Goodenough College

Artificial Intelligence and Machine Learning Society

Computer Science, Data Science, AI, Machine Learning...

experts, non-experts, beginners, anyone who has interest is welcome!!

GC AI & ML Society sharing session 10/2024

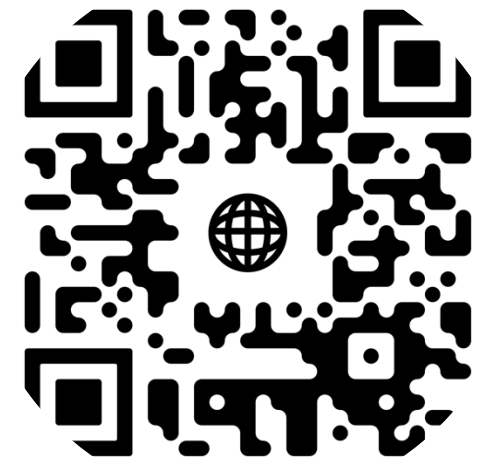# How to Use AI Technology to Create Picture

Presenter: Binglun Wang

Self-introduction

- GC member from Jan 2023 in WGH

PhD student in Computer Science at UCL;

Visiting Researcher at Adobe Research and Max Plank UCL Center

Problem Statement

- Need to design pictures every where. e.g. prepare slides our GC logo.

- Creating picture by AI is powerful and can save your time from 2022 due to diffusion model technology.

OpenAI DallE SoRA etc. are all based on diffusion model technology.

Problem Statement

- Existing tools are not flexible to users needs or need cost
e.g. Chatgpt & Adobe Photoshop

- Coding is flexible but require years-training skills and time cost
e.g. 1000 lines code for a mature text2img in diffuser


```
1040        image = self.vae.decode(latents / self.vae.config.scaling_factor, return_dict=False, generator=generator)[
1041            0
1042        ]
1043        image, has_nsfw_concept = self.run_safety_checker(image, device, prompt_embeds.dtype)
1044    else:
1045        image = latents
1046        has_nsfw_concept = None
1047
1048    if has_nsfw_concept is None:
1049        do_denormalize = [True] * image.shape[0]
1050    else:
1051        do_denormalize = [not has_nsfw for has_nsfw in has_nsfw_concept]
1052
1053    image = self.image_processor.postprocess(image, output_type=output_type, do_denormalize=do_denormalize)
1054
1055    # Offload all models
1056    self.maybe_free_model_hooks()
1057
1058    if not return_dict:
1059        return (image, has_nsfw_concept)
1060
1061    return StableDiffusionPipelineOutput(images=image, nsfw_content_detected=has_nsfw_concept)
```
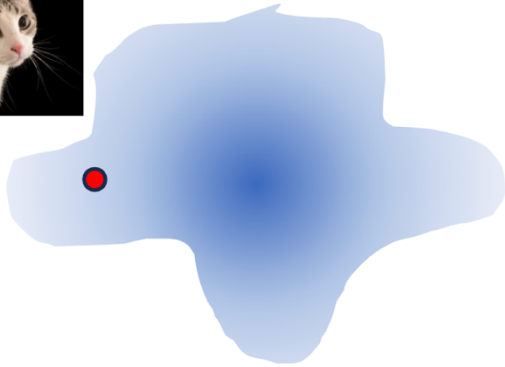
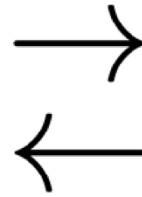- Trade off between flexible and professional skills

Problem Statement

- To day, we will introduce a totally free and a tradeoff middle point ComfyUI.

- Currently, a lot of artist used the ComfyUI

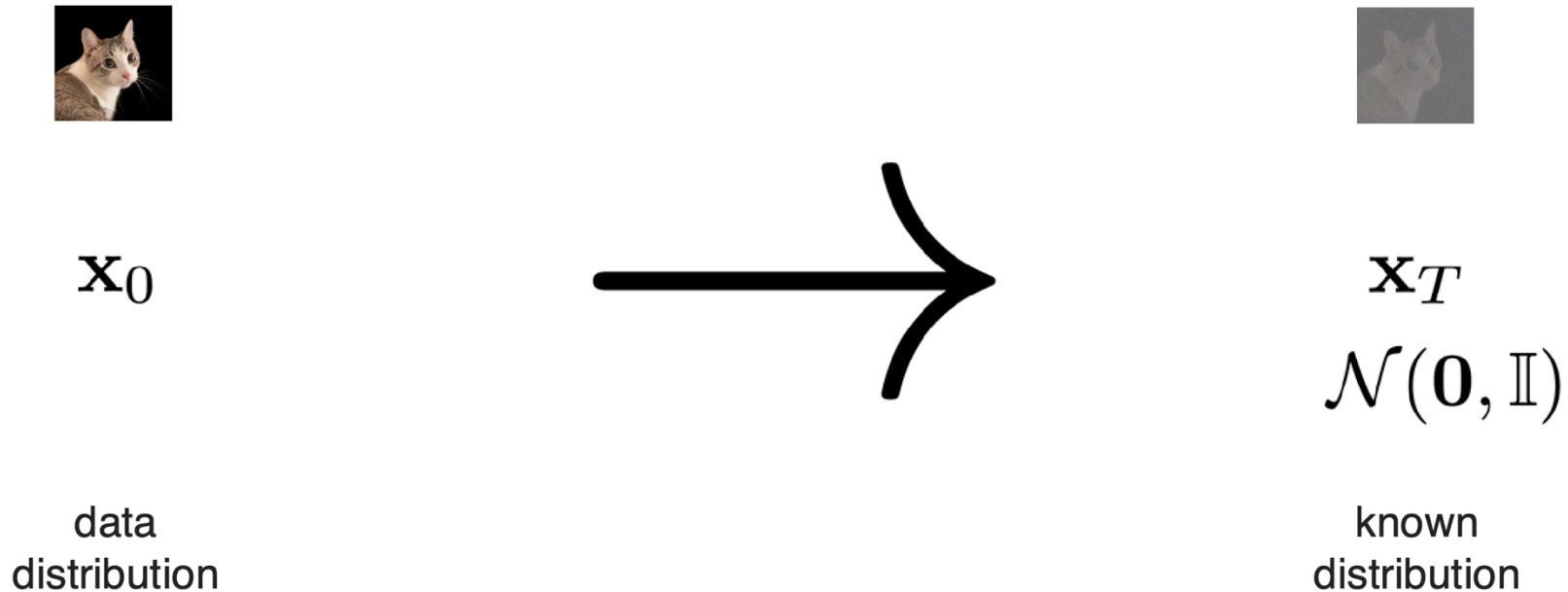# How it works

What is a Diffusion Process?

unknown map

(unknown) data distribution
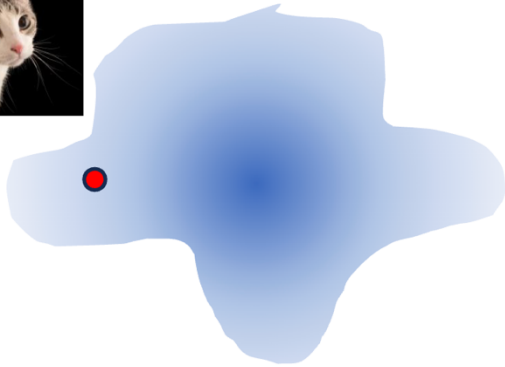
known distribution

Sampling ⟺ (Unconditional generation)

How it works

Mapping between Distributions

$$\mathbf{x}_0$$

data distribution

$$\mathbf{x}_T$$
$$\mathcal{N}(\mathbf{0}, \mathbb{I})$$
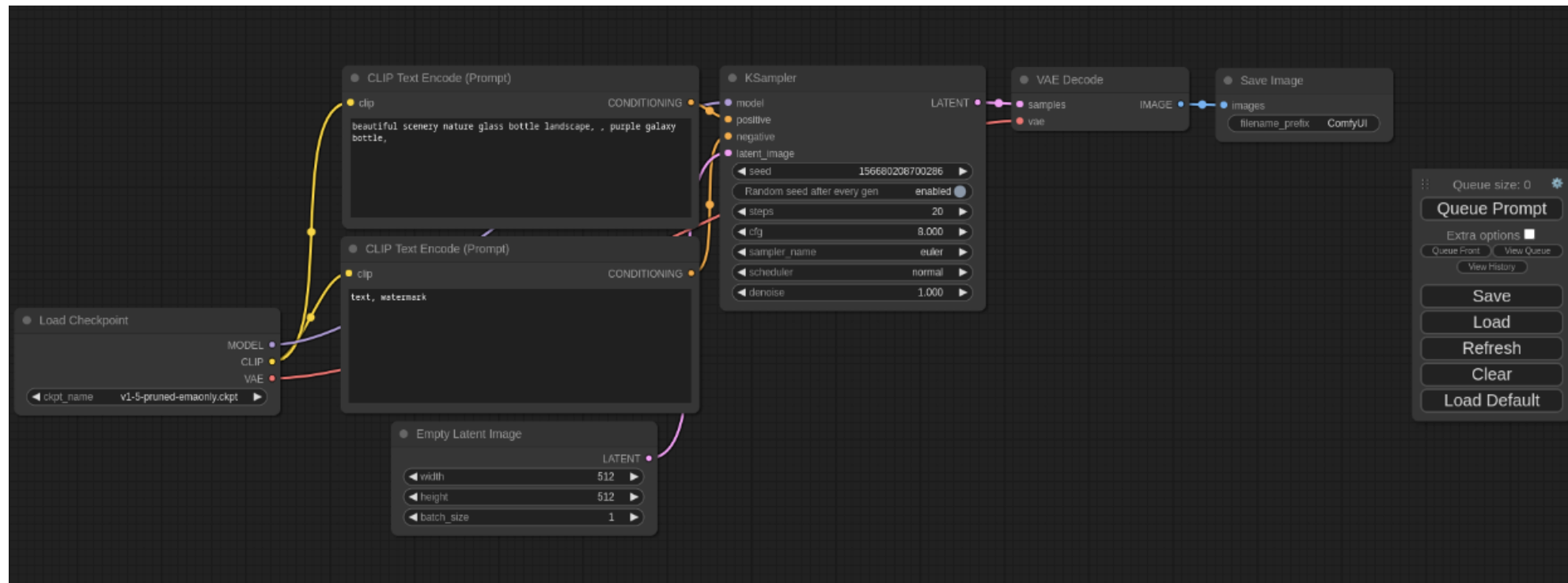
known distribution

How it works

How it works

And the most model is currently used latent diffusion
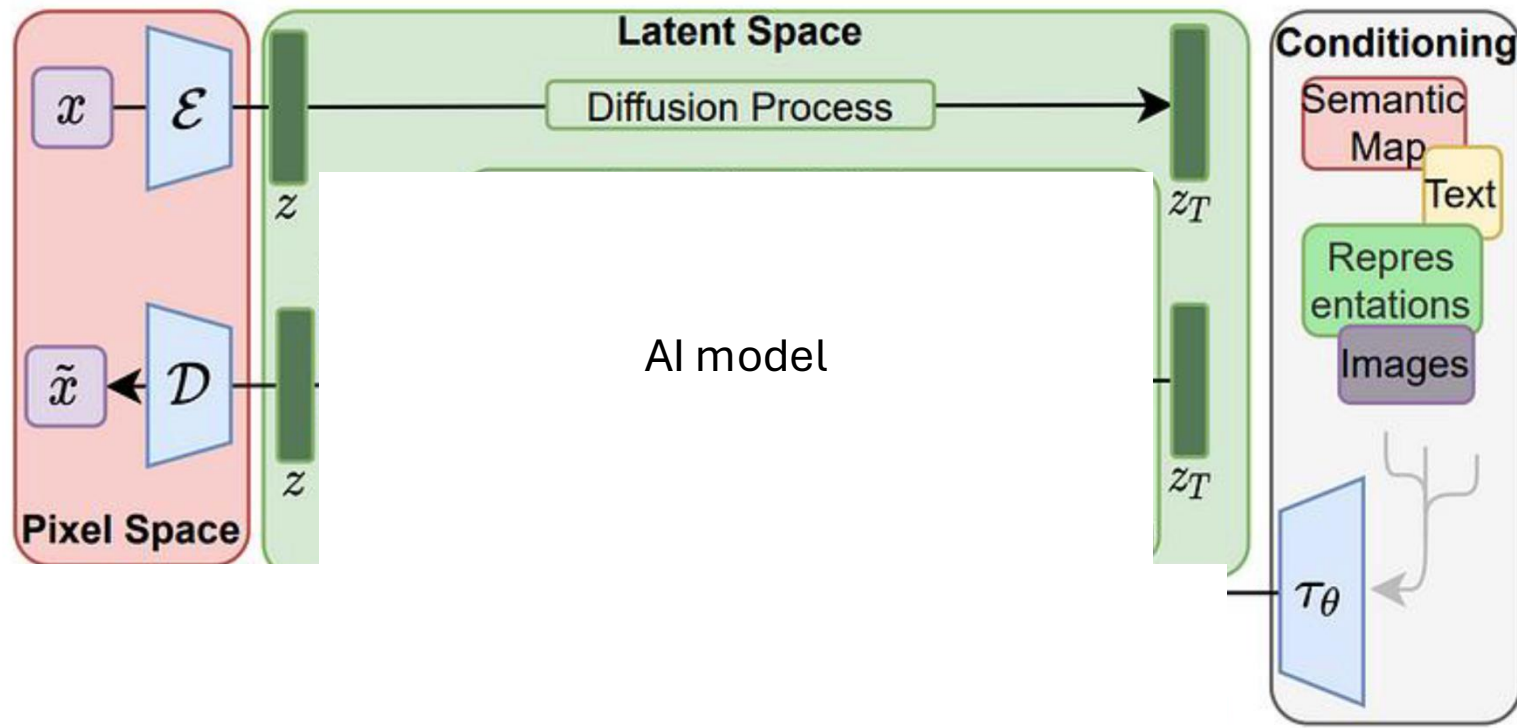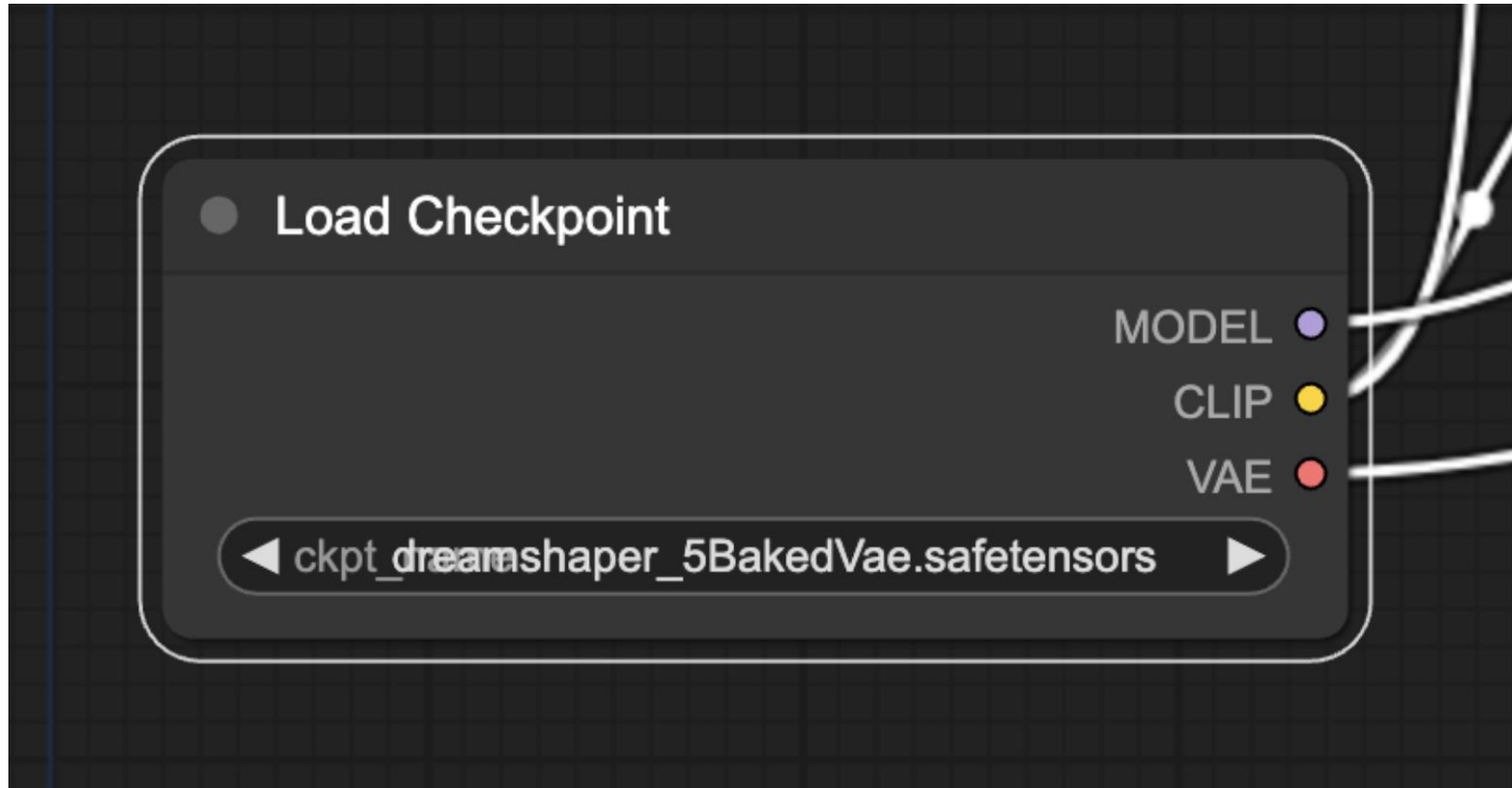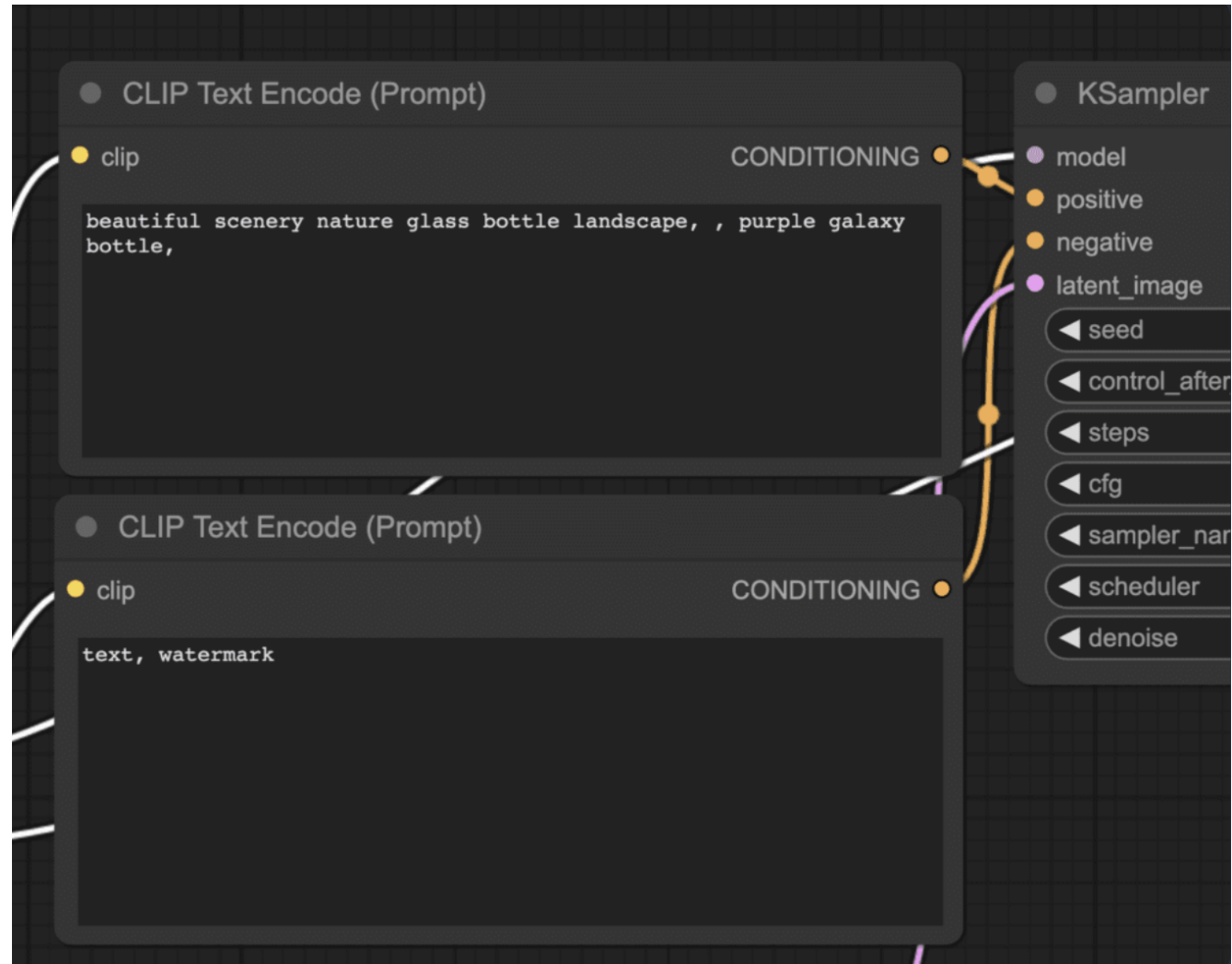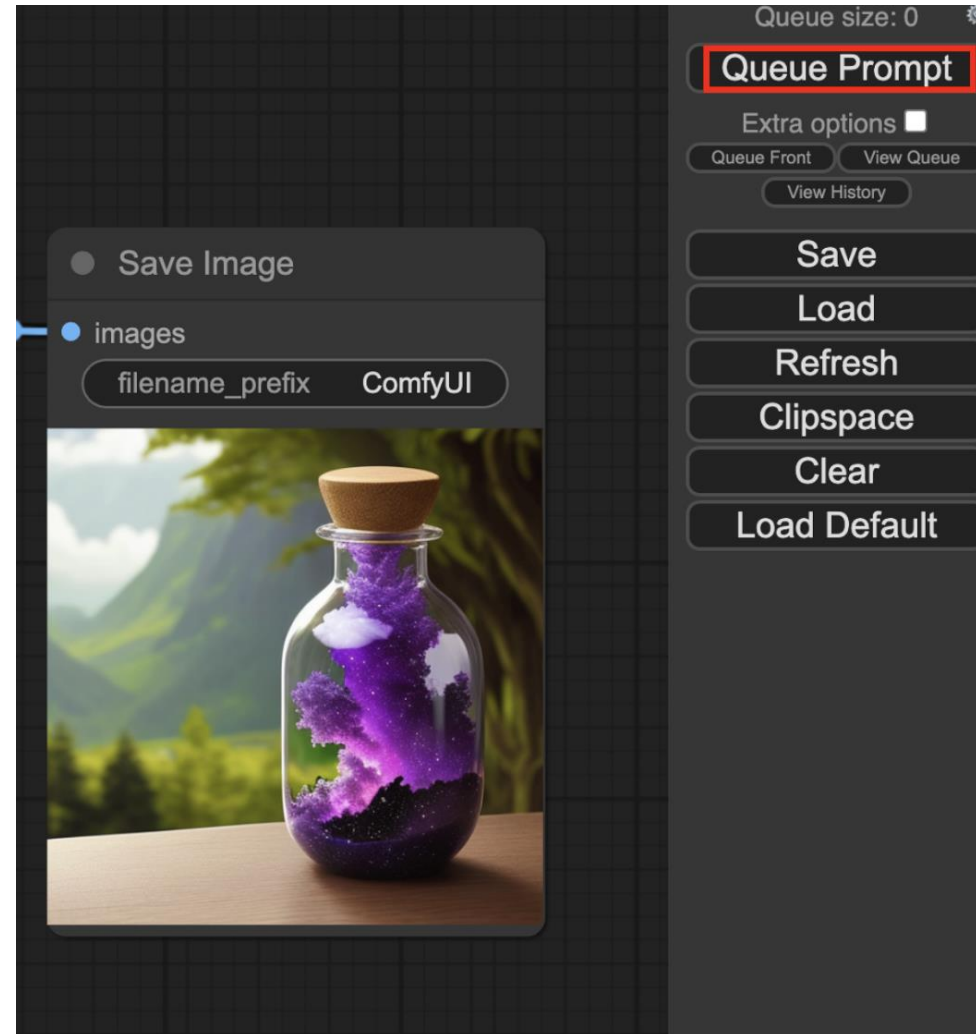
ComfyUI

- Consistent of a few blocks

1. Text 2 image Workflow
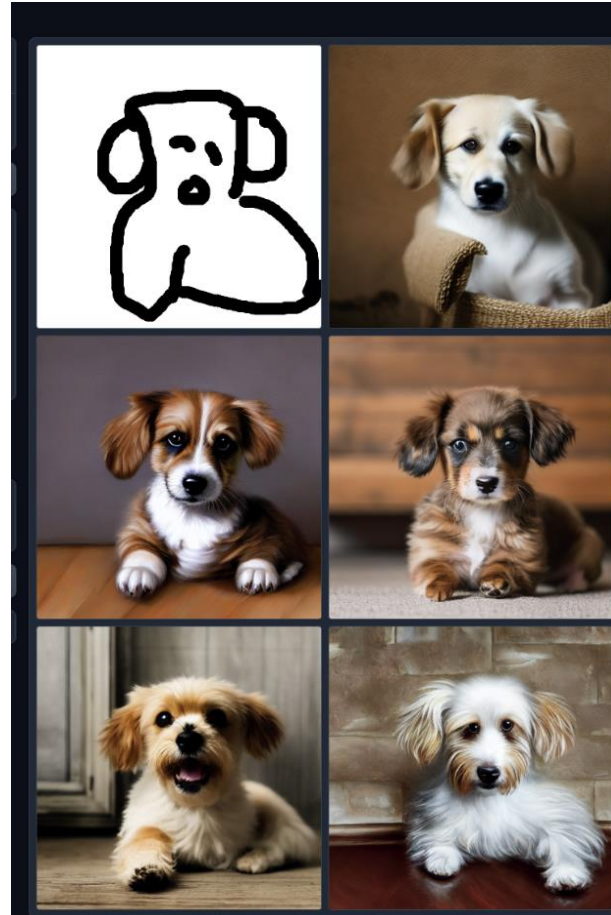
1. Text 2 image Workflow

1. Text 2 image Workflow

2. Img2Img – ControlNet
[Zhang et al., ICCV2023]

2. Img2Img – IPadapter
[Ye et al., arXiv2023]

Anime Model   Realistic Model   Structural Controls

Image Prompt

Variation

+ *wearing sunglasses*

Text Prompt

+

Inpainting   Realistic Model Anime Model
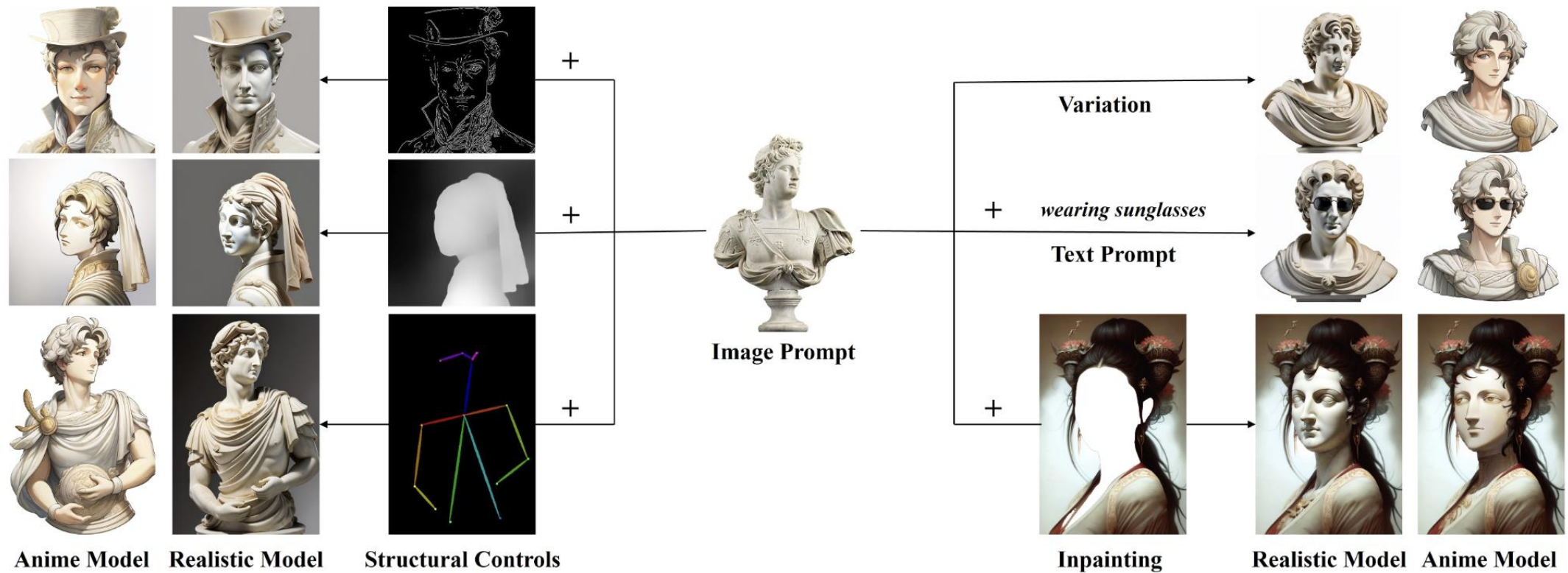
3. Others

- Different AI model: https://civitai.com/models

- Text2videos/Image2videos:

W/ AnimateDiff

https://animatediff.github.io/