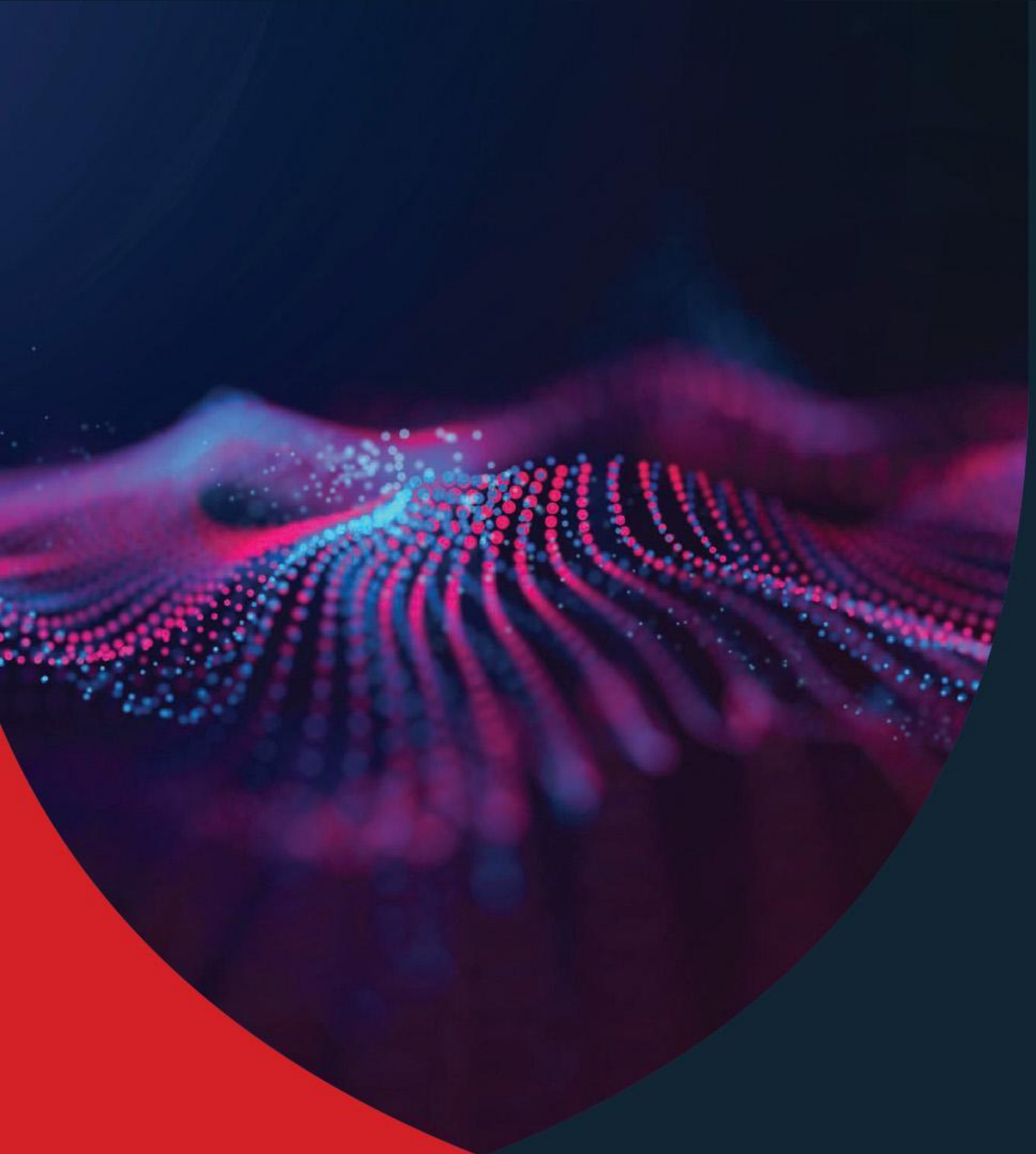


Project Pitch: Interactive AI Image Creation Platform

Supervisor: Dr Lingqiao Liu

Group 4-K3

THE UNIVERSITY of ADELAIDE

Content

Section	Presented by
Industry overview	Yuqing Chen
Analysis of current solutions	Yuqing Chen
Our project solution	Amber (Ngoc Anh Dao)
Analysis of our solution	Harry (Hoang Son Le)
Overview of the project plan & Conclusion	Harry (Hoang Son Le)

THE UNIVERSITY of ADELAIDE

AI Image Generation Industry Overview

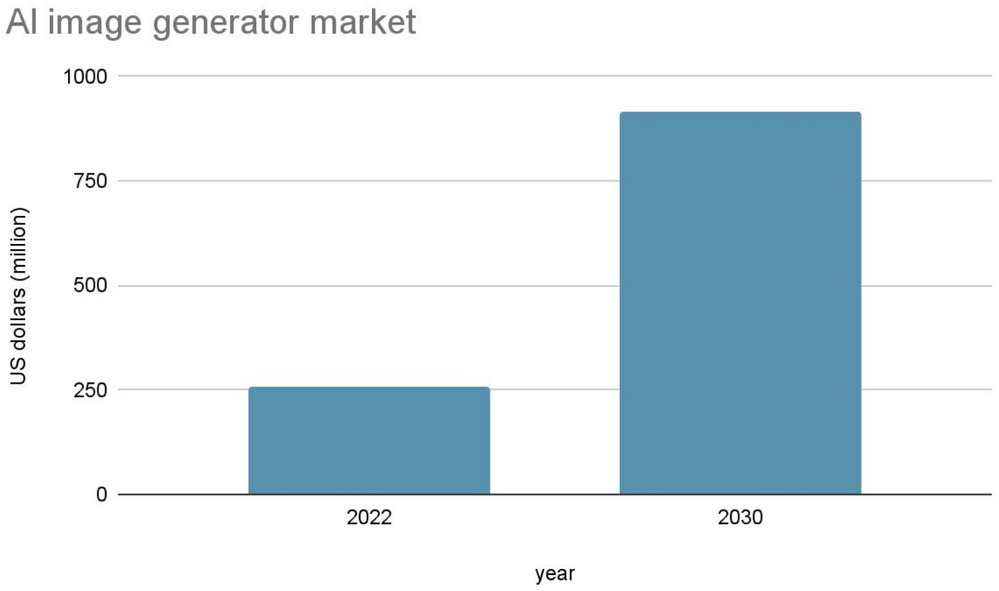
An AI image generator is a technology powered by Artificial Intelligence (AI) that can produce images, artworks, or visual content without direct human intervention.

THE UNIVERSITY of ADELAIDE

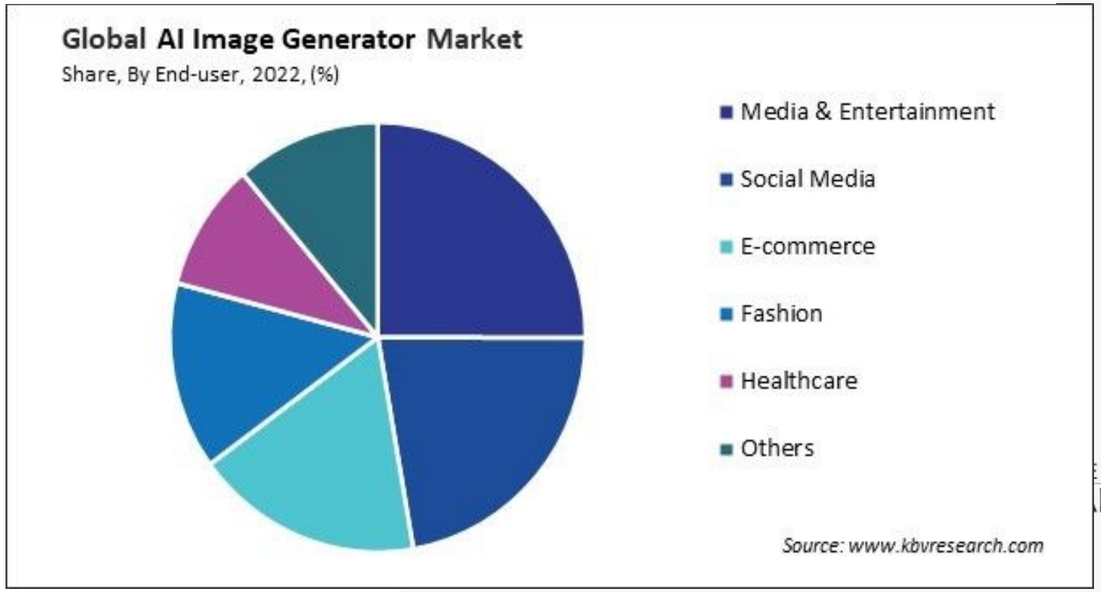
AI Image Generation Industry Overview

The AI image generator market has significant growth in recent years.

- **According to the Market Research Report, the AI image generator market was valued at US$257.175 million in revenue in 2022 and is anticipated to reach US$917.448 million by 2030[1]**

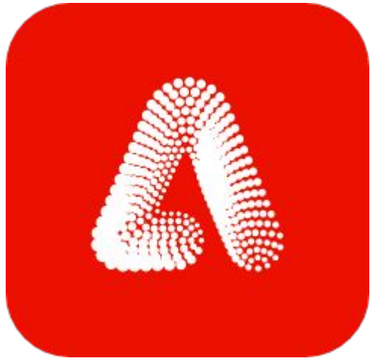


- **On the basis of end-user, the market is divided into media & entertainment, healthcare, fashion, social media, e-commerce, and others[1].**



Major players in the field

Currently, the main image generator tools on the market come from: Midjourney, ChatGPT, Adobe and Canva.



THE UNIVERSITY of ADELAIDE

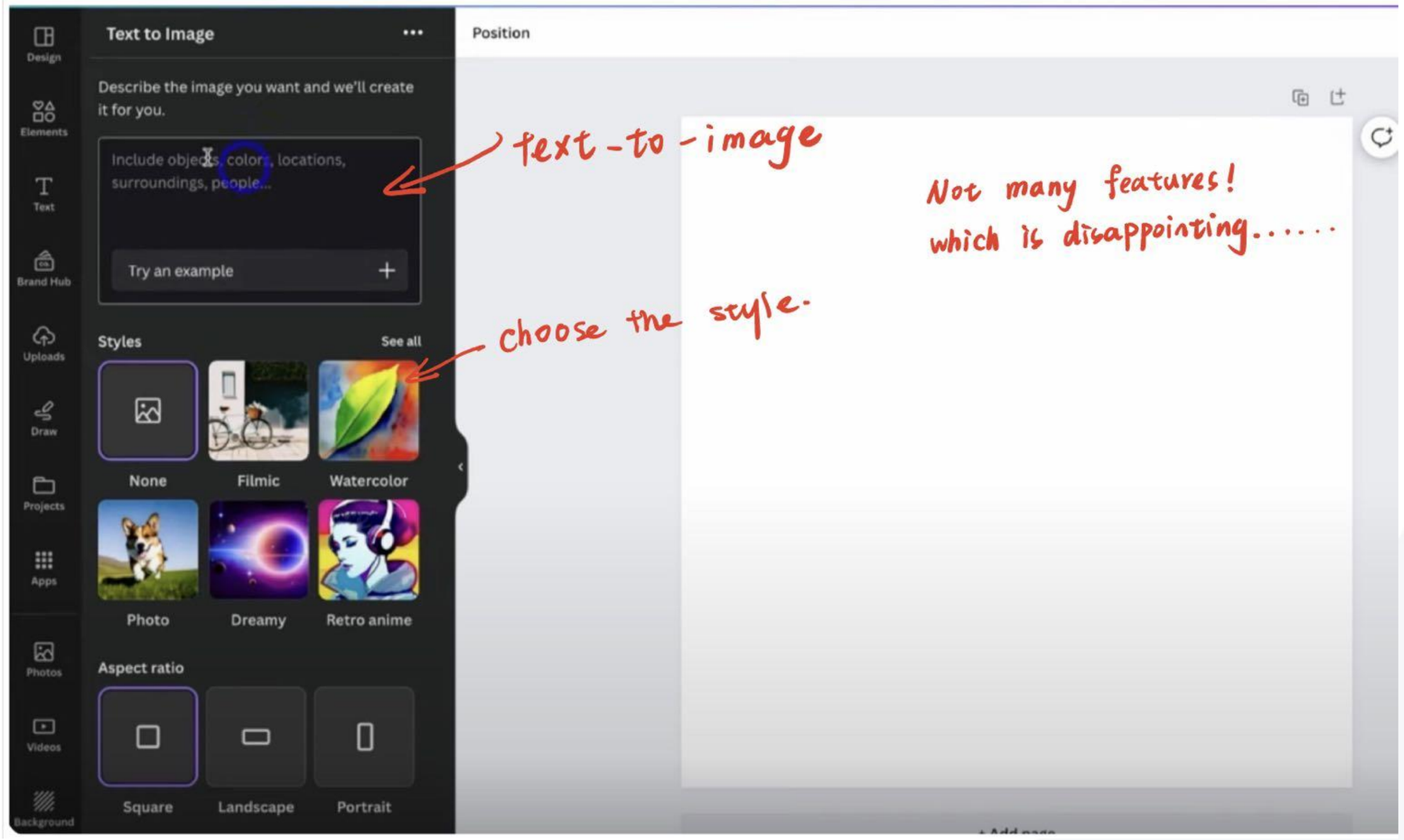
Market share by countries

North America still holds the major market share of ai image generator. According to Market Analysis Report, in 2022, North America has a market share of 36%[2].

THE UNIVERSITY of ADELAIDE

AI Image Generation in Australia

- **In Australia, Canva has an add-on for image generation but still in early stages.**



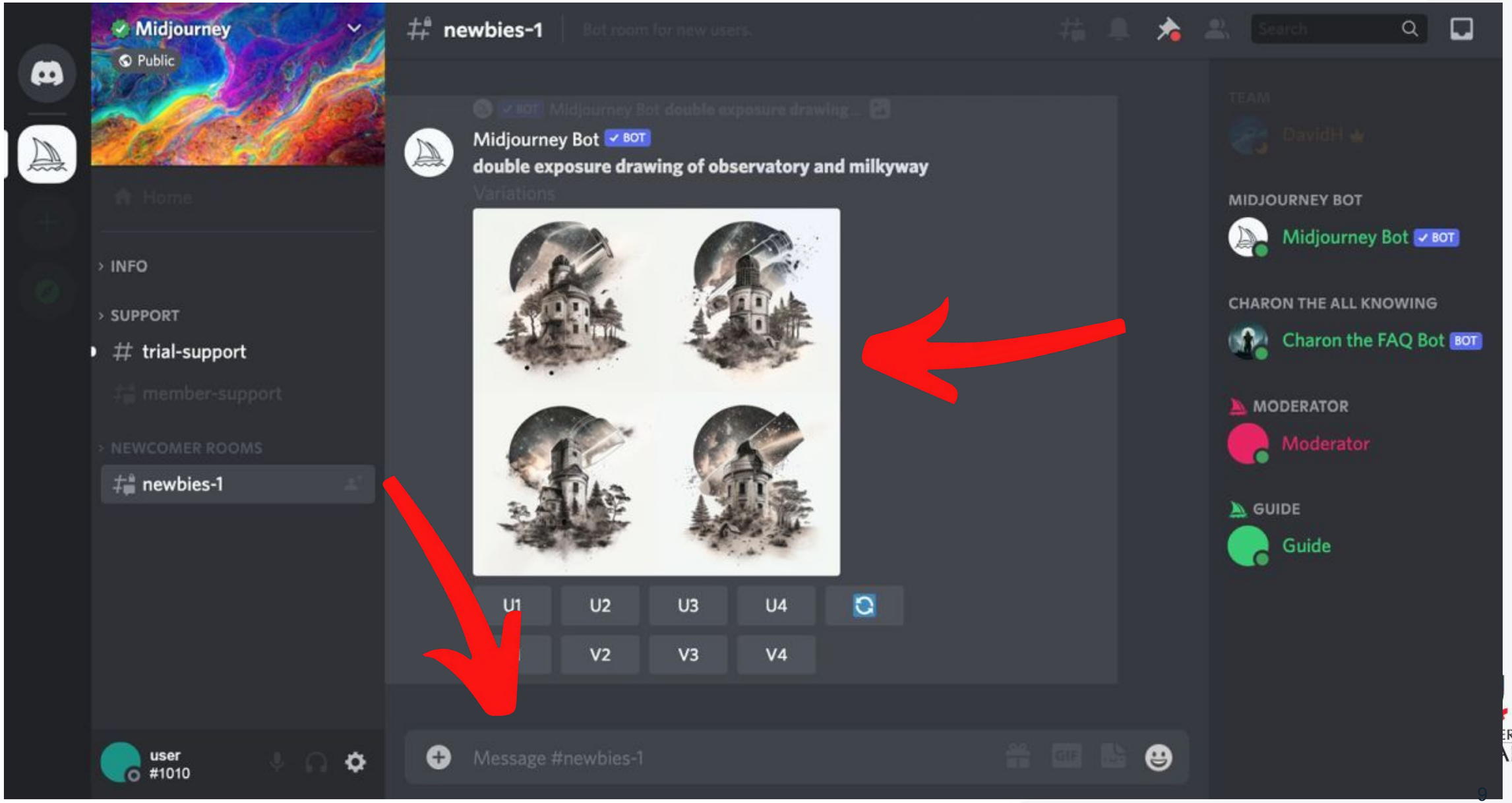
That is because building a Machine Learning model from scratch is expensive, not many companies can do it.

But we can explore other resources. Such as collaboration with universities.

That’s why we need this project!

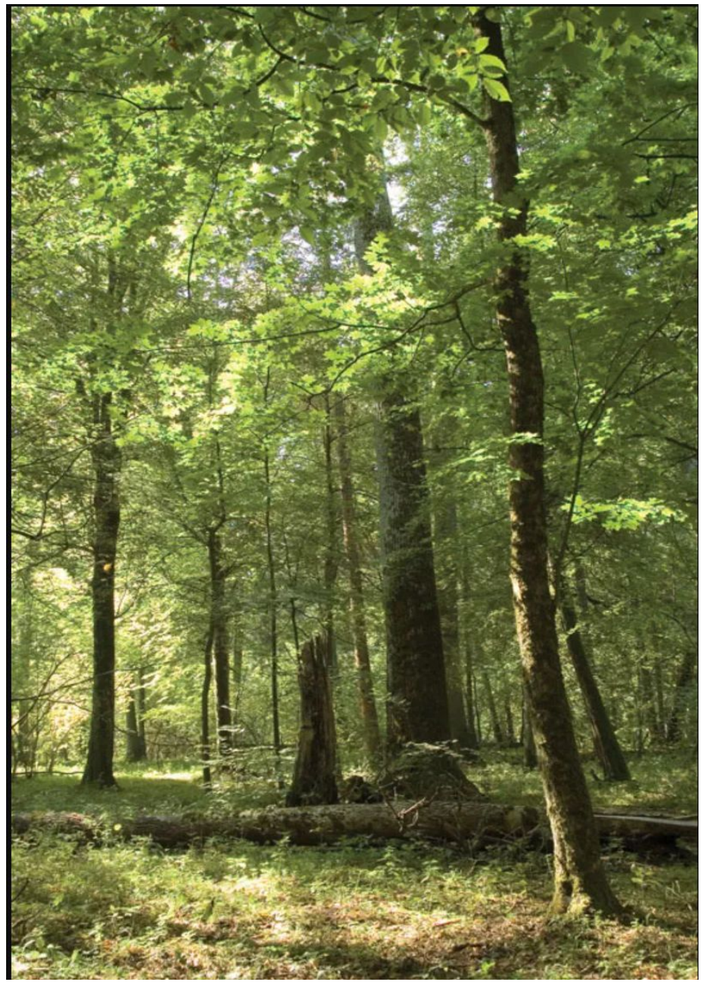
THE UNIVERSITY of ADELAIDE

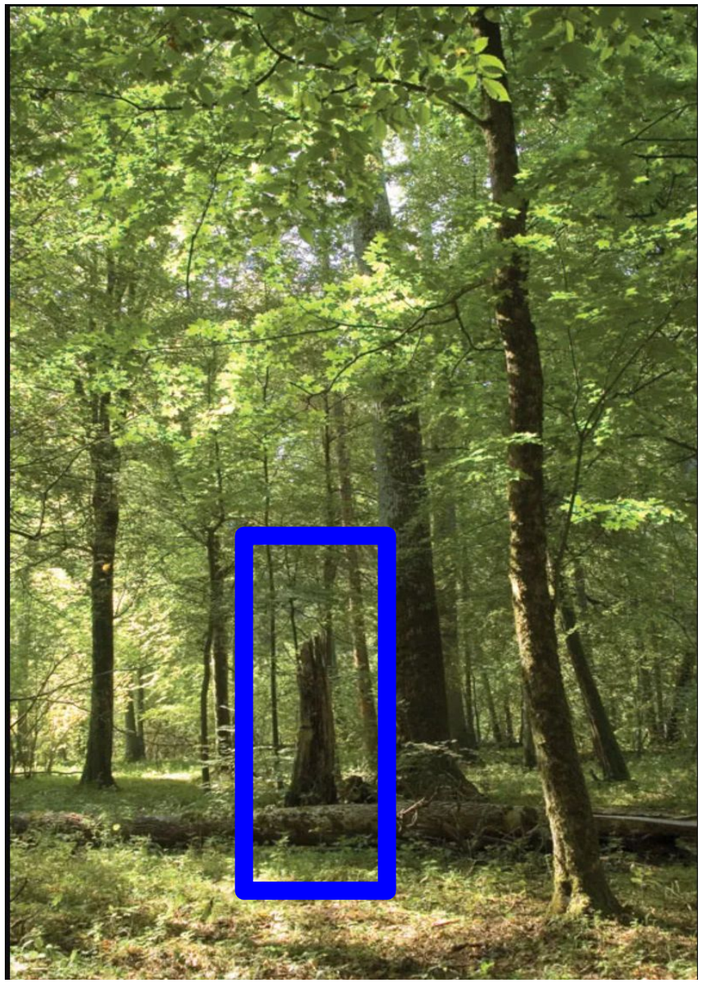
Analysis of current solutions



Limitations of Existing Systems

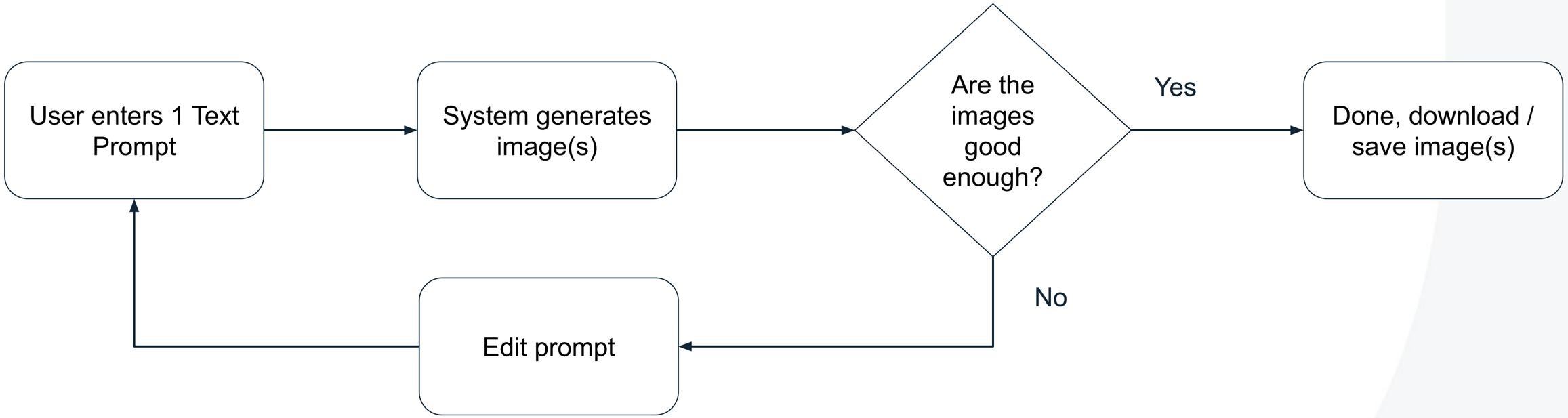
- Text prompt as the only means of interaction
- Lack of controllability - difficult to represent abstractions



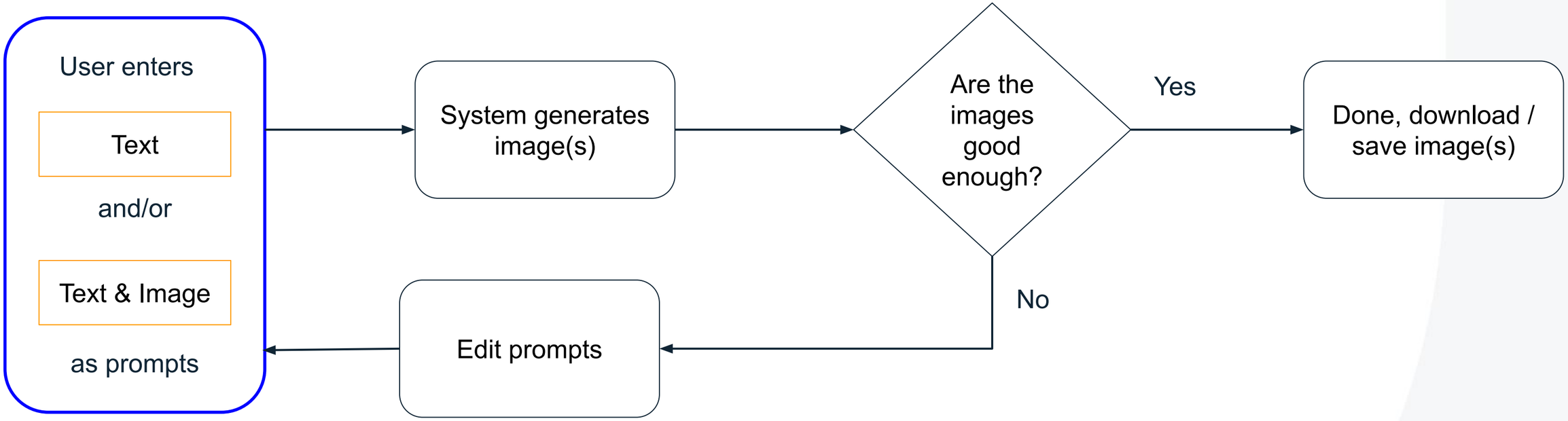


THE UNIVERSITY of ADELAIDE

Existing Solutions



Our Solution



Key differences compared to existing solutions:

1) Taking multiple prompts
2) Taking images as prompts which can be
   - Reference images
   - Sketches of the final image
   - Part of a previously generated image

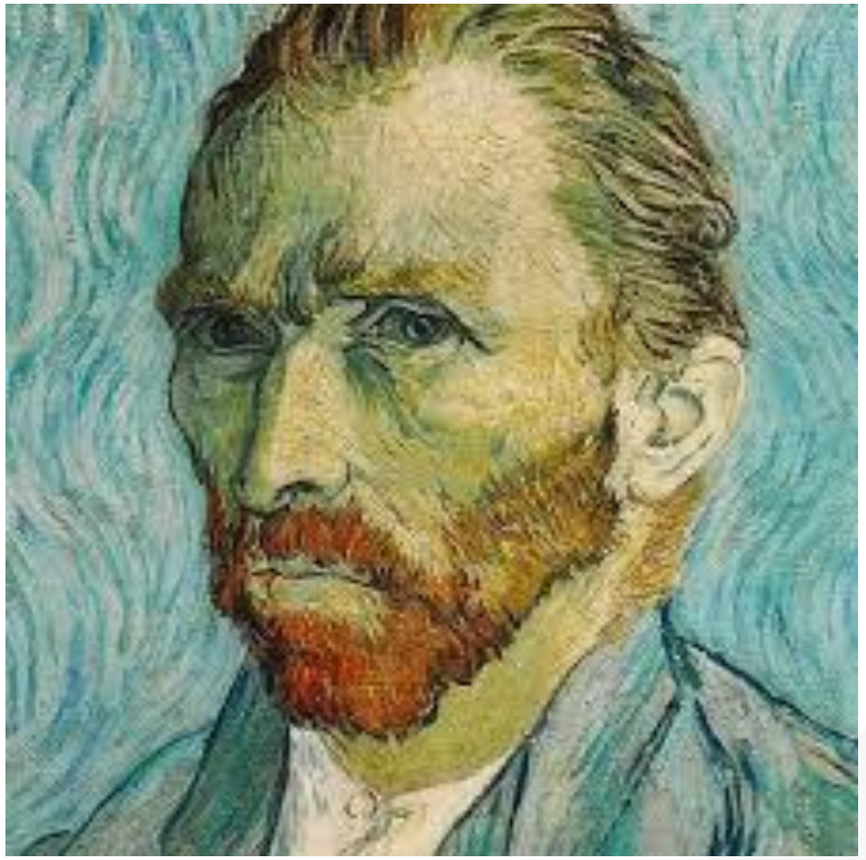
THE UNIVERSITY of ADELAIDE

Key Difference - Taking Images as Input

500 × 500

Reference Image

eg. "Convert this painting into a photo"



THE UNIVERSITY of ADELAIDE

Key Difference - Taking Images as Input

Sketches

eg. "Complete sketch to make an image of a walking man"

THE UNIVERSITY of ADELAIDE

Key Difference - Taking Images as Input

Part of a previously generated image

eg. "Replace the teddy bear selected by a puppy"

User can combine multiple prompts, such as providing the image of the said puppy

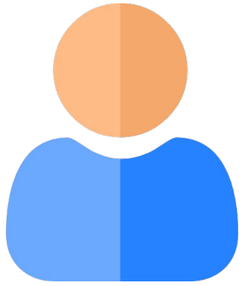
THE UNIVERSITY of ADELAIDE

Supporting features

They are features that already exist in current solutions - they are not key differentiators but still needed to bring our product on par with current solutions

Project organisation to separate different topics / creations

Hosted on cloud for easy access anywhere

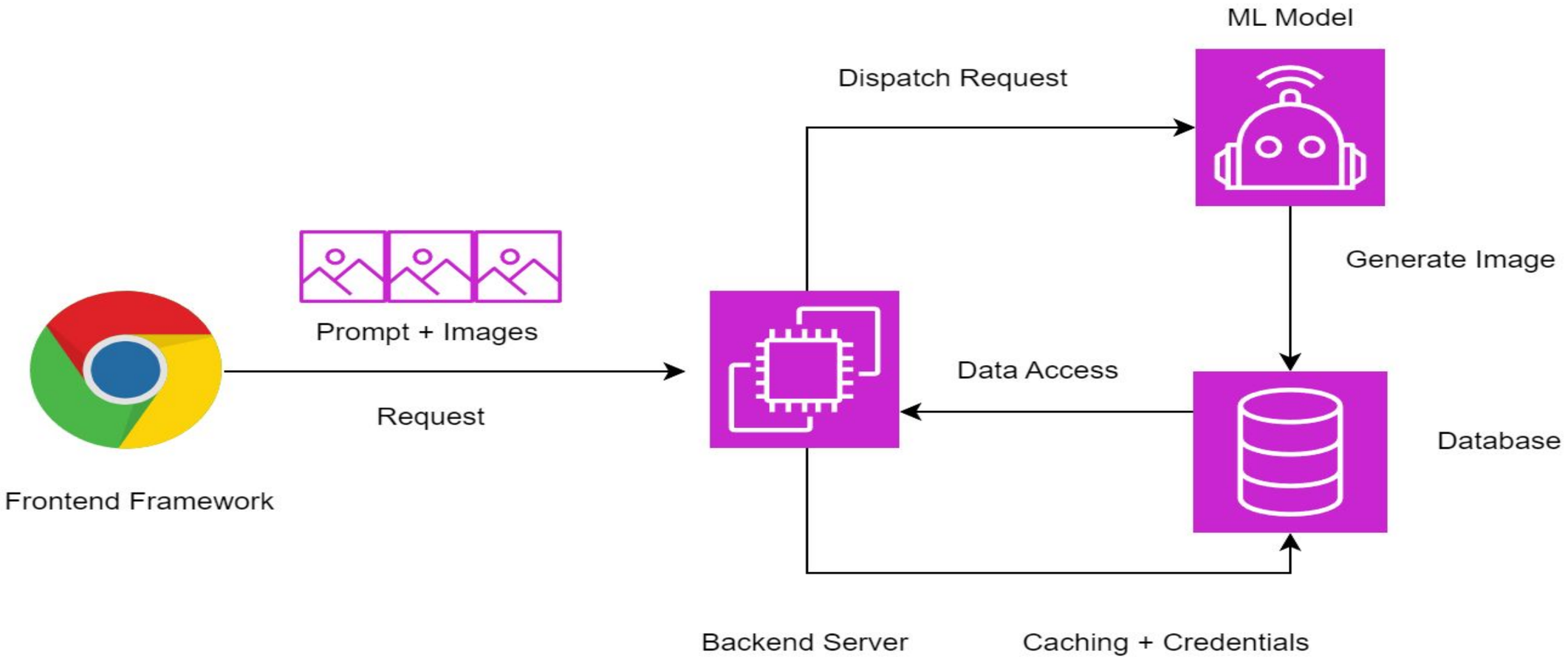


User account creation and login

Flexible options to download and share files

THE UNIVERSITY of ADELAIDE

Proposed System Architecture



Proposed Techstack

Frontend:

- React
- React-router
- TailwindCss

Backend

- Litestar
- Postgres
- File Storage (S3/GoogleDrive/OneDrive)

THE UNIVERSITY of ADELAIDE

Analysis of our solution

Incorporate features of existing solutions

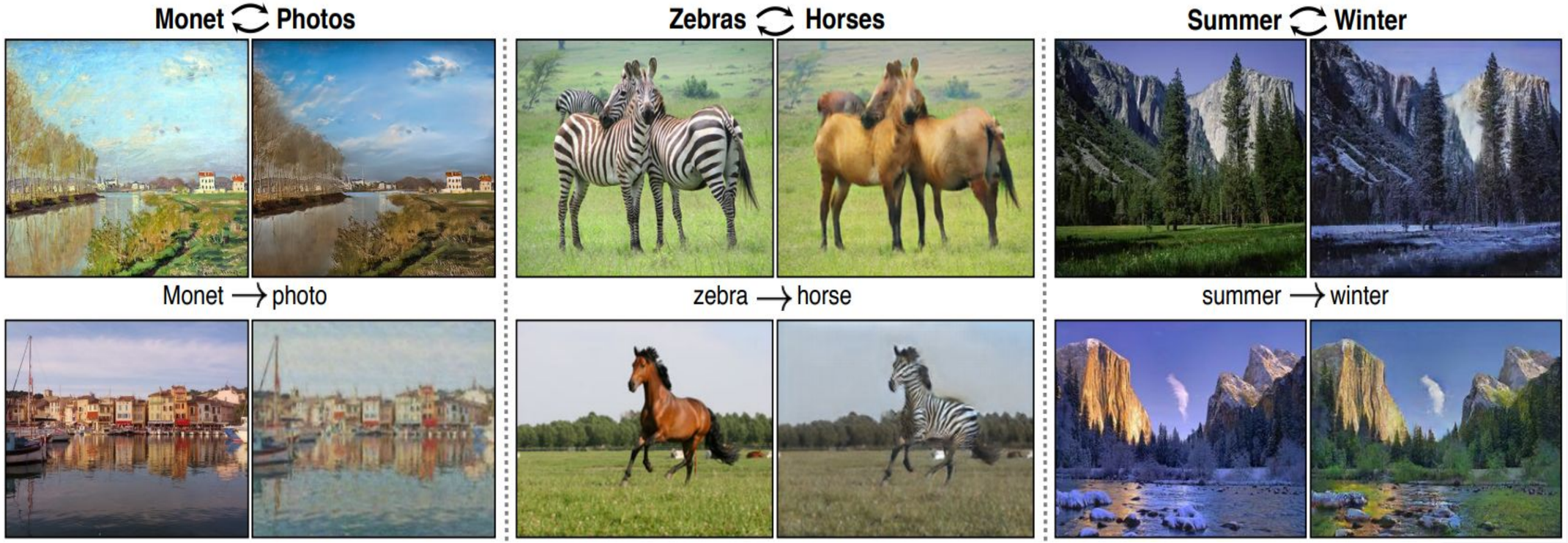
- Text to image generation
- Chat log/conversation log

Add features based on current research capabilities

- Image editing
- Masked inpainting
- Image to Image translation

THE UNIVERSITY of ADELAIDE

Image to Image Generation



Masked guided generation

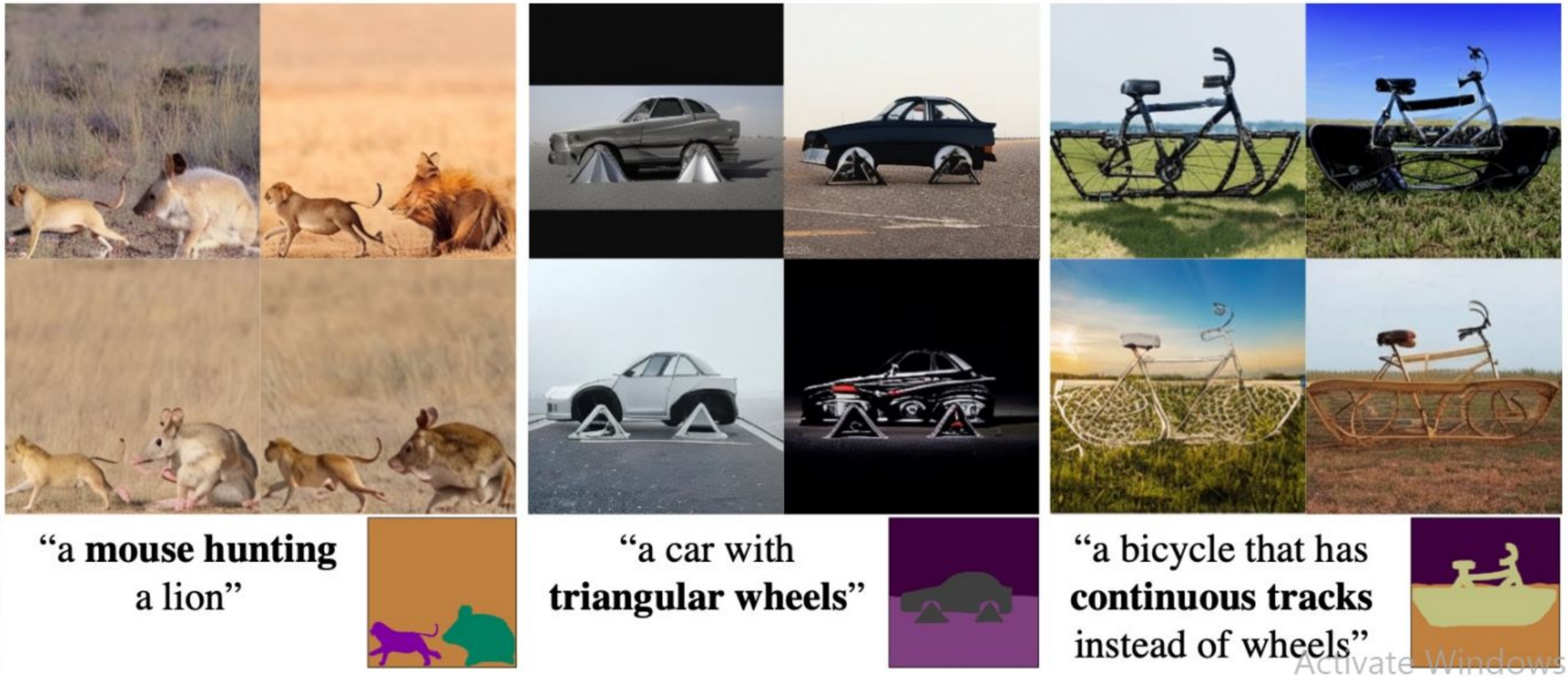
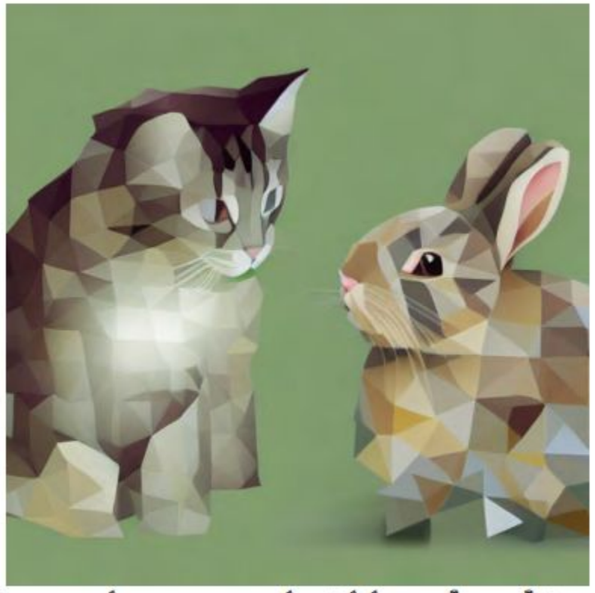


Image to Image translation

Input Real Image



"A polygonal illustration of a cat and a bunny"

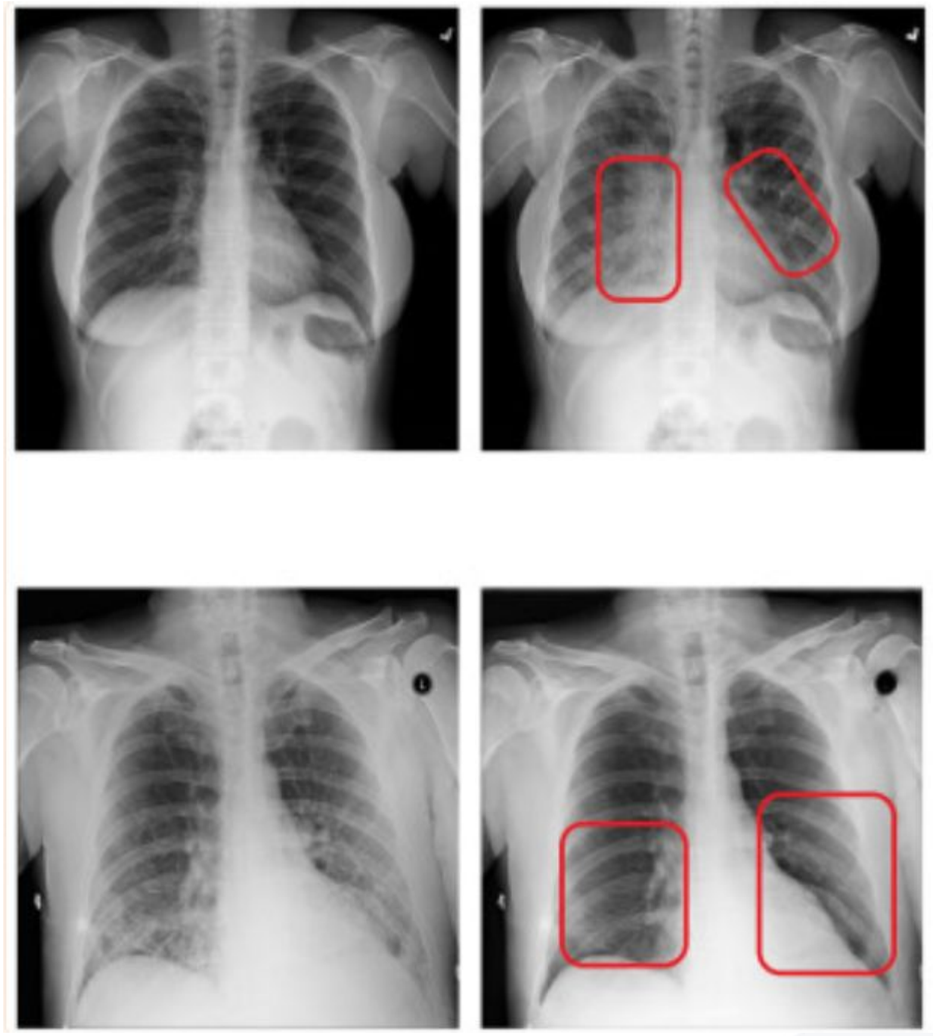
"A photo of bear cubs in the snow"

Possible applications - Medical

Counter-factual analysis:

Example:

“Generate a CT-Scan of a person not afflicted with pneumonia”



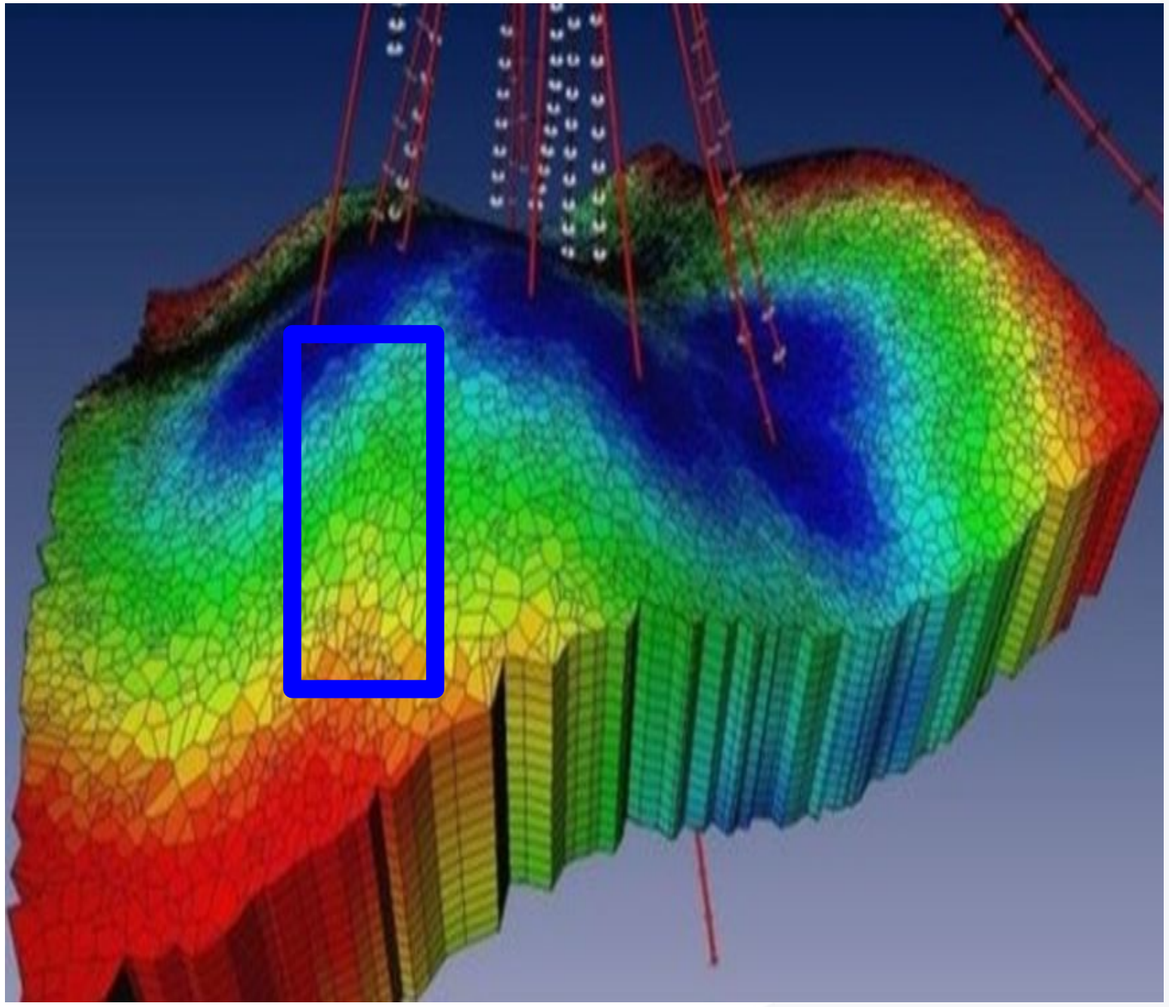
THE UNIVERSITY of ADELAIDE

Possible applications - Geology

Guided structural realisation based on human prior:

Example:

"Generate another realisation where there is a fault zone in the highlighted area"



THE UNIVERSITY of ADELAIDE

Overview of Project Plan

Milestone 1	Finish FE UI design
	Finish backend database system etc.
	Obtain pretrained image generation models
Milestone 2	Implement text to image generation interface
	Implement image editing interface + feature
Possible extension in the future	Login/credentials
	Deployment
	Scaling and Optimisation

THE UNIVERSITY of ADELAIDE

Conclusion / Summary

Text to image generation platforms exist but

- Lack controllability and flexibility
- Do not support editing features

Provide an open-source text to image generation platform

- Allow for image generation with better human-guided prior
- Enable image editing capability

THE UNIVERSITY of ADELAIDE

References

[1] 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, Fortune Business Insights, viewed March 2024, <https://www.fortunebusinessinsights.com/enquiry/request-sample-book/ai-image-generator-market-108604>

[2] AI Image Generator Market Size, Share & Trends Analysis Report By Component (Software, Services), By End-user (Media & Entertainment, Healthcare), By Region, And Segment Forecasts, 2023 - 2030, Great View Research, viewed March 2024, <https://www.grandviewresearch.com/industry-analysis/artificial-intelligence-ai-image-generator-market-report>

THE UNIVERSITY of ADELAIDE