

Getting started generating images with Stable Diffusion

Alex J. Bowers, Ph.D.

Professor of Education Leadership
Teachers College, Columbia University
Faculty Affiliate, Columbia University Data Science Institute
Faculty advisor to the TC student Data Science in Education Assoc.
New York, New York, USA

<http://www.tc.columbia.edu/faculty/ab3764/>

Email: Bowers@tc.edu

Twitter: [@Alex_J_Bowers](https://twitter.com/Alex_J_Bowers)

ORCID ID: 0000-0002-5140-6428

October 13, 2022



DSEA
Data Science and Education Association



TEACHERS COLLEGE
COLUMBIA UNIVERSITY
Graduate Student Life & Development

DSEA DATA WORKSHOP 01

AI Image Generation with Stable Diffusion

Speaker: Prof. Alex Bowers

Interested in producing digital images based on given text descriptions? Please join us and learn more about text-to-image models!



Date: Wednesday, Oct 12
Time: 3:00 - 4:45 PM (ET)
Location: GDH 277
Registration:
<https://bit.ly/stablediffusion1012>

Prompt: a photorealistic image of an anthropomorphic Jedi frog in white robes wielding a lightsaber, on a lily pad in a swamp, 8k, highly detailed, dynamic lighting, HD, bokeh, unreal engine

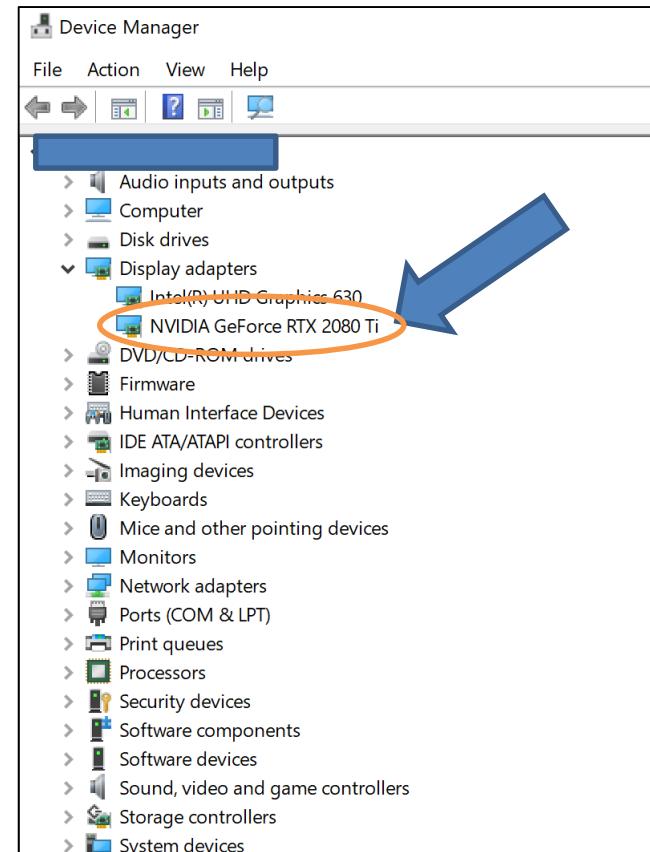
To request disability-related accommodations, contact OASID at oasid@tc.edu, or 212-678-3689, (646) 755-3144 video phone, as early as possible.

If you have a Windows laptop with an Nvidia graphics card

- Checkpoint and GFPgan files

- All others go to this website:

<https://huggingface.co/spaces/stabilityai/stable-diffusion>



Today's workshop session

- Overview of Stable Diffusion images and prompts.
- What are diffusion models?
 - DALLE2, Midjourney, Disco Diffusion, Stable Diffusion
- Stable Diffusion (public, open, and free!)
 - Github code in Python running in Gradio in the browser
 - Or just use the website
- Prompt engineering
 - Camera settings, graphic design, artists, art history movements
- Options in Automatic1111's Python Gradio implementation
 - Prompt matrix
 - Hyperparameter “Guidance Scale” settings, samplers, and more!
 - Image 2 Image
 - Negative prompts
 - Inpaint and outpaint
- Your turn! Making prompts and images in Stable Diffusion
- Discussion:
 - What is this?
 - You believe you made that don't you?
 - What does Machine Learning Image Generation mean for Art, Graphic Design, and Education?



Prompt:
a photorealistic image of an anthropomorphic Jedi frog in white robes wielding a lightsaber, on a lily pad in a swamp, 8k, highly detailed, dynamic lighting, HD, bokeh, unreal engine



Prompt:
a portrait of a cat with white
fur who is a knight dressed in
intricately carved shining
metal armor, highly detailed,
8k, HD, dynamic lighting



Prompt:
a photorealistic nice cute
cozy anthropomorphic
hedgehog in a tweed vest
with a cute bowtie, sitting in
a cozy chair, reading a
book, surrounded by
bookshelves, with a
Victorian lamp and a roaring
fireplace in the background,
with a glowing Victorian
stained glass window in the
background, bokeh, highly
detailed, 8k, HD



A velociraptor graduating from
Columbia University in blue
regalia

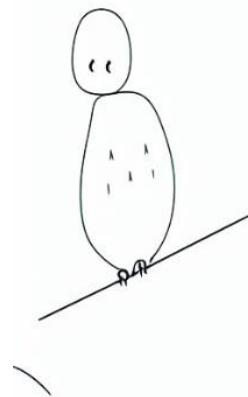
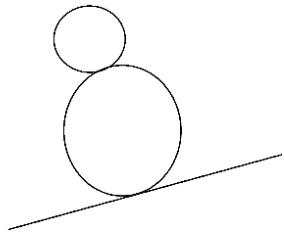
https://twitter.com/Alex_J_Bowers/status/1569684954740572162

Prompt: a stunning fantasy scene, with a magic owl reading a sparkling book in the foreground, highly detailed, with lens flare



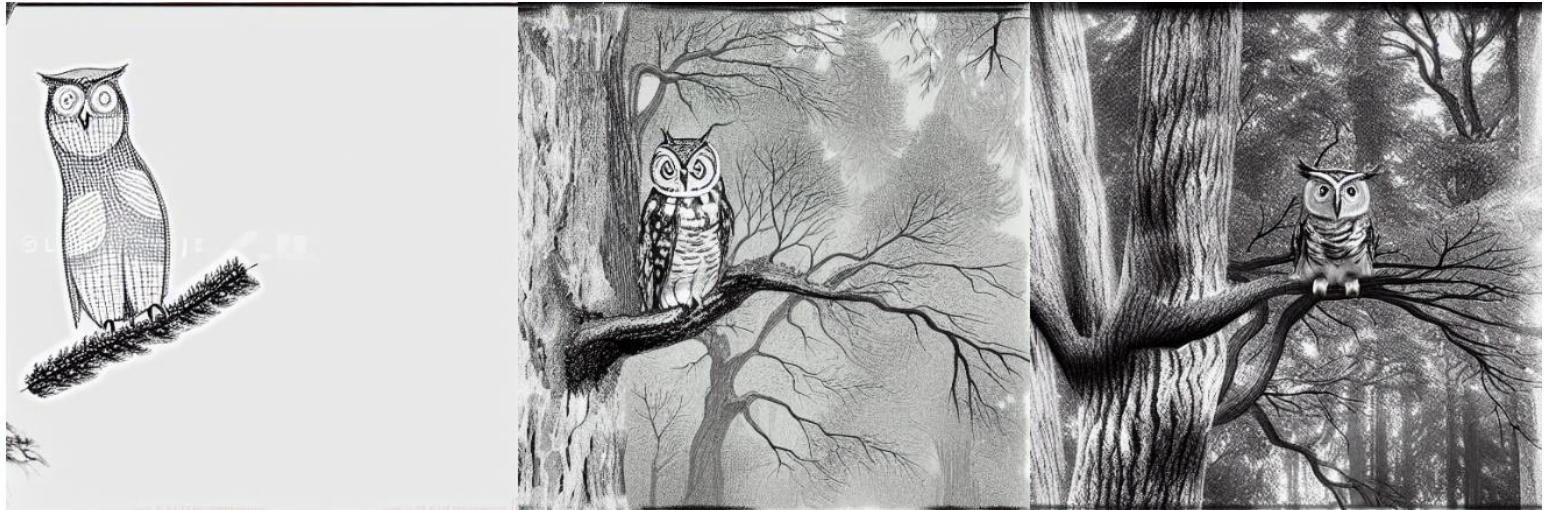
Prompt: a stunning fantasy landscape, with a castle far in the distance, highly detailed, with lens flare





Loopback on a drawing of 2 circles and a line:

Prompt: a photo of an owl sitting on a tree branch in a forest, highly detailed, 8k, unreal engine, bokeh



Prompt Matrix

Prompt: A photo of a stunning excellent executive office suite with many intricate decorations and a mirror and a fabulous view through large windows, dynamic lighting, highly detailed, 4k, HD, promotional media, trending on Pixiv, |in an ultra-modern style |in a Louis the XIV style |in a Cowboy West Texas style

~~in a Cowboy West
Texas style~~

in a Cowboy West
Texas style

~~in an ultra modern style
in a Louis the XIV style~~



~~in an ultra-modern style
in a Louis the XIV style~~



~~in an ultra-modern style
in a Louis the XIV style~~



Prompt Matrix

a 5 story castle estate in the country,
[as a 3D pop-up book page |as a photo
in the Scottish countryside] in the style
of cutaway French architectural
draftsmanship (1770)

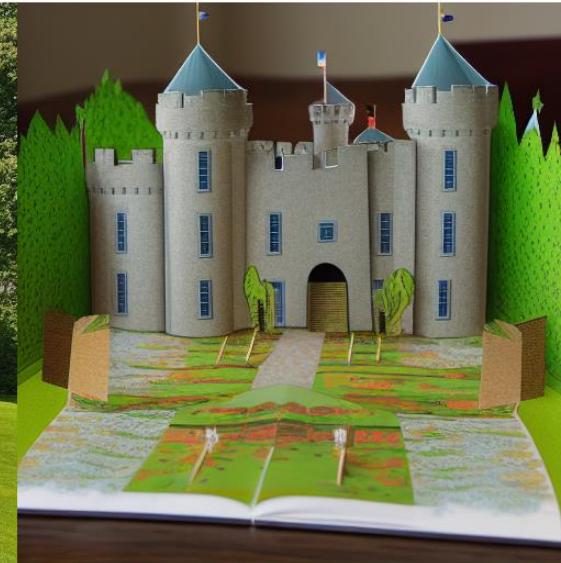
~~in the style of
cutaway French
architectural
draftsmanship
(1770)~~

~~in the style of
cutaway French
architectural
draftsmanship
(1770)~~

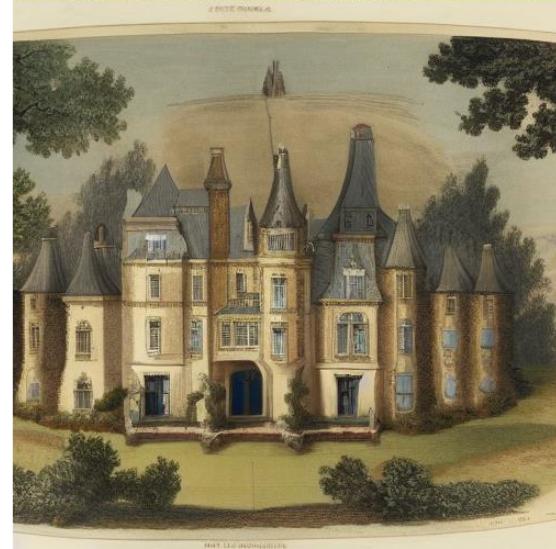
~~as a 3D pop-up book page
as a photo in the Scottish
countryside~~



~~as a 3D pop-up book page
as a photo in the Scottish
countryside~~



~~as a 3D pop-up book page
as a photo in the Scottish
countryside~~



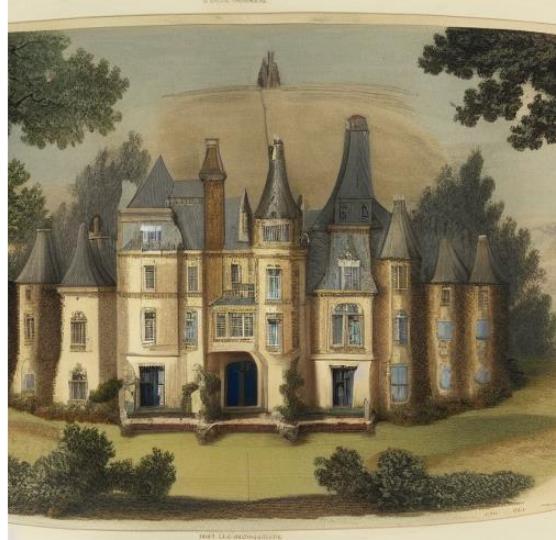
Prompt Matrix

a 5 story castle estate in the country,
[as a 3D pop-up book page |as a photo
in the Scottish countryside] in the style
of cutaway French architectural
draftsmanship (1770)



architectural
draftsmanship
(1770)

as a 3D pop-up book page
as a photo in the Scottish
countryside



as a 3D pop-up book page
as a photo in the Scottish
countryside



as a 3D pop-up book page
as a photo in the Scottish
countryside



Prompt Matrix

a 5 story castle estate in the country,
[as a 3D pop-up book page |as a photo
in the Scottish countryside] in the style
of cutaway French architectural
draftsmanship (1770)



as a 3D pop-up book page
as a photo in the Scottish
countryside



as a 3D pop-up book page
as a photo in the Scottish
countryside



as a 3D pop-up book page
as a photo in the Scottish
countryside

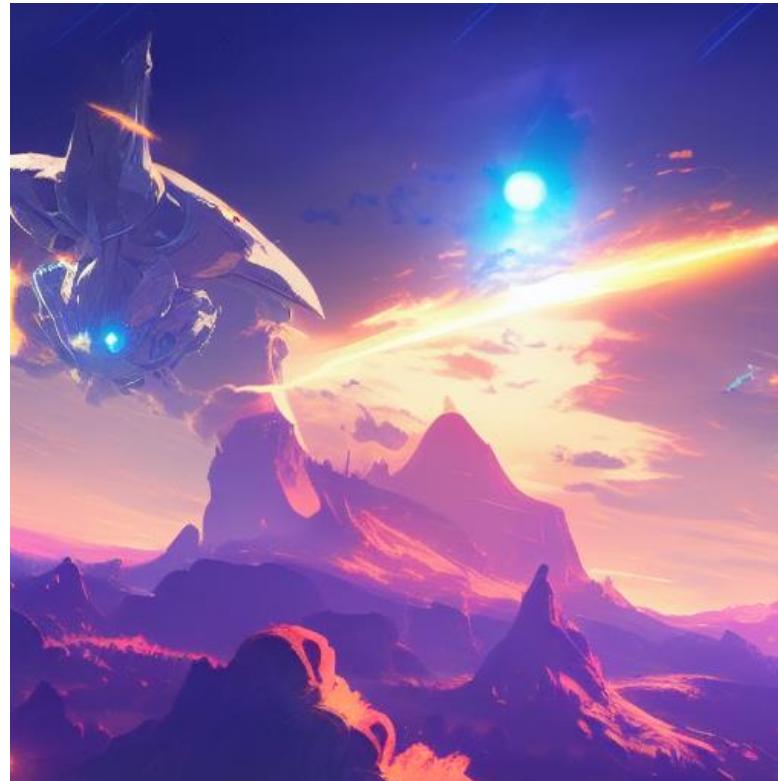


A Trick: Name a show, movie, or artist and you get all of the aesthetics all at once

A stunning fantastic starship battle, ((as seen from below on the ground over fantasy mountains)), dynamic lighting, highly detailed, 4k, HD, lens flare, trending on Pixiv, masterpiece, best quality,



in the style of "Star Wars"



from the videogame Legend of
Zelda Breath of the Wild



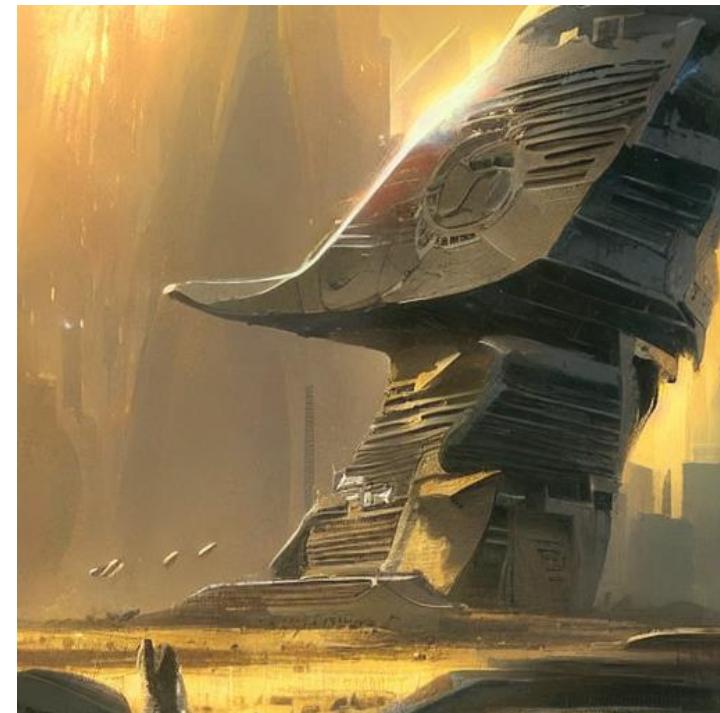
by Craig Mullins



a stunning fantasy landscape with a castle in the distance by Cory Loftis



a stunning digital painting of a floating medieval city by Cory Loftis



a lonely futuristic sci-fi tower by Craig Mullins



a stunning fantasy landscape with a castle in the distance by Bruce Pennington

Diffusion Models



DALL·E 2

DALL·E 2 is a new AI system that can create realistic images and art from a description in natural language.



Stable Diffusion

stability.ai



Disco Diffusion



Diffusion Models



DALL·E 2

DALL·E 2 is a new AI system that can create realistic images and art from a description in natural language.



Stable Diffusion

stability.ai



Disco Diffusion



Full Model made Public August 22, 2023



Search models, datasets, users...

Models Datasets Spaces Docs Solutions Pricing

≡

?

Hugging Face is way more fun with friends and colleagues! [Join an organization](#)

Dismiss this message

← Back to blog

Stable Diffusion with Diffusers

Published August 22nd, 2022.

[Update on GitHub](#)



[valhalla](#)
Suraj Patil



[pcuenq](#)
Pedro Cuenca



[natolambert](#)
Nathan Lambert



[patrickvonplaten](#)
Patrick von Platen

[Open in Colab](#)

Stable Diffusion

...using Diffusers

Stable Diffusion is a text-to-image latent diffusion model created by the researchers and engineers from [CompVis](#), [Stability AI](#) and [LAION](#). It is trained on 512x512 images from a subset of the [LAION-5B](#) database. LAION-5B is the largest, freely accessible multi-modal

To run the pipeline, simply define the prompt and call pipe.

```
prompt = "a photograph of an astronaut riding a horse"
```

```
image = pipe(prompt)["sample"][0]
```

```
# you can save the image with  
# image.save(f"astronaut_rides_horse.png")
```

The result would look as follows





Alex J Bowers
@Alex_J_Bowers

After a bunch of tweaking, I got the [#stablediffusion](#) [#python](#) code to run in [#Rstudio](#) [#Rmarkdown](#) using the code from [@StableDiffusion](#) [@huggingface](#) [@rstats](#). Also, yes Python runs in [@Rstudio!](#) Thank you [#Reticulate!](#)

2/n

huggingface.co/CompVis/stable...

```
Full Stable Diffusion Code in RMarkdown running on CPU only

```{r}
#setup
library(reticulate)
py_install("diffusers")
py_install("huggingface_hub")
```

Generate an image from prompt running Python in RStudio Markdown

```{python}

This is code modified from https://huggingface.co/CompVis/stable-diffusion-v1-4
Running in Rmarkdown in RStudio
I couldn't get the cuda code to work, so it runs on cpu with minimal edits to
base code from the main huggingface/github pages

from torch import autocast
from diffusers import StableDiffusionPipeline

Huggingface requires the code to pass a token.
First, get an account at huggingface.co
Then in settings, under tokens, get a token at https://huggingface.co/settings/tokens
The token will start with "hf_" followed by a string of characters
You must also agree to the stable diffusion license at (click the checkbox)
https://huggingface.co/Compvis/stable-diffusion-v-1-4-original

Copy/paste your full token here. Don't share your code as this is your personal
token hardcoded.

token = "hf_..."

pipe = StableDiffusionPipeline.from_pretrained(
 "CompVis/stable-diffusion-v1-4",
 use_auth_token=token
)

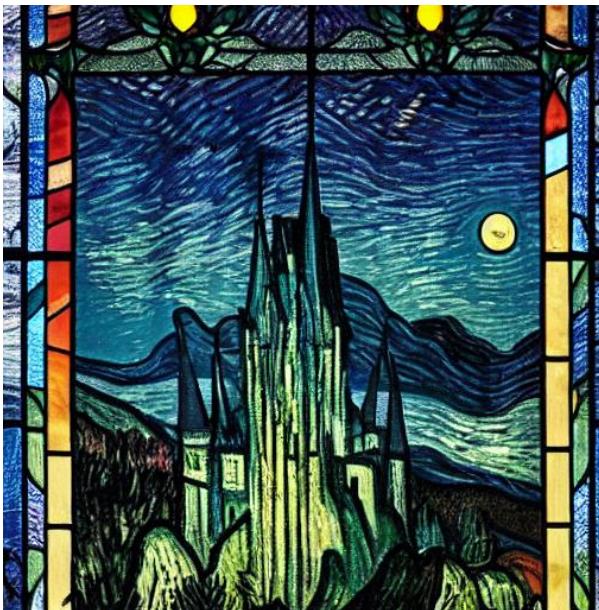
prompt = "Gothic Dracula's castle at night in a misty forest with the moon rising behind dark mountains"
with autocast("cuda"):
 image = pipe(prompt)[["sample"]][0]

image.save("drac01.png")
```

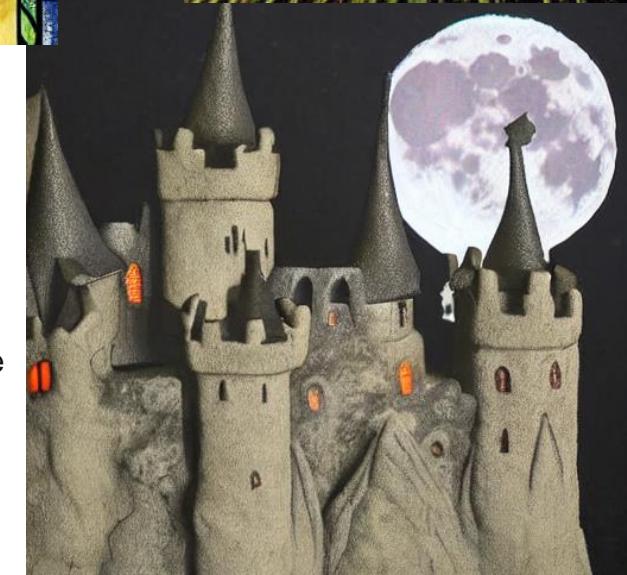
```

...

A stain glass window depicting Gothic Dracula's castle at night in a misty forest with the moon rising behind dark mountains in the style of "Wheat field with cypresses" by Vincent van Gogh (1889)

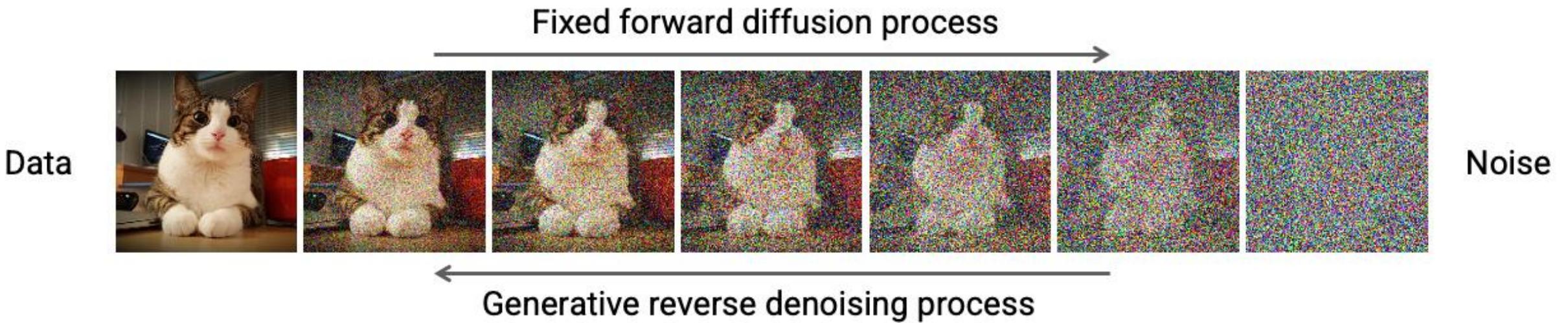


Gothic Dracula's castle at night in a misty forest with the moon rising behind dark mountains in the style of "Wheat field with cypresses" by Vincent van Gogh (1889)



Claymation art of Gothic Dracula's castle at night in a misty forest with the moon rising behind dark mountains, 100mm, extremely detailed

Diffusion Models



Diffusion Models: What the hyperparameters are doing

Prompt:

a fantastic glowing purple tree in
the middle of New York City at
sunset, dynamic lighting,
intense shadows, 4k, HD,
trending on artstation, trending
on Pixiv, best quality,
masterpiece



Classifier Free Guidance Scale
How strongly the image should
conform to the prompt. Lower
values produce more creative
results.



CFG Scale: 1.0

Steps for the sampling method. How many rounds of sampling. 

Steps: 1

Steps: 5

Steps: 20

Steps: 50



CFG Scale: 7.0

CFG Scale: 15.0

CFG Scale: 25.0



Diffusion Models: What the hyperparameters are doing

Prompt:

a fantastic glowing purple tree in
the middle of New York City at
sunset, dynamic lighting,
intense shadows, 4k, HD,
trending on artstation, trending
on Pixiv, best quality,
masterpiece

Different samplers can give
different results. Euler a is weird
as the more steps the more
different it becomes. All of the
others will usually converge on an
image but on somewhat different
rates.

Steps for the sampling method. How many rounds of sampling. →

Euler a



Euler



LMS



DDIM





Danielle Computer Images 🎵
@djaskin_images

...

there are two prompts inside you



**weathered 35mm double-exposure film
still of human-computer iteration
glimpsing into an infinity mirror, 1985,
Criterion Collection, by Gustave Doré, with
warm natural light, grainy, melancholy,
and a faded dusty color palette**

elmo at a KFC

9:21 PM · Jun 12, 2022 · Twitter Web App

Bowers, 2022 Getting started with Stable Diffusion

https://twitter.com/djaskin_images/status/1536156787329495040

| Artists | Modifiers |
|---|---|
| Fantasy Landscapes
- Tyler Edlin
- Mark Simonetti | 4k resolution
8k resolution
Unsplash photo contest winner
Trending on artstation |
| Fantasy Characters
- Justin Gerard
- Wayne Barlowe
- Victor Adame Minguez
- Jesper Ejsing
- Gerald Brom
- Greg Rutkowski | Deviantart
#pixelart
3d art
Digital art
Blender
Octane Render
Unreal engine
Watercolor |
| Classics
- DaVinci
- Pablo Picasso
- Van Gogh
- Winslow Homer
- M.C. Escher | Oil painting
Acrylic painting
Shot on film
35mm lens
Portrait photography
Portrait
Character design |
| Sci-Fi
- Jim Burns
- John Harris
- Dean Ellis
- H.R. Giger | Cgsociety
Mandelbulb 3D
Trending on Flickr.
Vaporwave |
| Anime
- Studio Ghibli
- Makoto Shinkai | (You can really put anything though, don't let this limit you) |

Build prompts using the labels the model was trained on

<https://sweet-hall-e72.notion.site/A-Traveler-s-Guide-to-the-Latent-Space-85efba7e5e6a40e5bd3cae980f30235f>

Albrecht Dürer

"a beautiful painting of a waterlily pond by Albrecht Dürer, Trending on artstation."



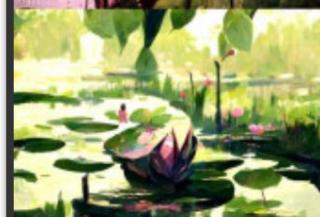
"a beautiful painting of a building in a serene landscape by Albrecht Dürer, Trending on artstation."



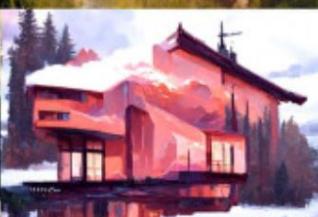
Albrecht Dürer by KyrickYoung

Alena Aenami

"a beautiful painting of a waterlily pond by Alena Aenami, Trending on artstation."



"a beautiful painting of a building in a serene landscape by Alena Aenami, Trending on artstation."



Alena Aenami by sureailabs

Alex Grey

"a beautiful painting of a waterlily pond by Alex Grey, Trending on artstation."



"a beautiful painting of a building in a serene landscape by Alex Grey, Trending on artstation."



Alex Grey by proximasan

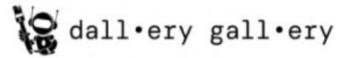
<https://dallery.gallery/wp-content/uploads/2022/07/The-DALL%C2%B7E-2-prompt-book.pdf>

Brought to you by...
dallery.gallery

Version 1.0
July 13, 2022

DALL·E 2 prompt book

Illustrations, 3D + textured



Isometric 3D ([source](#))



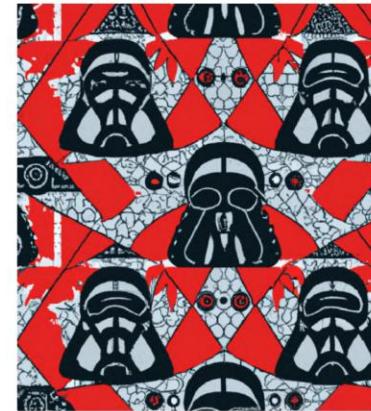
3D render, houdini 3D, octane 3D, ZBrush, Maya, Cinema 4D, Blender



Claymation, Aardman Animation, ([source](#))



Felt pieces ([source](#))



Fabric pattern ([source](#))



Black velvet, Edgar Leeteg ([source](#))



Scratch art, foil art, gold on black



Perler beads (Hama beads)



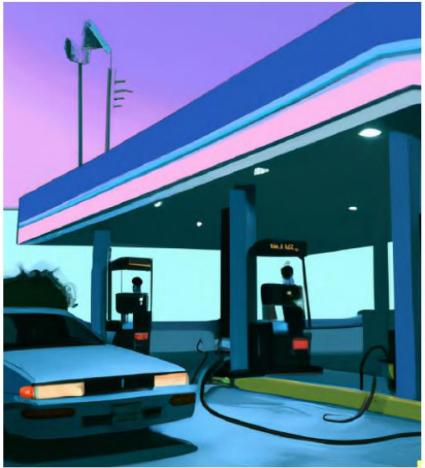
Screenshot of X from (game, real or imaginary, console, year) ([via](#))



Tattoo ([source](#))

Looks, vibes, -punks, -waves

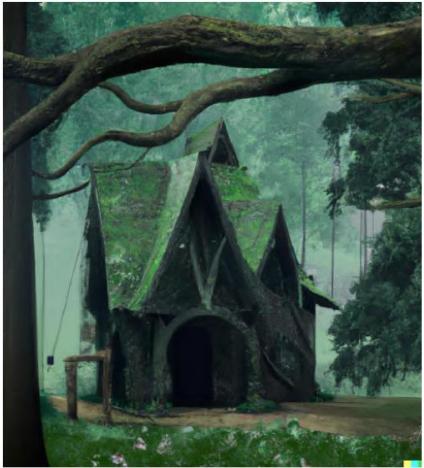
🤩 Check out [this huge list!](#)



Vaporwave: neon, pink, blue, geometric, futuristic, '80s.



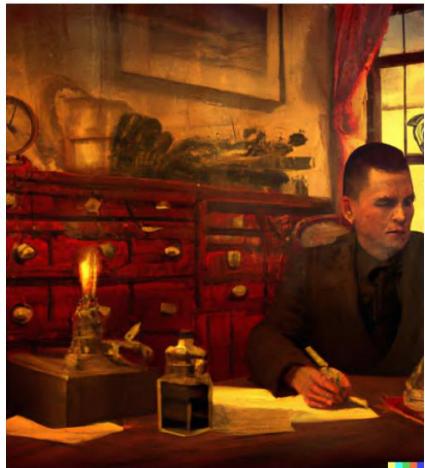
Post-apocalyptic: grey, desolate, stormy, fire, decay



Gothic, fantasy: stone, dark, lush, nature, mist, mystery, angular



Cybernetic, sci-fi: glows, greens, metals, armor, chrome



Steampunk: gold, copper, brass, Victoriana,



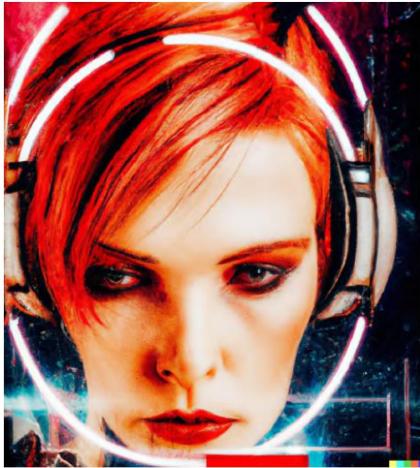
Memphis, Memphis Group, 1980s, bold, kitch, colourful, shapes



Dieselpunk, grimy, steel, oil, '50s, mechanised, punk cousin of steampunk



Afrofuturism: futuristic, and African!



Cyberpunk, 1990s, dyed hair, spiky, graphic elements .



Biopunk, organic: greens, slimes, plants, futuristic, weird

Faces are difficult for all diffusion models. Celebrities with lots of training data in the dataset give good results



A photo of Morgan Freeman standing next to a unicorn with a rainbow in the distance



A photo of Meryl Streep standing next to a unicorn with a rainbow in the distance

Stable Diffusion can also understand Idioms
And the dataset is extremely biased given what it was trained on



Prompt:
A fossil of a man watching TV

One-Click WebUI

- Must have Windows or Linux
- Must have an Nvidia graphics card
- Follow the directions to the letter (Python 3.6.10!!!!)

A screenshot of the GitHub repository page for `AUTOMATIC1111/stable-diffusion-webui`. The page displays the repository's structure, recent commits, and various files. Key commits shown include fixes for iterator bugs and improvements to the training template.

README.md

Stable Diffusion web UI

A browser interface based on Gradio library for Stable Diffusion.

The screenshot shows the `README.md` file for the `stable-diffusion-webui` repository. The file contains a brief description of the web UI and its purpose. Below the file is a screenshot of the web application interface. The interface includes a text input field for the prompt, a sidebar with various sampling and generation parameters, and a preview window showing the generated image of two dice. The bottom of the interface provides options to save the image or send it to other services.

Check the [custom scripts](#) wiki page for extra scripts developed by users.

Features

Detailed feature showcase with images:

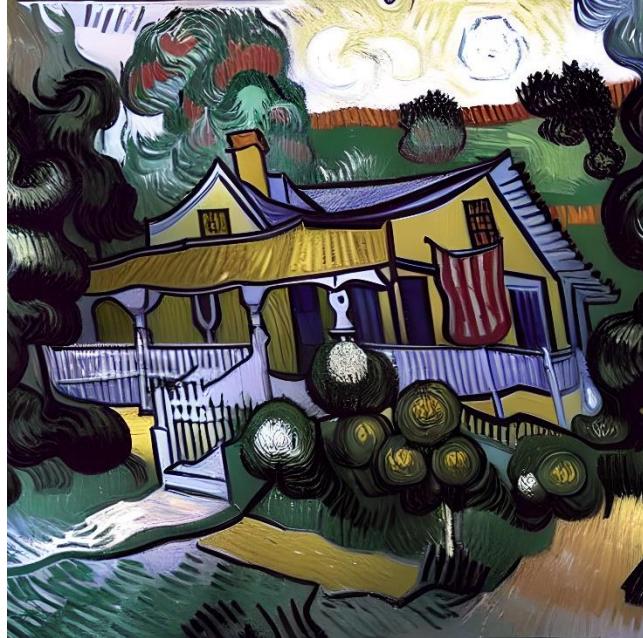
- Original txt2img and img2img modes
- One click install and run script (but you still must install python and git)

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

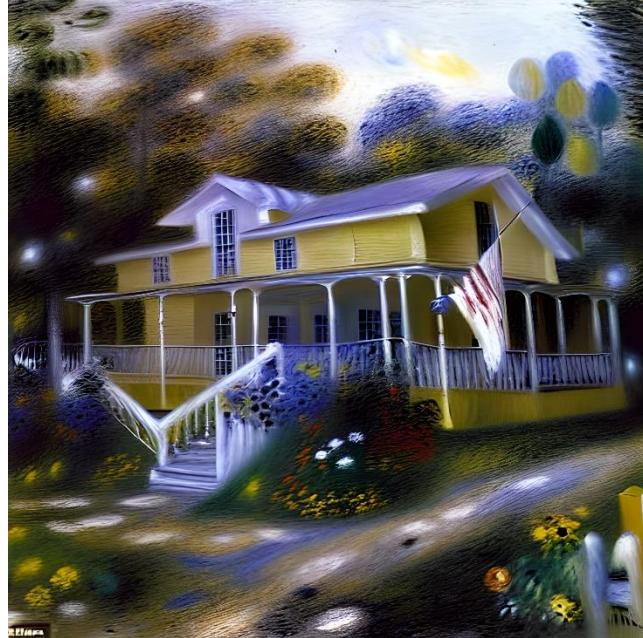
Image to Image (Img2Img)



Original photo I took of my parent's house in Michigan



in the style of Vincent van Gogh



in the style of Pierre-Auguste Renoir



in the style of Georges Seurat

an oil on canvas painting of a tall narrow light-yellow cottage in the woods with an American flag on the front porch and a porch with a white railing,

Img2Img Alternative



Original Prompt:
A man looking into the camera
with grey hair

A man looking into the camera
with brown hair

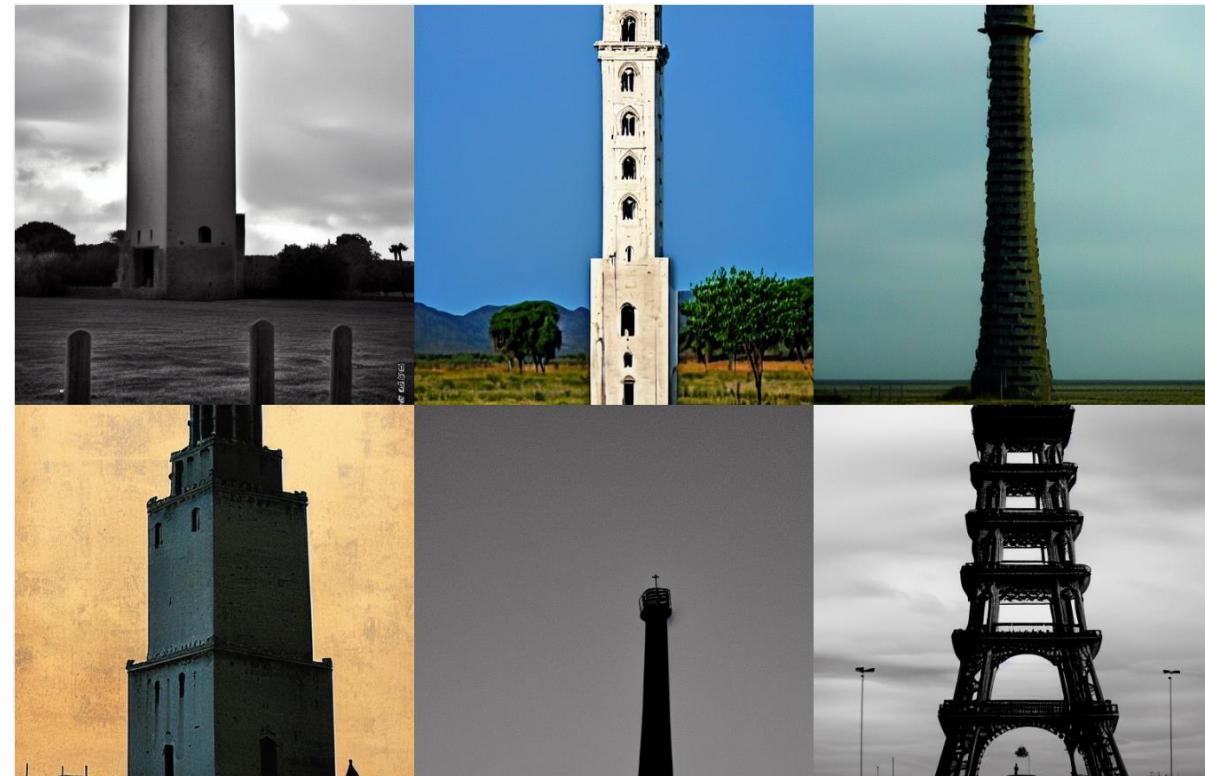
A man looking into the camera
with blonde hair

A man looking into the camera
with red hair

Negative Prompts



A lonely tower, by Pablo Picasso



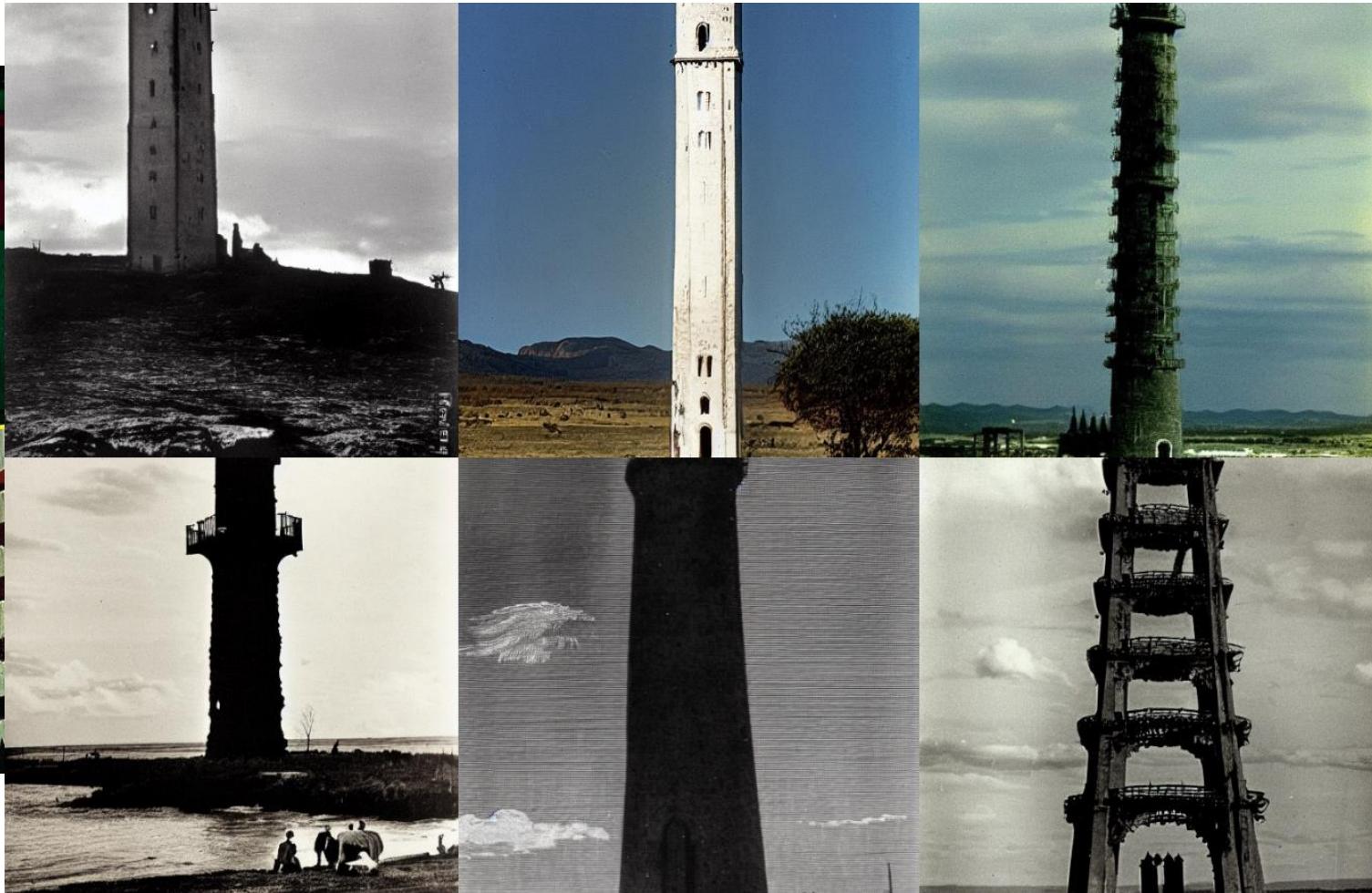
A lonely tower, by Pablo Picasso
Negative Prompt: Wassily Kandinsky

Steps: 100, Sampler: Euler a, CFG scale: 7.5, Seed: 2555614503, Size: 512x512

Negative Prompts



A lonely tower, by Pablo Picasso



A lonely tower, by Pablo Picasso
Negative Prompt: cubism art style

Steps: 100, Sampler: Euler a, CFG scale: 7.5, Seed: 2555614503, Size: 512x512

Your turn! Making prompts and images in Stable Diffusion

- Discussion and image generation groups of 2-3
 - Introduce yourself
 - Discuss why you came today to the session
 - Start building prompts together!
- Start by using the website:
 - If it gives an error just click “generate” again.
 - Open multiple tabs to get it cooking on many prompts.
 - Save your prompts and images to your computer!!!!
 - <https://huggingface.co/spaces/stabilityai/stable-diffusion>
 - Please upload images you wish to share to:
- If you have a computer with an Nvidia GPU let's get Automatic1111's Gradio WebUI working for you. Bowers and DSEA team will help!
- After you have made prompts and images together, we'll spend some time at the end of today discussing the implications.

Prompt Engineering Resources:

- Dalle-2 Prompt book:
 - <https://dallery.gallery/wp-content/uploads/2022/07/The-DALL%C2%B7E-2-prompt-book.pdf>
- 700+ artist comparison
 - <https://weirdwonderfulai.art/resources/disco-diffusion-70-plus-artist-studies/>
- Bowers Image Gardens:
 - https://twitter.com/Alex_J_Bowers/status/1569684954740572162
 - https://twitter.com/Alex_J_Bowers/status/1567190154430873601
 - https://twitter.com/Alex_J_Bowers/status/1564259191946190849

Open Issues and Unanswered Questions

- What is Stable Diffusion and machine learning image generation?
- Is it stealing?
- “You believe you made that don’t you?”
- What does Machine Learning Image Generation mean for Art, Graphic Design, and Education?

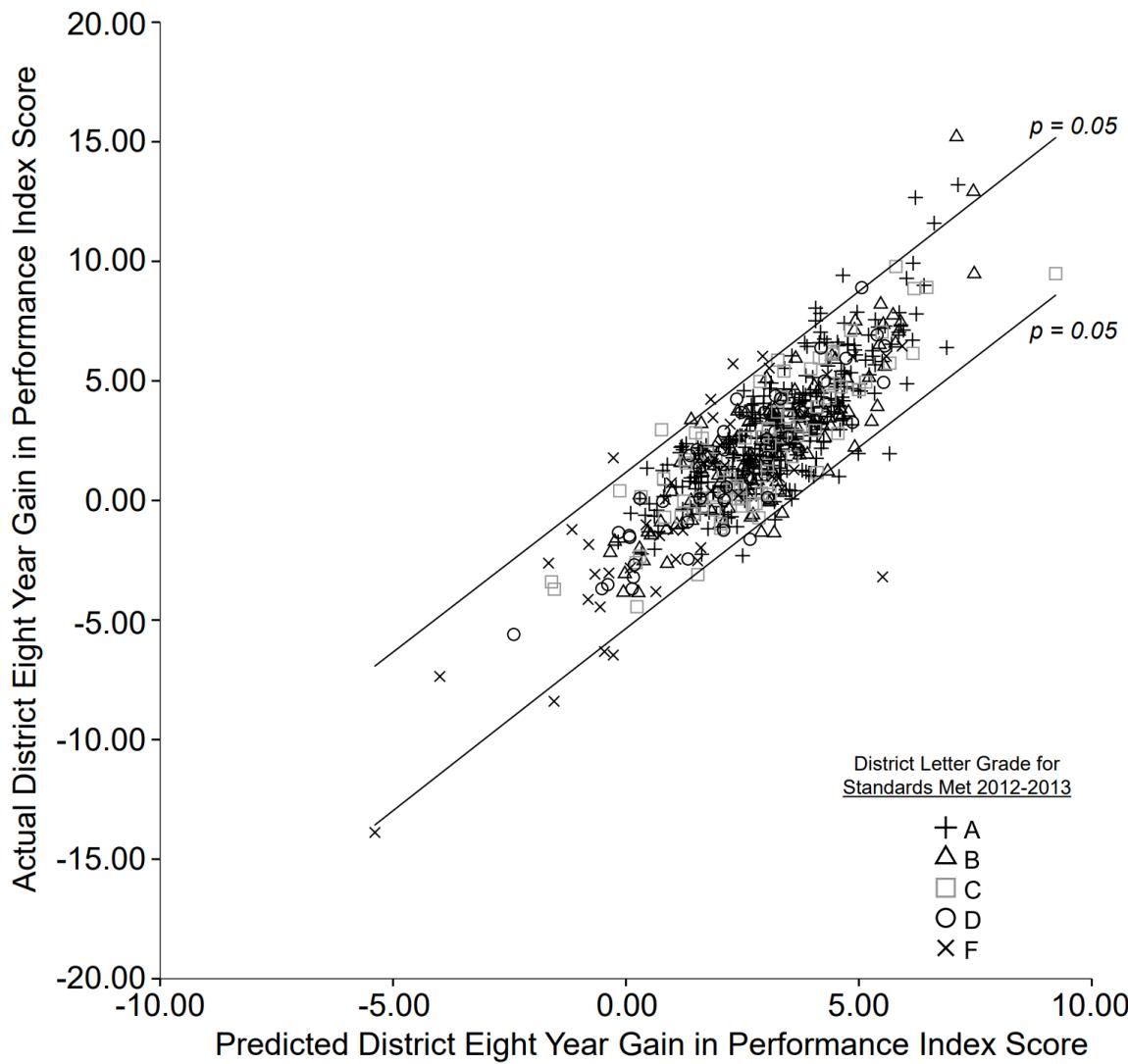


Figure 2: Comparison of predicted versus actual eight year Performance Index (PI) score gain in Ohio, 2005-06 – 2012-13. Hierarchical linear growth modeling predicted PI score gains are plotted on the x-axis for each of the 610 schools in Ohio over the time period, versus the actual gains from 2005-06 to 2012-13. The 95% confidence intervals are shown. Individual districts are plotted as symbols representing the Ohio Department of Education letter grade for adequate yearly progress on the number of standards met, with A being the highest (90-100% of standards met) and F the lowest (0-49%). Districts that are outside the 95% confidence interval are designated as significantly outperforming or underperforming their demographic and context variables.

Bowers, A.J. (2015). Site Selection in School District Research: A Measure of Effectiveness Using Hierarchical Longitudinal Growth Models of Performance. *School Leadership & Management*, 35(1) p. 39-61. <http://dx.doi.org/10.1080/13632434.2014.962500>

Thoughts on Machine Learning Image Generation:

This is Figure 2 from Bowers (2015).
Do I believe I made this? (Bowers asks some questions)

- I did not collect the data. I downloaded a public file.
- It's data about kids!
- I did not invent any of the statistics.
- I used SPSS, none of which I made.
- I clicked some buttons and selected options.
- Then this image appeared after some iterations.
- I published this with some words around it in a journal article.
- Do I believe I made this?