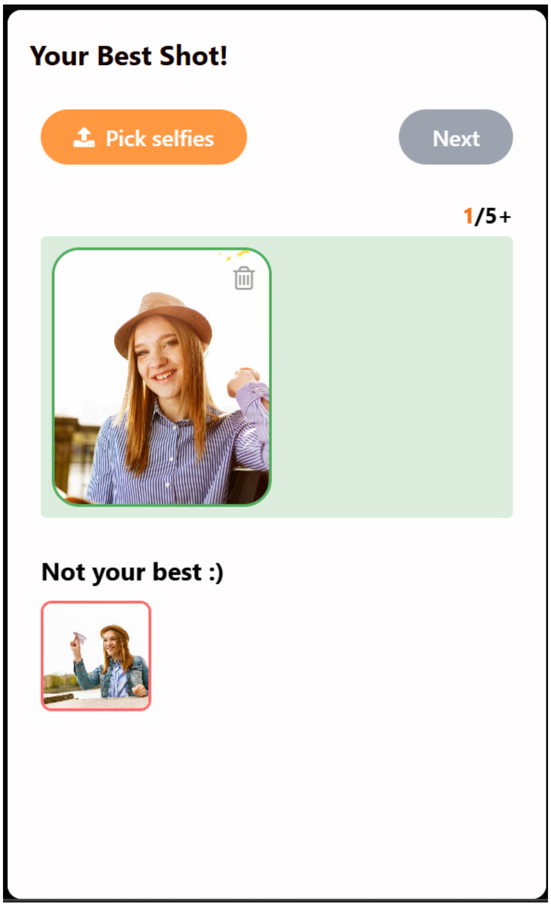
The participants were asked to think aloud while completing these tasks, which included freely verbalizing their thoughts while navigating the interface, articulating expectations when interacting with elements, and reading text aloud where helpful. Participants were instructed to proceed at their own pace and to imagine they were exploring the application independently in their usual environment. They were also advised to indicate explicitly when an activity had been completed or when they wished to stop, for example by stating "I'm done" or "I'd like to stop here." Throughout the session, it was emphasized that the study aimed to test the design itself rather than the participants' abilities, and participants were encouraged to share any difficulties openly. Participants were informed that notes would be taken during their interaction. It was explained that, although the facilitator might remain silent for extended periods, occasional questions or instructions could be provided to guide the session or to transition to the next activity.

6.1.1 Changes made

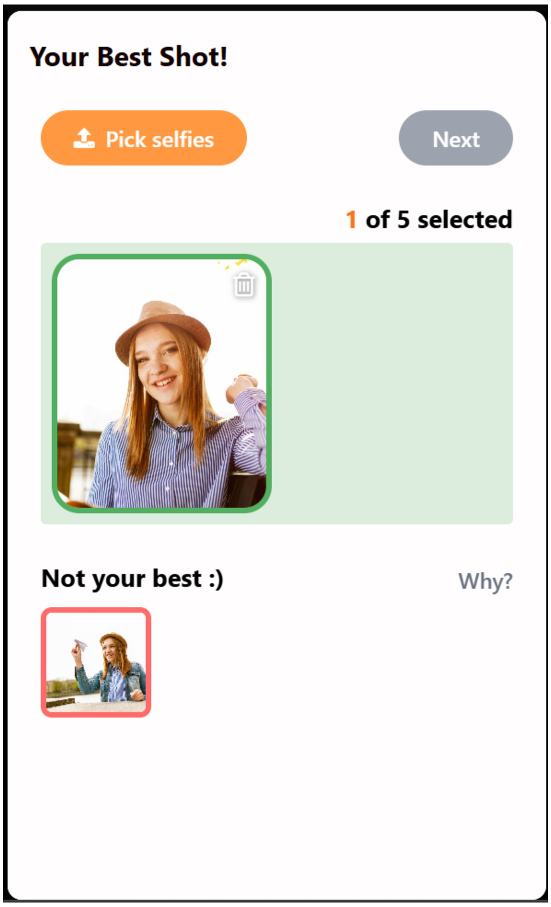
Based on the most frequently observed issues during the first round of usability testing, several targeted interface improvements were implemented between Version A and Version B:

- **Image Upload Screen:** Added clearer textual guidance and an explanatory help link to clarify image selection criteria (see Figure 6.1).
- **Training Feedback Display:** Included an estimated processing time indicator after image submission to manage user expectations (see Figure 6.2).
- **Profile Detail View:** Added a direct action button for initiating image generation from the profile screen (see Figure 6.3).
- **Image Generation Screen:** Introduced a zoom icon to allow closer inspection of generated images (see Figure 6.4).

These adjustments aimed to improve clarity, reduce cognitive load, and facilitate more efficient task completion.

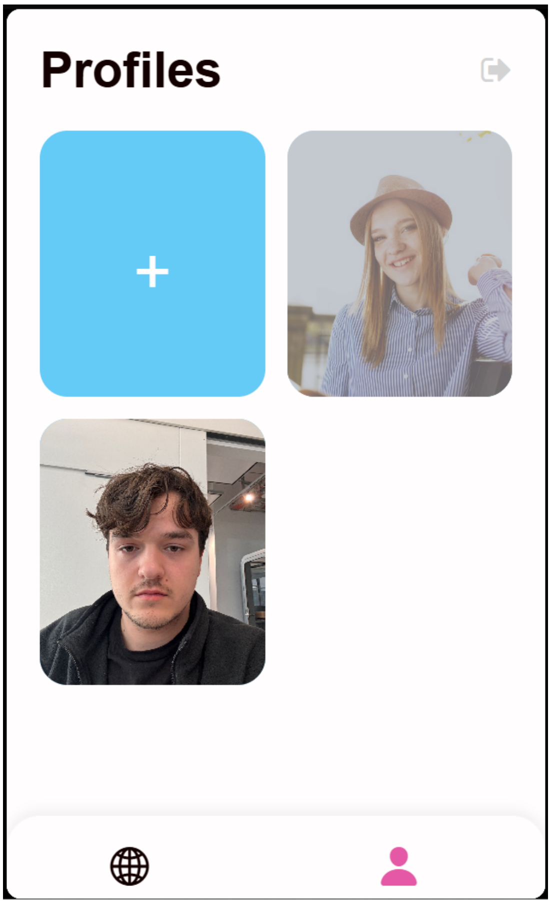


a: Version A: Image upload screen.

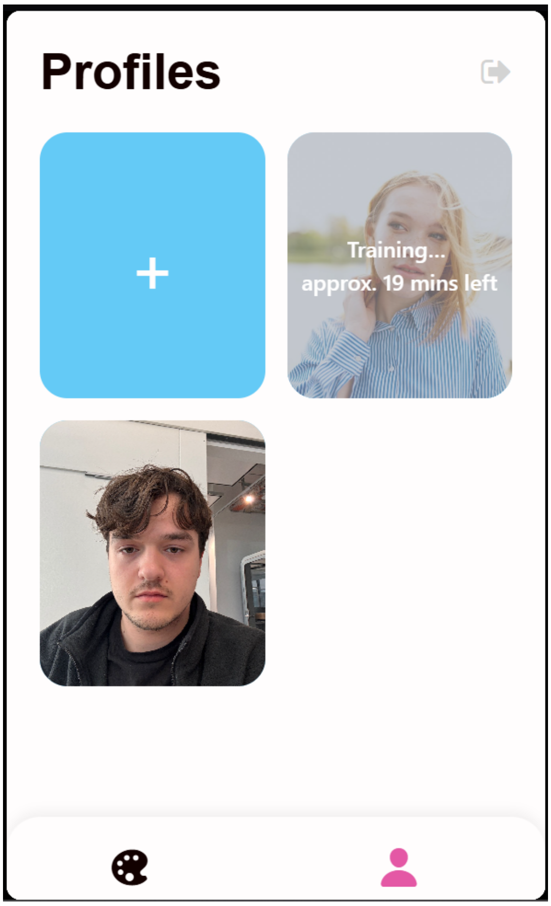


b: Version B: Image upload screen with clearer instructions and an explanatory help link.

Figure 6.1: Comparison of the image upload screen in Version A (left) and Version B (right), highlighting improvements to clarify image selection criteria.

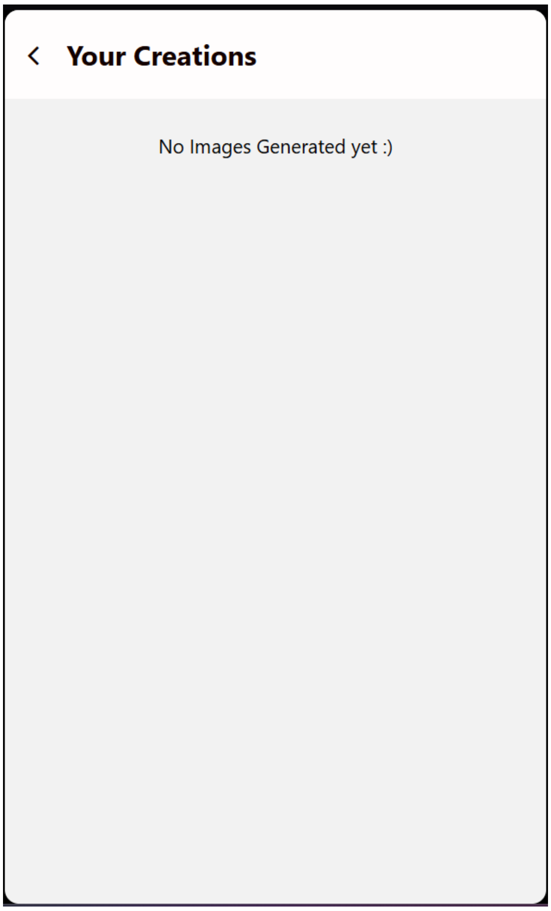


a: Version A: Training feedback display.

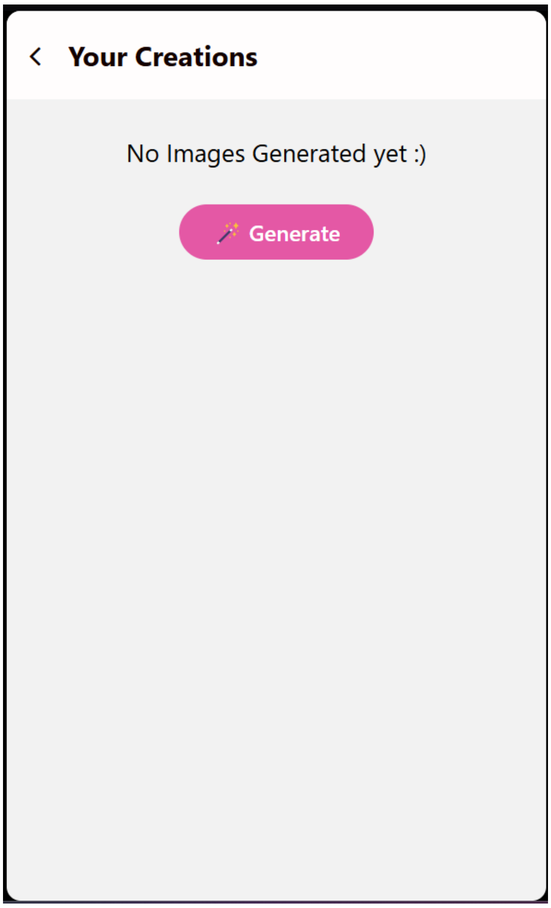


b: Version B: Training feedback display showing estimated remaining processing time.

Figure 6.2: Comparison of the training feedback display in Version A (left) and Version B (right), with the latter providing an estimated processing time after image upload.

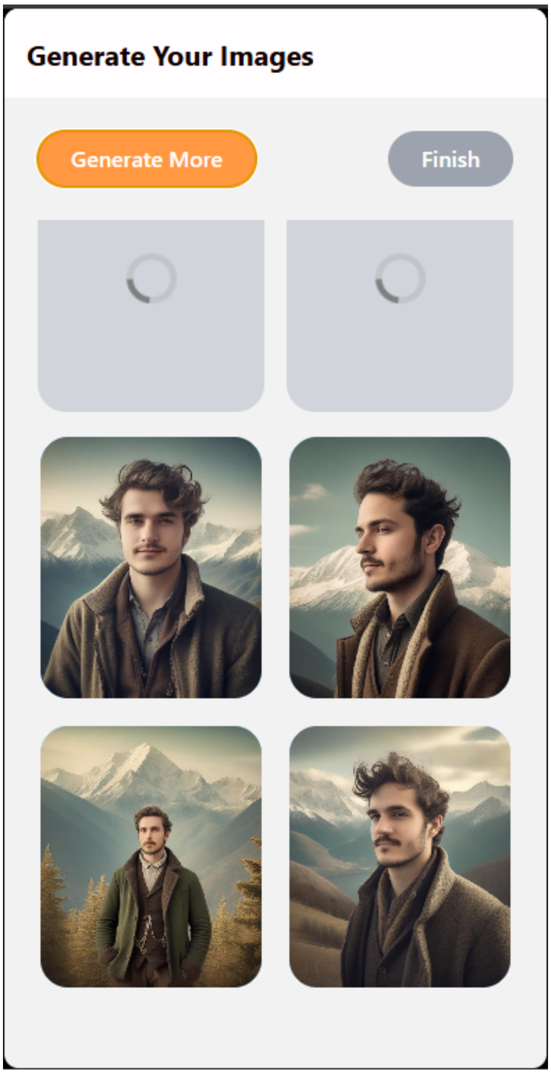


a: Version A: Profile detail view.

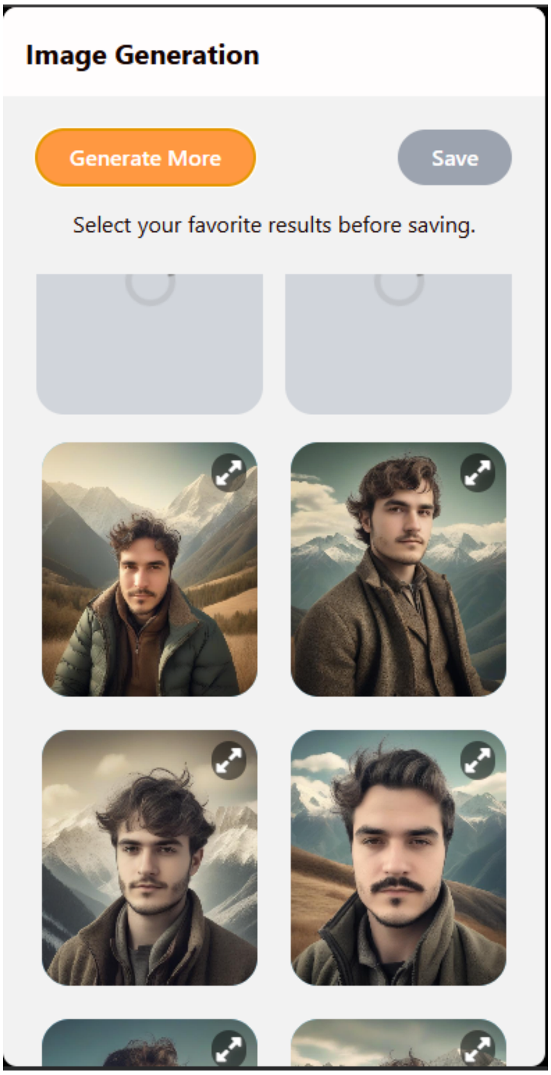


b: Version B: Profile detail view with a direct image generation button.

Figure 6.3: Comparison of the profile detail view in Version A (left) and Version B (right), showing the addition of a direct action button to initiate image generation.



a: Version A: Image generation screen.



b: Version B: Image generation screen with zoom function for generated images.

Figure 6.4: Comparison of the image generation screen in Version A (left) and Version B (right), showing the addition of a zoom option for detailed inspection of generated images.

6.1.2 Usability Testing Results

The following section summarizes key observations from the usability testing sessions. The findings are structured around core interaction moments within the app and reflect both verbal feedback and observed navigation behavior during the sessions.

Onboarding Flow and General Engagement: Both in Version A and B, participants generally found the onboarding flow intuitive and easy to understand. They were able to navigate through the initial setup without difficulty, and no major issues were reported during the introduction.

Image Upload Process: During the image upload process in Version A, **10 out of 13 participants** did not initially recognize the requirement to upload at least five images. After submitting a first image, several hesitated about the next step and repeatedly tried to click the `Next` button, noting that it appeared disabled: "I guess it is deactivated since it is greyed out, but I don't understand what to do."

Following adjustments to the interface and the slight adjustment inclusion of clearer instructions, in Version B, **9 out of 11 participants** understood the requirement to upload at least five images without additional prompting.

Post-Upload Navigation: In Version A, after uploading selfies participants were redirected to the profile overview, where the newly created profile appeared greyed out while training was pending. Despite being briefed at the start that training would take about 20 minutes, several participants attempted to click the inactive profile card and then moved back and forth between tabs before finding the prompt screen.

In Version B, clearer status messaging and labels (see Figure 6.2 and Figure 6.3) reduced confusion; participants verbalized their understanding (e.g., "Ah, now the training happens for 20 minutes, like you said"), and 10 of 11 participants navigated directly to the prompt screen without issue.

Sample Profile and Detailed View: Participants who explored the sample profile sometimes encountered confusion when accessing the detailed page. In Version A, **8 out of 13 participants** expected to generate images directly from the detailed profile page and were unsure how to proceed when no images were displayed.

Adjustments made in Version B included clearer guidance and improved navigation cues, resulting in **11 out of 11 participants** successfully navigating this section.

Consent Interaction Timing: Across all sessions, most participants spent fewer than five seconds on the consent screen before proceeding, indicating a tendency to quickly accept the terms without extended reading. Only **7 out of 24 participants** interacted with the additional information link (question mark icon) to access more detailed consent information. However, several participants who did not engage with the additional information later expressed uncertainty about how their images would be handled.

6.1.3 Task Timing Analysis

To further assess the efficiency of the user interface and the intuitiveness of navigation, task durations were analyzed retrospectively from the session recordings. The measurements focus exclusively on active user interaction; time spent waiting for system processes, such as image generation or other loading periods, was excluded, as these vary due to external factors (e.g., server response time, participant internet speed) beyond the scope of the interface design and thus do not directly affect the evaluation of user flow. Three time-based metrics were extracted, with all times reported in **minutes:seconds (mm:ss)** format.

1. **Task 1 Duration:** Time from entering the access code on the home screen (see Figure 5.1) to completing profile creation.
2. **Task 2 Duration:** Time from navigating to the prompts screen to saving the first set of generated images using the sample profile.
3. **Transition Time:** Time between completing Task 1 and successfully locating the correct interface element to begin Task 2.

These measurements allow for a comparative analysis between the first (Version A) and second (Version B) iterations of the application, and serve as an indicator of usability improvements made between testing phases.

Task 1: Selfie Upload and Profile Creation.

Participants were asked to upload a minimum of five selfies and create a profile. In Version A, the average time to complete this task was **01:49** (median **01:29**). Several participants expressed confusion regarding the minimum image requirement, which contributed to longer completion times. In Version B, the average completion time was reduced to **00:53** (median **00:48**).

Task 2: Image Generation with Sample Profile.
In the second task, participants generated images using the sample profile. In Version A, the average duration of this task was **00:47** (median **00:45**). In Version B, the average task time decreased to **01:07** (median **00:53**).

Transition Time Between Tasks.
A recurring issue observed in Version A was the confusion following the completion of profile creation. Participants often remained on the profile page, unsure of how to proceed to the next step. The average time between completing Task 1 and initiating Task 2 in Version A was **00:47** (median **00:46**), with some participants taking over a minute to locate the correct action. In Version B, the introduction of step indicators and redirect prompts reduced this delay significantly. The average transition time dropped to **00:05** (median **00:05**).

6.2 User Engagement Scale (Short Form)

All the 24 participants who took part in the usability test also completed the User Engagement Scale Short Form (UES-SF) questionnaire. The following tables present the results for both versions of the application.

6.2.1 Results for Version A

Table 6.1: UES-SF Results: Version A

Subscale	Mean	SD
Focused Attention	3.85	0.63
Perceived Usability	3.97	0.63
Aesthetic Appeal	3.51	0.65
Reward	4.28	0.45
Overall Engagement	**3.90**	**0.40**

6.2.2 Results for Version B

Table 6.2: UES-SF Results: Version B

Subscale	Mean	SD
Focused Attention	3.88	0.64
Perceived Usability	4.24	0.84
Aesthetic Appeal	3.55	1.15
Reward	4.24	0.88
Overall Engagement	**3.92**	**0.76**

6.3 System Causability Scale

From all the participants who were contacted, five submitted the completed SCS questionnaire. Table 6.3 shows the mean and standard deviation for each SCS item, along with the overall mean score. The scores suggest generally positive perceptions of the explainable AI feature, with certain areas performing stronger than others.

Table 6.3: System Causability Scale Results

Item	**Mean**	**SD**
SCS1: All relevant factors included	3.80	0.84
SCS2: Understood explanations in my context	3.60	0.89
SCS3: Appropriate level of detail	3.90	0.74
SCS4: No external help needed to understand	3.40	1.14
SCS5: Helped me understand causality	4.20	0.45
SCS6: Clear sequence of information	4.00	0.71
SCS7: Able to adjust detail level on demand	3.70	0.82
SCS8: Information was precise	3.60	0.55
SCS9: Information was complete	3.80	0.45
SCS10: Overall clarity and comprehensibility	4.10	0.32
Overall Mean	**3.81**	**0.26**

The calculated normalized SCS score was 0.762 (3.81/5).

7 Discussion

This chapter interprets the empirical results in relation to the project goals and RQ1–RQ3. The analysis synthesizes evidence from two rounds of usability testing (n=24), time-on-task metrics, UES-SF engagement scores, and SCS ratings for the explanation wizard. It contrasts Version A and Version B, connects observed effects to the literature, and discusses implications for consent design, explainable AI, and deployment in a public exhibition context. Limitations, contextual constraints, and priorities for further work are outlined to inform subsequent iterations.

7.1 Usability Testing

The usability testing was organized in a structured manner, though certain practical challenges arose, which are elaborated upon in Chapter 7.7. Initially, 26 participants were scheduled, but only 24 attended the sessions. The two who did not appear were contacted via follow-up email and upon receiving no response, they were excluded from the study.

In general, Version A of the application revealed more instances of participants becoming disoriented, clicking around more frequently, and asking additional questions for clarification. By contrast, in Version B, with the refinements, the sessions ran significantly more smoothly. Participants were able to complete both tasks in a more continuous flow with fewer interruptions and substantially fewer clarifying questions. Thus, the adjustments made between versions demonstrably improved the overall flow of interaction, as evidenced by the Usability Testing Results in Section subsection 6.1.

7.1.1 Time Analysis

In alignment with the usability improvements, the time taken to complete tasks was also notably reduced in Version B. As detailed in the previous section, participants experienced much less confusion and were able to transition seamlessly from Task 1 to Task 2, largely due to the introduction of a direct generation button on the profile detail page. This refinement reduced transition time from approximately 47 seconds to just 5 seconds.

While the duration for Task 2 was slightly longer in Version B than in Version A, this was attributable to participants utilizing the new zoom function to examine generated images in greater detail before making a selection. Thus, the additional time reflected more engaged interaction rather than confusion or inefficiency.

In conclusion, the refinements made between versions clearly contributed to a more intuitive user experience and a reduction in unnecessary delays. This is evidenced not only by the quantitative task time improvements but also by the qualitative feedback indicating smoother transitions and more confident user navigation.

7.2 Analysis of UES-SF Results

The UES-SF captures four facets of engagement on a 1–5 scale: Focused Attention, Perceived Usability, Aesthetic Appeal, and Reward, plus an overall score. Across both versions, overall engagement stayed stable (A: M=3.90, SD=0.40; B: M=3.92, SD=0.76), indicating a broadly positive experience and suggesting that the changes described in Chapter 6.1.1 did not disrupt what already worked well.

Focused Attention remained relatively stable (A: M=3.85, SD=0.63; B: M=3.88, SD=0.64). This points to an interaction that supports efficient task completion without demanding deep immersion, which is appropriate for short, public-session use.

Perceived Usability showed the clearest positive shift (A: M=3.97, SD=0.63; B: M=4.24, SD=0.84). The larger standard deviation in Version B indicates more varied experiences across participants. Qualitative feedback supports this: one participant noted (translated from German), "I have seen many apps and AI tools and am not easily impressed. Compared with python/diffusers, using this was simple, but it still does not have the smooth UX of deepai.org or civitai.com." This suggests that while many users found the system straightforward relative to code-based workflows, others benchmarked it against highly polished consumer services, which helps explain the greater variability in perceived usability. Consistent with these perceptions, the tracked time results in Chapter 6.1.3 show that participants navigated more quickly in Version B, indicating reduced operational friction.

Aesthetic Appeal stayed about the same (A: M=3.51, SD=0.65; B: M=3.55, SD=1.15). The larger standard deviation in Version B suggests more polarized reactions to the visual presentation. Notably, Aesthetic Appeal was the lowest-scored subscale overall, which may reflect individual preference rather than a systemic flaw. Qualitative feedback supports this interpretation: one participant commented ironically (translated from German), "Oh, I love pink and orange—those are my favourite colors," while others said the design felt "fresh" and that they "liked the vibes." Together, these remarks indicate taste-based divergence. To accommodate differing preferences without diluting the core visual identity, we will keep clean defaults and expose optional controls through progressive disclosure. In future iterations, a lightweight front-end customizer (e.g., theme and contrast presets or control density) could allow users to tailor the look and feel to their taste. This feature has been added to the development backlog.

Reward remained relatively the same (A: M=4.28, SD=0.45; B: M=4.24, SD=0.88), indicating that the core sense of satisfaction and the spirit of the app were preserved. The wider dispersion likely reflects differing expectations about generative outputs and the desired level of control. Fr example, one participant remarked, "I would have liked more control over the prompts and maybe the ability to mix a variety of prompts," while another said (translated from German), "I liked that the prompts are preselected and prepared." This points to offering an optional advanced mode for prompt editing while keeping a simple default. It should also be noted that participants were recruited via email and self-selected into the study; such interest-driven sampling can introduce a positive-interest bias, potentially inflating perceived reward (and usability) relative to a broader walk-up exhibition audience.

Overall, the engagement profile shows high perceived usability and reward, with adequate attention and acceptable aesthetics—appropriate for a public exhibition context. Version B exhibits greater variance across subscales, indicating heterogeneous user needs; subsequent iterations should preserve a simple default while adding optional guidance, advanced controls, and clearer affordances. Although the sample is modest ($n = 24$), these conclusions are corroborated by behavioral metrics: participants completed tasks faster and with fewer errors in Version B, indicating reduced operational friction. Accordingly, the priority in the future work will be to reduce between-user variability rather than to shift mean scores.

7.3 XAI Feature "Wizard" and Analysis of SCS

Design decisions in the wizard feature closely reflect the research findings discussed in Section 4.3. Specifically, the implementation was guided by recommendations to use multi-level explanations with progressive disclosure, step-by-step visualisations of model behaviour, and minimal but meaningful interactivity to support cause–effect understanding. The wizard aimed to strike a balance between accessibility and informativeness, especially for non-expert audiences in short, public sessions.

Content for each step—including titles, short and expanded texts, and media references—was managed through a lightweight JSON file loaded at runtime. While this approach does not allow for complex authoring logic or in-app editing, it provided a practical and efficient way to keep content modular and editable under tight development constraints. Despite its simplicity, this structure supported the key goals of clarity, maintainability, and flexibility for future expansion or localization.

7.3.1 Analysis of SCS Results

The original wording of the item ("I understood the explanations within the context of my work") reflects the System Causability Scale's initial design for professional and domain-specific settings, where explanations are evaluated in relation to a user's occupational tasks and knowledge. In the context of this study, the AI Selfie App was evaluated by a general public audience rather than a professional user group. Therefore, the item was rephrased to "I understood the explanations within the context of my everyday life" to ensure that the question was meaningful and relevant to participants without a specific occupational frame of reference.

The overall SCS score of 0.762 indicates a moderately high perceived quality of the explainable AI feature. The strongest aspects were users' understanding of causality (Mean = 4.20, SD = 0.45) and overall clarity (Mean = 4.10, SD = 0.32), aligning well with the intended step-by-step "wizard" design that uses illustrative images and concise text.

However, lower scores on items like understanding explanations in personal context (Mean = 3.60, SD = 0.89) and the absence of a need for external help (Mean = 3.40, SD = 1.14) suggest that some participants struggled to relate the explanations to their own understanding. This may be due to varying levels of AI knowledge or unfamiliarity with terminology such as "U-Net" or "guidance scale."

Similarly, the ability to adjust detail level on demand (Mean = 3.70, SD = 0.82) was rated below the top-performing items, indicating that the "show more" feature might not have been fully discovered or fully met participants' needs.

Given the small sample size (n = 5), these results should be interpreted with caution. Nonetheless, they offer preliminary insights into strengths such as clarity and causality explanation, as well as areas needing refinement, such as context relevance and user independence in understanding.

While the overall score is moderately positive, it is acknowledged that the sample size is limited. Therefore, a tracking mechanism is planned to gather more comprehensive data over time. This will allow for an evaluation of how many users actually engage with the explainable AI content. If future data indicate significant engagement, further refinements will be prioritized. For now, the primary next step is to implement this tracking tool to gain a clearer understanding of user interaction with the feature.

7.4 Consent Feature

The consent flow, implemented as a lightweight modal in Figure 5.4, avoids overwhelming the user at the moment of action while still providing access to more detailed information for those who seek it. No images are transmitted to the server unless explicit consent has been given. This ensures that no image data is collected without informed agreement and a fallback path is always available. Users who choose not to upload images can still explore the app using the sample profile, allowing them to participate without compromising their privacy.

Still, only 7 out of 24 participants opened the additional information tab, which is relatively low. Notably, some participants asked during or after the session what would happen to their images. This suggests that, while the consent mechanism was visible and functional, it was often bypassed in favor of continuing with the task flow. This observation highlights the challenge posed in **RQ2** (see chapter 1.4), as even a minimally disruptive design may not fully overcome the attention constraints of short, time-limited interactions. A possible improvement could be to offer users an option to revisit the consent information later (e.g., through a persistent menu link or a short post-interaction reminder).

7.5 Sample Profile

Including a pre-configured sample profile proved beneficial, enabling participants who lacked time or preferred not to upload personal images to experience the app. However, the current design made before–after comparisons salient primarily to the participant familiar with the individual depicted. For other participants, no direct visual reference to the sample profile's original images was visible during generation (see Figure 5.19). Moving forward, adjusting the interface to display the sample person's original selfies alongside the generated images would help all users appreciate the transformation, regardless of familiarity with the reference individual.

7.6 Context of use

The AI Selfie App is intended for deployment in a public exhibition setting, with a web-based approach chosen to minimize installation friction and broaden device compatibility. The environment typically includes FHNW staff available for guidance and support. Although an on-site introduction is often present, usability feedback indicated that at least one participant would not have understood certain details without a facilitator's introduction. While staff presence currently mitigates this risk, a brief in-app introduction covering purpose, image deletion policies, and operating hours has been recorded in the backlog to support unattended use.

7.7 Challenges

The AI model is hosted on the museum's infrastructure, limiting API availability to exhibition hours. During usability testing, occasional slower response times highlighted potential performance fluctuations. Communication with the hosting team is ongoing to improve reliability. However, final resolution lies partly outside direct project control. The recommended mitigation is close coordination with the hosting team to ensure stability during live deployment.

7.8 Further work

Several areas for further development have been identified and recorded in the backlog. One important enhancement involves providing persistent access to consent and data-use information without requiring users to initiate an upload. This could be implemented through a clearly labeled menu entry or a brief post-interaction summary that remains accessible throughout the session.

Additional improvements under consideration include the introduction of a lightweight customization feature, allowing users to adjust the visual appearance of the interface according to their preferences. A concise onboarding screen at first launch, outlining the application's purpose, image usage policy, and operational hours, is also planned to support independent use without staff assistance.

Finally, the integration of real-time tracking with Realtime from Supabase („Supabase Realtime Guide“, 2025) is proposed to gather data on user engagement with educational and consent-related content. This would provide valuable insights into which elements are most frequently accessed and help refine future iterations of the onboarding and consent flow.

8 Conclusion

This bachelor thesis explored the design, implementation, and evaluation of the AI Selfie App, a web-based application developed for use in public exhibition environments. The project focused on creating an engaging, accessible, and educational experience that enables users to generate AI-based self-portraits.

The following research questions guided the project:

- **RQ1**: What design principles and features, proven to maintain user engagement in mobile applications, can be successfully integrated and validated in the AI Selfie App?
- **RQ2**: How are consent process implementations in mobile applications affected by the constraints of busy and time-limited environments, and which techniques, when implemented, can effectively overcome these challenges?
- **RQ3**: How can the integration of explanation techniques within the AI Selfie App improve user understanding of AI image generation processes?

The primary goals of the project were to enhance usability for short, self-guided sessions, to develop a consent mechanism appropriate for public use and to improve user understanding of AI image generation through interactive explanation techniques. These goals were addressed through iterative design, prototype development, and empirical testing with 24 participants across two usability testing rounds.

In relation to **RQ1**, the findings indicate that targeted interface refinements such as clearer action cues, streamlined navigation, and improved status feedback—successfully improved perceived usability and reduced operational friction. This is supported by both qualitative observations from the usability testing sessions and quantitative data from the User Engagement Scale (Short Form), including significant reductions in task completion and transition times between the first and second testing iterations.

RQ2, which focused on consent mechanisms under time constraints, was addressed through a minimally disruptive modal design with optional access to further information. While this approach ensured no image data was processed without explicit consent, the low engagement with the additional information tab (7 out of 24 participants) revealed a limitation. Despite the mechanism functioning as intended, many users proceeded without fully understanding what they had consented to. This highlights the need for supplementary strategies, such as persistent access to privacy information or post-interaction reminders.

Regarding **RQ3**, the integration of the explanation feature was generally well-received. The System Causability Scale results indicated strengths in clarity and the explanation of causality. However, there was variation in users' ability to relate the content to their own understanding, especially among those with limited technical background. These results suggest that while the explanation model provided a good foundation, further personalization and improved discoverability of detailed information would increase its effectiveness.

Overall, the goals of the project were largely met. The app demonstrated strong usability, and delivered a positive user experience that aligns well with the requirements of a public exhibition setting.

Future work includes persistent, easily accessible consent and privacy information; a lightweight color customizer to accommodate visual preferences; a concise first-run introduction covering purpose and operating hours; and real-time instrumentation to monitor engagement with explanatory and consent content. These enhancements are expected to refine comprehension and flexibility over time.

In summary, the AI Selfie App provides a functional and engaging framework for playful interaction with AI-generated imagery while addressing key ethical and educational considerations. Although further improvements are planned, the current state of the application is ready for deployment and use in its intended exhibition context.

Bibliography

Anonymous sign-ins with supabase auth. (2025). Supabase. https://supabase.com/docs/guides/auth/auth-anonymous

Barbosa, M. G., de Saboya, L. A., & Bevilaqua, D. V. (2015). A smartphone headset for augmented-reality experiences in museums. *MW2015: Museums and the Web 2015.* https://mw2015.museumsandtheweb.com/paper/a-smartphone-headset-for-augmented-reality-experiences-in-museums/

Beck, K., Beedle, M., van Bennekum, A., Cockburn, A., Cunningham, W., Fowler, M., Grenning, J., Highsmith, J., Hunt, A., Jeffries, R., Kern, J., Marick, B., Martin, R. C., Mellor, S., Schwaber, K., Sutherland, J., & Thomas, D. (2001). Manifesto for agile software development. http://www.agilemanifesto.org/

Bussone, A., Stumpf, S., & O'Sullivan, D. (2015). The role of explanations on trust and reliance in clinical decision support systems. *2015 international conference on healthcare informatics*, 160–169.

Cai, C. J., Winter, S., Steiner, D., Wilcox, L., & Terry, M. (2019). "hello ai": Uncovering the onboarding needs of medical practitioners for human-ai collaborative decision-making. *Proceedings of the ACM on Human-computer Interaction*, *3*(CSCW), 1–24.

Calendly, LLC. (2025). *Calendly.* https://calendly.com

Charters, E. (2003). The use of think-aloud methods in qualitative research an introduction to think-aloud methods. *Brock Education Journal*, *12*(2).

Eas hosting - expo documentation. (2025). Expo. https://docs.expo.dev/eas/hosting/introduction

Ehsan, U., & Riedl, M. O. (2020). Human-centered explainable ai: Towards a reflective sociotechnical approach. *International Conference on Human-Computer Interaction*, 449–466.

Eiband, M., Buschek, D., Kremer, A., & Hussmann, H. (2019). The impact of placebic explanations on trust in intelligent systems. *Extended abstracts of the 2019 CHI conference on human factors in computing systems*, 1–6.

European Commission. (2021). Special eurobarometer 510: E-communications in the european union [Fieldwork conducted Feb–Mar 2021. 96% of EU respondents reported having mobile telephone access.]. https://ec.europa.eu/commfrontoffice/publicopinion/index.cfm/Survey/getSurveyDetail/instruments/SPECIAL/surveyKy/2155

Expo. (n.d.). *Expo documentation.* Retrieved July 14, 2025, from https://docs.expo.dev

Felt, A. P., Egelman, S., Finifter, M., Akhawe, D., & Wagner, D. (2012). How to ask for permission. *HotSec*, 0.

fenos. (2022, September 30). *Storage file deletion from sql.* Supabase. https://github.com/orgs/supabase/discussions/7067#discussioncomment-3773245

FHNW. (2024, March 11). *Ki-exponat der fhnw entführt museumsbesuchende in neue welten.* https://www.fhnw.ch/de/medien/newsroom/medienmitteilungen/medienmitteilungen-2024/ki-exponat-der-fhnw-entfuehrt-museumsbesuchende-in-neue-welten

Gillespie, N., Lockey, S., Curtis, C., Pool, J., & Akbari, A. (2023). *Trust in artificial intelligence: A global study* (tech. rep.). The University of Queensland and KPMG Australia. https://doi.org/10.14264/00d3c94

Google LLC. (2025a). *Google forms.* https://www.google.com/forms/about/

Google LLC. (2025b). *Google meet.* https://meet.google.com

Haque, A. B., Islam, A. N., & Mikalef, P. (2023). Explainable artificial intelligence (xai) from a user perspective: A synthesis of prior literature and problematizing avenues for future research. *Technological Forecasting and Social Change*, *186*, 122120.

Holdener, M., Gut, A., Angerer, A., et al. (2020). Applicability of the user engagement scale to mobile health: A survey-based quantitative study. *JMIR mHealth and uHealth*, *8*(1), e13244.

Holzinger, A., Carrington, A., & Müller, H. (2020). Measuring the quality of explanations: The system causability scale (scs) comparing human and machine explanations. *KI-Künstliche Intelligenz*, *34*(2), 193–198.

Hudon, A., Demazure, T., Karran, A., Léger, P.-M., & Sénécal, S. (2021). Explainable artificial intelligence (xai): How the visualization of ai predictions affects user cognitive load and confidence. In *Neurois retreat* (pp. 237–246). Springer.

Inc., F. (2025). *Figma: The collaborative interface design tool.* https://figma.com

Kim, S., Eun, J., Oh, C., & Lee, J. (2025). "journey of finding the best query": Understanding the user experience of ai image generation system. *International Journal of Human–Computer Interaction*, *41*(2), 951–969.

Krul, A. (2015, November 9). *Wie führt man einen Literatur-Review durch?* [Scribbr]. Retrieved July 15, 2025, from https://www.scribbr.ch/aufbau-und-gliederung-ch/wie-fuehrt-man-einen-literatur-review-durch/#literatur-sammeln

Lee, S., Hoover, B., Strobelt, H., Wang, Z. J., Peng, S., Wright, A., Li, K., Park, H., Yang, H., & Chau, D. H. P. (2024). Diffusion explainer: Visual explanation for text-to-image stable diffusion. *2024 IEEE Visualization and Visual Analytics (VIS)*, 96–100.

Lyons, L. (2009). Designing opportunistic user interfaces to support a collaborative museum exhibit. *CSCL (1)*, 375–384.

Maguire, M. C. (1999). A review of user-interface design guidelines for public information kiosk systems. *International journal of human-computer studies*, *50*(3), 263–286.

Manamendra, C., Manathunga, K., Perera, H., & Kodagoda, N. (2013). Improvements for agile manifesto and make agile applicable for undergraduate research projects. *Proceedings of the 8th International Conference on Computer Science and Education, ICCSE 2013*, 539–544. https://doi.org/10.1109/ICCSE.2013.6553969

Meta Platforms, Inc. (2024). *React native documentation.* Retrieved July 14, 2025, from https://reactnative.dev

Midjourney. (2025, June 18). *Introducing our v1 video model.* https://updates.midjourney.com/introducing-our-v1-video-model/

O'Brien, H. L., Cairns, P., & Hall, M. (2018). A practical approach to measuring user engagement with the refined user engagement scale (ues) and new ues short form. *International Journal of Human-Computer Studies*, *112*, 28–39. https://doi.org/10.1016/j.ijhcs.2018.01.004

O'Hara, K., Kindberg, T., Glancy, M., Baptista, L., Sukumaran, B., Kahana, G., & Rowbotham, J. (2007). Collecting and sharing location-based content on mobile phones in a zoo visitor experience. *Computer Supported Cooperative Work (CSCW)*, *16*(1), 11–44.

Pexels GmbH. (2025). *Pexels.* https://www.pexels.com

Rhee, B., & Choi, Y. (2015). Using mobile technology for enhancing museum experience: Case studies of museum mobile applications in s. korea. *International Journal of Multimedia and Ubiquitous Engineering*, *10*(6), 39–44.

Rinderknecht, E., von Winning, D., Kravchuk, A., Schäfer, C., Schnabel, M. J., Siepmann, S., Mayr, R., Grassinger, J., Goßler, C., Pohl, F., et al. (2024). Modification and validation of the system causability scale using ai-based therapeutic recommendations for urological cancer patients: A basis for the development of a prospective comparative study. *Current Oncology*, *31*(11), 7061–7073.

Solis, T. (2021). *So führst du eine Literaturrecherche für deine wissenschaftliche Arbeit durch* [Accessed: 2025-07-15]. Scribbr. https://www.scribbr.ch/aufbau-und-gliederung-ch/literaturrecherche/

Supabase cron – schedule recurring jobs in postgres. (2025). Supabase. https://supabase.com/modules/cron

Supabase functions documentation. (2025). Supabase. https://supabase.com/docs/guides/functions

Supabase Inc. (n.d.). Supabase — the postgres development platform [Accessed: 2025-08-12].

Supabase realtime guide [Accessed: 2025-08-13]. (2025). Supabase.

Tang, Y., Zhang, N., Ciancia, M., & Wang, Z. (2024). Exploring the impact of ai-generated image tools on professional and non-professional users in the art and design fields. *Companion Publication of the 2024 Conference on Computer-Supported Cooperative Work and Social Computing*, 451–458.

Warren, G., Keane, M. T., & Byrne, R. M. (2022). Features of explainability: How users understand counterfactual and causal explanations for categorical and continuous features in xai. *arXiv preprint arXiv:2204.10152*.

Wegdam, M., & Plas, D.-J. (2008). Empowering users to control their privacy in context-aware system through interactive consent.

Wiebe, E. N., Lamb, A., Hardy, M., & Sharek, D. (2014). Measuring engagement in video game-based environments: Investigation of the user engagement scale. *Computers in human behavior*, *32*, 123–132.

Windl, M., Laboda, P. Z., & Mayer, S. (2025). Designing effective consent mechanisms for spontaneous interactions in augmented reality. *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, 1–18.

W.I.R.E. & Mindfire. (2022). *Swiss ai report: Beyond the hype – on the use of artificial intelligence in swiss companies* [Contributors: Stefan Pabst, Dr. Stephan Sigrist, Dr. Daniela Suter]. Retrieved March 11, 2025, from https://www.mindfire.global/swissaireport

Declaration of Authenticity

I hereby declare that I have prepared the present bachelor's thesis independently and solely using the sources indicated. Passages taken verbatim or in substance from the listed sources are identified in the work as quotations or paraphrases. This bachelor's thesis has not yet been published, made available to other interested parties, or submitted to any other examination authority.

Thun, 14. August 2025

Name: Hava Fuga

Signature: 

A Excerpt from IP5 Paper

The following pages are reproduced from the previous semester's IP5 paper, which describes the initial implementation and technical setup of the AI Selfie App.

5.4 Development

This section outlines the technical implementation of the AI Selfie App, detailing the system architecture, database and storage integration, AI model interaction, image upload and validation process, and the overall app design. Each subsection delves into the specific components and methodologies employed to build the application.

5.4.1 System Architecture and Design

The system is implemented using a layered architecture that adheres to the Model-View-Controller (MVC) pattern, with the addition of the Repository and Service patterns. How these elements interact with each other is illustrated in Figure 30, which depicts the components and their relationships, with each component labelled by number.

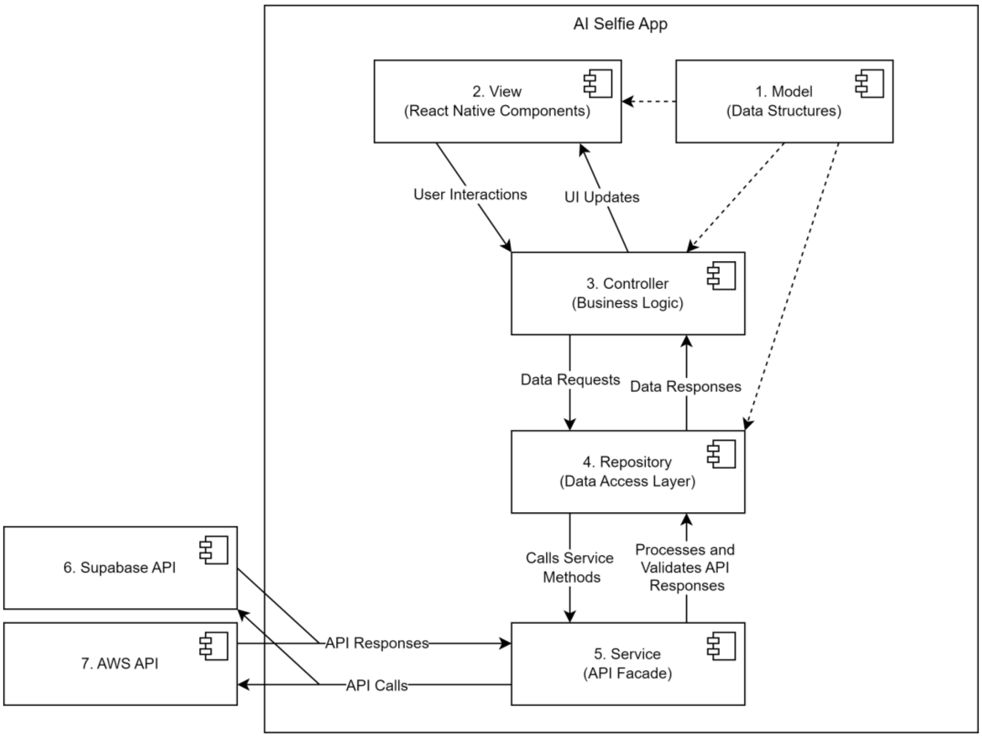


Figure 30: Component diagram – AI Selfie App

The View, as shown as number one in Figure 30, consists of React Native components. These components are responsible for the user interface and user interaction. They render the data received from the Controller and capture user inputs. These inputs are then passed to the Controller for processing. The View is the only layer directly visible to the end user and acts as the entry point for user interactions with the application.

The Model, represented as number two in Figure 30, defines the structure of the data used within the system. It is implemented using TypeScript types and interfaces, ensuring type safety and consistency. The Model remains passive and does not perform any operations. It serves as a schema for the data that flows between other parts of the architecture.

The Controller, represented as number three in Figure 30, serves as the central processing unit for the application. It handles all business logic and acts as a mediator between the View and the underlying data layers. For instance, when a user triggers an action in the View, the Controller processes the input, determines the necessary operations, and communicates with the Repository to fetch or modify data. After processing the data, the Controller sends the updated information back to the View for rendering.

The Repository, denoted as number four in Figure 30, provides an abstraction for data access. It includes methods for fetching, saving, and updating data. The Repository interacts solely with the Service layer, delegating all external API calls to it. This ensures that the data access logic is centralized and decoupled from the specifics of API communication.

The Service, shown as number five in Figure 30, acts as a facade for external API interactions. It encapsulates all the complexity of communicating with external systems, such as constructing API requests, managing authentication, and handling errors. These methods ensure that the Repository can work with a consistent interface, irrespective of the underlying API.

The External APIs, including Supabase as number six and AWS as number seven in Figure 30, provide the external backend functionalities required by the application. Supabase and AWS are treated solely as external API connections, and only the Service layer interacts directly with them.

5.4.2 Database and Storage with Supabase

The database and storage system were implemented using Supabase, with the schema comprising four primary tables: Prompt, Profile, Image, and Selfie. As illustrated in the schema diagram in Figure 31, with each table labelled by number.

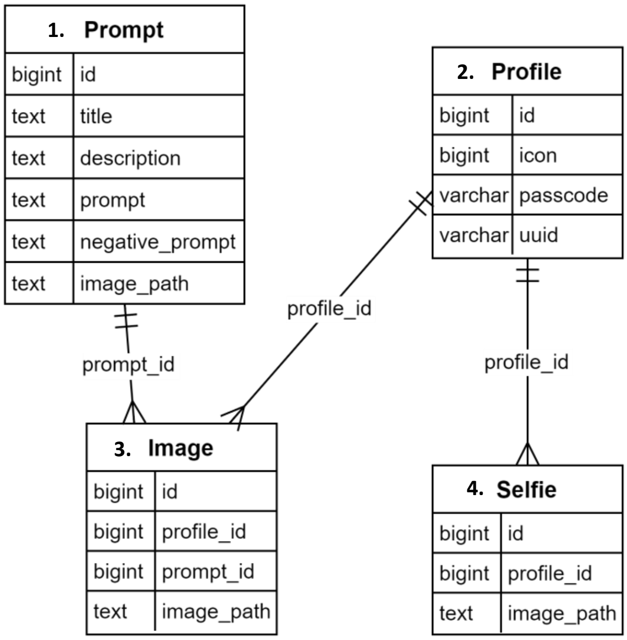


Figure 31: Database schema diagram

The Prompt table, labelled as number one in Figure 31, stores information related to predefined AI styles used in the application.

The Profile table, shown as number two in Figure 31, represents user-specific data and serves as the central entity linking user-related information across the schema.

The Image table, labelled as number three in Figure 31, handles metadata for AI-generated outputs, maintaining associations with both user profiles and prompts.

The Selfie table, labelled as number four in Figure 31, stores references to user-uploaded selfies and links these to the corresponding profiles.

Supabase's storage system complements the database by managing user-uploaded selfies and AI-generated images. Each file stored in the system is associated with a unique identifier that links it to its corresponding database entry, ensuring consistency and traceability throughout the application. The storage system's directory structure is logically organized to align with user profiles and their associated files, as illustrated in Figure 32. The storage system employs a hierarchical structure, with each profile ID serving as the root folder. Within each root folder, subfolders are designated for user selfies and AI-generated images. This organization facilitates efficient file management and retrieval, supporting seamless integration with the database.

```
Profile/
    Id/
        Selfies/
            [All user-uploaded selfies filename is 01
            and up since only once can selfie be up-loaded]
        Images/
            [All AI-generated images filename is date.now()]
```

Figure 32: Supabase Storage - Directory structure

Connections to both the database and storage systems are established using the Supabase API. The Supabase client is configured with the project's unique URL and public API key, ensuring secure and reliable communication. Interactions with the database and storage are separated into two dedicated files: supabase-database.ts and supabase-storage.ts. Both files utilize the Supabase client to perform operations specific to their respective domains. The implementation of the client is demonstrated in the code snippet from the file supabase-client.ts, as shown in Figure 33.

```
/**
 * Initializes the Supabase client with the specified configuration.
 */
const supabase: SupabaseClient<Database, 'ai_selfie_app'> = createClient(
    Config.supabaseUrl,
    Config.supabaseKey,
    { db: { schema: 'ai_selfie_app' } },
);
```

Figure 33: Code - supabase-client.ts

5.4.3 AI Model API

The integration of the AI Model API was implemented using a structure similar to the setup with Supabase. The API itself originates from the AI Selfie Box and serves as the foundation for handling AI-related operations.

The interaction with the AI Model API is facilitated through two files: aws-client.ts and aws-service.ts. The aws-client.ts file is responsible for configuring the API connection and handling low-level setup tasks. The aws-service.ts file provides a higher-level abstraction, encapsulating the REST operations required for interacting with the API.

As depicted in the component diagram in Figure 30, repositories have access exclusively to the aws-service.ts file. This design ensures that all API-related operations are mediated through the Service layer, maintaining a clean separation of concerns and abstracting the specifics of API interaction from the rest of the system.

B Design Prototype Specifications

The following PDF documents present the complete design specifications of the AI Selfie App prototype:

- **Prompt Overview.pdf** - Design of the prompt overview interface and navigation elements.
- **Image Generation.pdf** - User interface for the image generation process.
- **Profile Generation.pdf** – Screen design for the selfie upload.
- **Profile.pdf** – Layout of the user profile overview, including profile detail screen with and without generated images.
- **Colour & Typography.pdf** – Visual identity guidelines including color palette and typographic specifications.

The PDFs are included below as part of the appendix.

Home Screen Layout

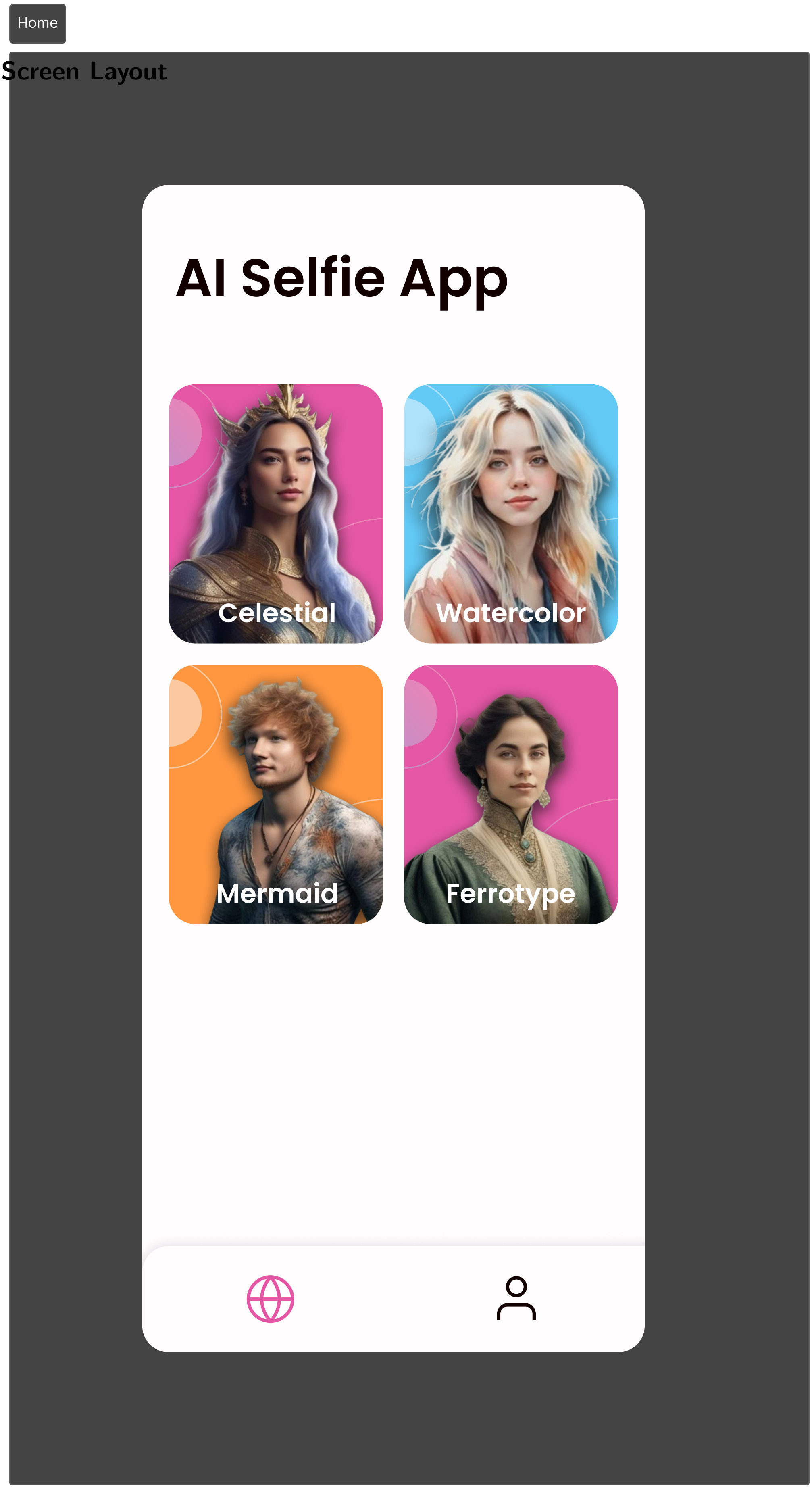
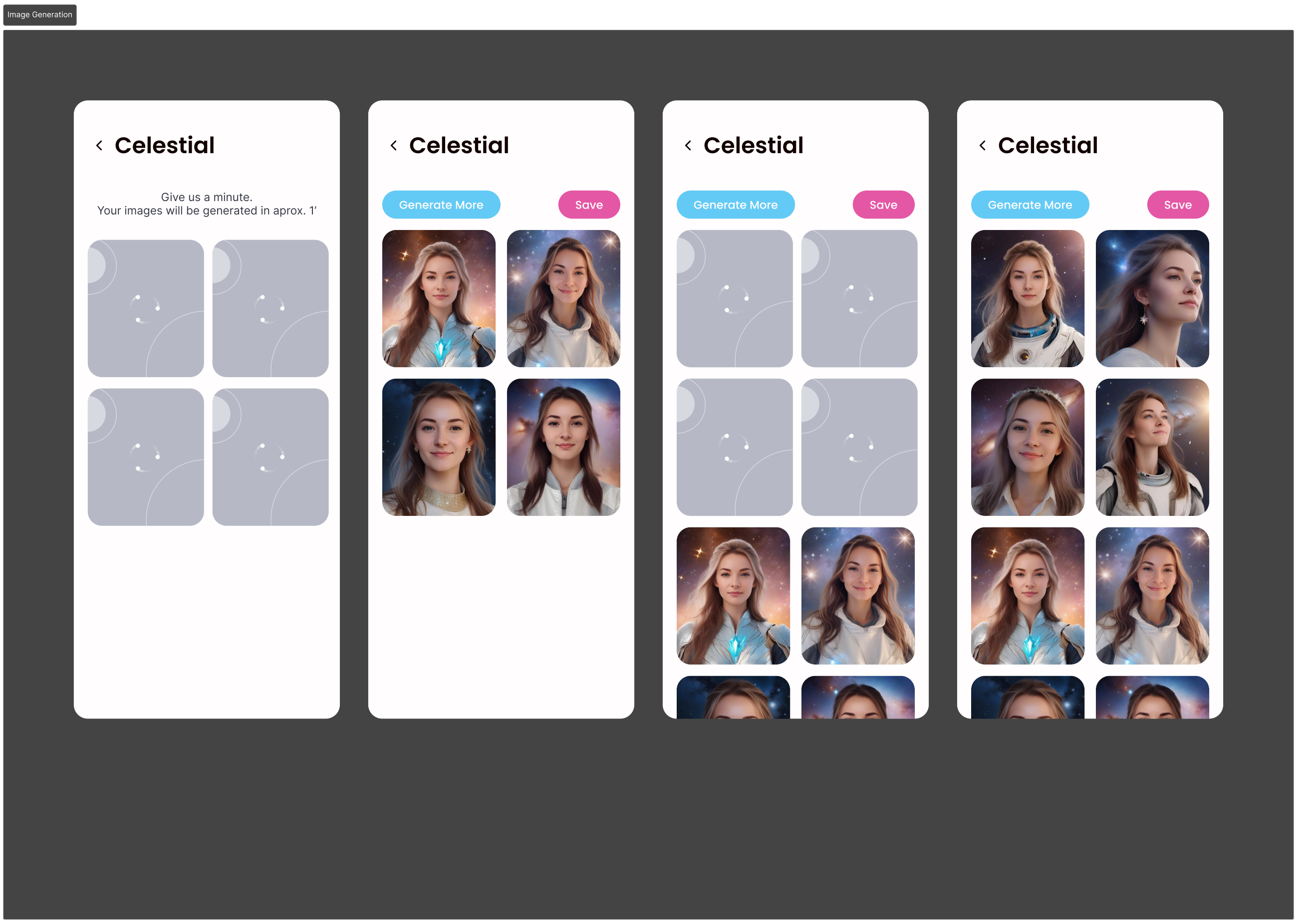
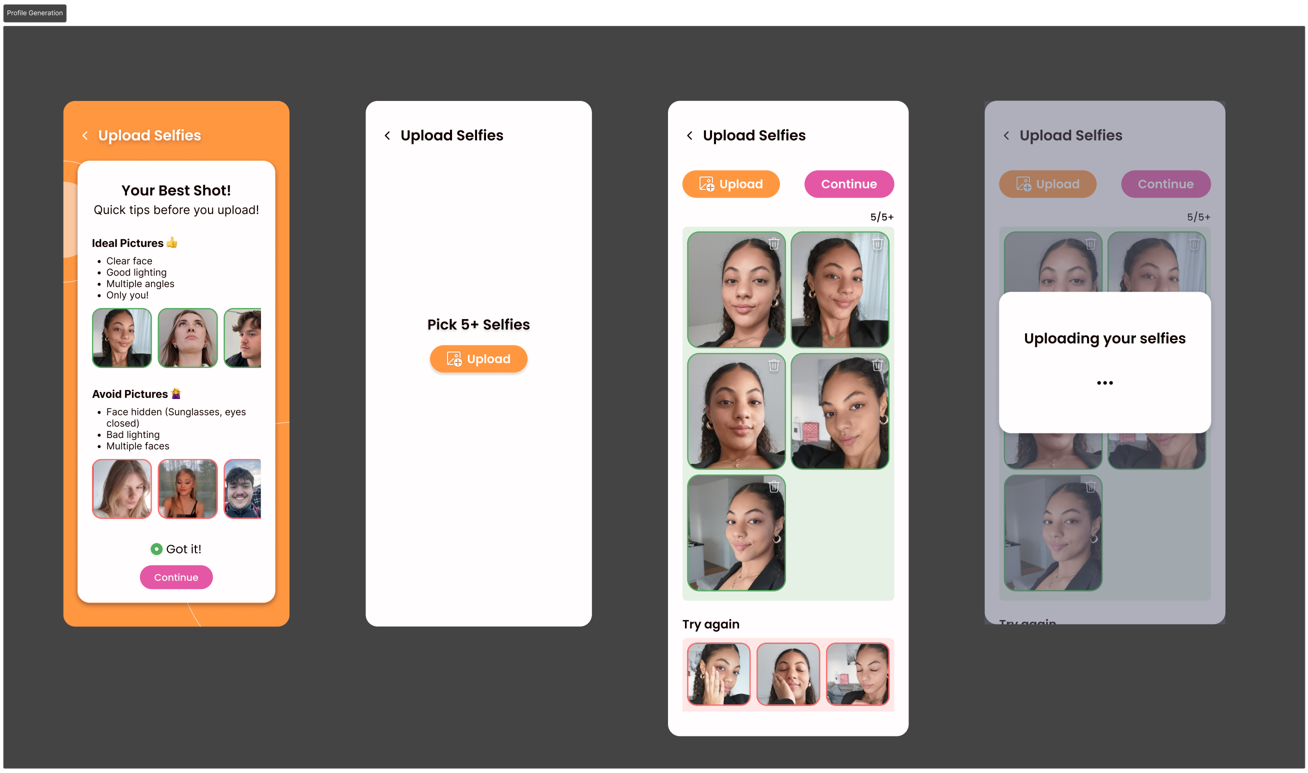


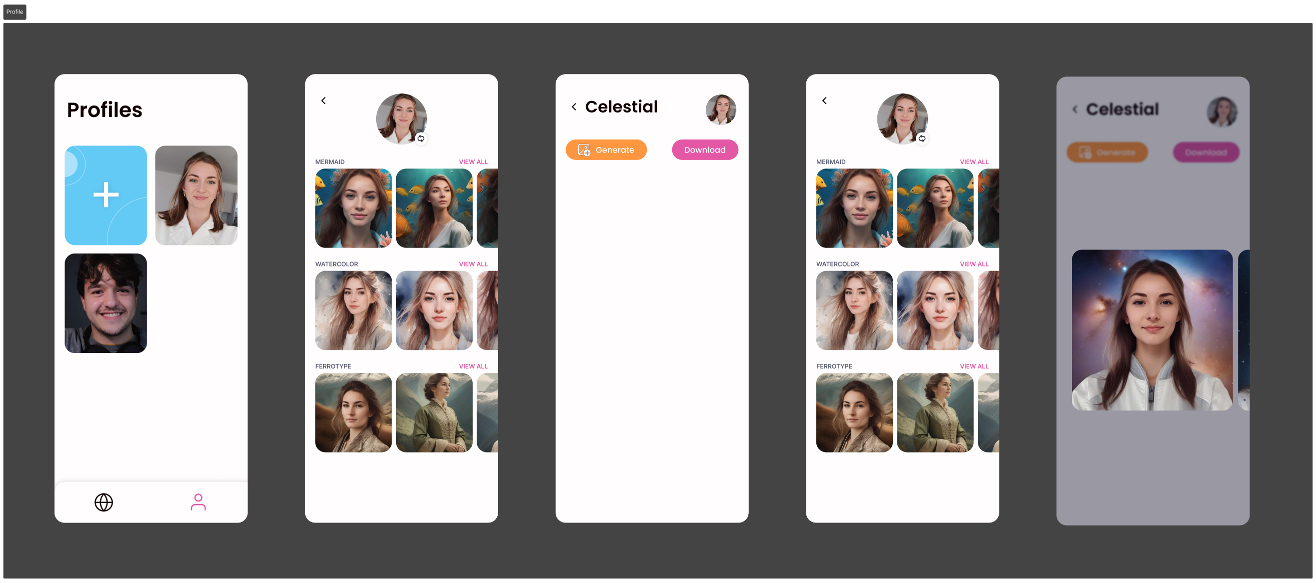
Image Generation Screen



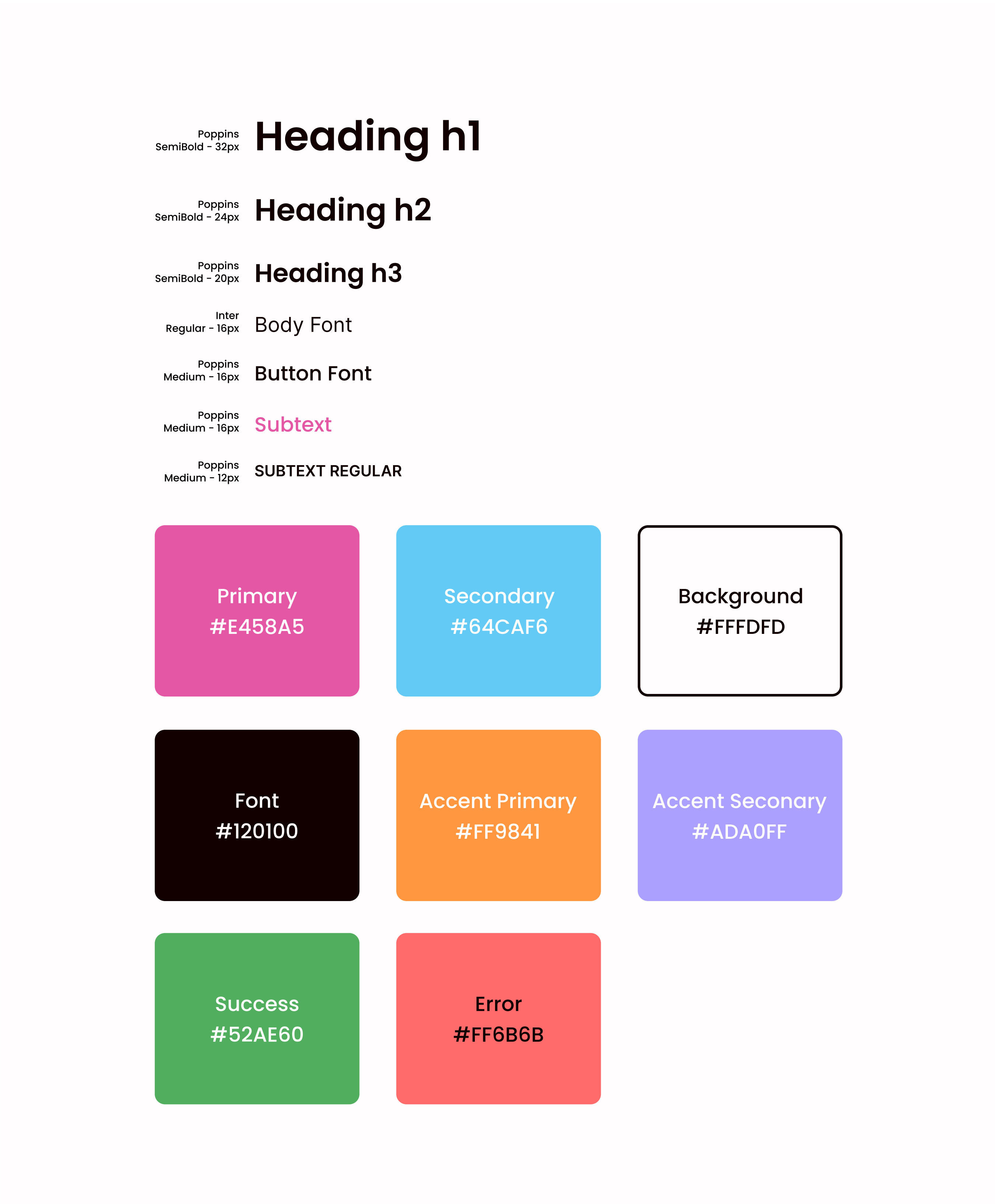
Profile Generation Screen



Profile Screen



Colour & Typography



C Example Selfies for Pre-session Email

The following images (from Pexels) were included in the pre-session email to illustrate suitable and unsuitable inputs for the AI Selfie App.

Figure C.1: Example selfies used in the pre-session email to illustrate suitable and unsuitable inputs.

D User Engagement Scale – Short Form (UES-SF) Items

The following table lists the 12 adapted items of the User Engagement Scale – Short Form (UES-SF) as presented to participants in this study. Responses were measured on a 5-point Likert scale: 1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly Agree.

#	**Statement**
1	I lost myself in this experience (in a positive way).
2	I felt frustrated while using this app.
3	This app was attractive.
4	Using this app was worthwhile.
5	The time I spent using this app just slipped away (in a positive way).
6	I found this app confusing to use.
7	This app was aesthetically appealing.
8	My experience was rewarding.
9	I was absorbed in this experience.
10	Using this app was taxing (= very tiring or draining).
11	This app appealed to the visual senses.
12	I felt interested in this experience.

Table D.1: User Engagement Scale – Short Form (UES-SF) items as used in the study.

E Participant Demographics

The following tables provide an overview of participant demographics for the study. To ensure anonymity, professions have been generalized into broad categories.

Age Distribution

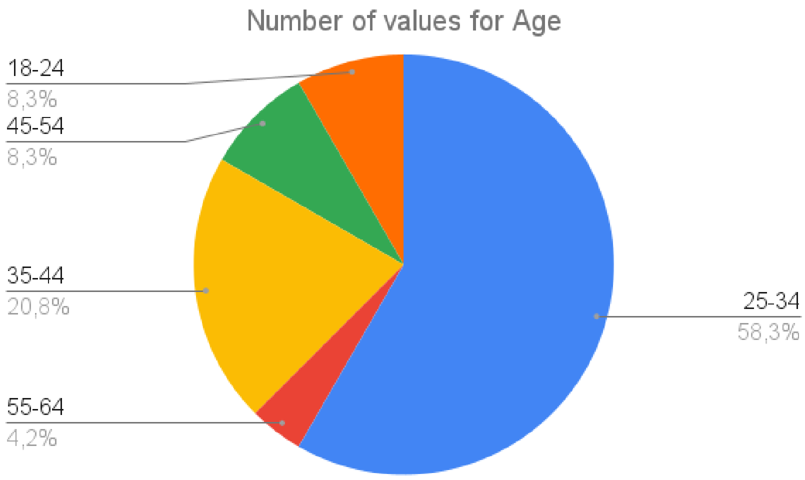


Figure E.1: Distribution of participants by age group.

Relation to Researcher

Relation	Count	%
External	12	50.0%
Co-Worker	9	37.5%
Friend	2	8.3%
Related	2	8.3%
Total	24	100%

Table E.1: Distribution of participants by relation to the researcher.

Professional Background

Category	**Count**	**%**
Project Management (IT/Media)	5	20.8%
Design (UI/UX)	1	4.2%
Administration	1	4.2%
Software Development	5	20.8%
Student (Design/Game Design)	2	8.3%
Academic Staff (HCI/XR)	2	8.3%
Intern (HCI/XR)	1	4.2%
Interaction Design	1	4.2%
Project Management (General)	1	4.2%
Systems Engineering	1	4.2%
Facility Management	1	4.2%
Academic Staff (Other)	1	4.2%
Mechanic	1	4.2%
Student (Education/Biochemistry)	2	8.3%
Total	24	100%

Table E.2: Distribution of participants by professional background.

F System Causability Scale (SCS) Items

The following table lists the 10 items of the System Causability Scale (SCS) as presented to participants in this study. Responses were measured on a 5-point Likert scale: 1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly Agree.

One item (originally "I understood the explanations within the context of my work") was adapted to "I understood the explanations within the context of my everyday life" to better suit the general public audience of the AI Selfie App.

#	**Statement**
1	I found that the explanations included all relevant known causal factors with sufficient precision and granularity.
2	I understood the explanations within the context of my everyday life.
3	The explanations were presented at an appropriate level of detail.
4	I did not need external support to understand the explanations.
5	The explanations helped me to understand the causality behind the results.
6	The sequence of information was logical and easy to follow.
7	I could adjust the level of detail in the explanations to my needs.
8	The explanations were precise and unambiguous.
9	The explanations were complete and contained all necessary information.
10	Overall, I found the explanations clear and comprehensible.

Table F.1: System Causability Scale (SCS) items as used in the study.

G Diffusion Explainer Steps (JSON)

The JSON file below defines the step content (IDs, titles, and explanatory text) used by the explainer feature.

```
[
  {
    id: 'noise',
    title: '1. Pure Noise',
    mini: 'We start with nothing but static. Pure random noise and just
        chaos.',
    more: (
      <Text>
        Diffusion models learn to undo noise: they practice on millions
            of examples where clean images are gradually
        noised, then learn the reverse steps. Generation starts from
            this noisy canvas and walks backward toward a
        picture that matches your prompt.
      </Text>
    ),
    getFrame: (p) => frame(p, 0)
  },
  {
    id: 'prompt',
    title: '2. Your Prompt',
    mini: 'Now here is where your input comes to play. What do you want
        to draw?',
    more: (
      <Text>
        Your words are turned into tokens like pieces of text, padded or
             truncated to the models limit, and encoded
        by a text encoder (for example, CLIP). The result is a vector
            that captures the meaning of your prompt. This
        vector steers every denoising step so the image drifts toward
            what you asked for.
        {'\n\n'}Learn more:{'\n'}
             <Link href="https://proceedings.mlr.press/v139/radford21a/
            radford21a.pdf" target="_blank">
        <Text style={{ color: '#007AFF' }}>CLIP paper</Text>
      </Link>
      </Text>
    ),
    getFrame: (p) => frame(p, 0)
  },
  {
    id: 'guess',
    title: '3. First Guess',
    mini: 'AI makes a rough sketch by wiping a little noise  a w a y blobs
         of colour start to appear.',
    more: (
      <Text>
        A U-Net predicts which parts of the image are "just noise".
            Remove a tiny bit, look again, repeat. Early steps
        reveal coarse layout and color blobs; fine details come later.
        {'\n\n'}Learn more:{'\n'}
             <Link href="https://arxiv.org/abs/1505.04597" target="_blank
            ">
```

```
        <Text style={{ color: '#007AFF' }}>U-Net architecture</Text>
      </Link>
      </Text>
    ),
    getFrame: (p) => frame(p, 10)
  },
  {
    id: 'refining',
    title: '4. Refining',
    mini: 'Predict     erase     repeat! Drag the slider to watch the
       picture sharpen step by step.',
    more: (
      <Text>
        The model balances two forces: staying faithful to your text and
            keeping the image natural. A knob called
        guidance scale (from classifier-free guidance) tilts that
           b a l a n c e higher values usually follow the prompt more
        strictly but can reduce variety.
        {'\n\n'}Learn more:{'\n'}
           <Link href="https://arxiv.org/abs/2207.12598" target="_blank
           ">
        <Text style={{ color: '#007AFF' }}>Classifier-Free Guidance
           paper</Text>
      </Link>
      </Text>
    ),
    getFrame: (p, s) => frame(p, s)
  },
  {
    id: 'polish',
    title: '5. Almost There',
    mini: 'Just a whisper of noise is left. The final brushstrokes bring
        textures and tiny details to life.',
    more: (
      <Text>
        In these last steps, the AI focuses on tiny details. It keeps
           checking your description as it clears the final
        bits of noise, so fine textures like fur strands, city lights,
           or brush strokes appear sharp and clear.
      </Text>
    ),
    getFrame: (p) => frame(p, 45)
  },
  {
    id: 'final',
    title: '6. Final Image',
    mini: 'Ta-da! A hi-res decoder enlarges the tiny internal picture
       into full glory.',
    more: (
      <Text>
        Stable Diffusion denoises in a compressed latent space for speed
           .
        A special 'image builder' turns the model's tiny internal
           picture into the full-size image you see and can save and
           share.
      </Text>
    ),
```

```
    getFrame: (p) => frame(p, 50)
  }
]
```