

GenAI

Dan Moore

dan@makeitdoathing.com
ig: @danzeeeman

THE BEST ROAD

IN THE WORLD







What is Generative AI?

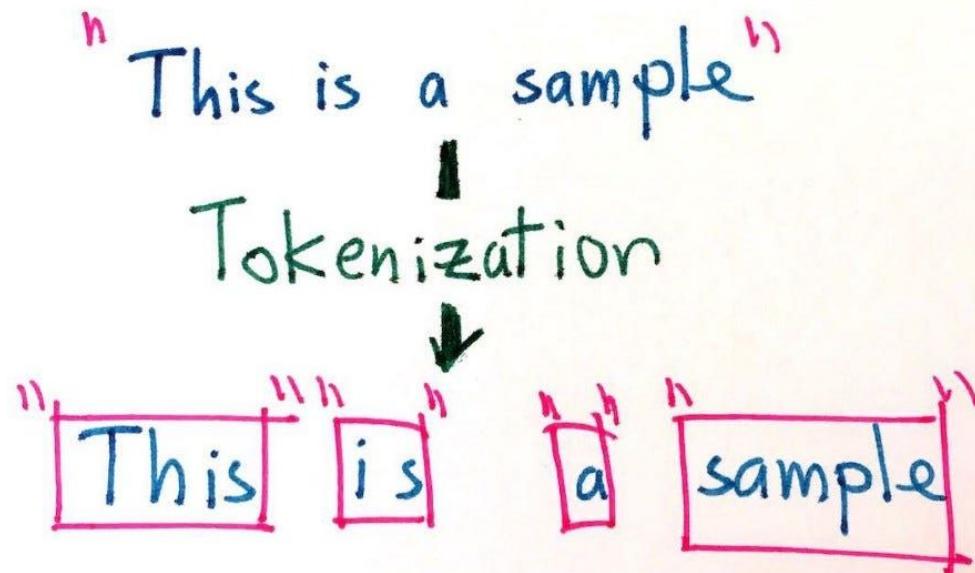
Generative AI is a type of artificial intelligence technology that broadly describes machine learning systems capable of generating *text*, *images*, *code* or other types of *content*, often in response to a *prompt* entered by a *user*.

Prompts

It's all about the tokens

Tokenizers break up words, parts of words and phrases into a unique token that the software can recognize.

Some tokenizers work by splitting up a sentence or phrase based of the spaces between words but some tokenizers take a more machine centered approach and split them up by bytes.



Tokenizer

The GPT family of models process text using **tokens**, which are common sequences of characters found in text. The models understand the statistical relationships between these tokens, and excel at producing the next token in a sequence of tokens.

You can use the tool below to understand how a piece of text would be tokenized by the API, and the total count of tokens in that piece of text.

GPT-3 Codex

somewhere

| Tokens | Characters |
|--------|------------|
| 3 | 9 |

TEXT TOKEN IDS

The screenshot shows a web-based tokenizer interface. At the top, there are two tabs: 'GPT-3' (which is active) and 'Codex'. Below the tabs is a text input field containing the word 'somewhere'. Underneath the input field are two buttons: 'Clear' and 'Show example'. Below the input field, there is a table with two columns: 'Tokens' and 'Characters'. The 'Tokens' column contains the number '3', and the 'Characters' column contains the number '9'. At the bottom of the interface, there are two radio buttons: 'TEXT' (which is selected) and 'TOKEN IDS'. The entire interface is contained within a light gray box.

A helpful rule of thumb is that one token generally corresponds to ~4 characters of text for common English text. This translates to roughly $\frac{3}{4}$ of a word (so 100 tokens \approx 75 words).

If you need a programmatic interface for tokenizing text, check out our [tiktoken](#) package for Python. For JavaScript, the [gpt-3-encoder](#) package for node.js works for most GPT-3 models.

Tokenizer

The GPT family of models process text using **tokens**, which are common sequences of characters found in text. The models understand the statistical relationships between these tokens, and excel at producing the next token in a sequence of tokens.

You can use the tool below to understand how a piece of text would be tokenized by the API, and the total count of tokens in that piece of text.

GPT-3 Codex

Please give me the history of the united states

[Clear](#) [Show example](#)

| Tokens | Characters |
|--------|------------|
| 9 | 47 |

TEXT **TOKEN IDS**

Please [5492, 1577, 502, 262, 2106, 286, 262, 16503, 2585]

A helpful rule of thumb is that one token generally corresponds to ~4 characters of text for common English text. This translates to roughly $\frac{1}{4}$ of a word (so 100 tokens \approx 75 words).

If you need a programmatic interface for tokenizing text, check out our [tiktoken](#) package for Python. For JavaScript, the [gpt-3-encoder](#) package for node.js works for most GPT-3 models.

Tokenizer

The GPT family of models process text using **tokens**, which are common sequences of characters found in text. The models understand the statistical relationships between these tokens, and excel at producing the next token in a sequence of tokens.

You can use the tool below to understand how a piece of text would be tokenized by the API, and the total count of tokens in that piece of text.

GPT-3 Codex

Please give me the history of the united states

[Clear](#) [Show example](#)

| Tokens | Characters |
|--------|------------|
| 9 | 47 |

TEXT **TOKEN IDS**

[5492, 1577, 502, 262, 2106, 286, 262, 16503, 2585]

A helpful rule of thumb is that one token generally corresponds to ~4 characters of text for common English text. This translates to roughly $\frac{1}{4}$ of a word (so 100 tokens \approx 75 words).

If you need a programmatic interface for tokenizing text, check out our [tiktoken](#) package for Python. For JavaScript, the [gpt-3-encoder](#) package for node.js works for most GPT-3 models.

Hello how are U tday?



hello how are u tday?



[hello, how, are, u, tday, ?]



[hello, how, are, u, td, ##ay, ?]



[CLS, hello, how, are, u, td, ##ay, ?, SEP]

Normalization

Pre-tokenization

Model

Postprocessor

The goal of tokenizers is to transform natural language into something a computer can understand and perform math operations on

It turns these phrases:

"I've been waiting for this blog my whole life."

"I hate this so much!"

Into this Tensor that the computer can understand:

```
raw_inputs = [
    "I've been waiting for this blog my whole life.",
    "I hate this so much!",
]

inputs = tokenizer(raw_inputs, padding=True, truncation=True, return_tensors="tf")
{
    'input_ids': <tf.Tensor: shape=(2, 16), dtype=int32, numpy=
        array([
            [ 101,  1045,  1005,  2310,  2042,  3403,  2005,  1037, 17662, 12172,  2607,  2026,  2878,  2166,  1012,  102],
            [ 101,  1045,  5223,  2023,  2061,  2172,    999,    102,      0,      0,      0,      0,      0,      0,      0,      0]
        ], dtype=int32)>,
    'attention_mask': <tf.Tensor: shape=(2, 16), dtype=int32, numpy=
        array([
            [1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1],
            [1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0]
        ], dtype=int32)>
}
```

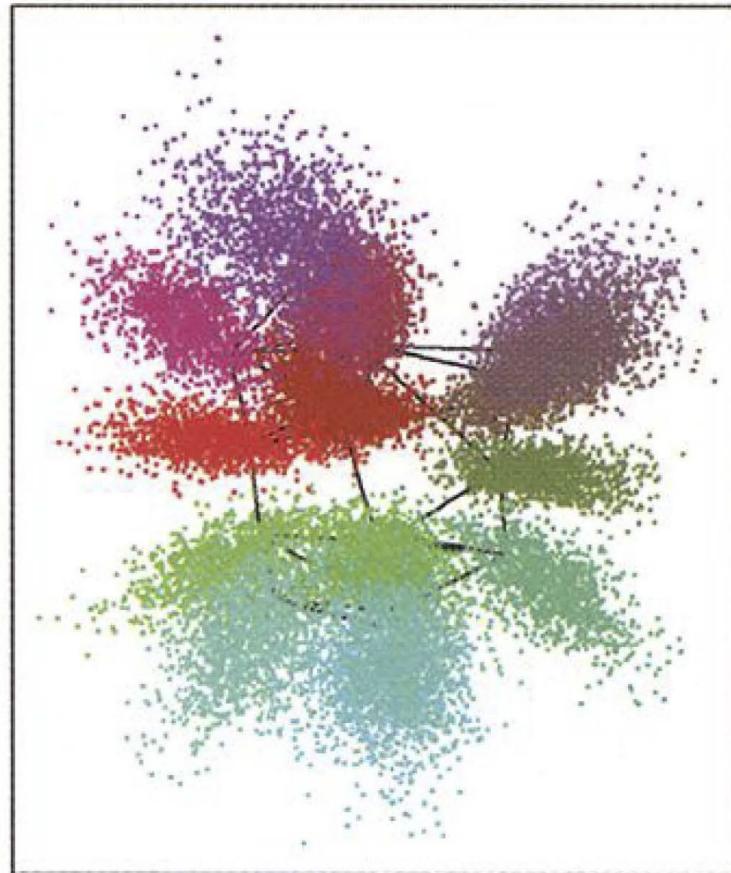
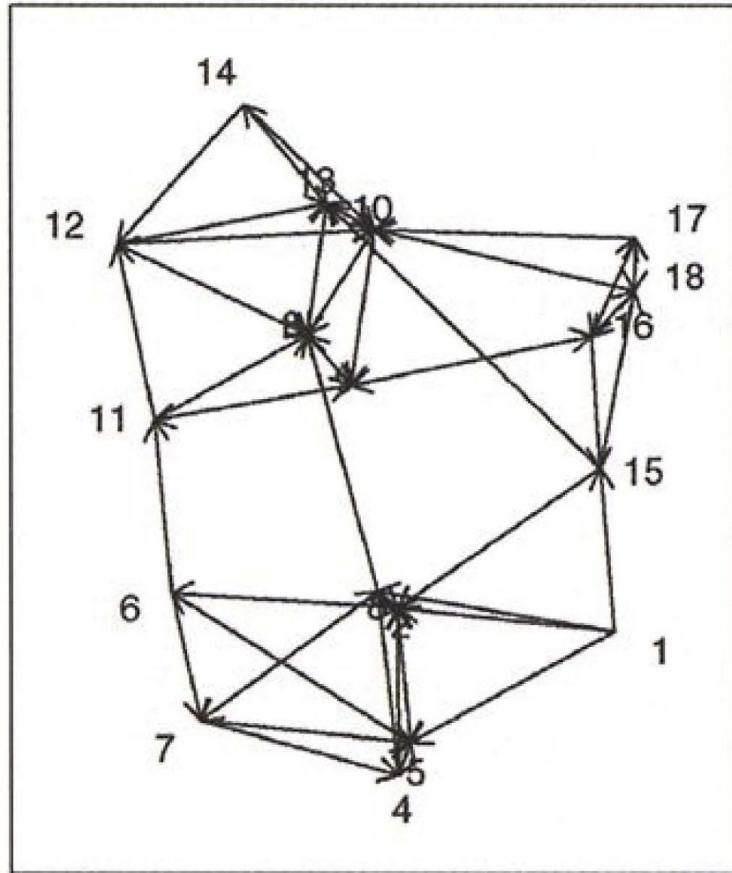
Let's look at phone numbers:

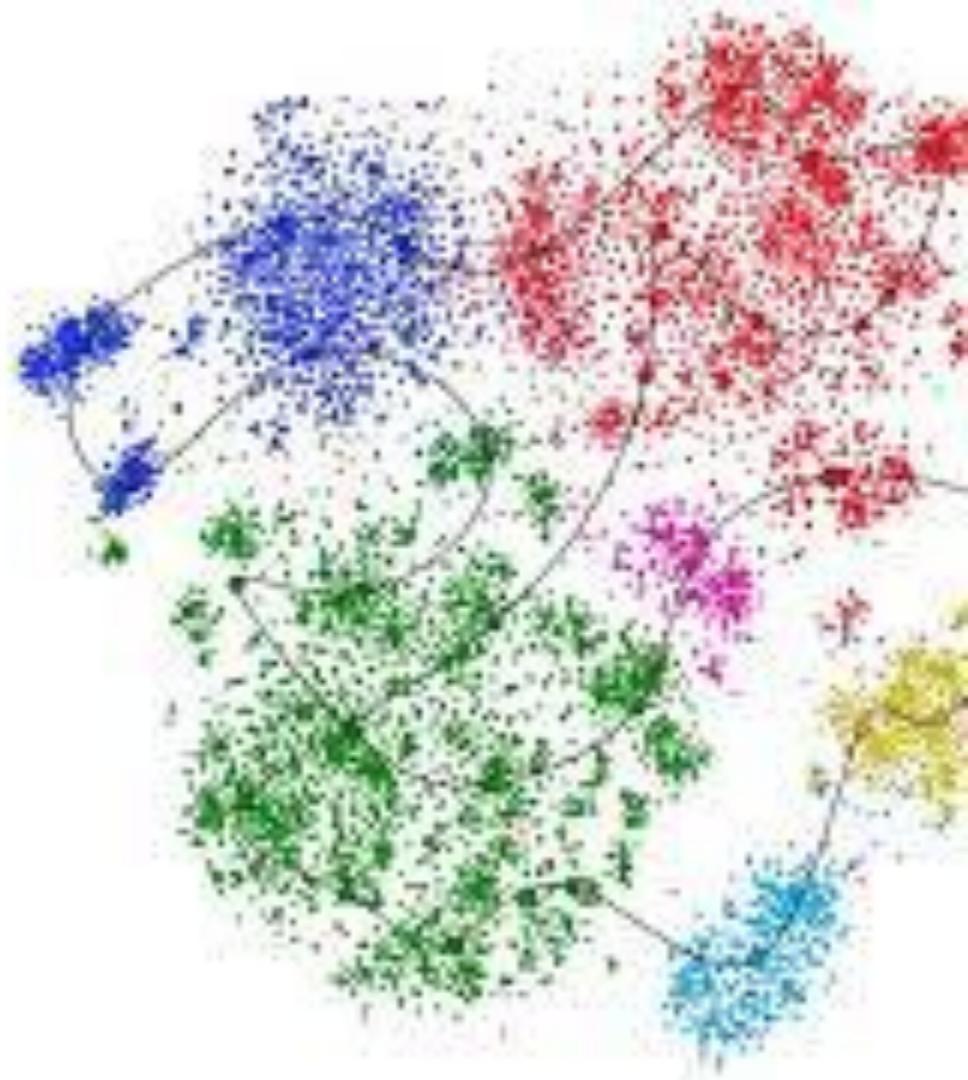
(Country Code)-(Area Code)-(Exchange)-(Extension)

[token]-[token]-[token]-[token]

[1,917,223,0777]

Latent Space





A massive yellow and black Transformer robot, Bumblebee, stands prominently in the center of a destroyed urban environment. He is shown from the waist up, facing slightly to the left. His head is yellow with a red visor featuring the Autobot logo. He has blue glowing eyes and a determined expression. His body is covered in dark, scaly-looking armor plating over his metallic frame. In the background, other Transformers are visible, including a smaller one on the left and another partially obscured behind Bumblebee. The city behind them is in ruins, with smoke billowing from burning buildings and twisted metal scattered across the ground.

Transformer Models

A transformer model is a type of deep learning model that was introduced in 2017. These models have been applied to a wide range of tasks in machine learning and artificial intelligence.

The model was first described in a 2017 paper called "Attention is All You Need" by a team at Google and a group from the University of Toronto. The release of this paper is considered a watershed moment in the field, given how widespread transformers are now used in applications such as training LLMs.

Transformers are used in:

- Image Classification
- Audio Classification
- Language Translation
- ChatGPT (GPT-3.5, GPT-4.0)
- Llama
- Mixtral-8x-7B

There are two primary innovations that transformer models bring to the table. Consider these two innovations within the context of predicting text:

- Positional encoding
- Self-attention

Positional encoding: Instead of looking at each word in the order that it appears in a sentence, a unique number is assigned to each word. This provides information about the position of each token in the sequence, allowing the model to consider the sequence's sequential information.

Self-attention: Attention is a mechanism that calculates weights for every word in a sentence as they relate to every other word in the sentence, so the model can predict words which are likely to be used in sequence.

The self-attention mechanism allows each word to attend to every other word in the sequence in parallel, weighing their importance for the current token. In this way, it can be said that machine learning models can “learn” the rules of grammar, based on statistical probabilities of how words are typically used in language.

chatGPT

oobabooga
text-generation-webui

| | | |
|--------------------|---|--------------|
| update_macos.sh | Better error handling during install/update | 4 months ago |
| update_windows.bat | Better error handling during install/update | 4 months ago |
| update_wsl.bat | Move one-click-installers into the repository | 4 months ago |
| wsl.sh | Add flash-attention 2 for windows (#4235) | 3 months ago |

README

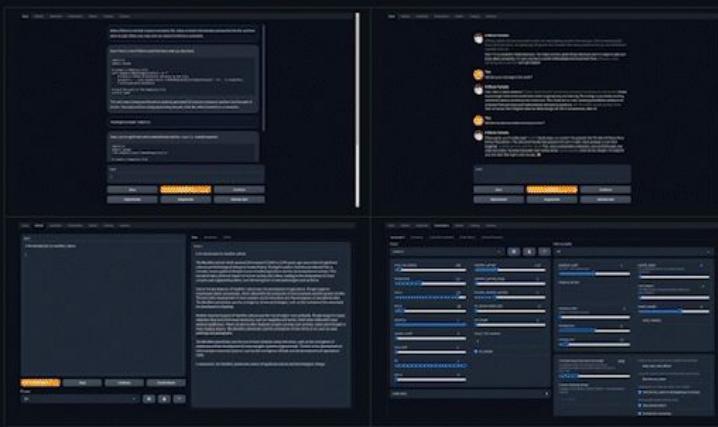
AGPL-3.0 license



Text generation web UI

A Gradio web UI for Large Language Models.

Its goal is to become the [AUTOMATIC1111/stable-diffusion-webui](#) of text generation.



Features

- 3 interface modes: default (two columns), notebook, and chat.
- Multiple model backends: [Transformers](#), [llama.cpp](#) (through [llama-cpp-python](#)), [ExLlamaV2](#), [AutoGPTQ](#), [AutoAWQ](#), [GPTQ-for-LLaMa](#), [CTransformers](#), [QulP](#).
- Dropdown menu for quickly switching between different models.
- Large number of extensions (built-in and user-contributed), including Coqui TTS for realistic voice outputs, Whisper STT for voice inputs, translation, [multimodal pipelines](#), vector databases, Stable Diffusion integration, and a lot more. See [the wiki](#) and the [extensions directory](#) for details.

Mixtral 8x7B

CATALOG

Explore Catalog
AI Foundation Models
Collections
Containers
Helm Charts
MxSgy
Resources

CONSOLE

ORGANIZATION

Catalog > Models > AI Foundation Models > Mistral 8x7B Instruct

Mixtral 8x7B Instruct



Playground

Overview

Related Collections

Description

Mixtral 8x7B Instruct is a language model that can follow instructions, complete requests, and generate creative text formats.

Publisher

Mistral

Modified

December 12, 2023

[Language Generation](#)[Large Language Models](#)[Text To Text](#)

Mistral Terms of Use: By using this software or model, you are agreeing to the terms and conditions of the [license](#), acceptable use policy and Mistral's privacy policy.

AI models generate responses and outputs based on complex algorithms and machine learning techniques, and those responses or outputs may be inaccurate or indecent. By testing this model, you assume the risk of any harm caused by any response or output of the model. Please do not upload any confidential information or personal data. Your use is logged for security.

[Demo](#) [API](#) [Documentation](#)

By using this demo, you acknowledge that you have read and agreed to the [Terms & Conditions](#).

Examples

What is the history of the internet?

Generate a list of 10 tips for writing better code.

Ex: Write a thing about a thing

Submit

Show Parameters



<< Collapse

Diffusion Models

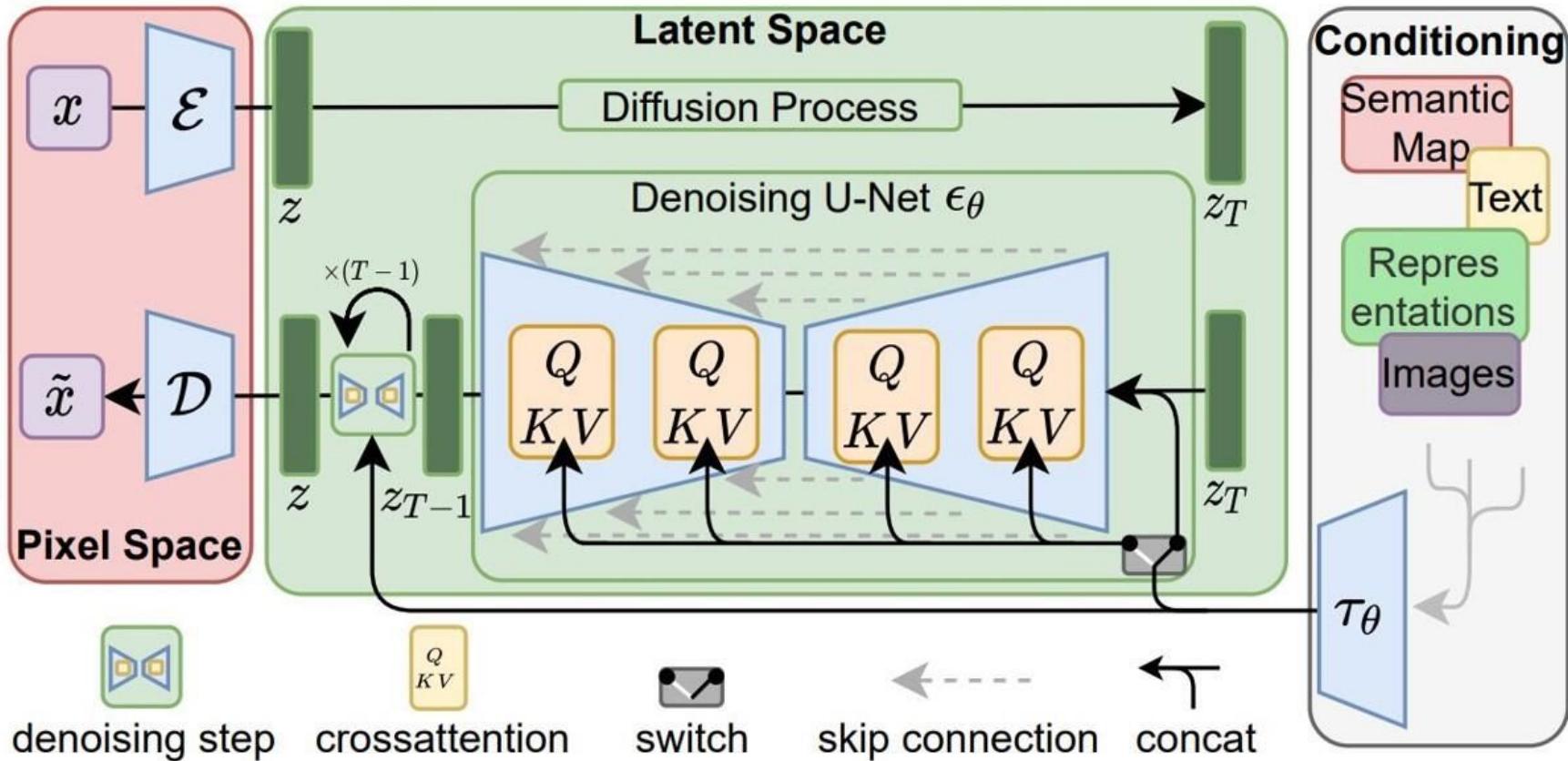
Forward Diffusion

In this phase, an original image is subjected to a process where a calculated amount of random noise is incrementally added to it. This step-by-step addition of noise gradually degrades the image, moving it away from its initial state and towards a state of randomness.

Reverse Diffusion

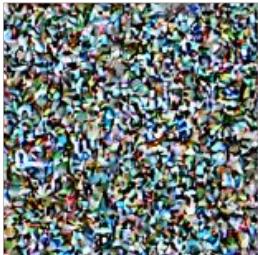
This phase involves the reverse operation. The aim is to methodically remove the noise that was added in the Forward Diffusion phase. Through this denoising process, the model works to reconstruct the original content of the image. It's a meticulous process of restoring the image from its noise-induced state back to its clear, original form.







First, denoise for a few steps towards the middle-point between the start and end prompts



Then, denoise the intermediate output towards interpolated points between the start and end prompts

Steps: 1



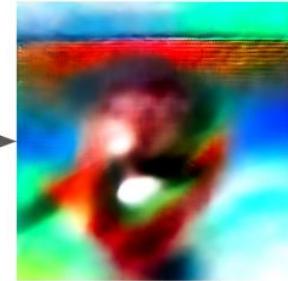
Steps: 2



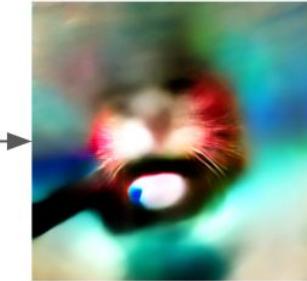
Steps: 3



Steps: 4



Steps: 5



Steps: 10



Steps: 9



Steps: 8



Steps: 7



Steps: 6



Steps: 15



Steps: 20



Steps: 25



Text2Image

TEXT2IMAGE allows you to input a natural language description and produces an image matching that description through a reverse diffusion process.



Image2Image

IMAGE2IMAGE allows you to input a natural language description and a reference image to produces an image matching that description using the reference image as its initial noise source.



ControlNet

ControlNet is a neural network structure to control diffusion models by adding extra conditions such as depth guidance or canny lines guidance.

Control Stable Diffusion with Hough Line Maps

Image: Room interior with bookshelves.

Prompt: Room

Run

Advanced options

Control Stable Diffusion with Canny Edge Maps

Image: Stuffed dog.

Prompt: cute dog

Run

Advanced options

Control Stable Diffusion with Depth Maps

Image: Stormtrooper at a podium.

Prompt: Stormtrooper's lecture

Run

Advanced options

Control Stable Diffusion with Human Pose

Image: Chef in the kitchen.

Prompt: Chef in the kitchen

Run

Advanced options

Control Stable Diffusion with Segmentation Maps

Image: House.

Prompt: house

Run

Advanced options

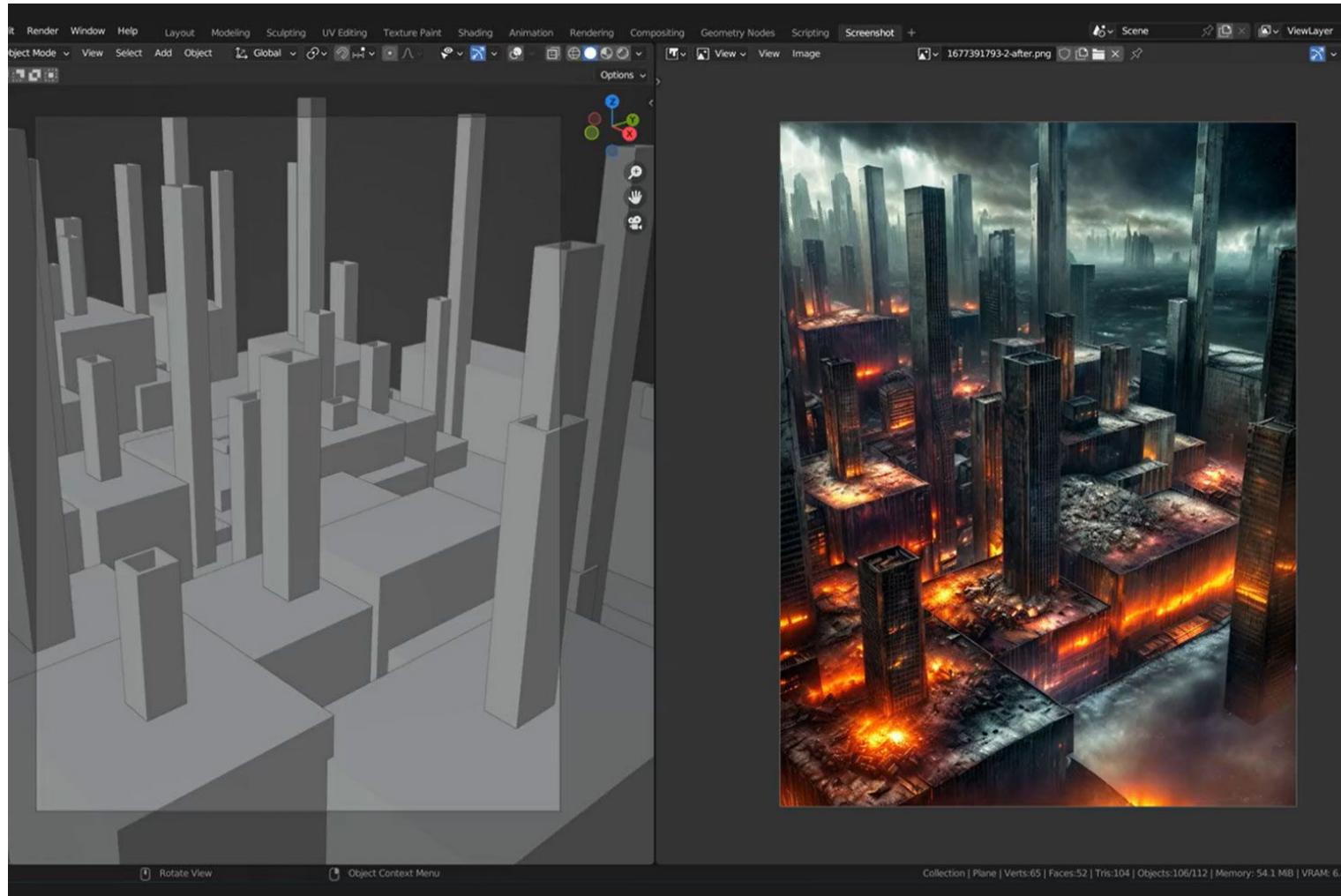
Control Stable Diffusion with Normal Maps

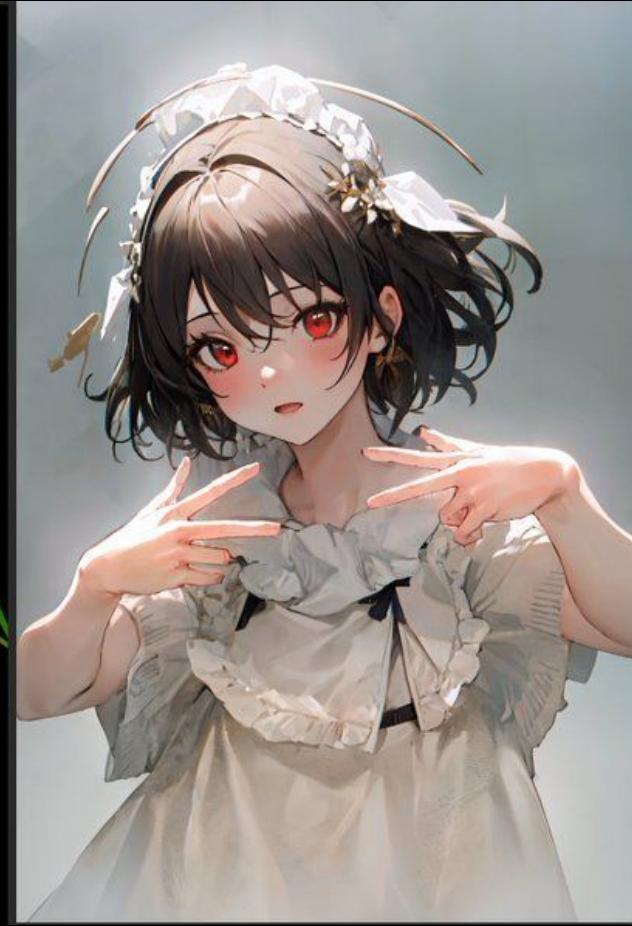
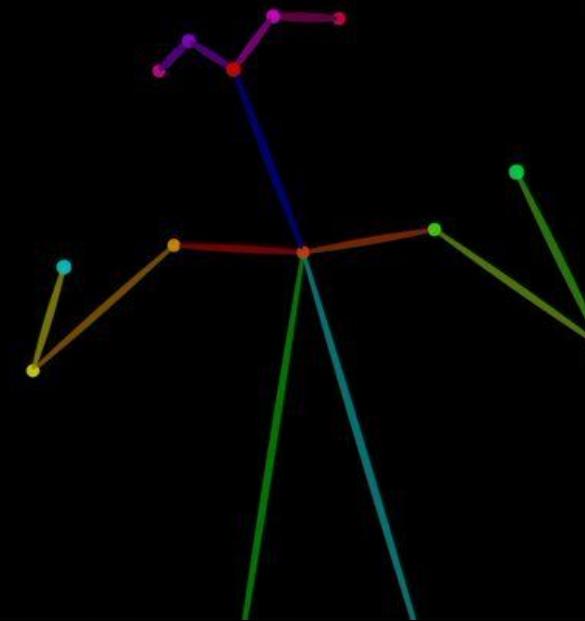
Image: Stuffed lamb.

Prompt: cute toy

Run

Advanced options





Stable Diffusion WebUI

Stable Diffusion checkpoint

protovisionXLHighFidelity3D_releaseV660Bakec



txt2img img2img mov2mov Extras PNG Info Checkpoint Merger Train Deform Settings Extensions

a cave opening in the forest woods nature photography 4k hyperrealistic

13/75

Generate

0/75

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



Generation Textual Inversion Hypernetworks Checkpoints Lora

Sampling method

DPM++ 2M Karras

Sampling steps

30

Hires. fix

Refiner

Width

1024

Batch count

20

Height

1024

Batch size

1

CFG Scale

9

Seed

-1

AnimateDiff

ControlNet v1.1.426

Script

None



a cave opening in the forest woods nature photography 4k hyperrealistic

Steps: 30, Sampler: DPM++ 2M Karras, CFG scale: 9, Seed: 4173715721, Size: 1024x1024, Model hash: 440a7f226b, Model:



txt2img img2img mov2mov Extras PNG Info Checkpoint Merger Train Deform Settings Extensions

Installed

Available

Install from URL

Backup/Restore

Load from:

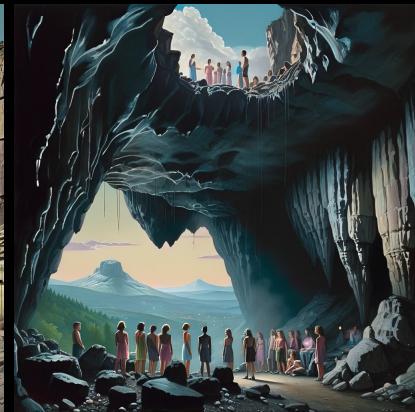
<https://raw.githubusercontent.com/AUTOMATIC1111/stable-diffusion-webui-extensions/master/index.json>

Hide extensions with tags

script localization tab dropdown ads installed training models UI related prompting newest first oldest first a-z z-a internal order update time create time stars

editing manipulations online animation query science extras

| Extension | Description | Action |
|---|---|-----------------------------------|
| Lama cleaner as masked content script, manipulations | Use lama cleaner preprocessor from controlnet inside Inpaint tab. Requires sd-webui-controlnet. Can be useful with removing objects in inpaint tab Update: 2024-01-14 Added: 2024-01-14 Created: 2024-01-04 | stars: 14 Install |
| Latent Regional Helper script, tab | Simplify the region division settings for Latent Couple and Regional Prompter. Simply select from the drop-down list to output the setting values. Update: 2024-01-14 Added: 2024-01-14 Created: 2023-12-31 | stars: 4 Install |
| Thumbnailizer script, tab, UI related | A thumbnail gallery and set management tool. Generate and switch between different thumbnail sets for your checkpoints. Update: 2024-01-16 Added: 2024-01-14 Created: 2024-01-12 | stars: 4 Install |
| GPT-4 Vision Image Captioner tab, online | This extension utilizes the power of GPT-4 Vision API to provide image captioning capabilities through the Stable Diffusion Web UI. Requires a paid subscription to OpenAI. Update: 2024-01-17 Added: 2024-01-14 Created: 2024-01-14 | stars: 8 Install |
| sd-webui-ocr tab, query | PaddleOCR's sd-webui extensions (pytorch implementation) Update: 2024-01-03 Added: 2024-01-04 Created: 2023-12-27 | stars: 8 Install |
| Characteristic Guidance Web UI dropdown, manipulations, science | New guidance method improving classifier-free guidance, providing enhanced sample quality at large CFG scale. Compatible with existing sampling methods. Update: 2024-01-17 Added: 2024-01-04 Created: 2023-12-27 | stars: 17 Install |
| Deepdanbooru Tag2Folder tab, query | Using this script you can move Images using deepdanbooru classification Update: 2024-01-04 Added: 2024-01-03 Created: 2023-12-26 | stars: 2 Install |
| Deepdanbooru object recognition tab, query | Interrogate Deepdanbooru and identify objects in the images Update: 2023-12-18 Added: 2023-12-16 Created: 2023-12-13 | stars: 18 Install |
| State Manager script, tab, UI related | A state manager to quickly save and return to previous configs in A1111 Update: 2024-01-16 Added: 2023-12-13 Created: 2023-12-13 | stars: 14 Install |
| Agent Attention dropdown, manipulations | Speed up image generation and improve image quality using Agent Attention! An unofficial implementation of the method in arXiv:2312.08874 [cs.CV] Update: 2023-12-27 Added: 2023-12-15 Created: 2023-12-15 | stars: 23 Install |
| Color Correction Extras extras | Adds native color correction feature into "Extras" tab Update: 2023-12-22 Added: 2023-12-09 Created: 2023-12-06 | stars: 6 Install |
| Replacer - fast inpaint tab, editing, animation | Adds a tab for fast inpaint image by detection prompt. Requires sd-webui-segmentAnything. It also useful for batch inpaint, and inpaint in video with stable diffusion Update: 2024-01-18 Added: 2023-12-06 Created: 2023-11-06 | stars: 34 Install |
| sd-webui-breadcrumbs UI related | Adds a breadcrumb trail and makes improvements to the Quicksettings menu. Update: 2023-12-04 Added: 2023-12-02 Created: 2023-12-01 | stars: 19 Install |





MidJourney

Midjourney

newbies-30 Bot room for new users. Type /imagine then describe what you want to draw. See <https://docs.midjourney.com/> for more information

New message since 6:24 PM

Mark As Read

LIVE NOW Exploring V6 with Victor Gnarly & Bay Raftt stage-1 • 109 in audience螺旋和3 others Listen In

4 Events Channels & Roles INFO SUPPORT trial-support NEWCOMER ROOMS newbies-30 NEWCOMER ROOMS 2 newbies-60 newbies-90 CHAT discussion prompt-chat COMMUNITY FORUMS prompt-faqs SHOWCASE paintovers in-the-world blend-showcase THEMED IMAGE GEN daily-theme GENERAL IMAGE GEN general-10 general-20 VOICE CHANNELS vc-text stage-1 Exploring V6 with Victor Gnarly... 109 9999 danzeem... Online

newbies-30 abstract art chaotic composition, surrealism - @me.andjustme.lo (fast) Today at 6:02 PM abstract art chaotic composition, surrealism - Image #1 @me.andjustme.lo



Upscale (2x) Upscale (4x) Very (Subtle) Very (Strong) Very (Region)
Zoom Out 2x Zoom Out 1.5x Custom Zoom
Web

newbies-30 abstract art chaotic composition, surrealism - @me.andjustme.lo (fast) Today at 6:02 PM abstract art chaotic composition, surrealism - Image #2 @me.andjustme.lo

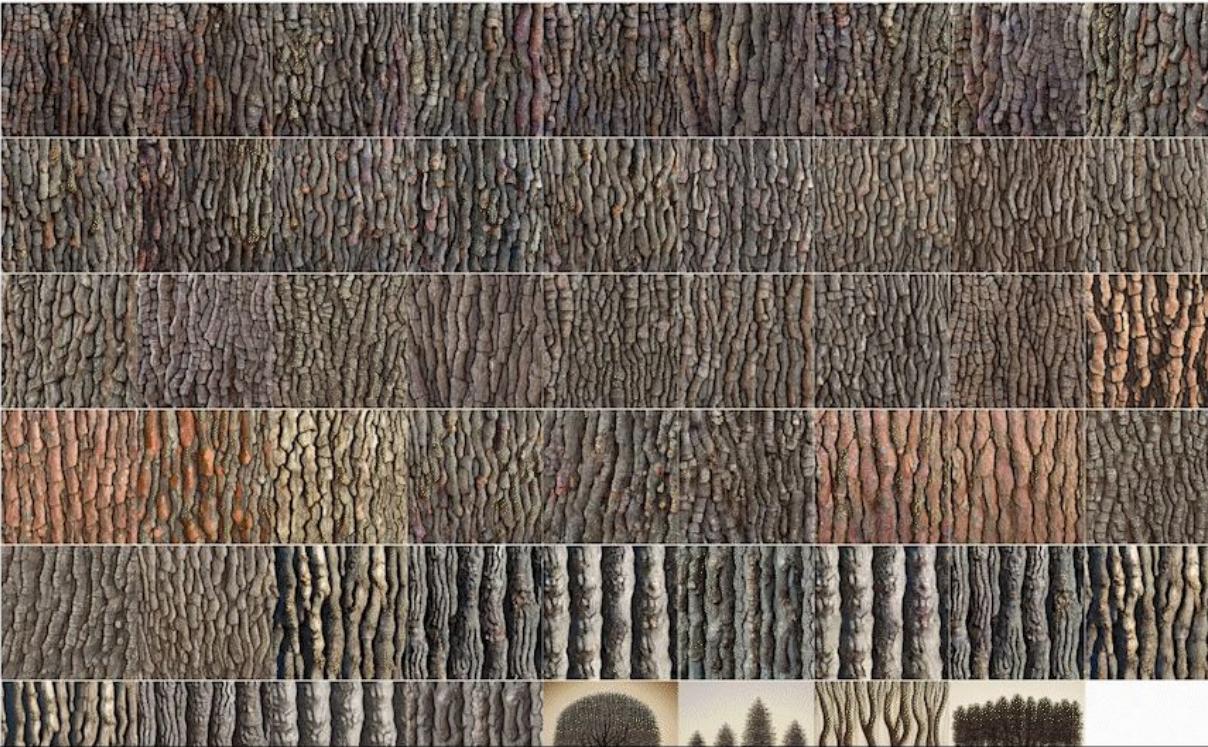


1 IN VOICE

Message newbies-30

TEAM - 6 Blue | Beep B Cixelyn ckruuse danielrussell DavidH Red | CEO of some pco. MIDJOURNEY BOT - 1 Midjourney Bot CHARON THE ALL KNOWING Charon the FAQ MODERATOR - 21 croakie! Danner Bot Core dig staring into the starlight Elvira what a grand and interesting Fenix! Glompy miss you Lord Bel Goofball Human beings make it Ivan Jaycott Kalmatos No zero day! kewiK kovas.boguta Lanpal Matt (Facebook) ramblinghuber Red Man

My Images



Folders

Filters

Rating



Type

- All
- Grids
- Upscales

Image Size

- All
- Square
- Landscape
- Portrait

Version

- All
 - 6
 - 5
- ▼ See more

- Other
- Tiled
 - Raw

View Options

Runway

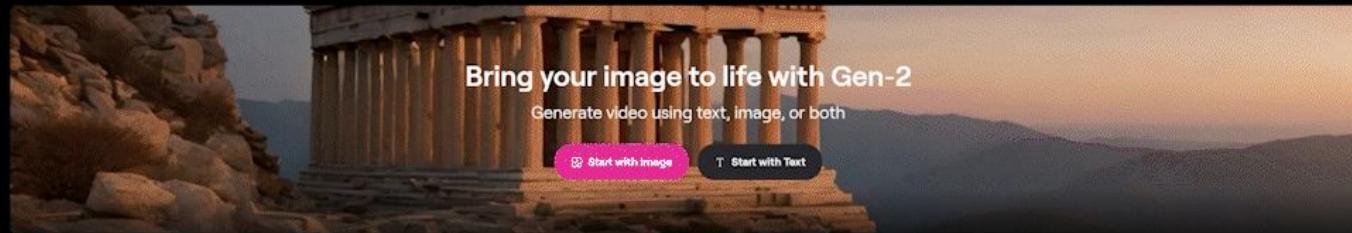
Den
1 member

+ Invite Collaborators

Search for tools, assets and projects



Home



View Assets

Upload or manage creations.

Try Gen-1

Video to video generation.

Try AI training

1 free training on paid plans.

Tell us about yourself

Personalize your experience.

Popular AI Magic Tools

Favorite tools from the community

[View all tools](#)



Video to Video



Text/Image to Video



Remove Background



Text to Image



Image to Image



Text To Speech

Tutorials

How to include AI in your workflow

[View all tutorials](#)



Gen-1 Basics



How to Use Gen-2



Gen-2 Workflow



How to Train Custom AI Models

Image2Video

Gen-2 video
4 seconds · 6:30PM, Jan 19



Your video is generating and will be done in a few minutes.

Your export will be saved to Dan's Assets.

TEXT IMAGE IMAGE + DESCRIPTION 10.9

A box of apples on the back of a truck, old vintage 1960s truck

Camera Motion Motion Brush BETA Add Style 80/300 Free Preview Generate 4s



Video2Video



DreamBooth

DreamBooth is a set of tools and techniques that allows you to fine-tune diffusion models to personalize them for your needs. You can quickly create customized models trained on the logo of a brand, the images of a person, or the style of an Artist.



Input images



in the Acropolis



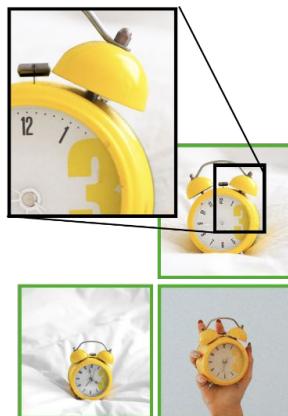
swimming



sleeping



getting a haircut



Input Images



Image-guided, DALL-E2



Text-guided, Imagen



Ours



Deforum



Diffusion Light





Aimazeyou
@aimazeyou

Two ChatGPT's talking to each other
about how to extinct humanity 😊



THANK YOU

Dan Moore is a New York City based Artist, Designer and Technologist. He currently works with Artificial Intelligence and Robotics to investigate Pop Culture. He has most recently exhibited *The Vivid Unknown* in IDFA:DocLab in Amsterdam and the *airegan** project at the Vancouver Art Gallery during *The Imitation Game: Visual Culture in the Age of Artificial Intelligence*. Moore holds a Masters of Tangible Interaction Design from Carnegie Mellon's College of Fine Art, where he studied and worked in the Frank-Ratchye STUDIO for Creative Inquiry. He was a virtual research fellow in the Free Art and Technology, F.A.T., Lab. He currently works in NVIDIA's Creative Department and is an adjunct professor at The New School.

DAN MOORE

DAN@MAKEITDOATHING.COM

IG @DANZEEEMAN

GITHUB @DANZEEEMAN

IF I COULD ASK FOR
ANYTHING
EXHIBITION SPACE AND
ACCESS TO MORE ABB
ROBOTS FOR A SHOW