# AI Art

Evan Lucas and Tim Havens

Overview

- History of AI art and image generation models
- Early generative models
- Birth of aesthetic and adjustable generative models
- Review of some interesting AI art models
- Interactive AI art model exploration
- Show and tell
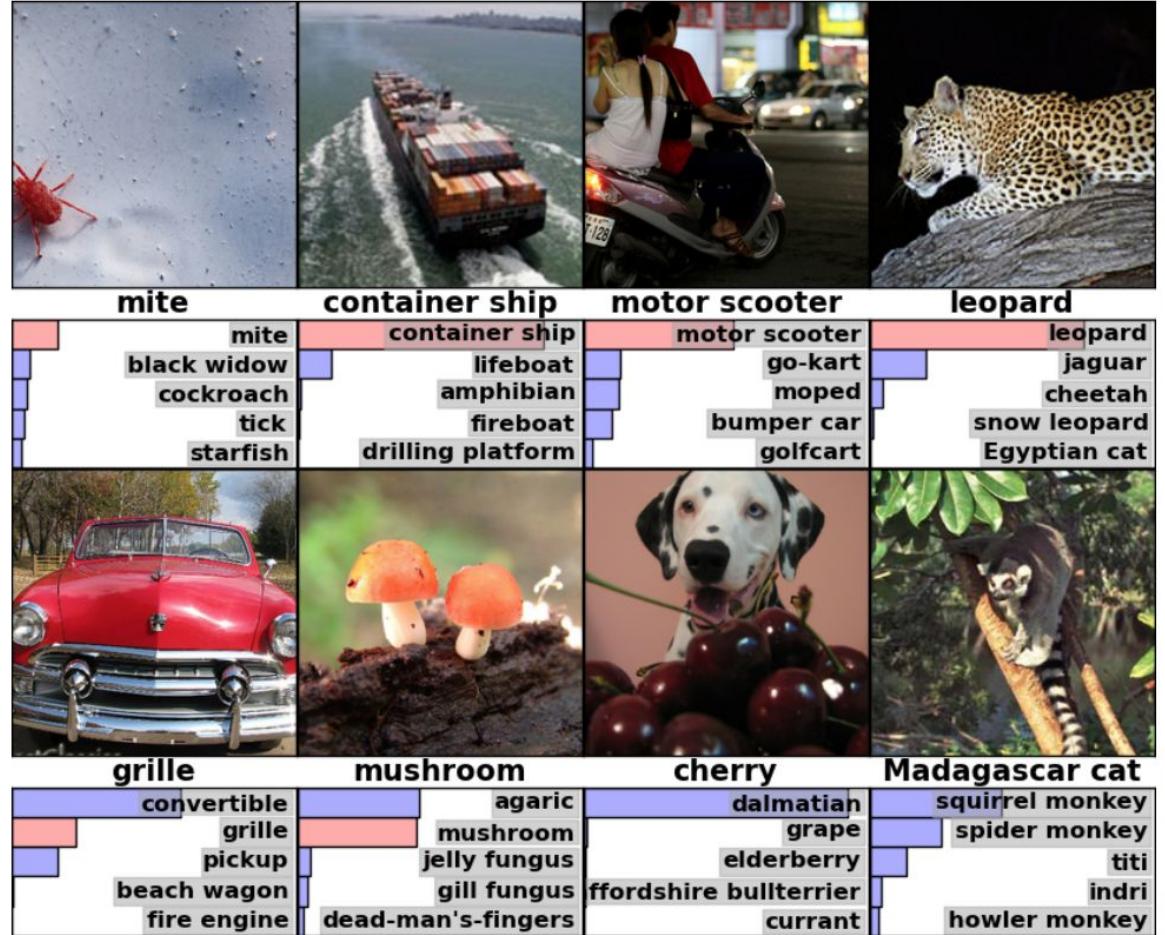
Michigan Tech

The origins of image generation

First steps: making and understanding good classifiers of images

ImageNet (2015) was a huge step forward in labeled image data sets and modern ways to classify them

Michigan Tech

Deep Convolutional Generative Adversarial Networks (DCGAN) 2015

First usable image generation model

Able to make images "within distribution" of training data - ie. it was good at making images of bedrooms if you trained it on images of bedrooms

Researchers learned to tweak it for some modifications, ie. with or without glasses

Michigan Tech

Sidebar: How do GANs work?

Start by training a model to say whether an image is a real image of something (the discriminator)

Is this a cat?

Use that model to train a model that generates images (the generator)

Iterate between improving discriminator and generator

Michigan Tech
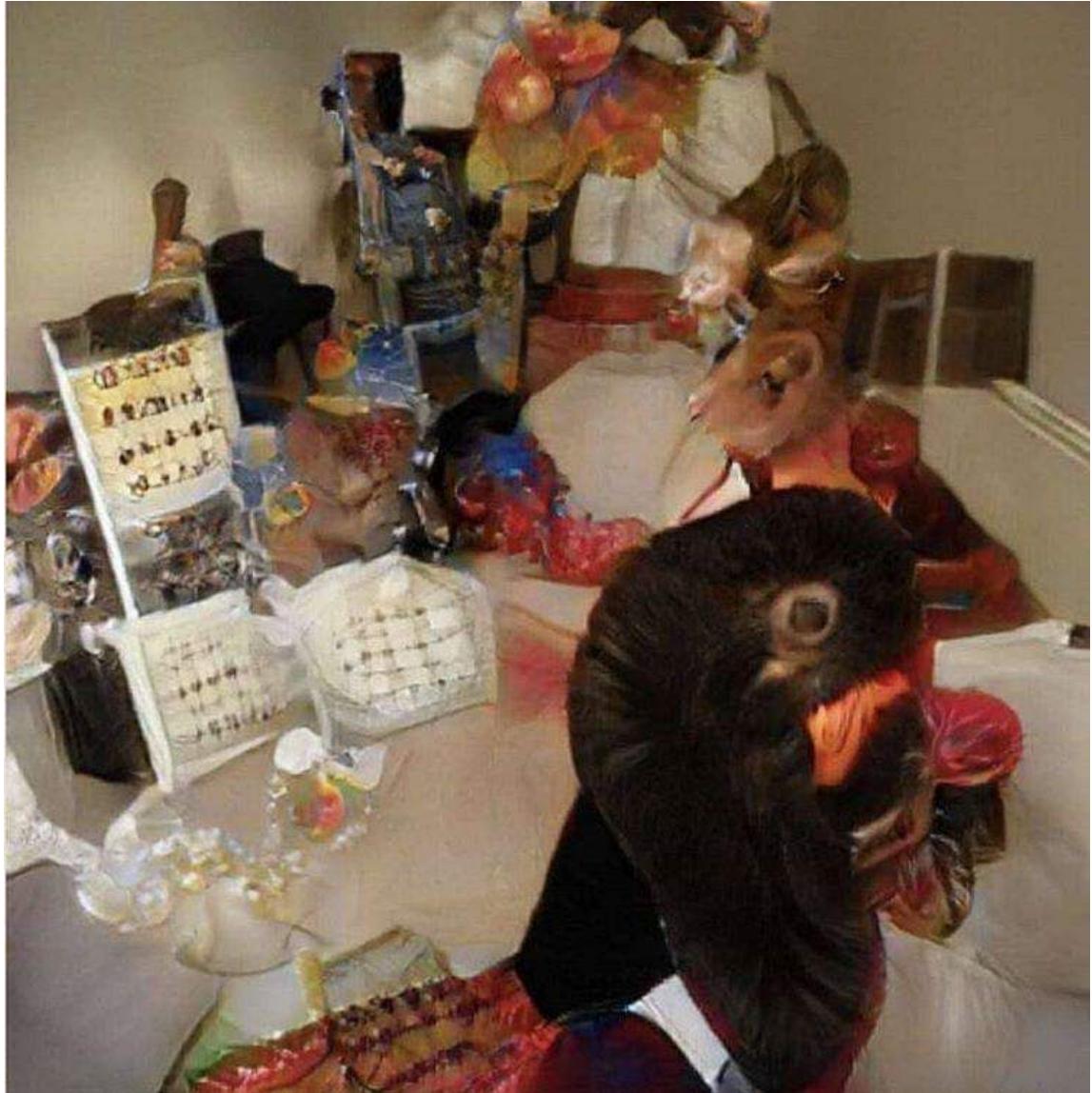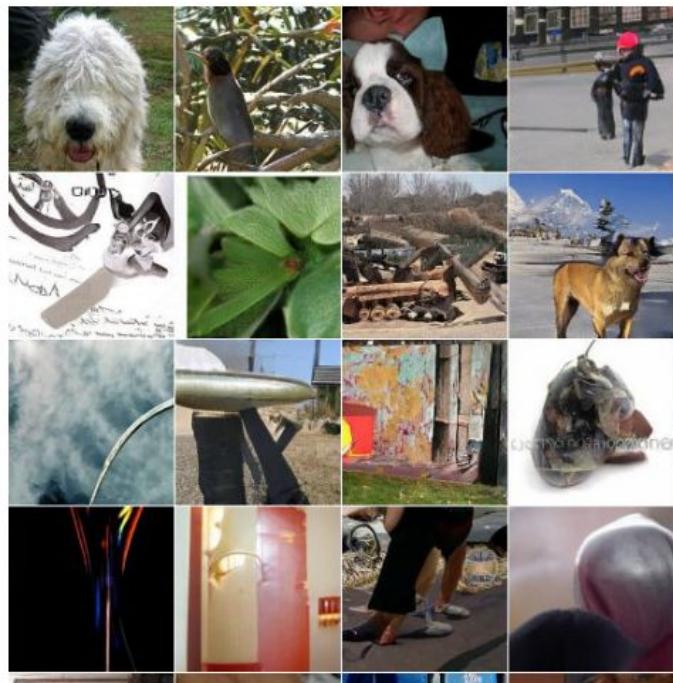
StyleGAN (2019) and StyleGAN2 (2019)

New explorations into making bigger images and generating different objects with the same model

Made famous by thispersondoesnotexist.com (now offline)

https://arxiv.org/abs/1912.04958

Michigan Tech

BigGAN (2019)

An exploration into
 trying bigger models

Michigan Tech

Vector Quantized GAN (VQ-GAN) (2020)

Big step forward in GAN image quality

Michigan Tech

What was the limitation at this point? (2020)
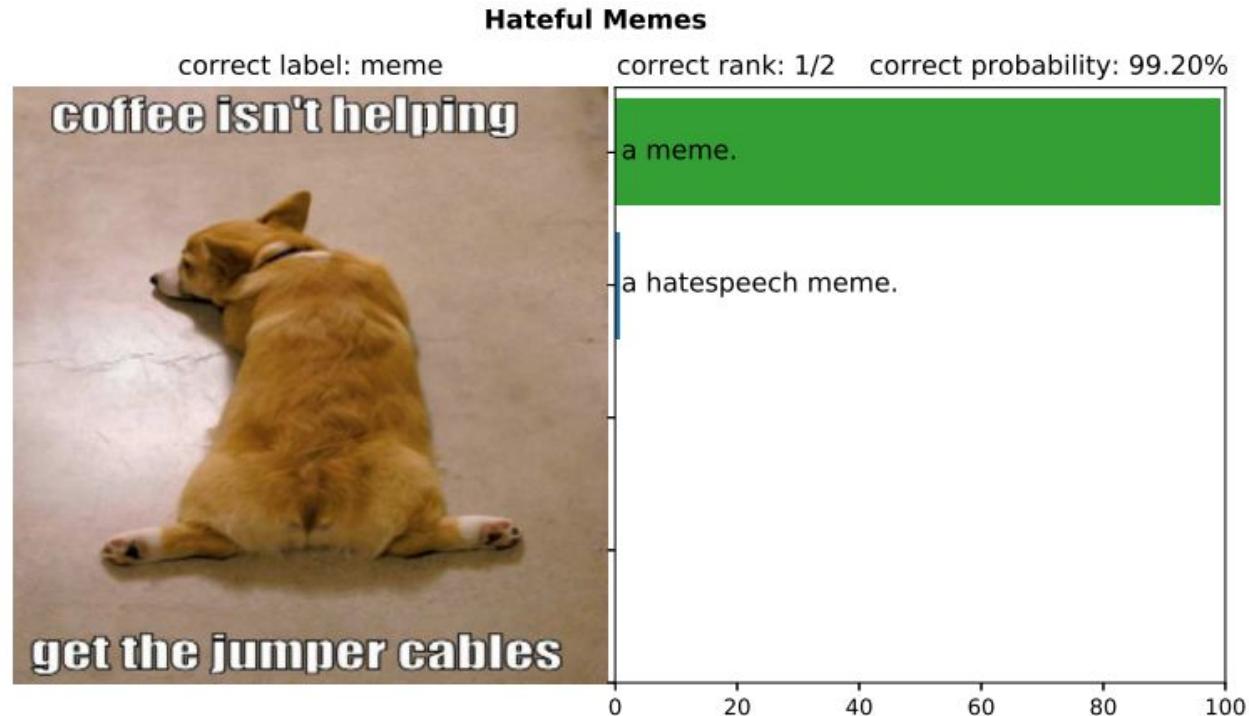
- Models were focused on generating a single object/scene
- Changes to the image generated were limited and often manually engineered

Michigan Tech

Contrastive Language-Image Pre-training (CLIP) (2021)

- Model trained on text-image pairs rather than just image object name
- Allowed a user to query the model about possible candidate image descriptions

Hateful Memes

correct label: meme   correct rank: 1/2   correct probability: 99.20%

coffee isn't helping

a meme.

a hatespeech meme.

get the jumper cables

Michigan Tech

And so it begins

By having a model that can assess whether a given image matches a description, a GAN or other generative model can use this model as the discriminator model.
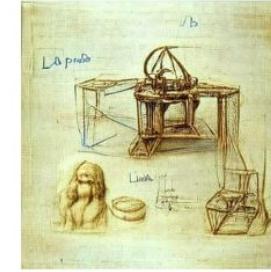
Pay attention to which models interest you in the following slides and try experimenting with them in a few minutes

Michigan Tech

VQ-GAN + CLIP (2021)

Combination of VQ-GAN that makes high quality images with CLIP, which allows any input sentence creates the first real AI art model (in my opinion)

(e) sketch of a 3D printer by Leonardo da Vinci

(f) an autogyro flying car, trending on artstation

(h) Baba Yaga's house + fantasy art

(i) pickled eggs, tempera on wood

(k) the Tower of Babel by J.M.W. Turner

(l) a futuristic city in synthwave style

Michigan Tech

Prompt engineering

a)   b)   c)   d)

Images generated with VQGAN+CLIP [9], 175 iterations, CLIP model ViT-B/32, VQGAN model wikiart_16384,
seed 6087304447281500163, text prompts:
a) "ufo landing"
b) "ufo landing, daguerreotype"
c) "ufo landing, daguerreotype, trending on /r/art"
d) "ufo landing, daguerreotype, by greg rutkowski, trending on /r/art"

Michigan Tech

Prompt engineering, continued

Try adding modifiers with commas. Ex:
- A husky eating an apple

Could become
- A husky eating an apple, Edward Gorey, pen and ink, trending on /r/art, moody

Types of modifiers people have used:
- Adjectives
- Artist names
- Art styles
    - Medium, camera descriptions, perspective, historical period, etc.
- Movies or video games
- Rendering engines
    - Unreal engine, octane, vray, blender, etc.
- Effects
    - Glowing, caustics, raytracing, lava, etc.
- "Trending on …" artstation, /r/art, etc.

Michigan Tech
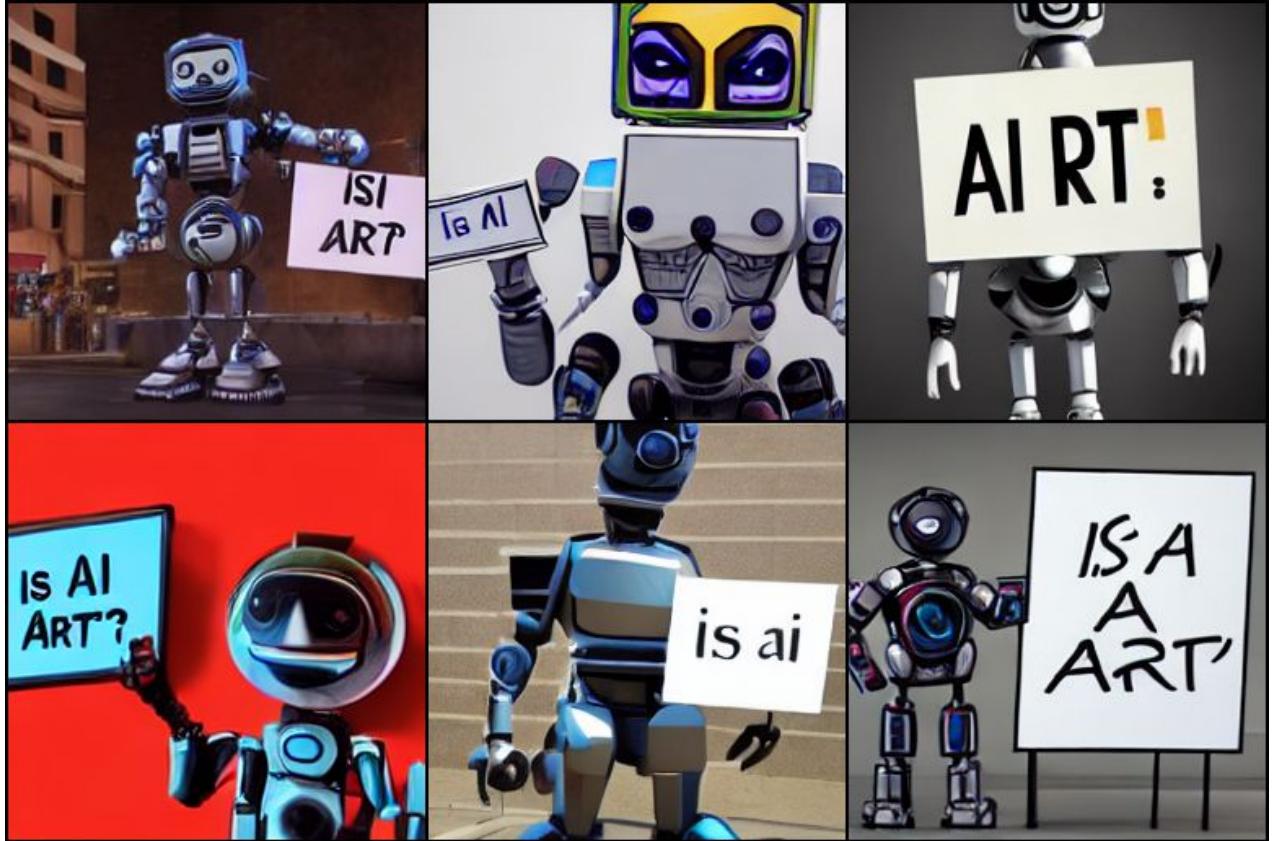
Latent Diffusion (2021)

First publicly available diffusion model guided by CLIP

Playing pianos filled with flames, watercolor

A robot holding a sign that says "is AI art art?"

Michigan Tech

Diffusion models

Teach models what an image looks like as it gets noisier and tells them to do that in reverse

"Denoising"

Generally preferred over GANs in modern image models

Michigan Tech

GLID-3

Based on GLIDE [1], CLIP, and Latent Diffusion

Focused on photo-realism

An oil painted dragon in the woods

A panda holding a beer

[1] https://arxiv.org/abs/2112.10741

Michigan Tech

Stable Diffusion (2022)

Based on Latent Diffusion

Most high-profile image model in the news - sued by artists and stock image companies for training on potentially copyrighted images
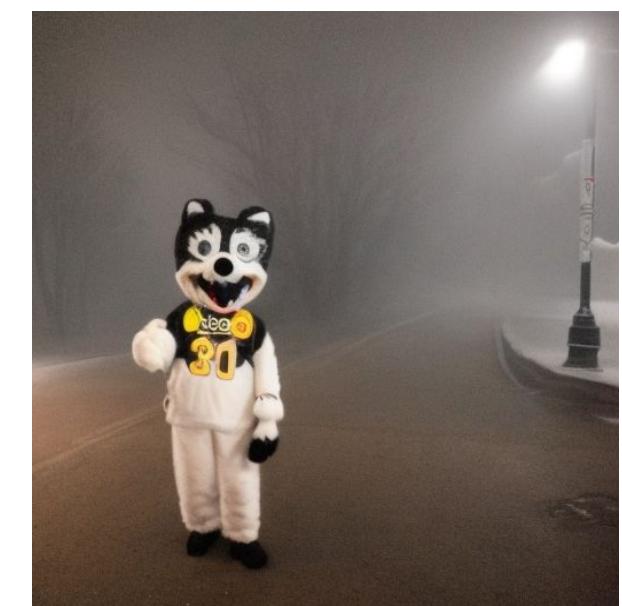
Michigan Tech

Dreambooth (2022)

Technique to focus an image model on a likeness of a person, animal, or thing

I used Blizzard T. Husky

Michigan Tech

Craiyon (formerly DALLE-Mini)

A husky that is part goose

Online platform - easy to use:
https://www.craiyon.com/

A goose that is part husky
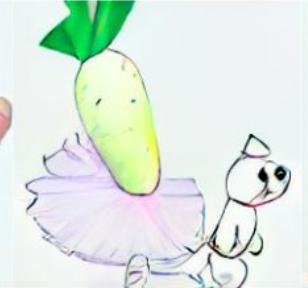
Michigan Tech

DALLE-Mega

Homebrew DALLE continues development with DALLE-Mega

Report on project here:
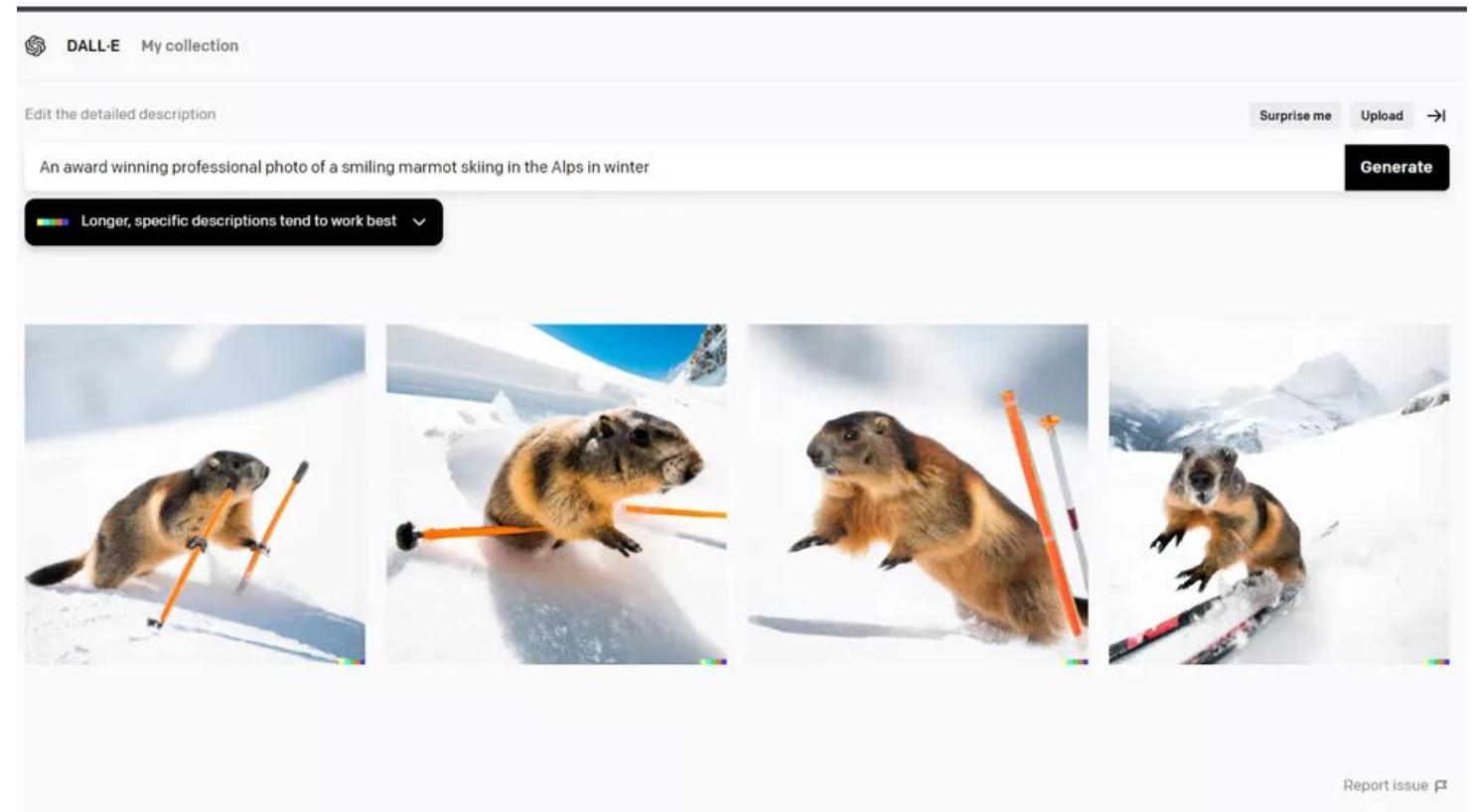
https://wandb.ai/dalle-mini/dalle-mini/reports/DALL-E-mini-Generate-images-from-any-text-prompt--VmlldzoyMDE4NDAy

Michigan Tech

DALL-E 2

Fairly high quality outputs

Only available through website/API

https://openai.com/product/dall-e-2

Michigan Tech

DALL-E 2

Current demo includes more than just image generation - images can be selectively recreated

https://openai.com/product/dall-e-2

Midjourney

Discord based demo,
art focused model

https://www.midjourney.com/

Michigan Tech

Is AI Art, Art?

https://www.instagram.com/averyseasonart/

Michigan Tech

Now it's your turn!

Links to tools available at:
https://tinyurl.com/sa2ftuh6

If you want to share an image at the end during "Show and Tell", email it to me with a short description (model, prompt, thoughts, etc.) at :

redacted
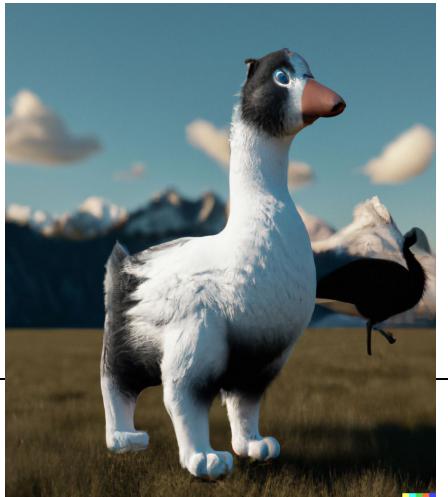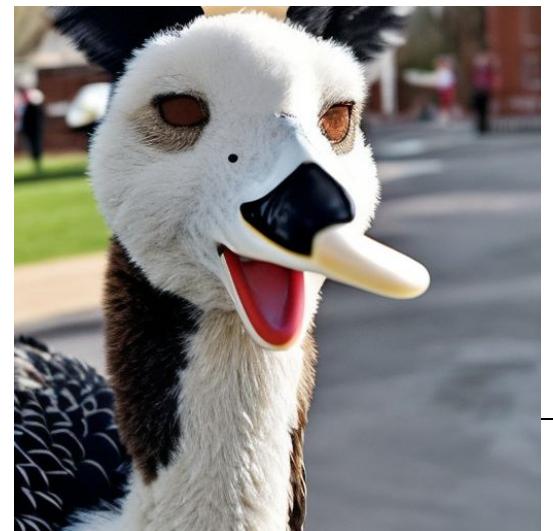
Michigan Tech

Show and tell time!

Some goose-husky creatures

Goosse Gusge

Michigan Tech

**orange and white cat playing a Gibson guitar while riding a dragon**
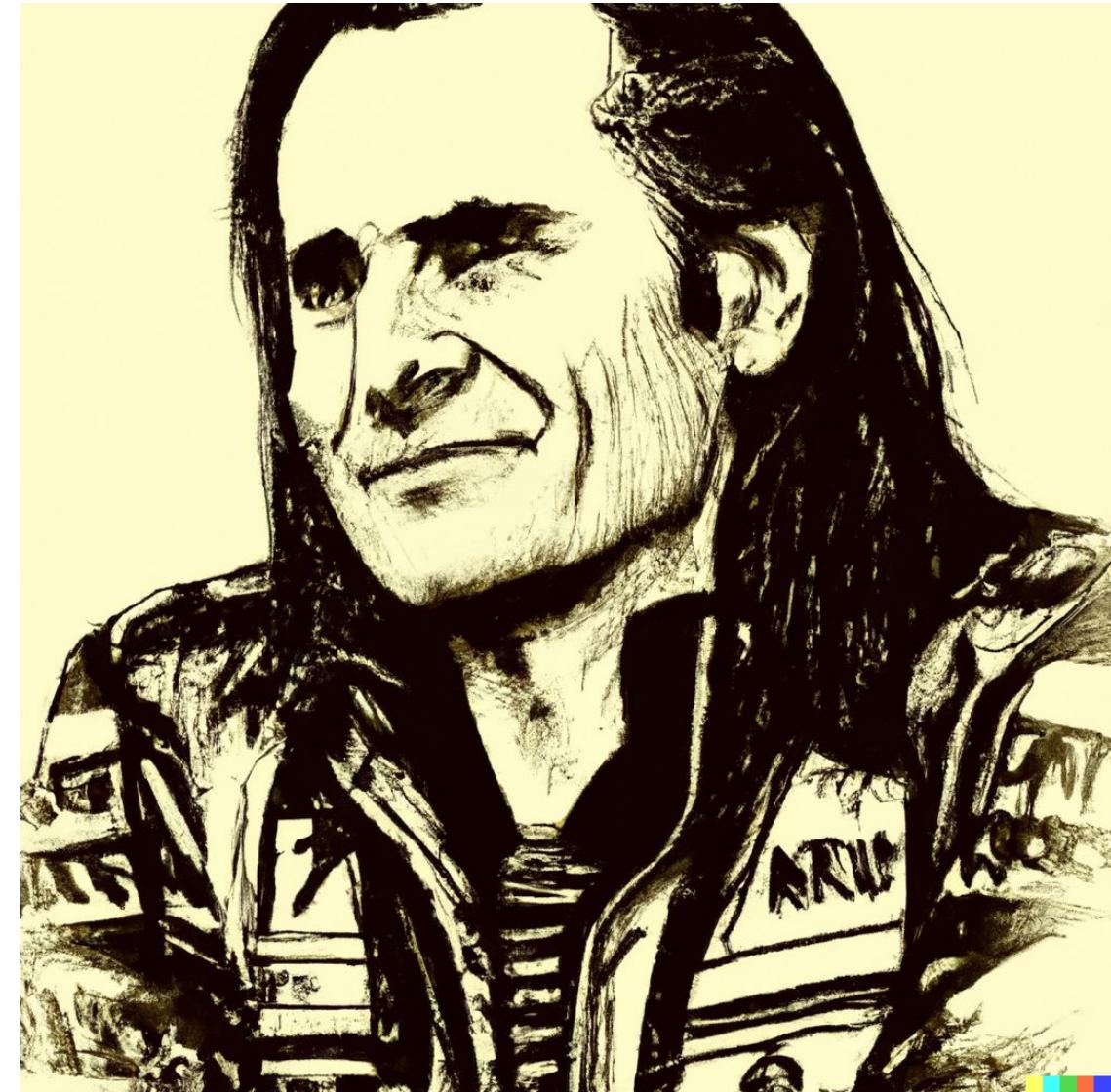
Michigan Tech

**subway station, underwater, forested, photorealistic, postapocalyptic**
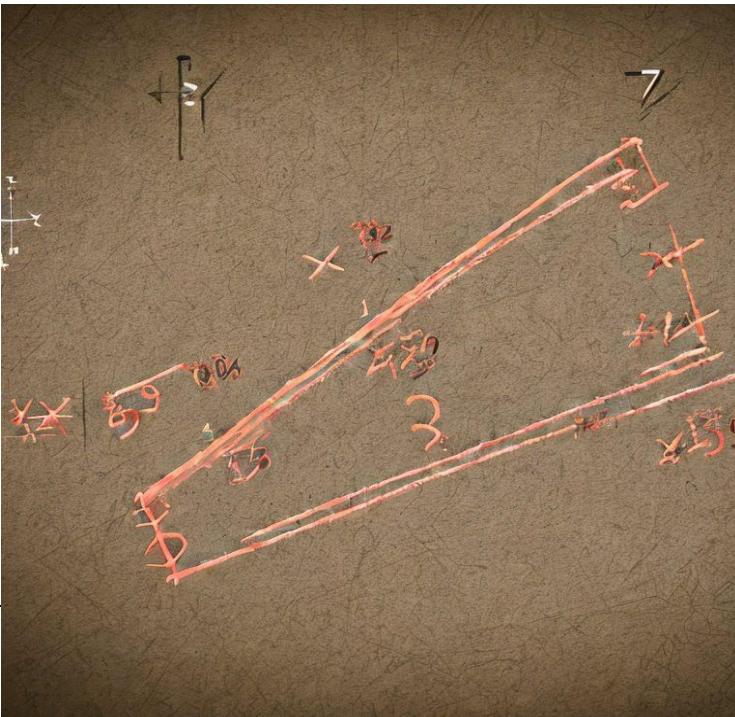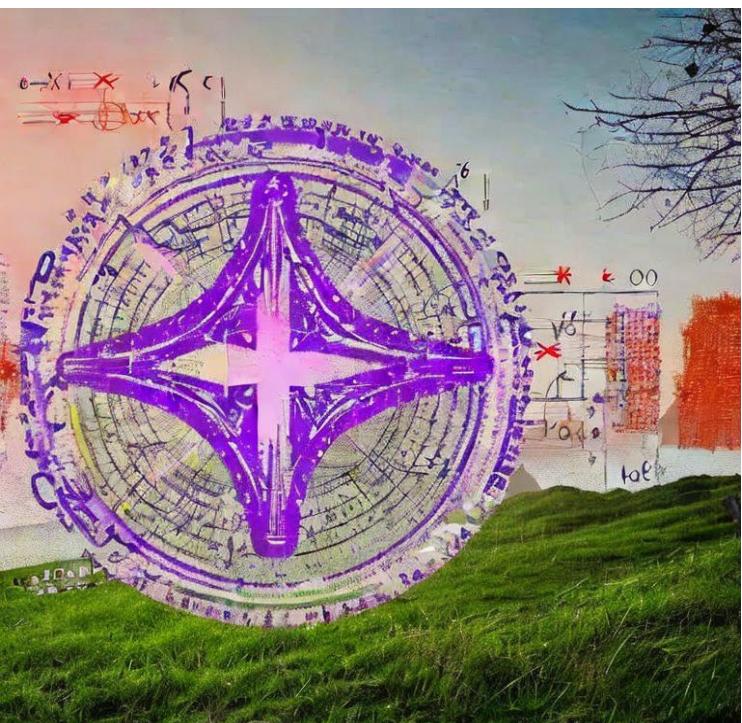
Michigan Tech

James May, Jeremy Clarkson, and Richard Hammond from Top Gear, futuristic neon, reflective puffy coat, decorated with traditional Japanese armaments, by ismail inceoglu dragan bibin hans thoma greg rutowski alexandros pyromallis nekro rene maritte illustrated, perfect face, fine details, realistic shaded, fine-face, pretty face

Mathematical formulas that don't exist, in the style of Ramanujan

Michigan Tech

AI Workshop - Bob Ross as a Knight - DALL-E mini
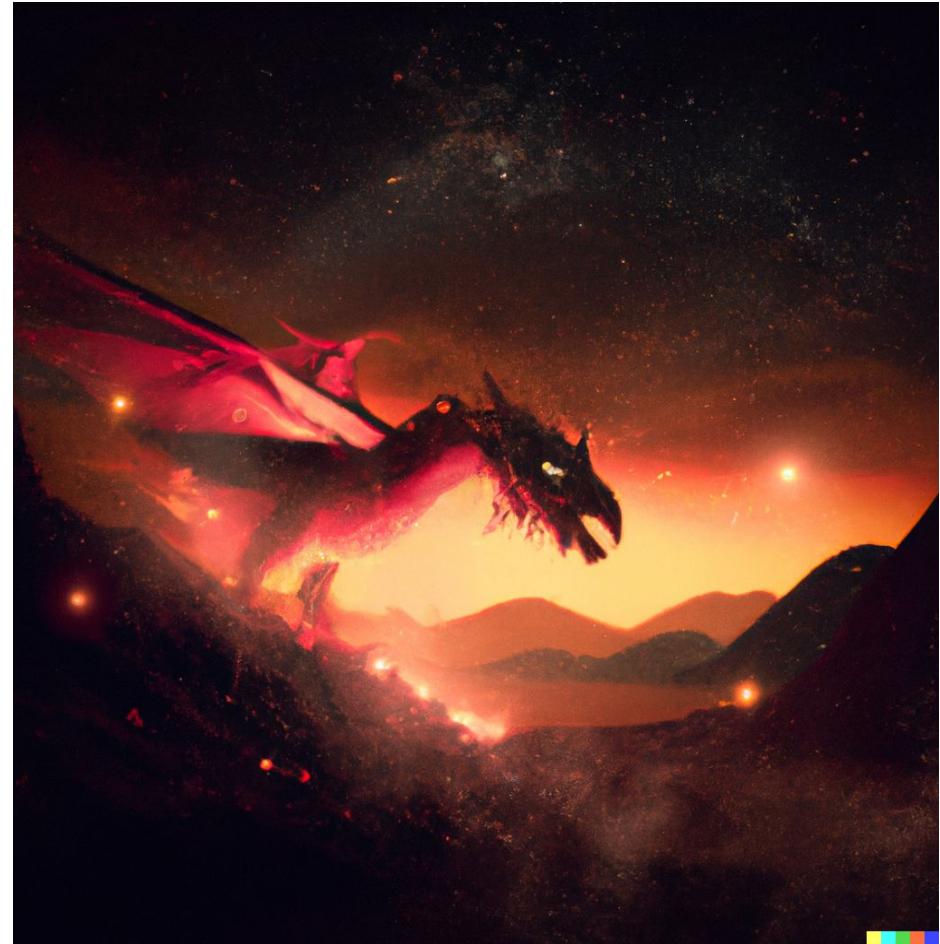
# A husky graduating from Michigan Tech
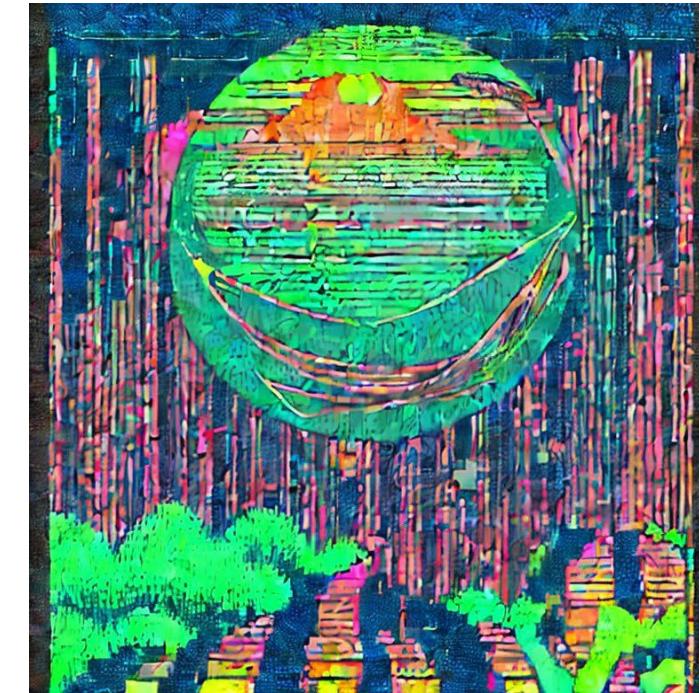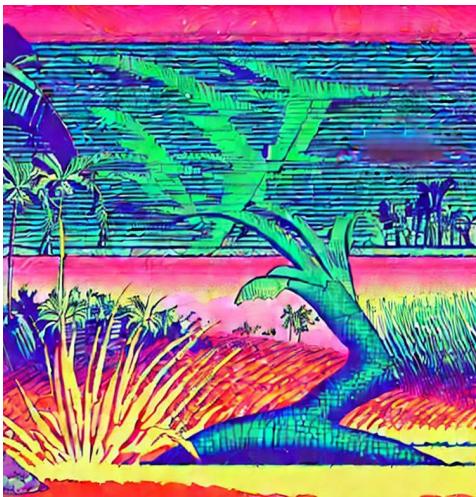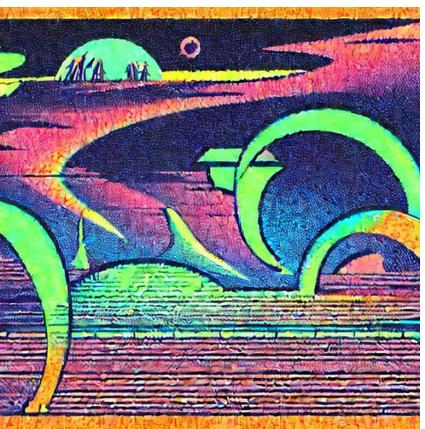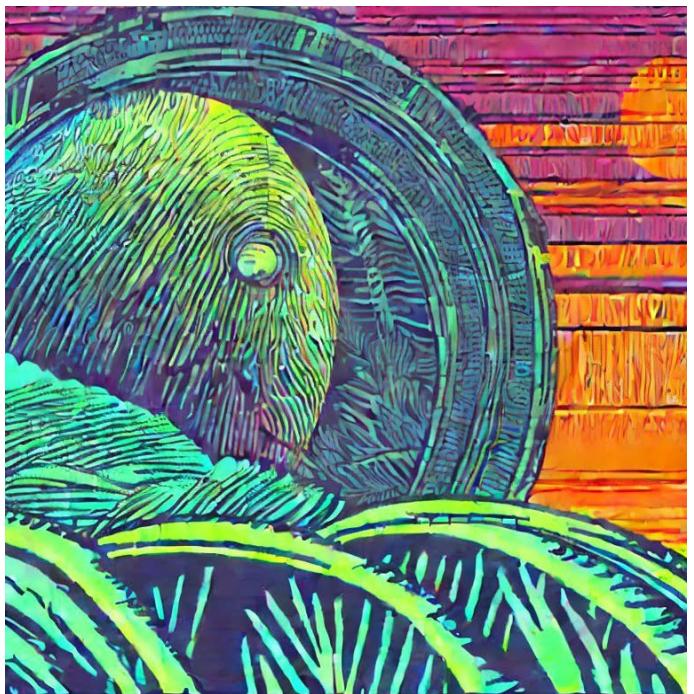
a powerful healing potion, and a mysterious black dragon. Both are supposed to be done in a D&D-like style.

Michigan Tech

Tropical alien planet, 1980s vector graphic style, VHS tape tracking error filter, one art deco feature in the foreground
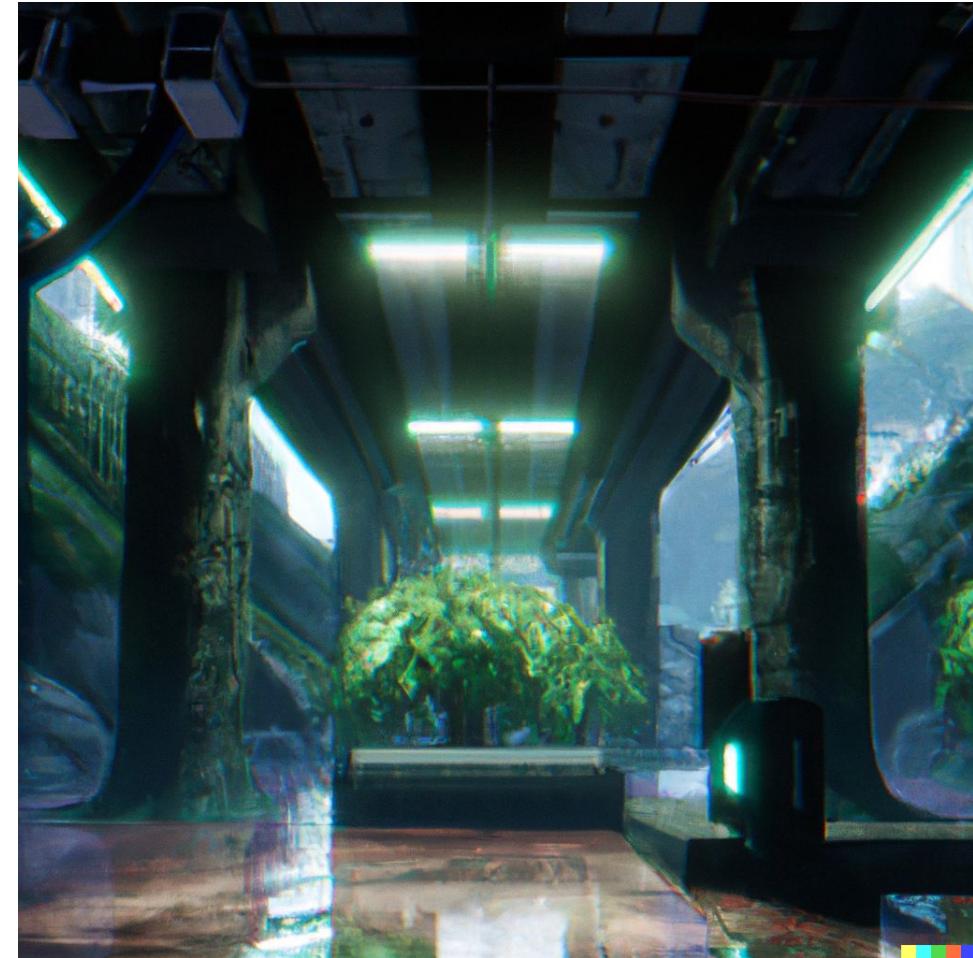

Michigan Tech

samurai with four hands walking through clouds, peaks of pagodas slightly obscured, all black armour with gold trim, katana in each of four hands, intense, dark, matte, painting Kentaro Miura, painting by Gustave Dare, in the style of The Great Wave off Kanagawa
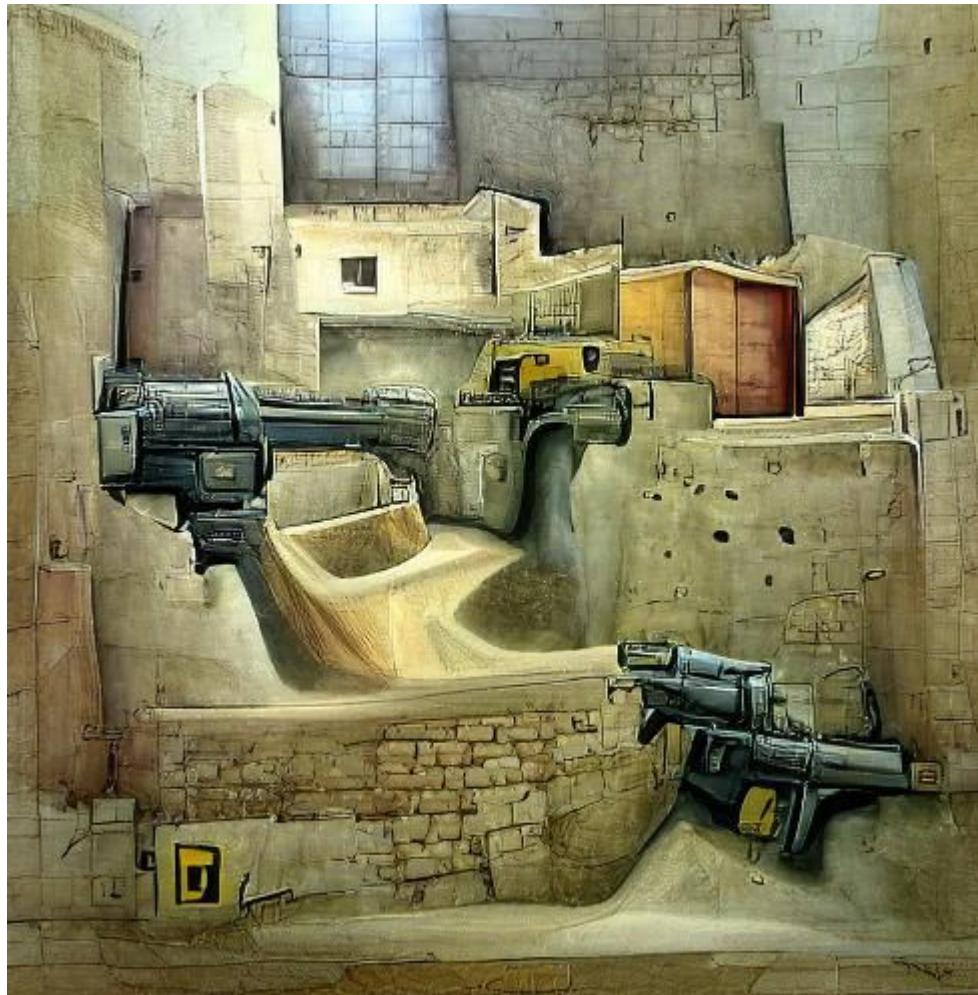
Michigan Tech

subway station, underwater, forested, photorealistic, postapocalyptic, murky

de_dust2 by James Gurney

Michigan Tech

"A space giraffe in a spacesuit riding a unicycle on the moon"

Live, Laugh, Love | horror
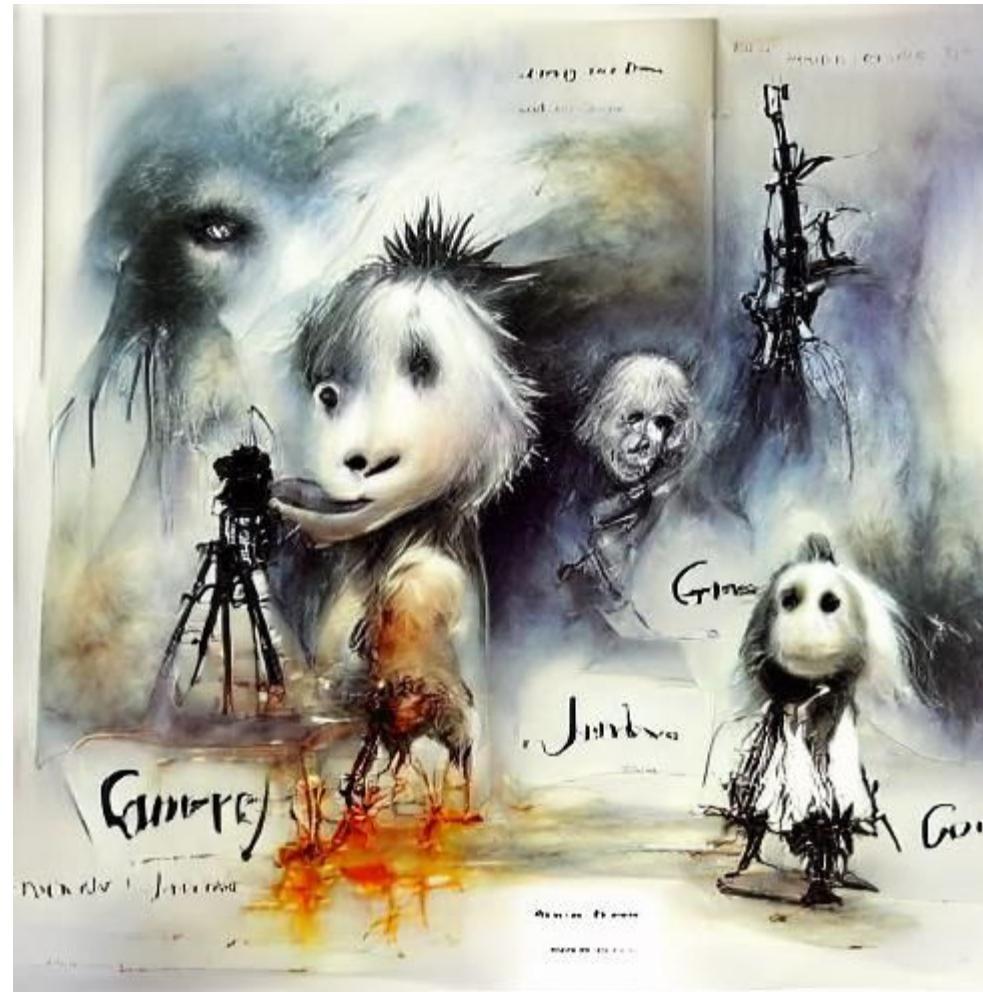
AI-generated art of the future

Stephen Gammell and James Gurney collaboration

It's risky to be with husky when its drinking whiskey

References and further reading

https://zentralwerkstatt.org/blog/ten-years-of-image-synthesis

Michigan Tech