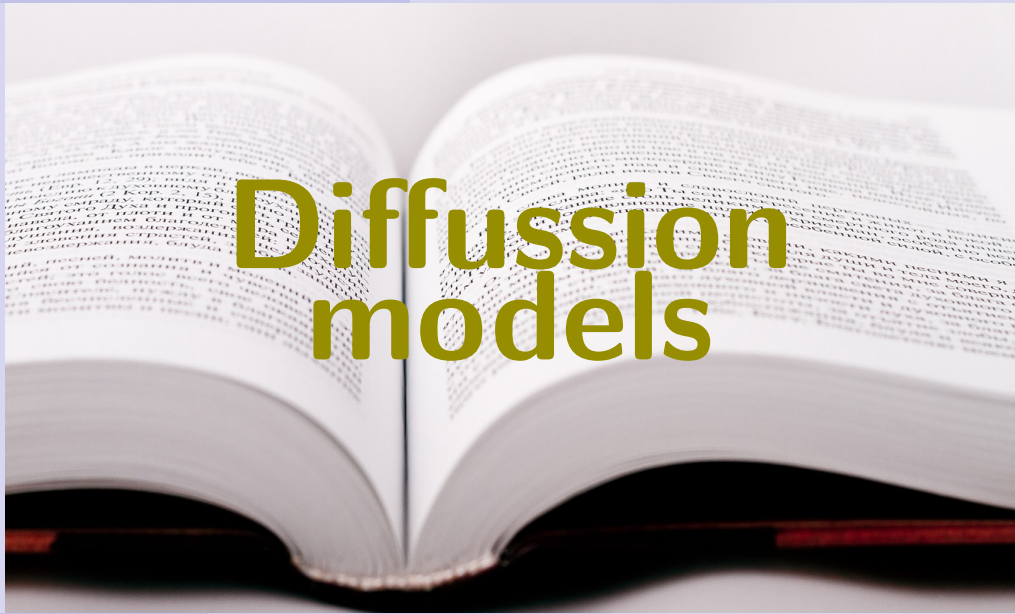
# Simple Introduction to Stable Difussion

Jesus M. Gonzalez-Barahona

Universidad Rey Juan Carlos
https://floss.social/@jgbarah
https://jgbarah.github.io/presentations

Machine Learning Spain
Madrid, Spain, December 1st 2022

# The plot



# Diffussion models

# What is FOSS(*)?

# Stable Diffussion release

# Stable Diffusion first release

Released on August 22nd

Licensed: Creative ML OpenRAIL-M

https://stability.ai/blog/stable-diffusion-announcement
https://colab.research.google.com/github/huggingface/notebooks/blob/main/diffusers/stable_diffusion.ipynb

# Dream Studio

Social site to give Stable Diffusion a try

Some gratis credit

USD 10 for 5,000 images

https://beta.dreamstudio.ai

# One week is just one week

## Dream Studio Beta

https://multimodal.art/news/1-week-of-stable-diffusion

# Stable Diffusion 2

Announced last week

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

Spain football team, winners of the World Cup in Qatar 2022, celebrating

Japan football team, winners of the World Cup in Qatar 2022, celebrating

Speaker presenting at Machine Learning Spain (25, 50)

Speaker presenting at Machine Learning China

# Extensions, integrations

# Integration in toolchains

# In-painting

# Out-painting

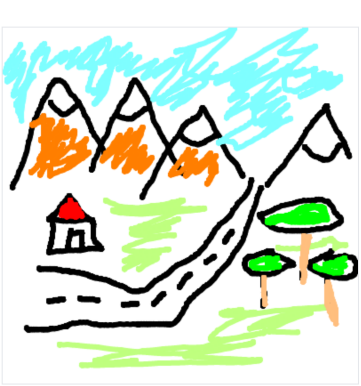
https://github.com/lkwq007/stablediffusion-infinity

# Image to image

Image + prompt produces an image

Even just with CPU!

https://huggingface.co/spaces/fffiloni/stable-diffusion-img2img




https://huggingface.co/spaces/huggingface-projects/diffuse-the-rest

# 3D assets

https://nv-tlabs.github.io/GET3D/

# 3D assets

## MotionDiffuse: Text-Driven Human Motion Generation with Diffusion Model

Mingyuan Zhang[1]* Zhongang Cai[1,2]* Liang Pan[1] Fangzhou Hong[1] Xinying Guo[1] Lei Yang[2] Ziwei Liu[1]+

[1]S-Lab, Nanyang Technological University [2]SenseTime Research

*equal contribution +corresponding author

| play the guitar | walk sadly | walk happily | check time |
|---|---|---|---|

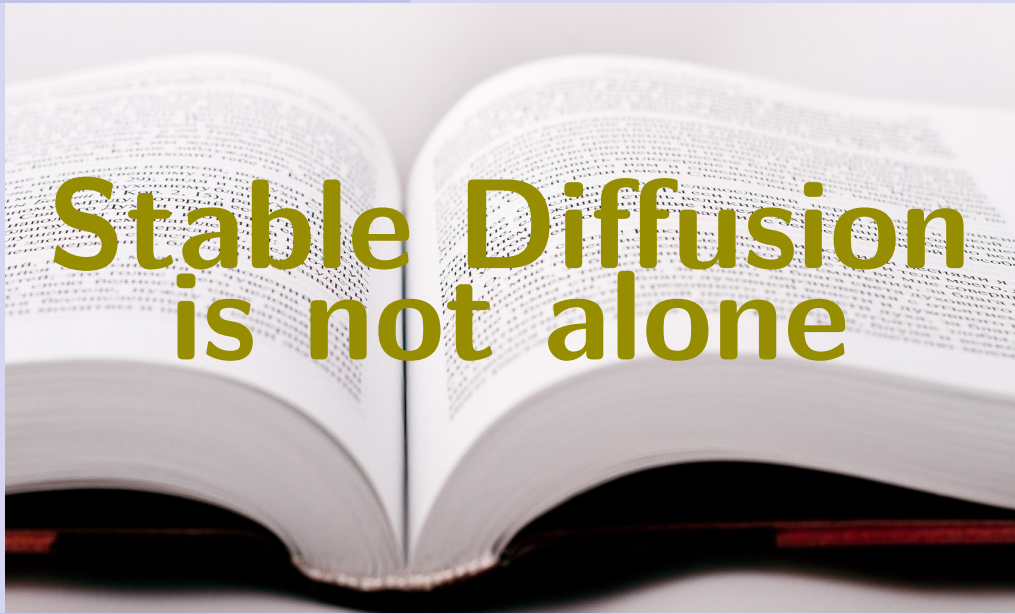
https://github.com/mingyuan-zhang/MotionDiffuse

# Videos

https://github.com/nateraw/stable-diffusion-videos

https://phenaki.github.io/

https://aiart.dev/posts/sd-music-videos/sd_music_videos.html



# Stable Diffusion is not alone

# Whisper

## Introducing Whisper

We've trained and are open-sourcing a neural net called Whisper that approaches human level robustness and accuracy on English speech recognition.

READ PAPER

VIEW CODE

VIEW MODEL CARD

https://openai.com/blog/whisper/ https://github.com/openai/whisper

License: MIT (open source)

# BLOOM

The World's Largest Open Multilingual Language Model

176 billion parameters

46 natural languages and 13 programming languages

https://bigscience.huggingface.co/blog/bloom

https://huggingface.co/bigscience/bloom

# Multilingual AI Assistant

Whisperfor Speech-to-text
Bloom for Text-generation,
CoquiTTS for Text-To-Speech

https://huggingface.co/spaces/ysharma/Talk_to_Multilingual_AI_WhisperBloomCoqui

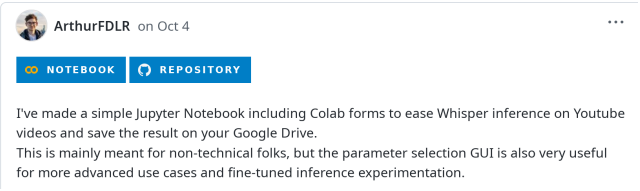
# Whisper to Stable Diffusion

https://huggingface.co/spaces/fffiloni/whisper-to-stable-diffusion

# Whisper fro Toutube captions

Easy to use Jupyter Notebook for Youtube video inference #239

ArthurFDLR started this conversation in **Show and tell**




https://github.com/openai/whisper/discussions/239

# Toonification of faces

From picture to toonified picture

From video to toonified video

https://huggingface.co/spaces/PKUWilliamYang/VToonify

https://github.com/williamyang1991/VToonify

License: S-Lab License 1.0 (non-commercial)

# Musika

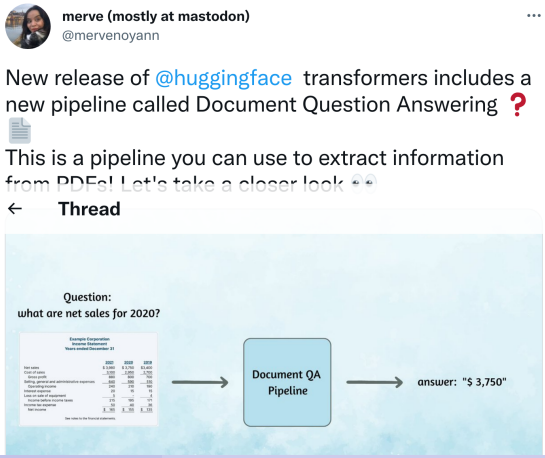
Blazingly Fast 44.1 kHz Stereo Waveform Music Generation of Arbitrary Length

https://arxiv.org/abs/2208.08706

https://huggingface.co/spaces/marcop/musika

License: MIT (open source)

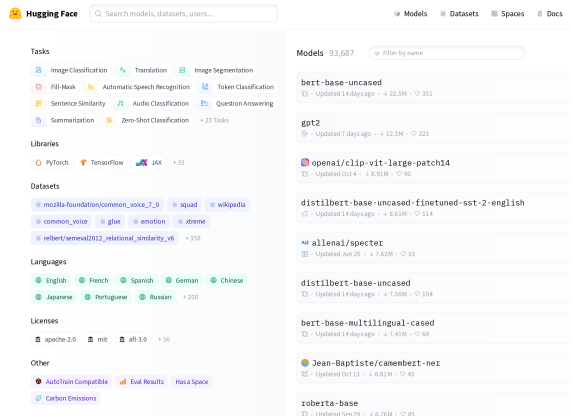
# Queries to documents




https://twitter.com/mervenoyann/status/1572168848622907393

# Infrastructure to play, to share

# Hugging Face




"GitHub for ML"

https://huggingface.co

# Gradio


https://gradio.app/

License: Apache 2.0

# Diffusers

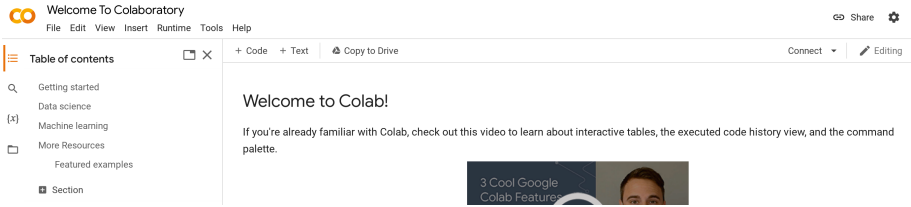

Pretrained diffusion models (vision, audio, etc.)

Modular toolbox for inference & training of diffusion models

https://github.com/huggingface/diffusers

License: Apache 2.0

# Collab




Python in the browser, zero configuration

Access to GPUs & easy sharing

https://colab.research.google.com/

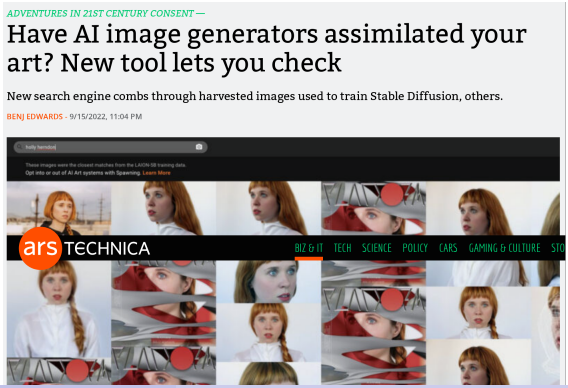
# Jupyter


Python in the browser, easy

https://jupyter.org/

# Many issues raised

# Intelectual property (training set)




https://haveibeentrained.com/

https://arstechnica.com/information-technology/2022/09/have-ai-image-generators-as

# Intelectual property (results)

Impact of Technology Deep Dive Report I

STUDY ON THE IMPACT OF ARTIFICIAL INTELLIGENCE ON THE INFRINGEMENT AND ENFORCEMENT OF COPYRIGHT AND DESIGNS

https://euipo.europa.eu/tunnel-web/secure/webdav/guest/document_library/observatory/documents/reports/2022_Impact_AI_on_the_Infringement_and_Enforcement_CR_Designs/2022_Impact_AI_on_the_Infringement_and_Enforcement_CR_Designs_FullR_en.pdf

# Licenses

| Model | Model License | Description | Link to License |
|---|---|---|---|
| GPT-2 | MIT License + generated output disclaimer | Permissive open source license | https://github.com/openai/gpt-2/blob/master/LICENSE |
| GPT-3 | Exclusive | Licensed to | Microsoft |
| YOLO | YOLO License | Public domain license | https://github.com/pjreddie/darknet/blob/master/LICENSE |
| DALLE-pytorch | MIT License | Pytorch implementation of DALLE created by individual researcher | https://github.com/lucidrains/DALLE-pytorch/blob/main/LICENSE |
| Stable Diffusion | CreativeML Open RAIL-M | Open & Responsible AI License (RAIL) created by Stability.ai and adapted from the BLOOM RAIL license, including use-based restrictions (see attachment A) | https://huggingface.co/spaces/CompVis/stable-diffusion-license |
| OPT | OPT-175B License | Meta restrictive license enabling use of the model weights for research purposes while establishing a set of use-based restrictions, which could be considered a RAIL | https://github.com/facebookresearch/metaseq/blob/main/projects/OPT/MODEL_LICENSE.md |
| BigScience | BigScience OpenRAIL-M | Open & Responsible AI License (RAIL) created by BigScience and adapted from the BLOOM RAIL license, including use-based restrictions (see attachment A) | https://huggingface.co/spaces/bigscience/license |
| Tsinghua University | GLM-130B license | Restrictive license enabling use of the model weights for research purposes | https://github.com/THUDM/GLM-130B/blob/main/MODEL_LICENSE |

https://hackmd.io/@jending12/HyvMU8sJo

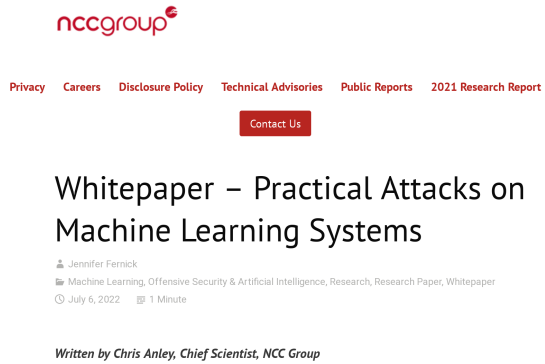
https://thegradient.pub/

machine-learning-ethics-and-open-source-li

# Bias


Mugshot of a technical speaker, machine learning expert, smiling, long hair, big eyes [t-shirt, curly hair]

# Security




https://research.nccgroup.com/2022/07/06/whitepaper-practical-attacks-on-machine-learning-systems/

https://simonwillison.net/2022/Sep/12/prompt-injection/

# Impact on professionals

# Prompt engineers

# What is true?

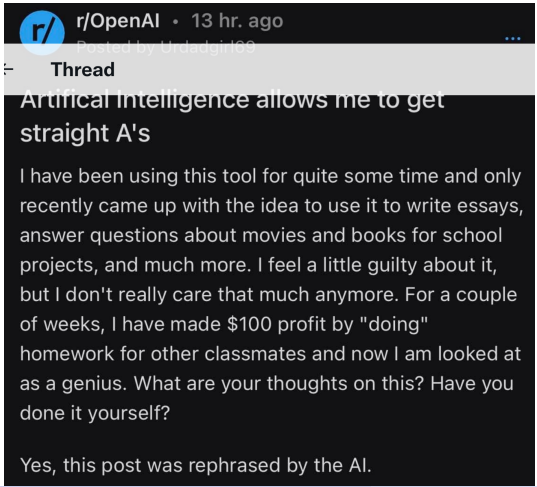
## Make-A-Video

Make-A-Video research builds on the recent progress made in text-to-image generation technology built to enable text-to-video generation. The system uses images with descriptions to learn what the world looks like and how it is often described. It also uses unlabeled videos to learn how the world moves. With this data, Make-A-Video lets you bring your imagination to life by generating whimsical, one-of-a-kind videos with just a few words or lines of text.


https://makeavideo.studio/

# The future

# Assignments



# Programming


https://arxiv.org/abs/2207.14502

# Programming

**Self-Programming Artificial Intelligence Using Code-Generating Language Models**

*Anonymous*

22 Sept 2022 (modified: 26 Oct 2022) ICLR 2023 Conference Blind Submission Readers: Everyone Show Bibtex Show Revisions

**Keywords:** Self-programming AI, NLP, code generation, AutoML

**TL;DR:** We develop and experimentally validate the first practical implementation of a self-reprogramming AI system.

**Abstract:** Recent progress in large-scale language models has enabled breakthroughs in previously intractable computer programming tasks. Prior work in meta-learning and neural architecture search has led to substantial successes across various task domains, spawning myriad approaches for algorithmically optimizing the design and learning dynamics of deep learning models. At the intersection of these research areas, we implement a code-generating language model with the ability to modify its own source code. Self-programming AI algorithms have been of interest since the dawn of AI itself. Although various theoretical formulations of generalized self-programming AI have been posed, no such system has been successfully implemented to date under real-world computational constraints. Applying AI-based code generation to AI itself, we develop and experimentally validate the first practical implementation of a self-programming AI system. We empirically show that a self-programming AI implemented using a code generation model can successfully modify its own source code to improve performance and program sub-models to perform auxiliary tasks. Our model can self-modify various properties including model architecture, computational capacity, and learning dynamics.

https://keras.io/examples/generative/random_walks_with_stable_diffusion/

# Summarizing

# The future just started

# References, credits, license

- Transformers-Tutorials

  https://github.com/NielsRogge/Transformers-Tutorials

- Vision Transformers

  https://cameronrwolfe.substack.com/p/vision-transformers

- A walk through latent space with Stable Diffusion

  https://keras.io/examples/generative/random_walks_with_stable_diffusion/

- How Open Source is eating AI

  https://lspace.swyx.io/p/open-source-ai

- Awesome Diffusion Models

  https://github.com/heejkoo/Awesome-Diffusion-Models

# Credits

Book, by NikolayFrolochkin, Pixabay.
License: Creative Commons CC0