Design Document

FaceCraft

Ethan Stanks:	dev.stanks@proton.me
Temitayo Shorunke:	analyticalgeniusking@gmail.com
William Hoover:	Billhoover070@gmail.com

Contents

1. **Introduction**
   - 1.1. Purpose
   - 1.2. Goal
   - 1.3. Definitions
2. **Overall Description**
   - 2.1. Product Perspective
     - 2.1.1. User Interface Hardware
     - 2.1.2. Hardware interfaces & Memory Constraints
     - 2.1.3. Software interfaces
     - 2.1.4. Library Dependencies
     - 2.1.5. 3rd Party Data and APIs
   - 2.2.
3. **Specific Requirements**
   - 3.1. MVP
   - 3.2. Alpha
   - 3.3. Beta
4. **Use Cases**
   - 4.1. User characteristics
   - 4.2. User interface
5. **Development Environment**
   - 5.1. Developer Tools
   - 5.2. Content-Generation Tools
   - 5.3. Data Tools
   - 5.4. 3rd-Party Dependencies and APIs
6. **System Architecture**
   - 6.1. ERD Class Diagram
7. **Production Plans**
   - 7.1. Code Review Plan
   - 7.2. Integration Plan
   - 7.3. Testing Plans
8. **Folder Hierarchy**

1. Introduction

Purpose:

Creating highly realistic faces through artificial intelligence represents a significant advancement in technology with a broad range of applications, particularly in marketing and design. This innovation enables companies to produce appealing and diverse imagery without the need for traditional photoshoots, which can be costly and time-consuming. By leveraging AI-generated faces, brands can effortlessly tailor their visual content to appeal to various demographics, enhancing engagement and personalization in advertising campaigns. This approach not only reduces costs but also

allows for creative flexibility, as these synthetic faces can be adjusted for different contexts, expressions, and even cultural backgrounds, making marketing efforts more inclusive and resonant with a global audience.

Furthermore, the use of synthetic image data for further innovation extends beyond the realms of marketing and design. In product development, for instance, realistic faces can be used to test facial recognition technologies, improving their accuracy and reliability. In the entertainment industry, they can enhance the realism of video game characters and virtual reality experiences, creating more immersive environments for users. These applications demonstrate the versatility of AI-generated imagery, showcasing its potential to revolutionize various sectors by offering practical solutions to complex challenges.

Supporting the development of technology that generates highly realistic faces also contributes to the broader field of AI research. By providing synthetic image data, researchers can refine machine learning algorithms, improving their ability to understand and replicate human facial features accurately. This not only advances the development of more lifelike AI-generated faces but also contributes to the understanding of AI's potential limitations and ethical considerations. As this technology evolves, it promises to open new avenues for creativity and innovation, further blurring the lines between the virtual and the real world, and paving the way for future breakthroughs in AI and computer vision.

Goal: What problem is this solving?

Cost Efficiency

Traditional methods of creating visual content, such as photoshoots, are often expensive and resource-intensive. This software mitigates these issues by generating high-quality, realistic faces without the logistical challenges and expenses associated with traditional content creation.

Diversity

A notable challenge in marketing is creating content that resonates with a global and diverse audience. This AI technology enables brands to produce imagery that reflects a wide range of demographics, making advertising campaigns more inclusive and appealing across different cultures and regions.

Personalization

Customizing visual content to match specific campaign needs or consumer preferences can be cumbersome with traditional photography. The AI-generated faces offer unparalleled flexibility, allowing brands to adjust images for various contexts, expressions, and cultural backgrounds, enhancing personalization in marketing efforts.

Existing Solution:

Is this based on a previous project? If so, this is the place to describe the current state of the project and to link to documentation related to that project, including repo, overview videos etc.
Name: This person does not exist
Current State: On the market.
Documentation Link: Here
Repo: N/A

Definitions:

- GANs (Generative Adversarial Networks): A type of artificial intelligence where two programs compete to improve their performance, often used to create realistic images or videos.
- Computer Vision: A field of technology that enables computers to interpret and understand visual information from the world, similar to how humans use their eyes.
- Synthetic Data: Artificially created information used for training machines or models, allowing them to learn without using real-world data.
- Software Engineering: The practice of designing, developing, testing, and maintaining software applications or systems.

2. Overall Description

2.1. Product Perspective

2.1.1. User interface Hardware

Keyboard and mouse.

2.1.2. Hardware interfaces & Memory Constraints

Enough memory for a chrome tab.

2.1.3. Software interfaces

Web browser.

2.1.4. Library Dependencies

Python, HTML5, CSS PyTorch, Matplotlib, OpenCV, Pandas, and Numpy

2.1.5. Database

No information stored.

2.1.6. 3rd Party Data and APIs

Llama v2, HuggingFace

3. Specific Requirements

3.1. Milestones

3.1.1. Minimum Viable Product (8 A Features)

- Text Input Interface: The user will be able to enter and submit textual prompts for face generation.
- Realistic Image Generation: The user will receive a realistic face image
- Prompt Interpretation Engine: The user's input will be interpreted accurately to reflect intended attributes in the generated image.
- Basic Customization Options: The user will be able to specify basic attributes like age, gender, and emotion for the face generation.
- Output Display: The user will be able to view the generated face image immediately after creation.
- Save/Download Function: The user will be able to save or download the generated face images.
- User Instructions: The user will have access to clear instructions on how to use the web application effectively.
- Responsive Design: The user will be able to interact with the web application seamlessly across various devices.

3.1.2. Alpha (7 B Features)

- Advanced Customization Options: The user will be able to specify detailed facial attributes for more personalized image generation on the web application.
- Feedback Mechanism: The user will be able to provide feedback or rate the images directly on the website.
- Social Sharing Options: The user will be able to share their generated images on social media platforms directly from the web application.
- Prompt Suggestions: The user will have access to and use suggested prompts for image generation on the website.
- Batch Processing: The user will be able to input and process multiple prompts at once, receiving a set of generated images through the web application.
- Refresh Button: The user will be able to re-run the prompt they just ran to see different results from the model.
- Random Image Generation: The user will be able to randomly generate an image.

3.1.3. Beta (5 C Features)

- Themes: The user will benefit from being able to switch between a dark or light theme.
- Language Selection: The user will be able to select their preferred language for the web application's interface, making it more accessible and user-friendly for a diverse audience without storing any user data.
- Contact Form: The user will be able to contact the developers.
- Performance Metrics Display: The user will be able to see basic performance metrics, like generation time and server response time, providing transparency about the web application's efficiency and responsiveness without overwhelming them with technical details.
- Customizable Font Sizes: The user will be able to adjust the font size of the text within the web application, including the input and output areas, to enhance readability according to their visual preferences.

4. User Experience

4.1. Use Cases

Preconditions: User has access to a web-browser.
Main Flow:
1. User gets onto the website.
2. User clicks on the prompt box.
3. User types the prompt for the face they want generated.
4. User presses enter.
5. Website presents the generated image to the user.
6. User downloads the image to the computer.

Sub Flow (Suggested Prompt):
1. User gets onto the website.
2. User clicks on the suggested prompt.
3. Website presents the generated image to the user.
4. User downloads the image to the computer.

Sub Flow (Shuffle Suggested Prompt):
1. User gets onto the website.
2. User clicks on the shuffle suggested prompt button.
3. User clicks on the suggested prompt.
4. Website presents the generated image to the user.
5. User downloads the image to the computer.

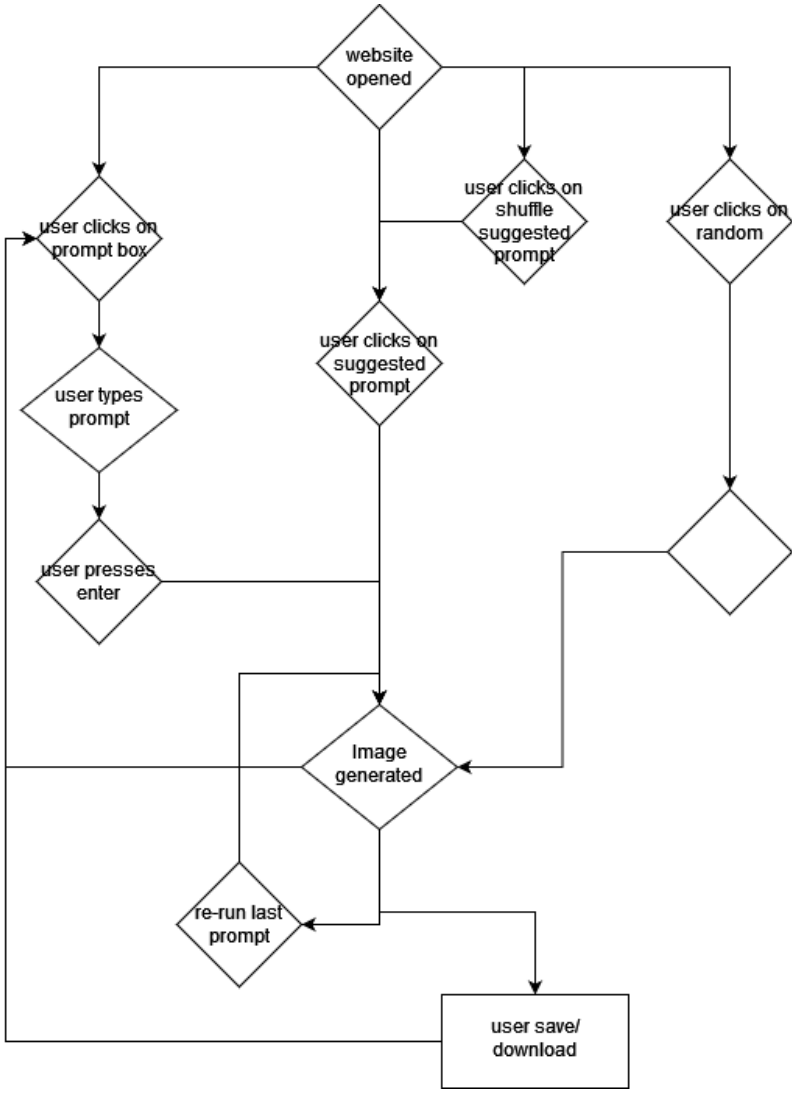
Sub Flow (Random Image):
1. User gets onto the website.
2. User clicks on the random button.
3. Website presents the generated image to the user.
4. User downloads the image to the computer.

Sub Flow (Re-Run Prompt):

1. User gets onto the website.
2. User clicks on the prompt box.
3. User types the prompt for the face they want generated.
4. User presses enter.
5. Website presents the generated image to the user.
6. User click on the re-run prompt button.
7. Website presents the newly generated image to the user.
8. User downloads the image to the computer.

Sub Flow (Not Saving):

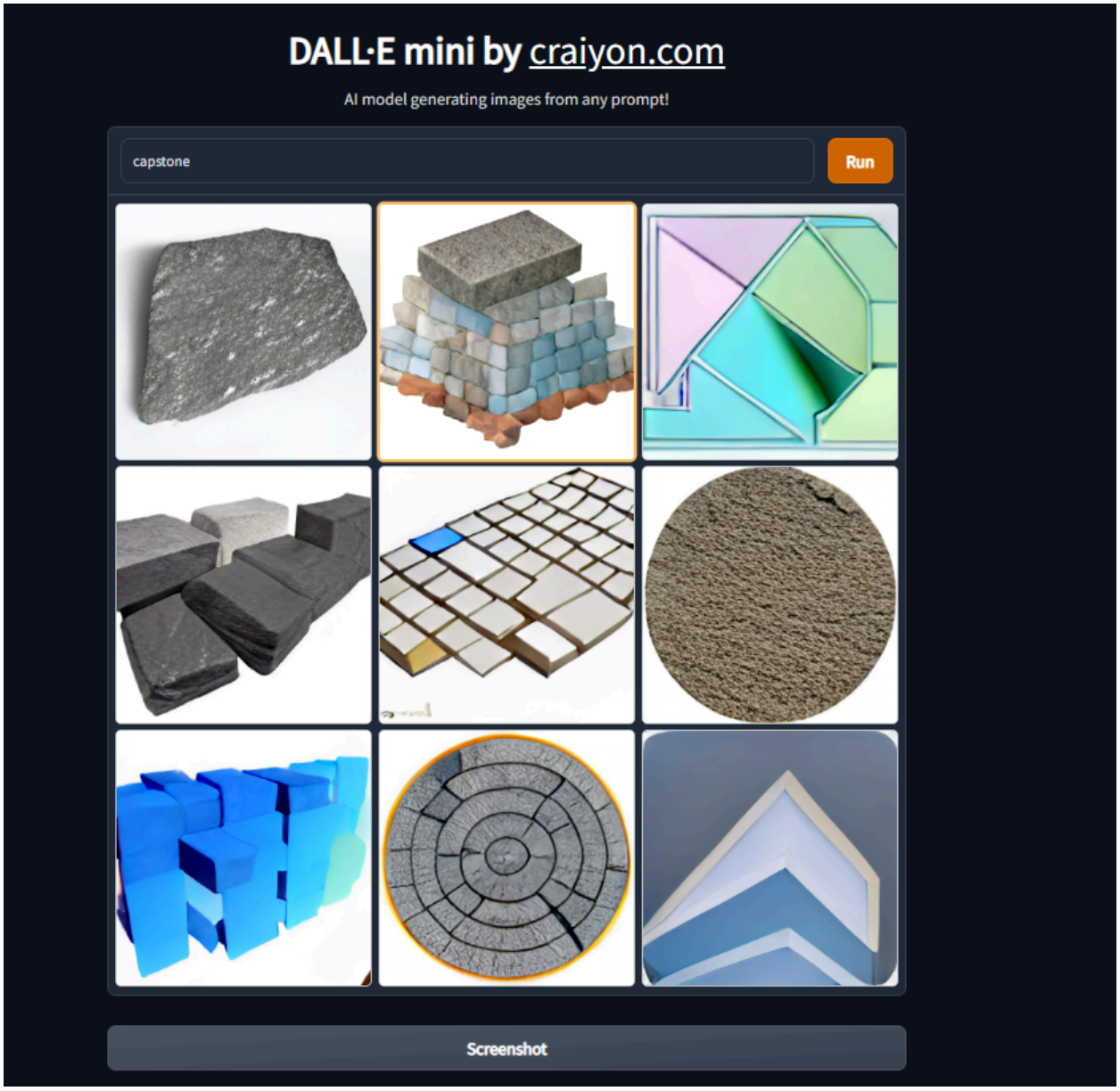
1. User gets onto the website.
2. User clicks on the prompt box.
3. User types the prompt for the face they want generated.
4. User presses enter.
5. Website presents the generated image to the user.
6. Goes back to number 2.



4.2. User characteristics

- Marketing Professionals and Agencies.
- Graphic Designers and Creative Agencies
- Product Developers.
- Video Game Developers and Animators.

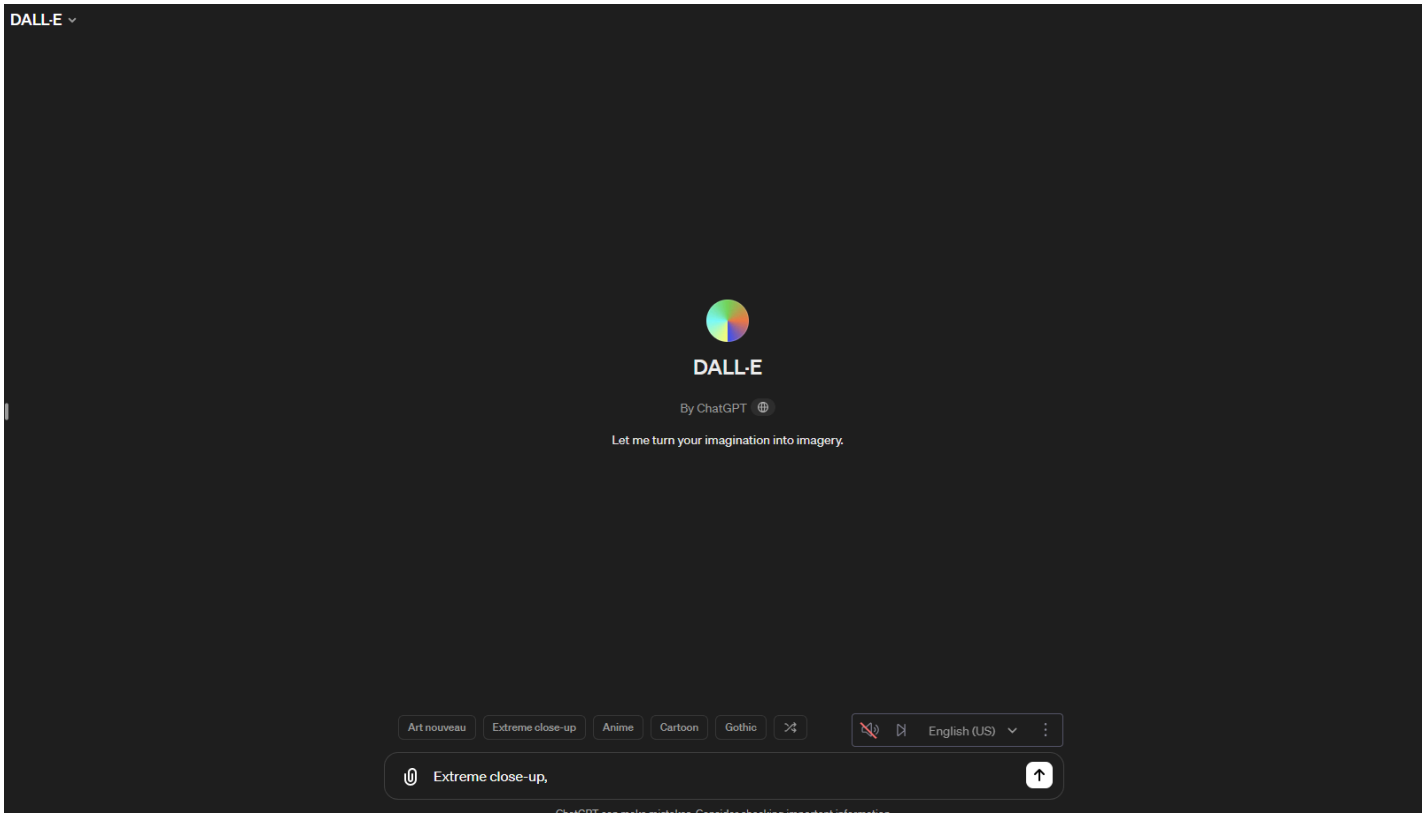
4.3. User interface



Users are able to enter a prompt that they want in the prompt box, but must follow guidelines. These prompt guidelines include:

Do not create images or use text prompts that are inherently disrespectful, aggressive, or otherwise abusive. Violence or harassment of any kind will not be tolerated.No adult content or gore. Please avoid making visually shocking or disturbing content.

The website generates multiple images from what the user enters after they click the generate button. The user can click the Screenshot button to save the images the website presented. Website: https://huggingface.co/spaces/dalle-mini/dalle-mini



https://chat.openai.com/
This design choice is effective in drawing the user's attention directly to where they will interact with the application.

Beneath the input box, we see a selection of buttons such as "Art nouveau," "Extreme close-up," "Anime," and others, which serve as filters or stylistic guides to fine-tune the image generation process. Users can presumably click these to quickly apply specific styles to their image requests, streamlining the experience and adding layers of customization.

The footer of the interface is neat and includes a language selection option, hinting at multilingual support, which caters to a global user base. Overall, the interface design conveys a user-centric approach, providing a straightforward and intuitive pathway for transforming one's imagination into digital imagery. It embodies a modern web aesthetic that balances functionality with a clean and unobtrusive layout, likely to appeal to a wide range of users from casual browsers to creative professionals.

5. Development Environment

Developer Tools:

- WSL <version 22.04.5>
- Visual Studio Code <version 1.87.2>
- Anaconda <version 23.9>
- Python <version 3.11.3>
- Docker <version 25.0>
- Git <version 2.44.0>
- HTML5
- CSS

Content Generation Tools:

n/a

Data Tools:

n/a

3rd Party Dependencies and APIs:

- PyTorch <version 2.2>
- Matplotlib <version 3.5.1>
- OpenCV <version 4.9.0>
- Pandas <version 2.2.1>
- Numpy <version 1.26.4>

Entity Relational Diagram

n/a

6. Production Plans

Code Review Plan

Requisites for Review

- All functionality is complete and working
  - Has been tested on end device
  - Needs no further revision
- Meets all coding standard requirements
- Functions are properly commented
- Test as a team BEFORE Build Reviews.
- All completed work needs to be present on the Development branch unless there is a good reason for otherwise.
- Meet the day before the milestone due dates.

Integration Plan

Source Control

- Git will be utilized for continuous development, with a main branch for the presentation ready code. Each team member will work off a feature branch for features they are implementing, and once that feature is complete and ready to merge we will all merge together to take care of any issues with the merge. Each feature will have its own branch.
- Docker will be used to backup the development of the program. When we have successfully all merged our feature branches we will then back the program up in docker.

Hosting

- The backend will be hosted on HuggingFace as a private model.
- The frontend will be hosted on HuggingFace as well under a private space to ensure only invited users can access it.

Testing Plans

Weekly Code Reviews

- The Day before Build Reviews with Mentor/Producer:
  - 1-2hrs - priority one
    - All team members will test all code completed during the work week
  - +15-30 min complete/incomplete discussion

Bug Database

- Maintained and updated by the team members
- Software
  - Trello

Outside Testing

- Month 4 of Capstone

- 5 Testers selected
- Cross Testing will be expected on a monthly basis with other teams in CS and GD
- A Testing Report Compiled
    - Scenarios planned
    - Findings
    - Surprises
    - Planned Solutions

7. Folder Hierarchy