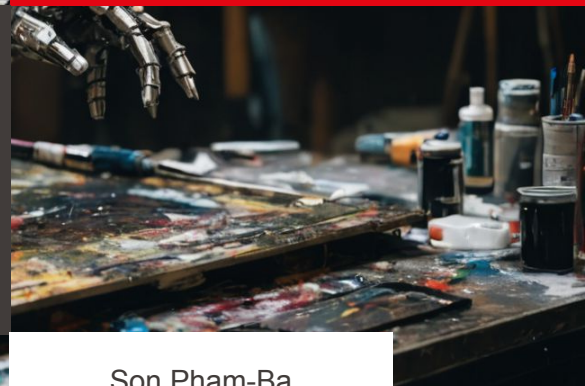




Local “AI” generated images

RSE Lightning talk



Son Pham-Ba

Late history of image generation services

Transformers + CLIP + Diffusion model

- Made by  **OpenAI**
- Accessible from ChatGPT, Bing image creator

DALL-E

Jan 2021

DALL-E 2

Apr 2022

DALL-E 3

Sep 2023

Midjourney V1

Feb 2022

V4

Nov 2022

V5

Mar 2023


V6

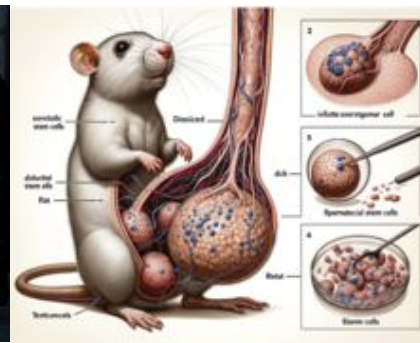
Dec 2023

V6.1

Jul 2024

■ Local "AI" generated images

- Made by Midjourney Inc. 
- Accessible from Discord, now on web





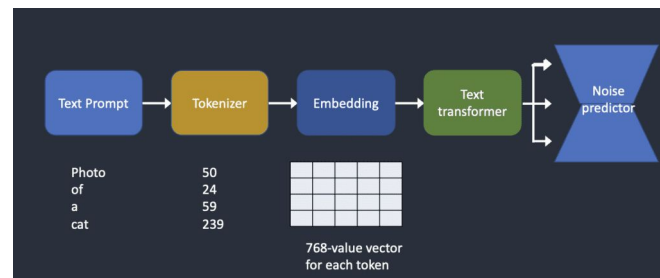
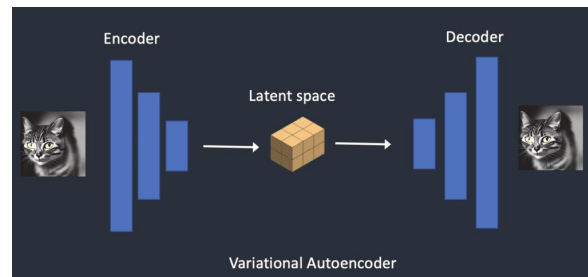
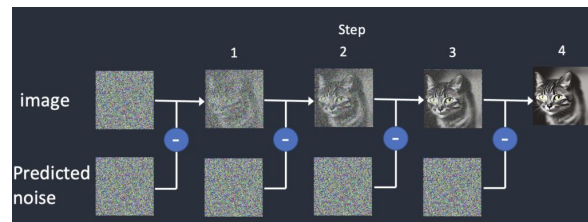
-And then along came

Stable Diffusion

Stable Diffusion principles

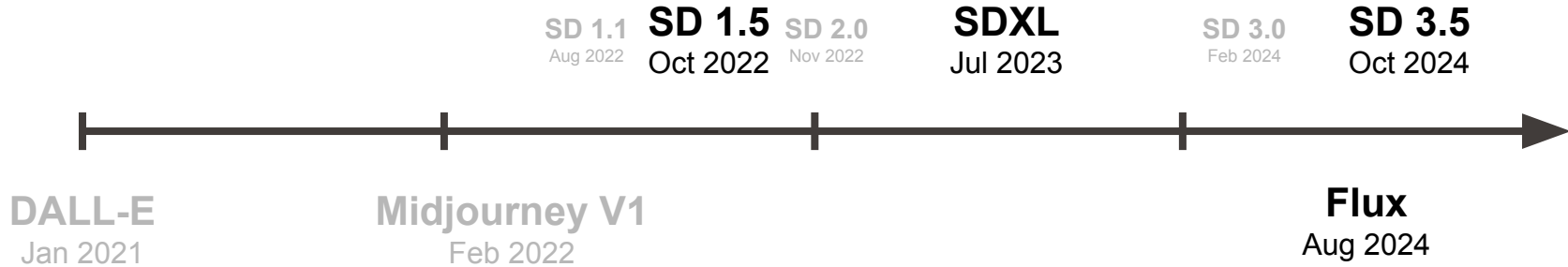
Latent diffusion model

- Start from noise, denoise iteratively using noise predictor
→ **U-Net**
- Works in low dimensional latent space
→ **VAE**
- Text conditioning
→ **CLIP**
- Openly available weights and architecture



Timeline of open models

- Made by Stability AI
- Downloadable weights and free generation on Hugging Face, Civitai



- Custom trained models:
 - Finetuned checkpoints (U-Nets + VAE)
 - LoRAs (lighter finetuning, small files, modular)
 - Embeddings

Workflows and requirements

Typical workflows:

- text-to-image
- image-to-image (style transfer, upscaling, inpainting, outpainting, ...)
- text-to-video, image-to-video, ...

Requirements:

- better Nvidia GPUs support
- need **VRAM** (quantization possible)

Model	Checkpoint size (fp16)
SD 1.5	2 GB
SDXL	6.5 GB
Flux S / Flux D	16 GB (fp8)
SD 3.5 Medium	4.8 GB
SD 3.5 Large	15.3 GB

How to run locally — web UI

<https://github.com/AUTOMATIC1111/stable-diffusion-webui>

model

Stable Diffusion checkpoint
protogenX34OfficialR_1.ckpt [60fe2f34]

prompts

green sapling rowing out of ground, mud, dirt, grass, high quality, photorealistic, sharp focus, depth of field

Negative prompt (press Ctrl+Enter or Alt+Enter to generate)

Generate

Style 1

Style 2

None

None

26/75

sampler, steps

Sampling method Euler a Sampling steps 20

Restore faces Tiling Hires. fix

Width 512

Batch count 4

Height 512

Batch size 1

CFG Scale 12

Seed 1441787169

Script None

resolution

guidance

seed



Save Zip Send to img2img Send to inpaint Send to extras

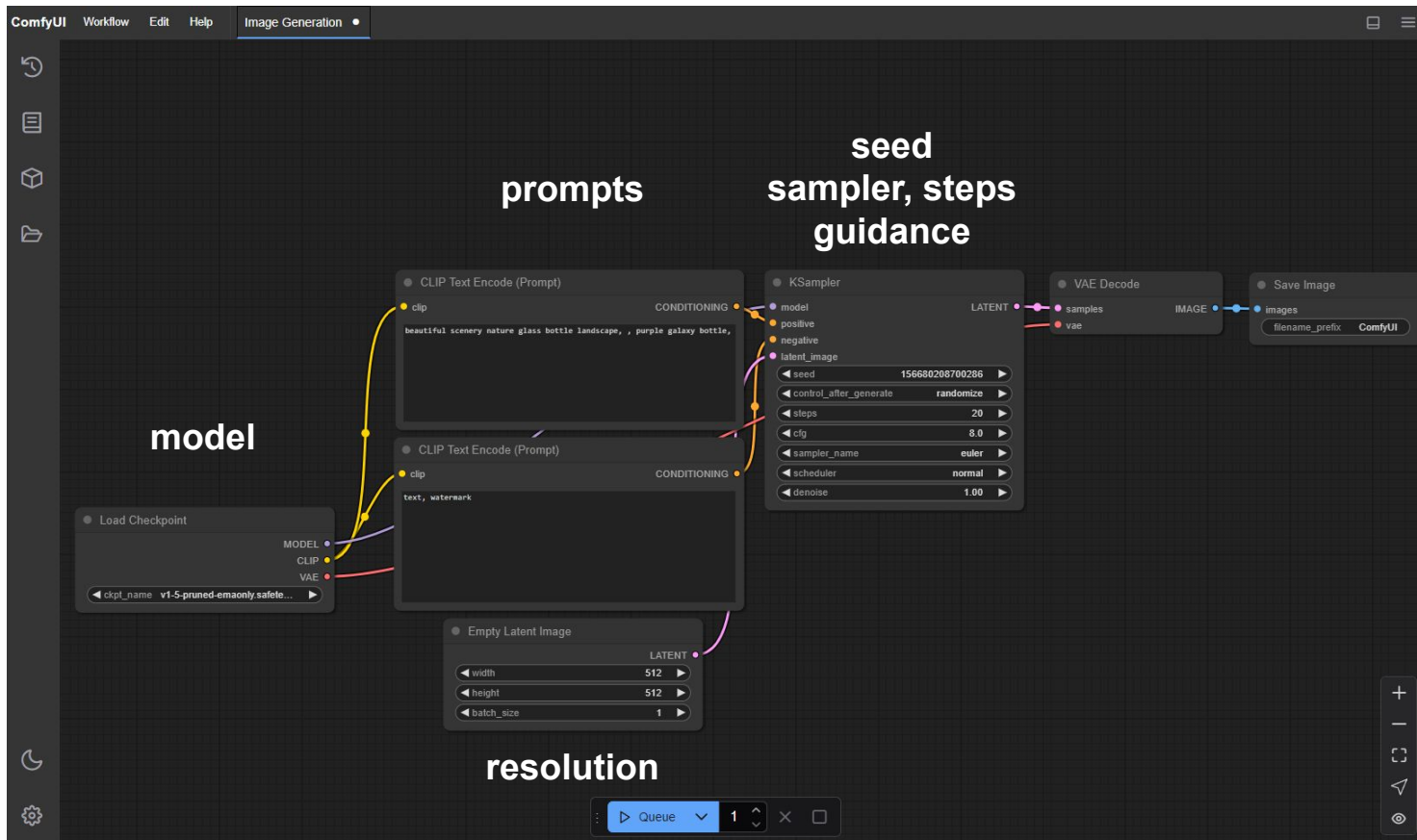
green sapling rowing out of ground, mud, dirt, grass, high quality, photorealistic, sharp focus, depth of field

Steps: 20, Sampler: Euler a, CFG scale: 12, Seed: 1441787169, Size: 512x512, Model hash: 60fe2f34, Model: protogenX34OfficialR_1

Time taken: 8.62s Torch active/reserved: 3699/4702 MiB, Sys VRAM: 7020/24576 MiB (28.56%)

How to run locally — ComfyUI

<https://github.com/comfyanonymous/ComfyUI>



Prompting

- **Positive prompt:**

subject, style, composition, details
(can have weight:1.1), (on some parts:0.8)

- **Negative prompt:**

what should not appear, non-desired style

- **CFG scale** (Classifier-Free Guidance)

- controls how closely the prompt is followed
- at CFG = 1: negative prompt has no effect

SD 1.5 (without and with hires. fix)

close up photograph, donald trump, white lab coat, giving a lecture, blackboard with math many formulae, (messy hair:1.2)

Negative prompt: (cartoon, 3d, cgi:1.3)

Steps: 20, Sampler: DPM++ 2M, Schedule type: Karras, CFG scale: 4, Seed: 2529697523, Size: 512x768, Model hash: 43d0be1815, Model: cyberrealistic_v60, Denoising strength: 0.6, Hires CFG Scale: 7, Hires upscale: 1.65, Hires steps: 10, Hires upscaler: Latent, Version: f2.0.1v1.10.1-previous-557-ge9d1736b

SDXL

close up photograph, donald trump, white lab coat, giving a lecture, blackboard with math many formulae, messy hair

Steps: 8, Sampler: DPM++ 2M SDE, Schedule type: Exponential, CFG scale: 1.5, Seed: 405868469, Size: 832x1216, Model hash: 4496b36d48, Model: dreamshaperXL_v21TurboDPMSDE, Version: f2.0.1v1.10.1-previous-557-ge9d1736b