

AILYAH - An Interactive AI Image Generator

Business Case and Draft Plan

Group 4-K3

1 Business case

1.1 Executive Summary

Image generation using artificial intelligence (AI) has been on the rise in recent years. While there are many platforms that provide this tool, users are often constrained by the limitations of having just a single text input box. Our project addresses this challenge by developing a user-friendly web interface for AI image generation with two key advantages. Firstly, it accepts images as input prompts, providing users with enhanced creative freedom. Secondly, it enables users to input multiple prompts simultaneously. By facilitating the expression of complex requirements, we believe that our solution will significantly enhance the user experience, ultimately contributing to greater satisfaction and usability.

1.2 Project Motivation

In recent years, there has been a significant rise in the market demand for artificial intelligence, including its applications in the field of AI image generation. According to Fortune Business Insights (2024), the global market for AI image generators has experienced substantial growth, with a valuation of approximately USD 257 million, and a projected escalation to nearly USD 1 billion by 2030.

Although the market is huge, it is still most dominated by a handful of companies, most of them are based in the USA. Notably, Midjourney, Dall-E, NightCafe, Starryai, and Stable Diffusion take an overwhelming 99.94% of the global market share (Thormundsson 2023). Although these companies offer outstanding image generation results, they still have several disadvantages that need to be addressed. For example, their interfaces remain simple, lacking the capability for users to input multiple prompts simultaneously or to incorporate various input types such as images and texts in combination.

Therefore, there is a need for a more interactive platform that empowers users to freely express their requirements, thus addressing the limitations of the existing solutions. Moreover, unlike proprietary image generators, which usually charge users fees, our platform uses the power of free and open-source machine learning algorithms. This approach democratises AI image generation, making it more accessible to more people.

1.3 Project Activities

Our project, named AILYAH, is a cloud-based AI image generator platform. Through AILYAH, users gain the flexibility to input text, images, or a combination of both as prompts. Unlike the existing platforms, users are not confined to a single prompt; they can input multiple prompts to refine their requirements as needed. In addition, users have the option to upload images or select from previously generated ones, providing them with flexible control over their image generation process.

Upon inputting prompts, AILYAH's backend seamlessly transfers the data to our machine learning model (provided by Dr Lingqiao). Leveraging advanced algorithms, the model generates images tailored to meet the users' specifications. Subsequently, these generated images are presented to the users for review and can undergo further iterations as desired.

1.4 Benefits

Our project will bring about several benefits to users

- Enhanced expression of requirements: By facilitating the simultaneous input of multiple prompts, our algorithm can better understand and generate results that closely align with user expectations.
- Accessibility and affordability: By harnessing free and open-source image generation algorithms, we democratise access to AI image generation. This approach not only eliminates financial barriers but also opens up opportunities for users across diverse sectors and industries, whether for personal or professional use.
- By taking in images as inputs, our platform can support various applications such as image to image generation and masked guided generation
- Versatile applications: By accepting images as inputs, our platform can support various applications, ranging from image-to-image generation to masked guided generation. This versatility enables users to explore various creative and practical uses, from artistic expressions to practical problem-solving in industries such as healthcare.

1.5 Goals

We aim to achieve these goals by the end of the semester:

- To develop an easy-to-use user interface design that facilitates the interaction between the user and the image generation algorithm
- To build a website that is cloud-based, allowing users to access anywhere
- To seamlessly integrate our website with the AI image generation model, enabling the transmission of prompt inputs and the display of image outputs in real-time
- To build a scalable database that can record the input and the outputs, facilitating the iterations on the generated images

1.6 Future Directions

If the timeline is longer, we would like to build more features to make this website more on par with full-fledged commercial products:

- Let users create accounts, storing login credentials and maintain the users list
- Add in more flexible options for image download and sharing
- Provide suggested keywords for users to easily generate prompts
- Filter prompts generated by users to avoid both receiving and creating inappropriate and harmful images
- Make our user interface more accessible to users with disability, such as adding alternate texts for voice-over, or ensuring that all functionality and interactive elements can be operated using a keyboard alone

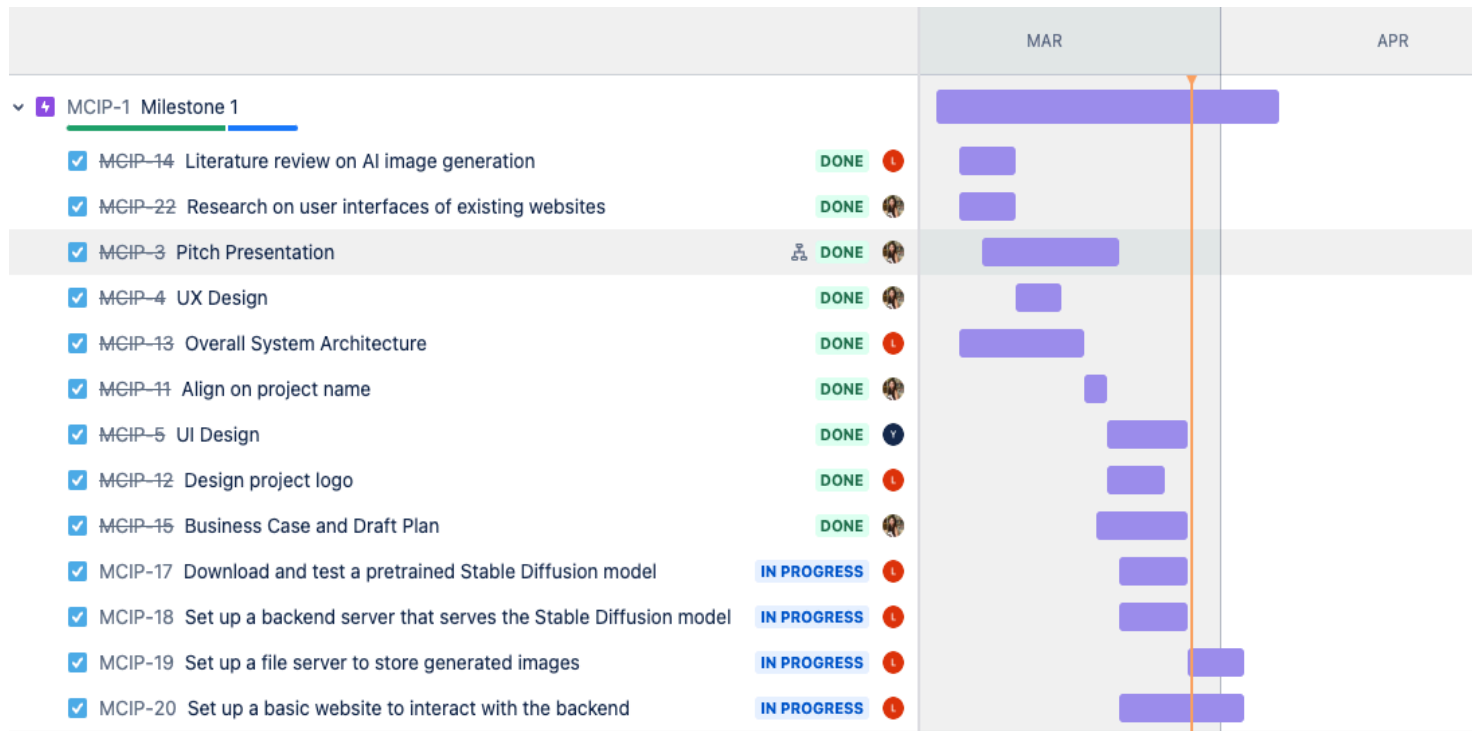
2 Draft Plan

2.1 Team Roles

- Project manager is Amber who will track the progress of the projects, host meetings, ensuring that the project moves on according to the plan. She will also be the main contact point of communication between the team and the client, ensuring that our product is aligned with the expectations of the client.
- Developer: Harry is our tech lead and main developer in this project. Harry will design the main system architecture and take the majority of the coding work and assign coding tasks to others.
- Designer: Amber is responsible for the UX while Yuqing is in charge of the UI. Having two persons working on the design ensures that our platform delivers an intuitive and smooth user experience as it is the key value of our project.
- Tester: Yuqing is our main tester for this project. She will prepare the test cases, conduct testing and verify the fixes to ensure that our website is robust, reliable, and free from defects.

2.2 Schedule

Below is our draft schedule for milestone 1.



Milestone 1	Activities	Projected Outputs	Assignee
<p>The main goals of our milestone 1 are confirming the system's overall architecture, finishing design of the front end and completing a basic version of the backend.</p>	Background research		
	Literature review on AI image generation	List of capabilities of current algorithms	Harry
	Research on user interfaces of existing AI image generators	Comparison of existing systems interfaces	Amber
	Design		
	User experience (UX) design	Document of UX flows on Figma	Amber
	User interface (UI) design	Document of UI specifications on Figma	Yuqing
	Logo design	Logo image	Harry
	Development (Backend)		
	Overall system architecture	Diagram of system architecture	Harry
	Download and test a pre trained diffusion model	The model works and can generate images	Harry
	Set up a backend server that serves the Stable Diffusion model	Backend server is set up and running	Harry
	Set up a file server to store generated images	File server is set up and running	Harry
	Development (Frontend)		
	Set up a basic website to interact with backend	Finished the frame of the website	Yuqing
	Reporting and Assignment Submissions		
	Pitch presentation	Presentation video is done and submitted	All
	Business case and draft plan	Document is completed and submitted	All

2.3 Communication Plan

Below is our communication plan and tools used.

Communication with	Type	Medium	Frequency
Team members	Text message	Whatsapp	Daily
Team members	Meeting	Zoom	Weekly
Team members	Task management	Jira	As needed
Team members	Program codes	Github	As needed
Team members	Project documents	Github	As needed
Team members	Design documents	Figma	As needed
Supervisor	Project updates	Email	Weekly
Supervisor	Meeting confirmation	Email	As needed
Supervisor	Meeting	In person	As needed
Supervisor	Project document	Github wiki	As needed

Internal Communication:

- Daily communication: Utilise WhatsApp for quick and informal daily updates and discussions among team members.
- Weekly meetings: Conduct weekly meetings every Wednesday at 7pm to review individual progress, discuss project-related issues, and align on next steps.
- Project management tool: Utilise Trello for task management, assigning tasks to team members with due dates and tracking progress by moving cards through different statuses.
- Document repository: Store project documents, including requirement documents, meeting minutes, assignments, and test cases, on GitHub for easy access and version control.
- Design files storage: Use Figma to store all design files, taking advantage of its advanced features and seamless conversion to codes for efficient design collaboration.

External Communication:

- Email updates: Provide weekly email updates to the project supervisor/client, outlining progress, milestones achieved, and any relevant updates.
- Meeting: Schedule in-person meetings every 2-3 weeks for more detailed discussions, addressing questions, and ensuring alignment with client expectations.

3 References

Fortune Business Insights 2024, *AI Image Generator Market Size, Share & COVID-19 Impact Analysis, By Application (Personal and Enterprise), By End-user (Advertising, Healthcare, Gaming, Fashion, E-commerce and Others), and Regional Forecast, 2023-2030*, viewed 28 March 2024, <<https://www.fortunebusinessinsights.com/ai-image-generator-market-108604>>

Thormundsson, B 2023, *Leading generative artificial intelligence (AI) powered image tools market globally in 2023*, viewed 28 March 2024, <<https://www.statista.com/forecasts/1423989/world-generative-ai-image-tool-market-share>>