CULTURAL AUTOMATION WITH MACHINE LEARNING

SESSION 07: GENERATIVE VIDEO

Parag K. Mital
UCLA DMA

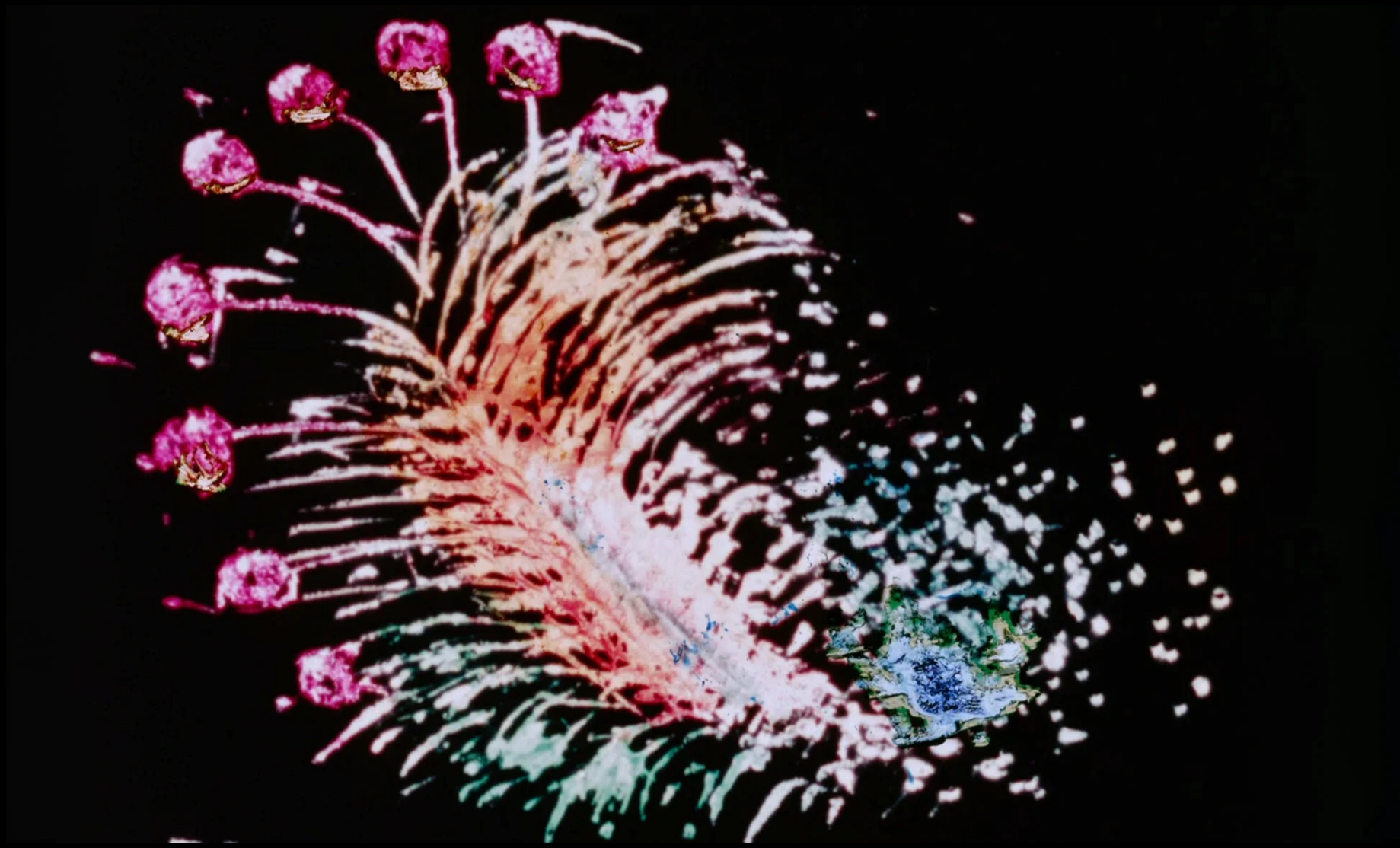
CULTURAL CONTEXT + HISTORY OF GENERATIVE VIDEO

Oskar Fischinger (1930's -40's)

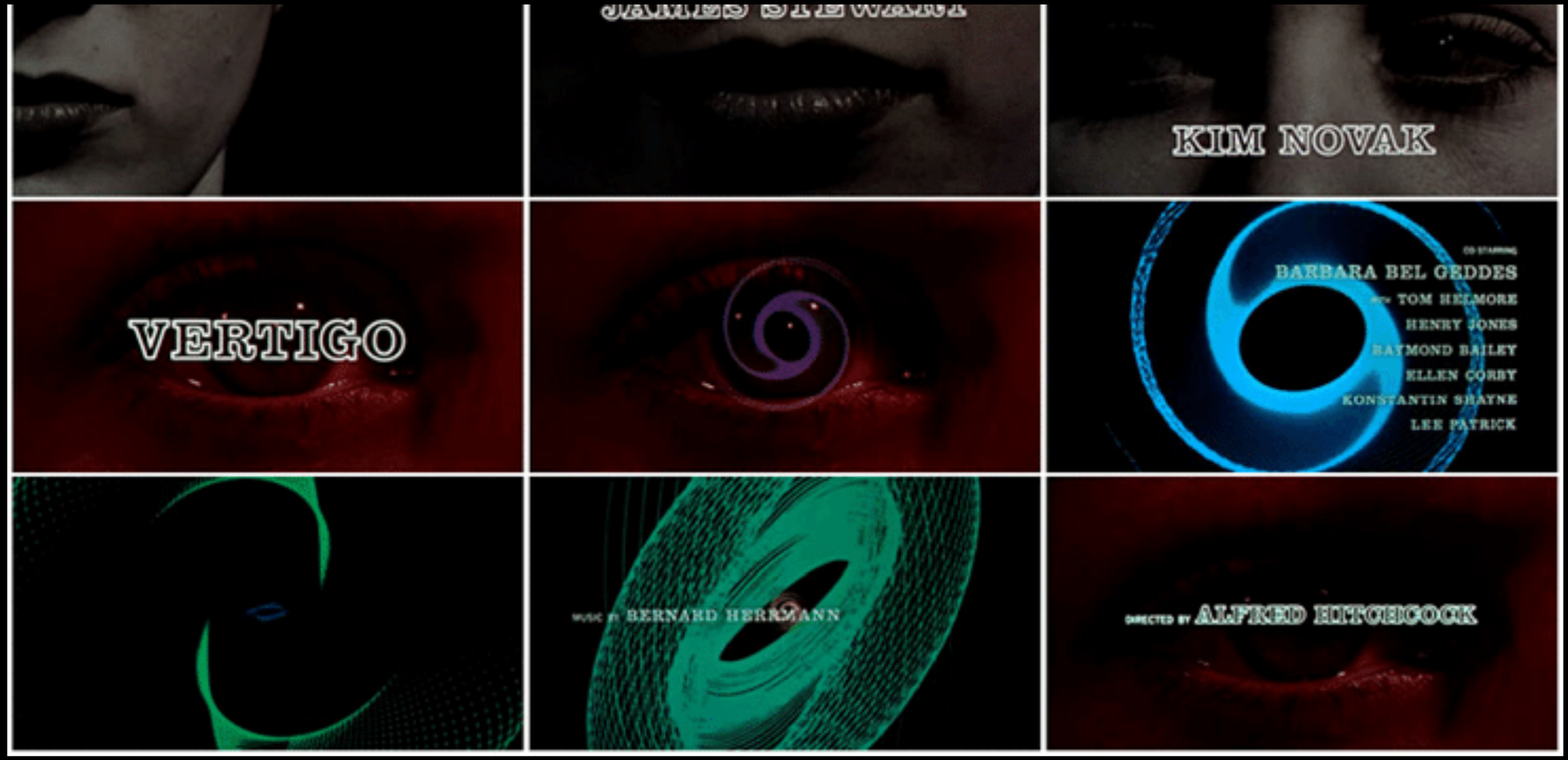
https://www.acmi.net.au/education/school-program-and-resources/the-art-of-visual-music/

https://www.google.com/logos/2017/fischinger/fischinger17.html

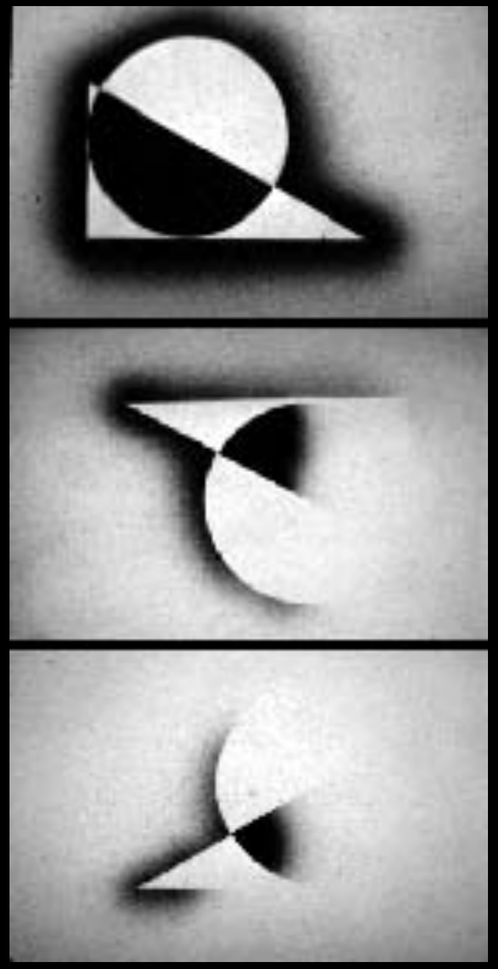
Mary Ellen Bute (1930’s - 40’s)

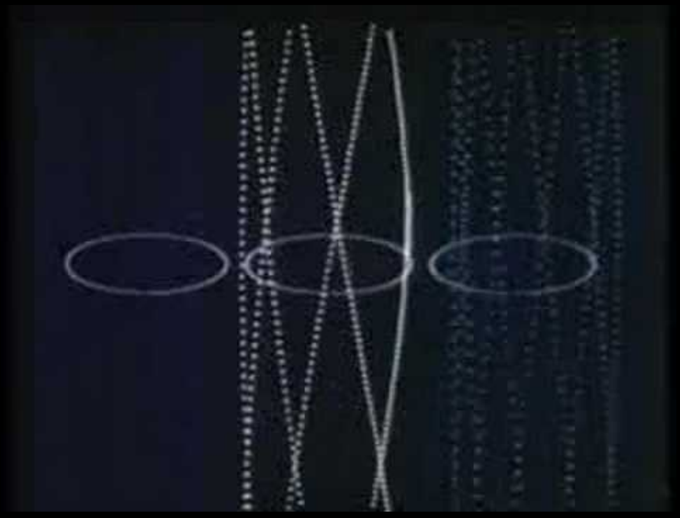


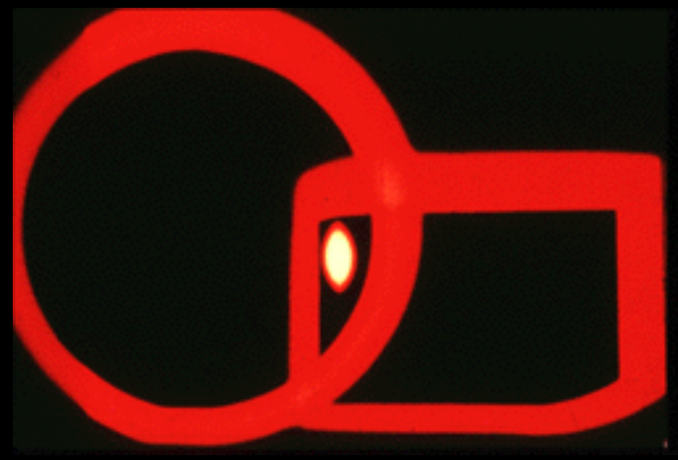
Norman McLaren (1940’s -50’s)

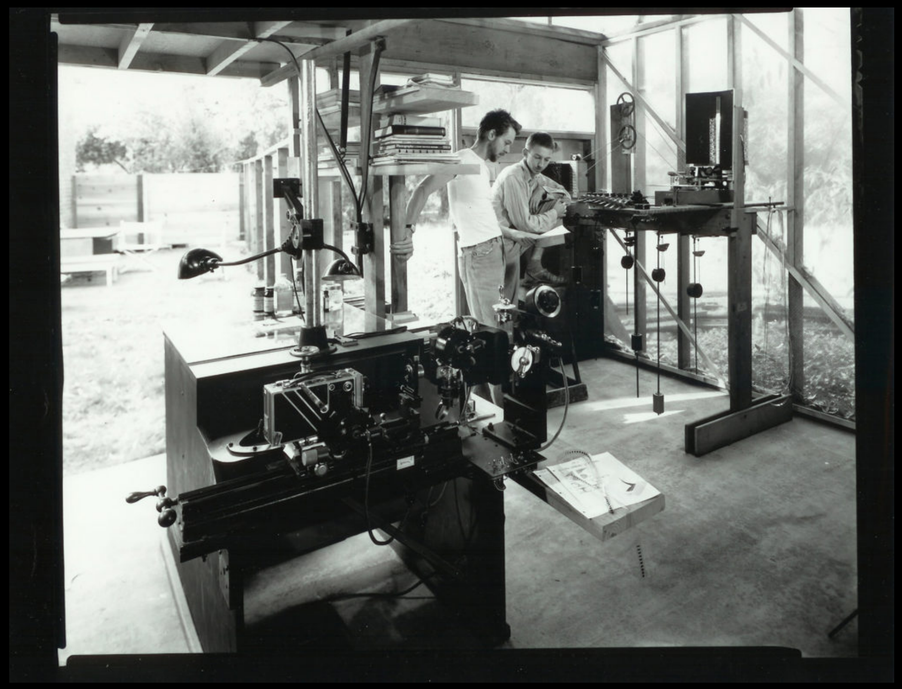


John Whitney Sr. + Saul Bass “Vertigo” (1958)

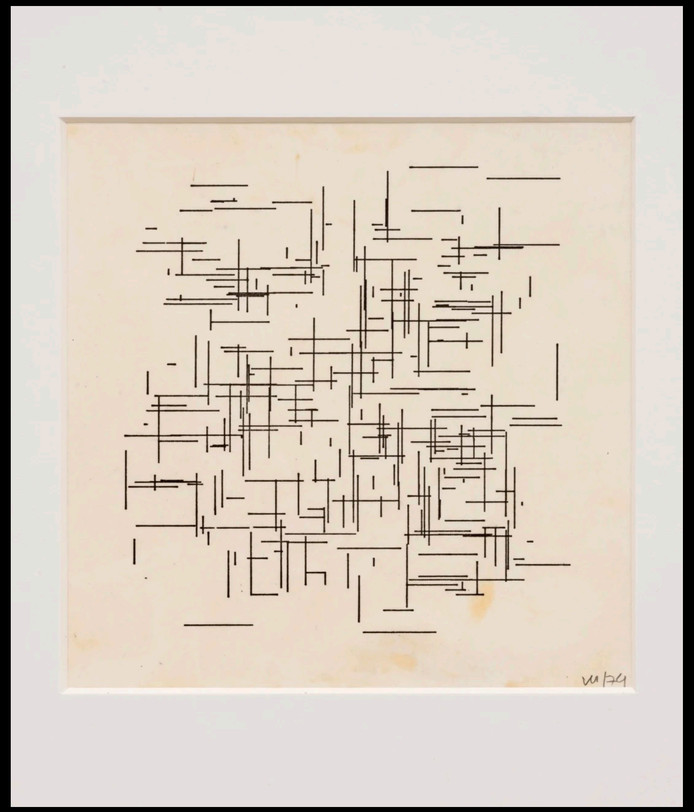


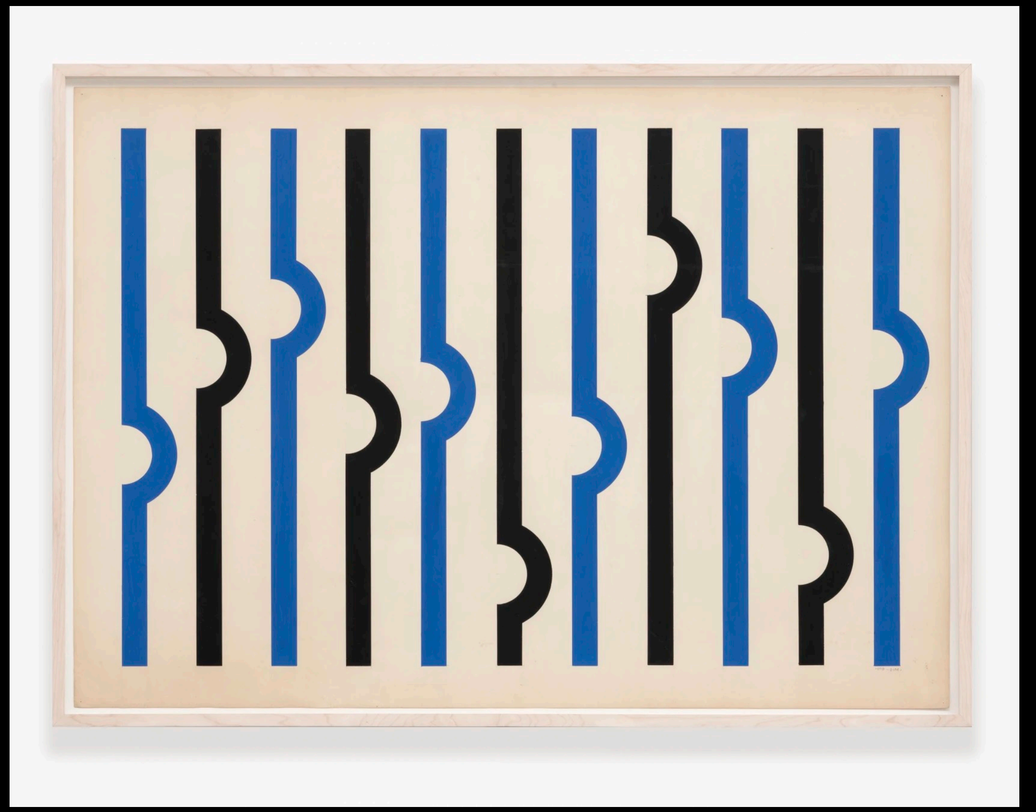




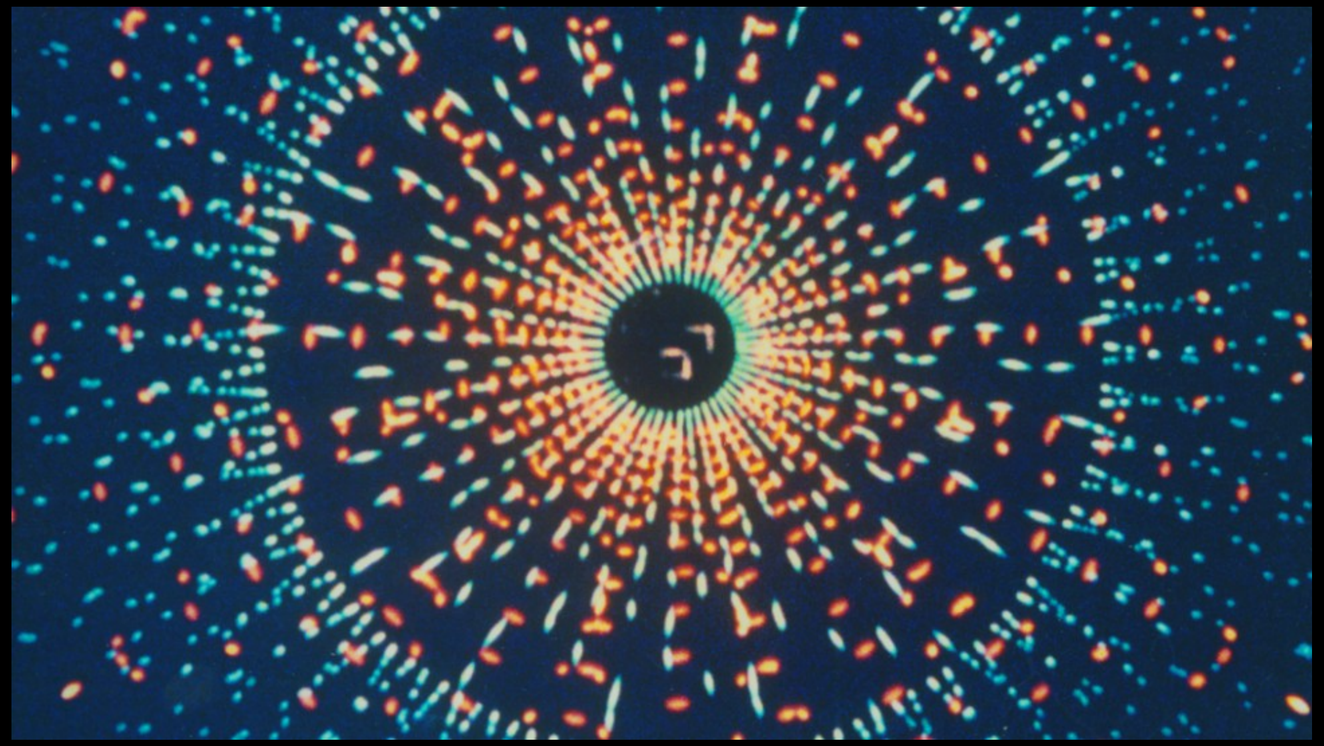


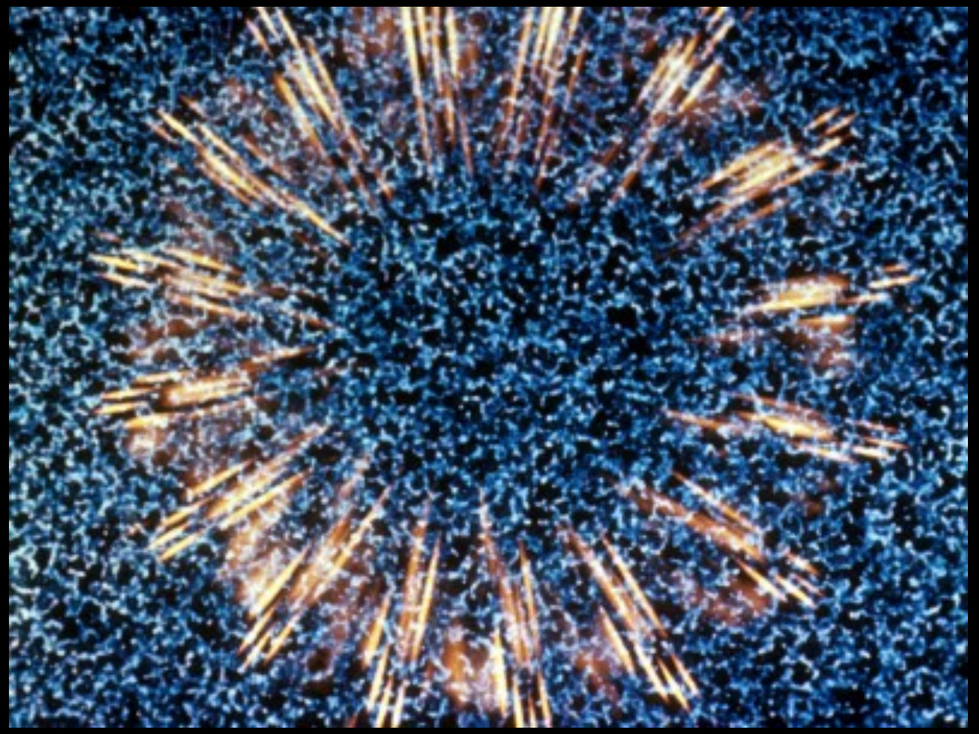
John Whitney (1950’s -60’s)





Vera Molnar (1960’s)

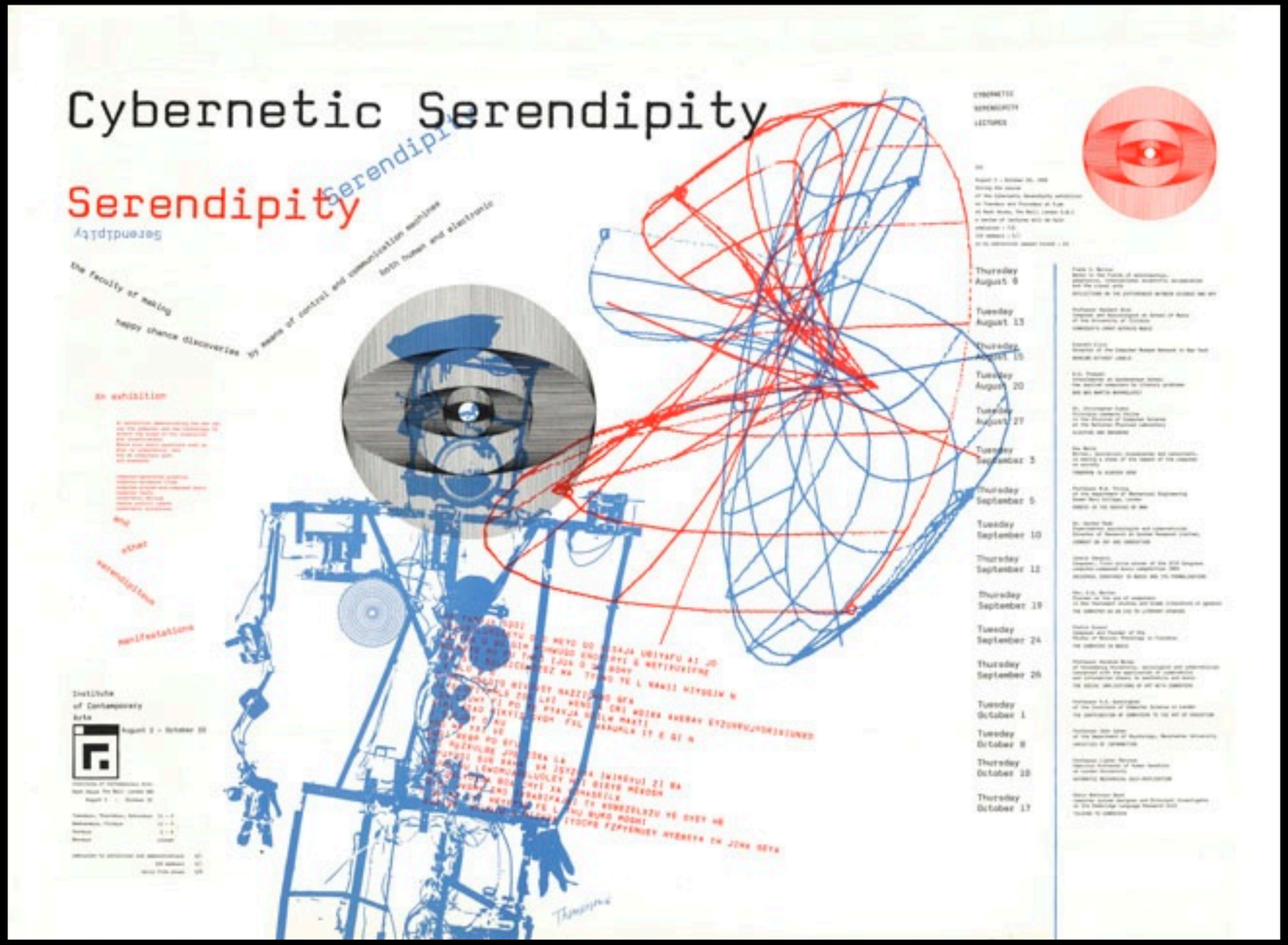




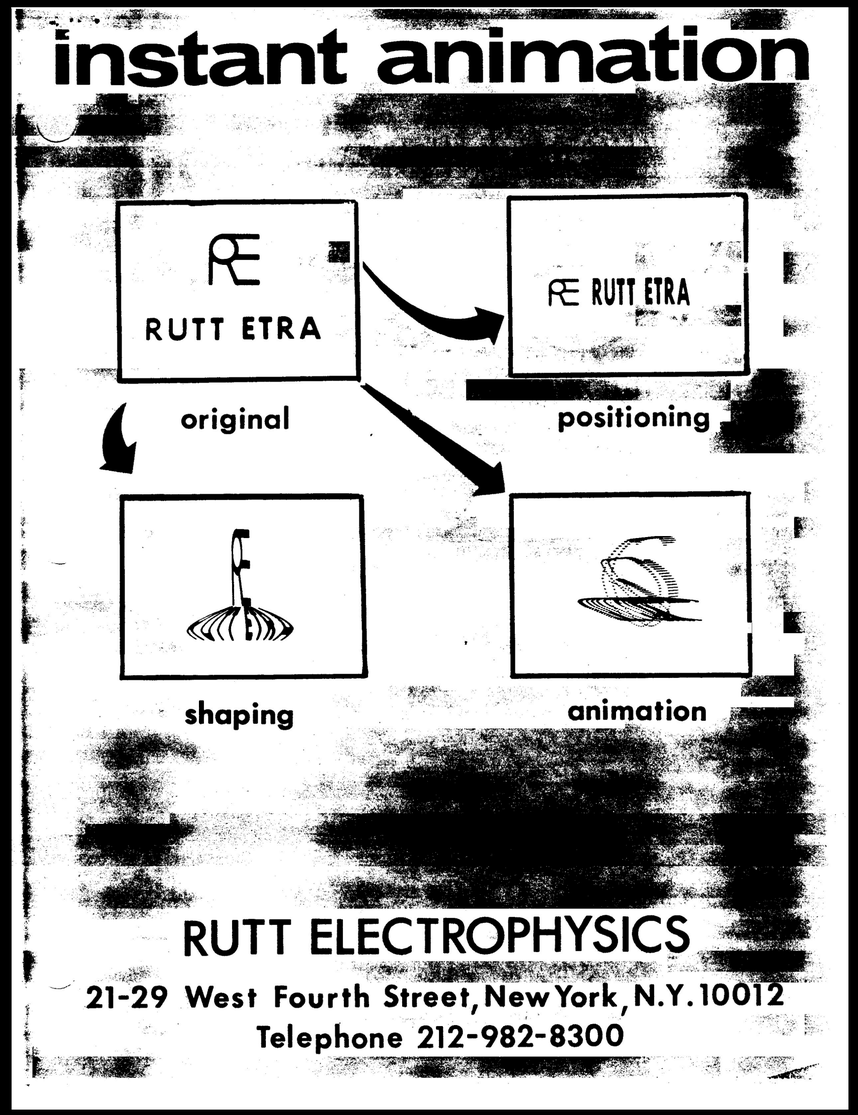
Jordan Belson (1960’s -70’s)

https://www.youtube.com/watch?v=oQMD7oufO4s

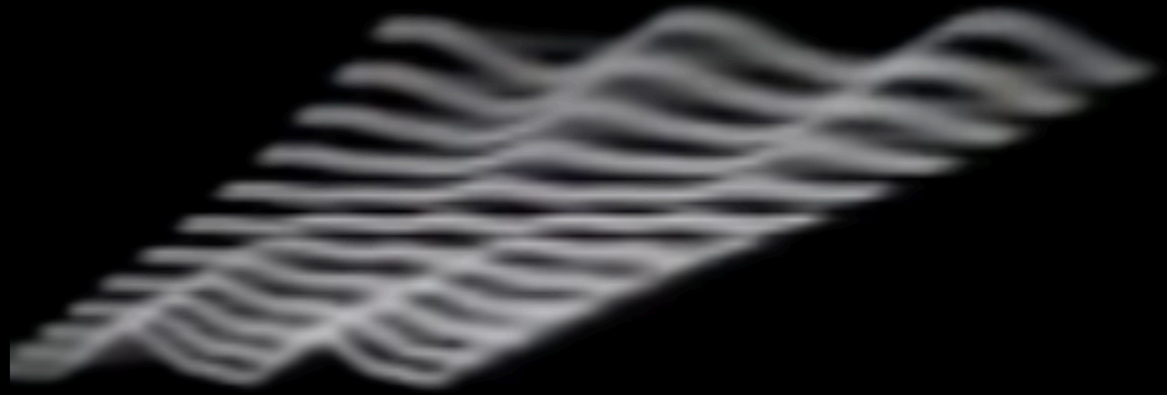
https://www.youtube.com/watch?v=sebqkS_EKBE



https://www.youtube.com/watch?v=oSwovB28B34



https://www.vasulka.org/archive/Artists2/Etra,Bill/general.pdf



https://www.youtube.com/watch?v=EZQiHuTbnes&t=138s

Rutt-Etra-Izer

This demo emulates the Rutt-Etra video synthesizer.
More info.

Controls

- Drop an image, or load a sample image.
- Click and drag to rotate.
- Mousewheel to scale.
- 'S' key to save image.

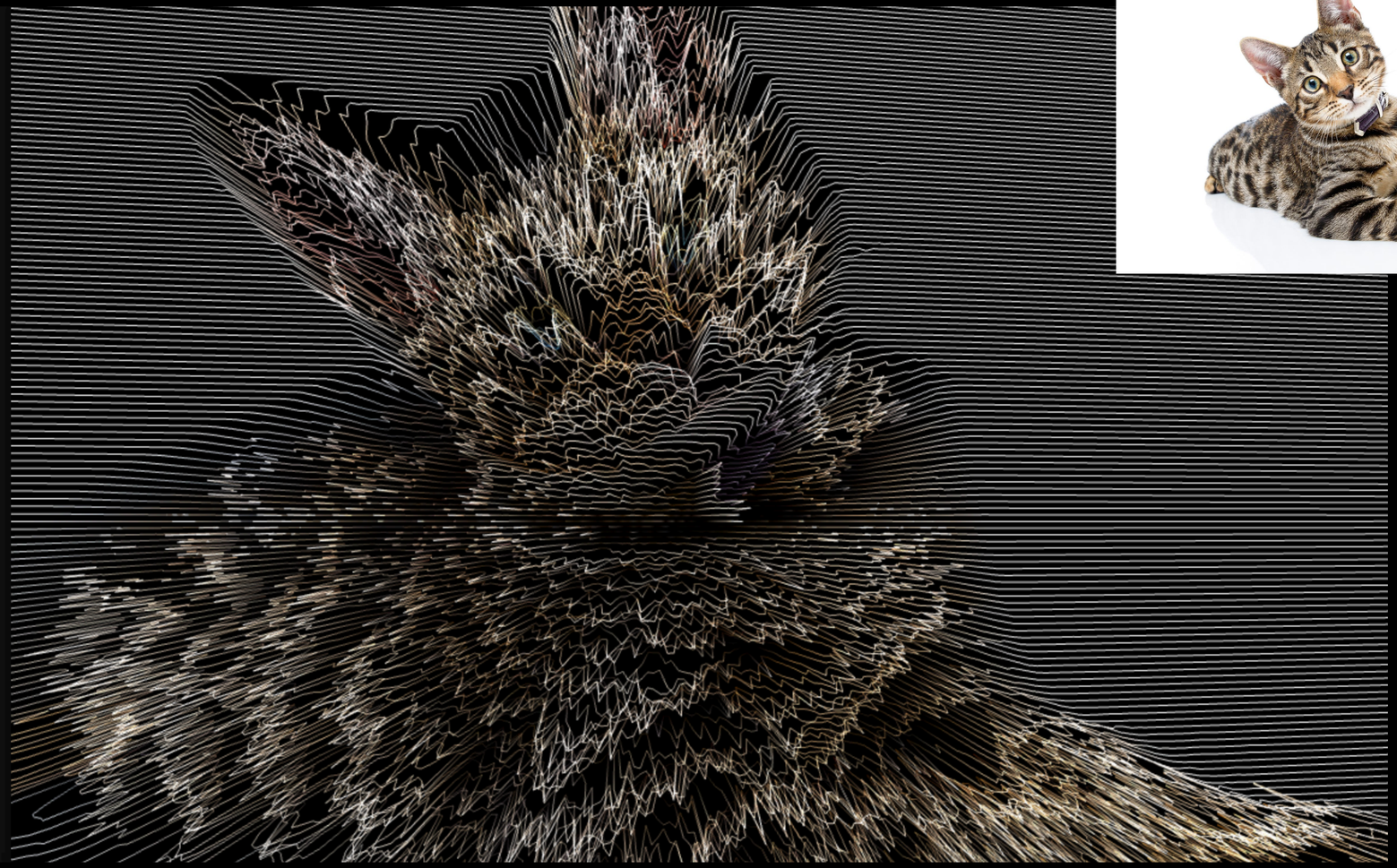
Options

Scale	2
Line Separation	9
Line Thickness	9.9
Max Line Depth	153
Brightness	0.9
Stage Size	1
Load Image	
Save Image	

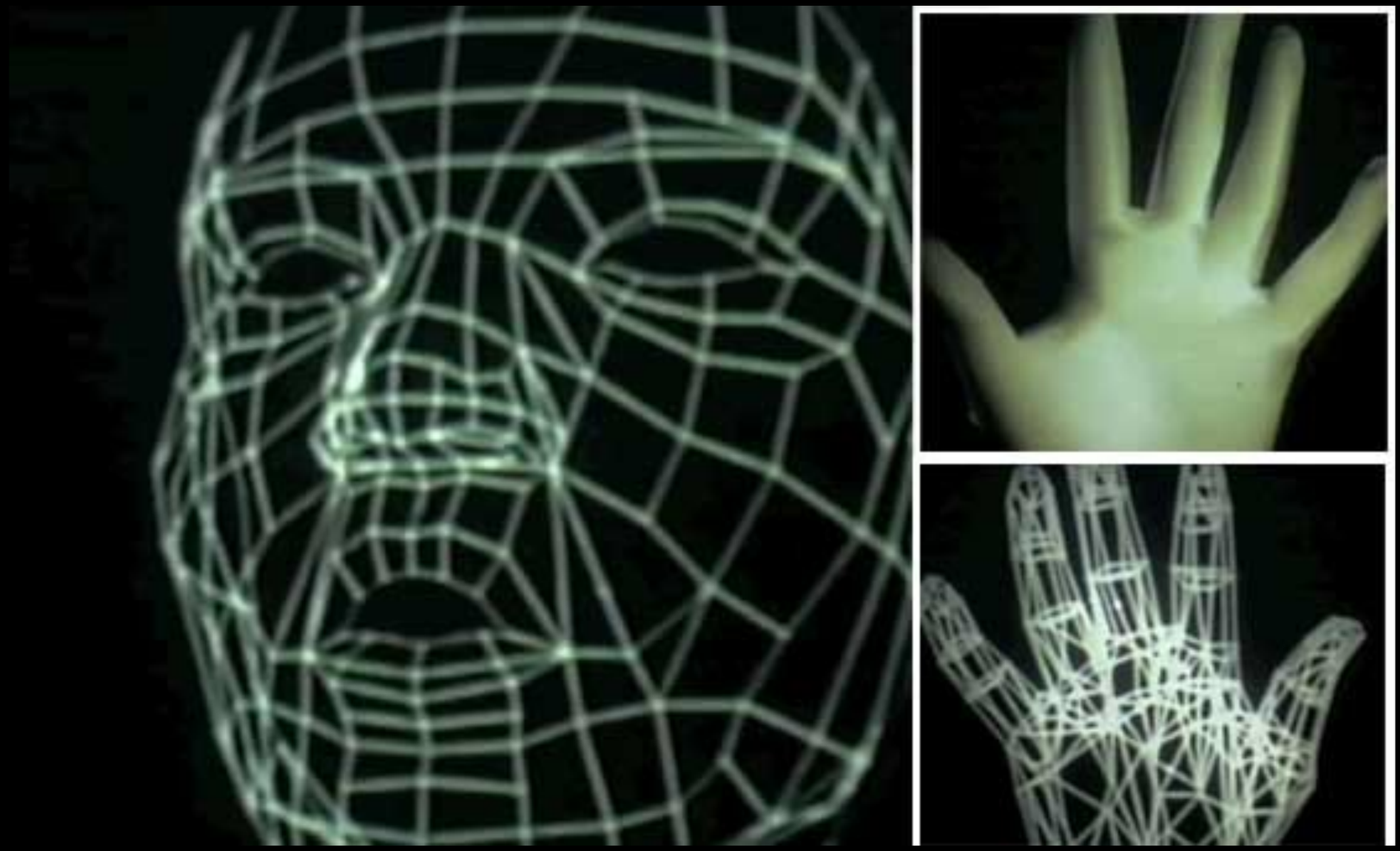
Close Controls

FPS

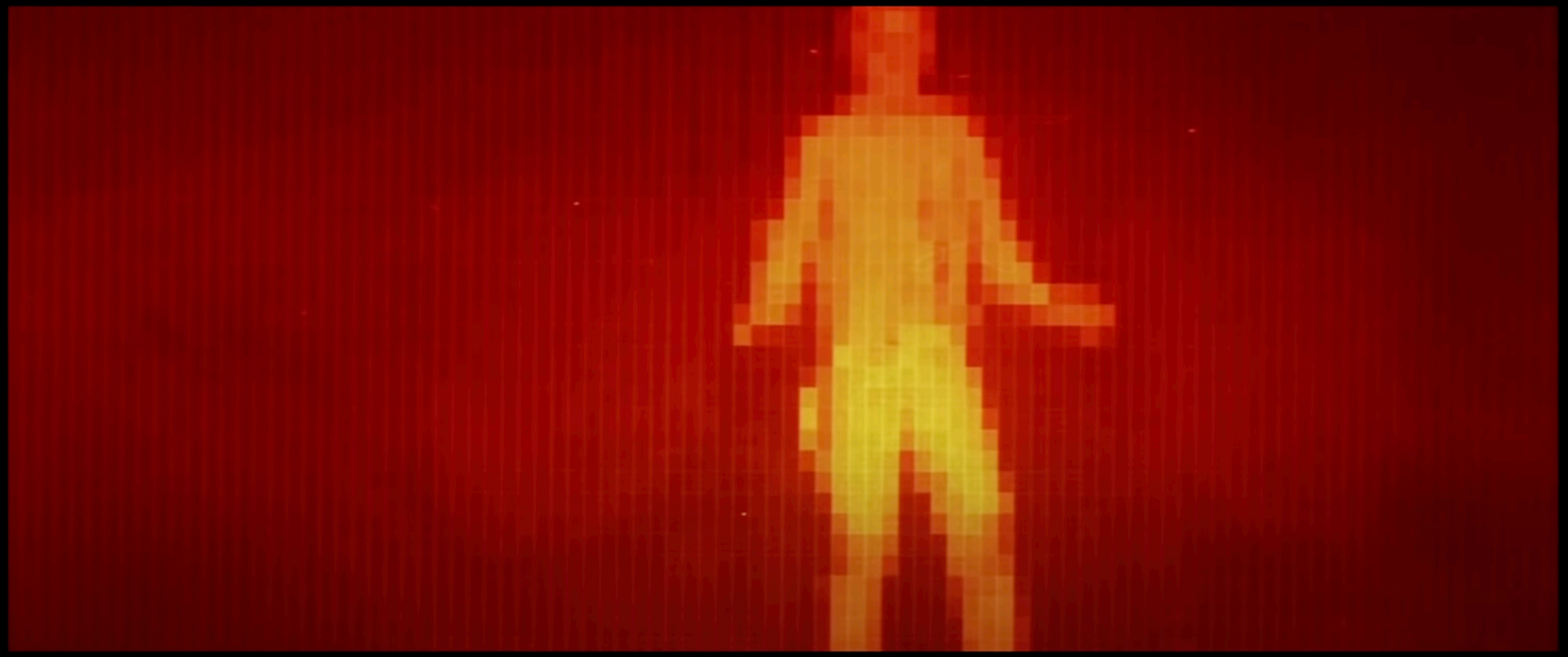
60 FPS (52-62)



https://airtightinteractive.com/demos/js/ruttetra/



Ed Catmull (1972)



https://www.youtube.com/watch?v=nILKFlpOZi0

ILM (1975)

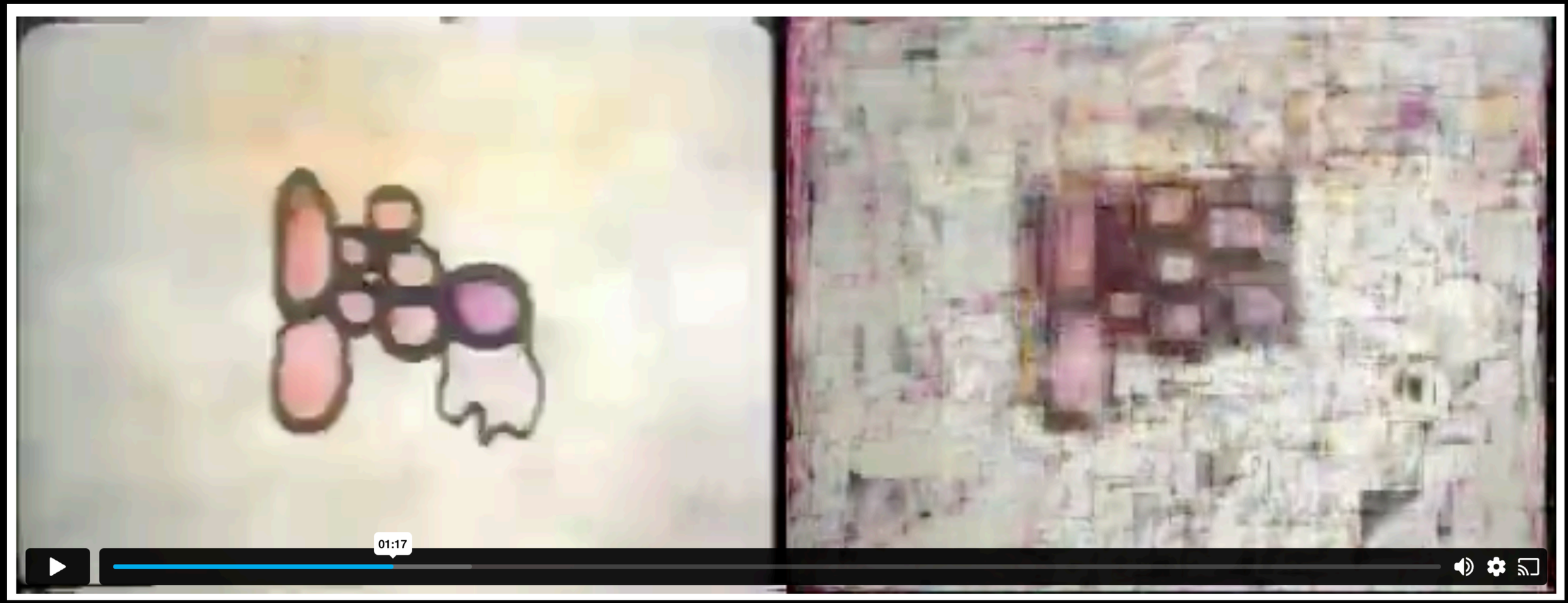
https://www.youtube.com/watch?v=578Xm6bgMdQ

https://vimeo.com/212875761

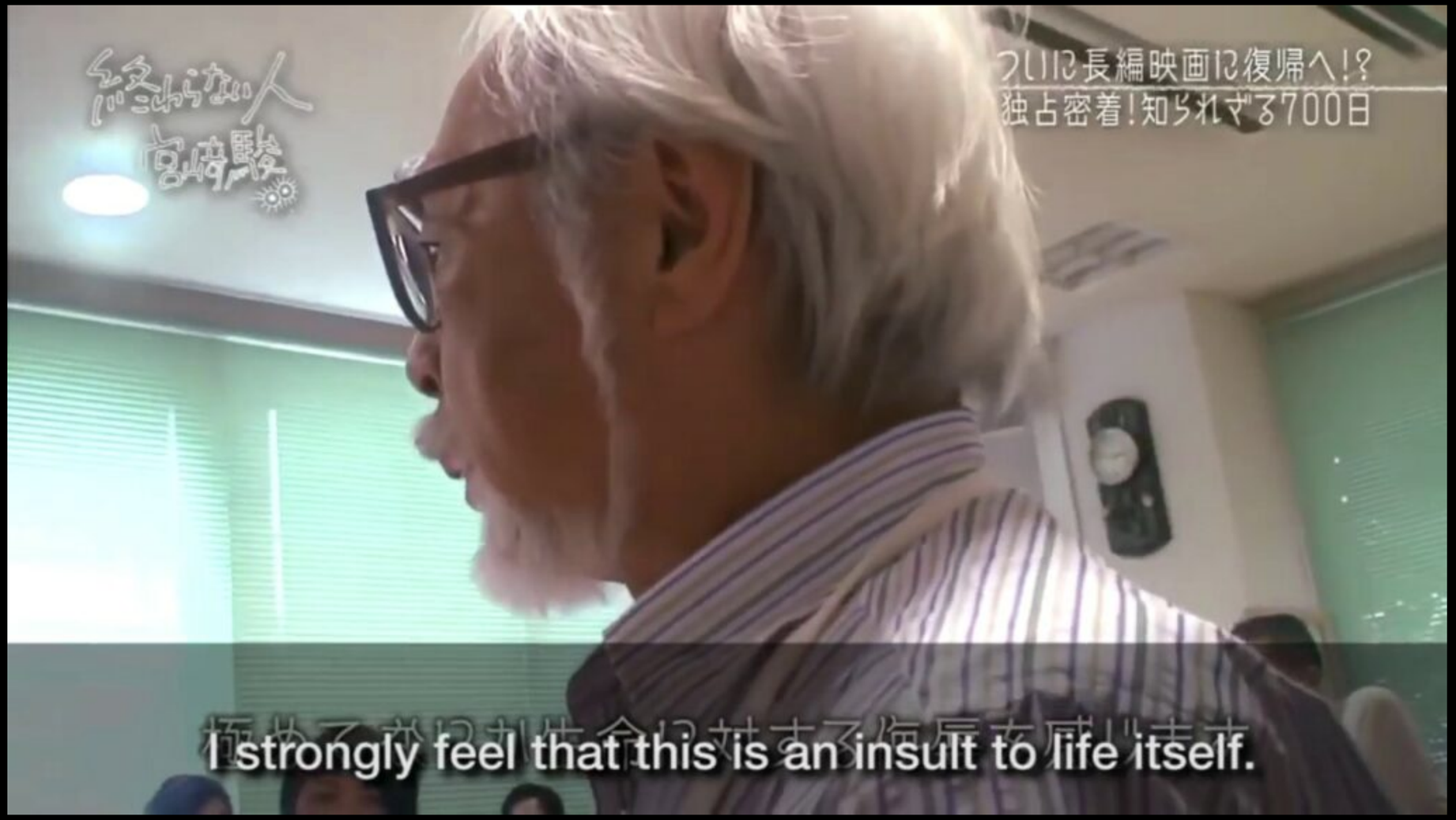
Toy Story (1995)

RJD2 - 1976

https://www.youtube.com/watch?v=Lp-xN-o6W90



https://pkmital.com/home/works/early-visual-synthesis-works/



https://www.youtube.com/watch?v=ngZ0K3lWKRc

"Sunspring" https://www.youtube.com/watch?v=LY7x2Ihqjmc

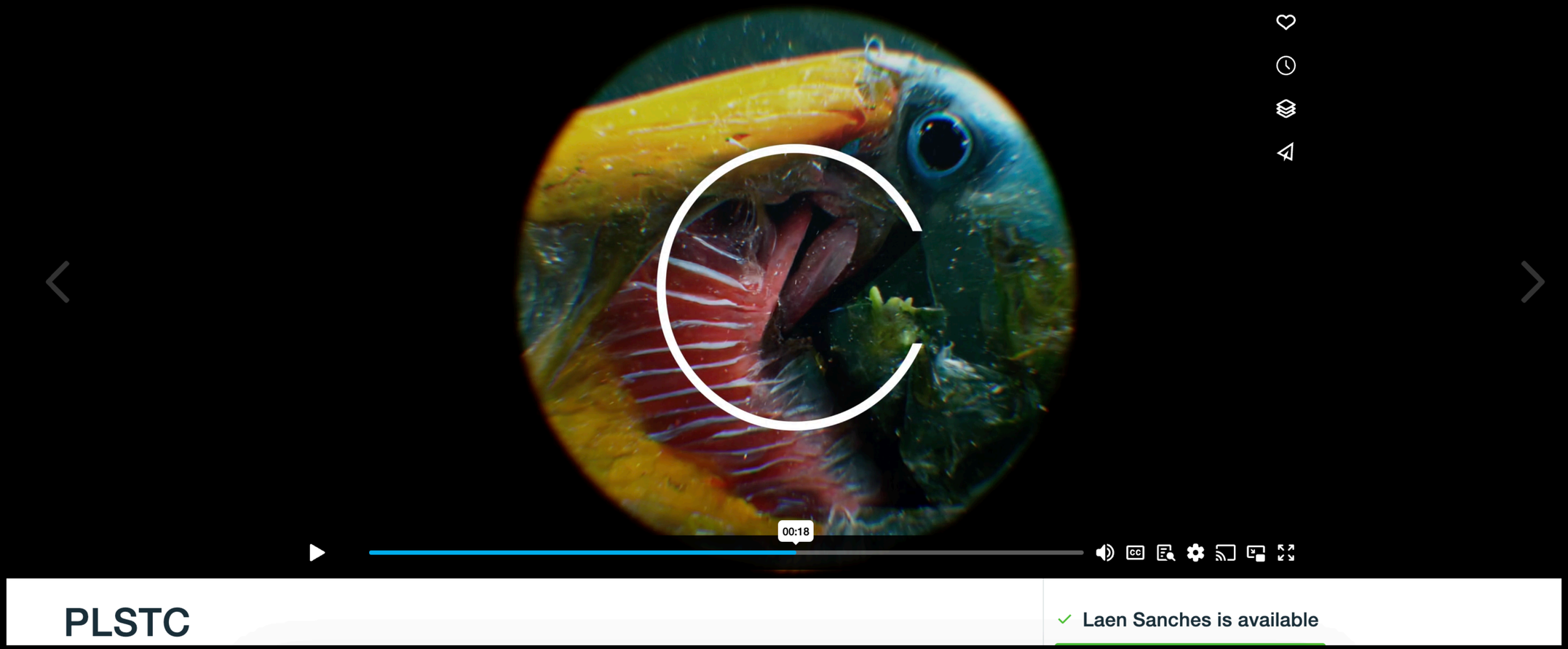
https://www.technologyreview.com/2023/06/01/1073858/surreal-ai-generative-video-changing-film/

THANK YOU FOR NOT ANSWERING

Paul Trillo is available

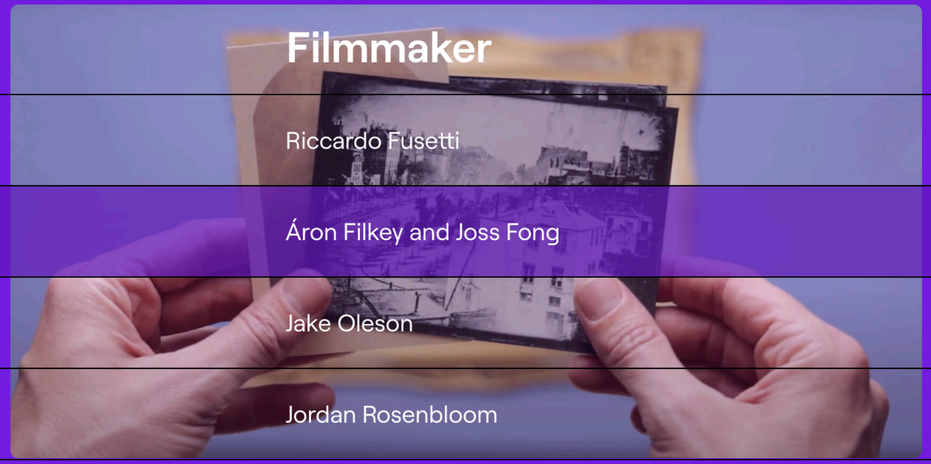
https://vimeo.com/821101511

https://www.newyorker.com/culture/screening-room/an-ai-generated-film-depicts-human-loneliness-in-thank-you-for-not-answering



https://vimeo.com/723429393

Title	Filmmaker	Award
Generation	Riccardo Fusetti	Grand Prix
Checkpoint	Áron Filkey and Joss Fong	Gold
Given Again	Jake Oleson	Silver
Original Voice	Jordan Rosenbloom	Honoree
PLSTC	Laen Sanches	Honoree
Landscape	Kyle Goodrich	Merit
I want 1000 Rabbits	Shan He	Merit
AI ARTIST	Jessey Nelson and Dani Barker	Merit



https://aiff.runwayml.com/2023#winners

The 2nd Annual International AI FILM FESTIVAL

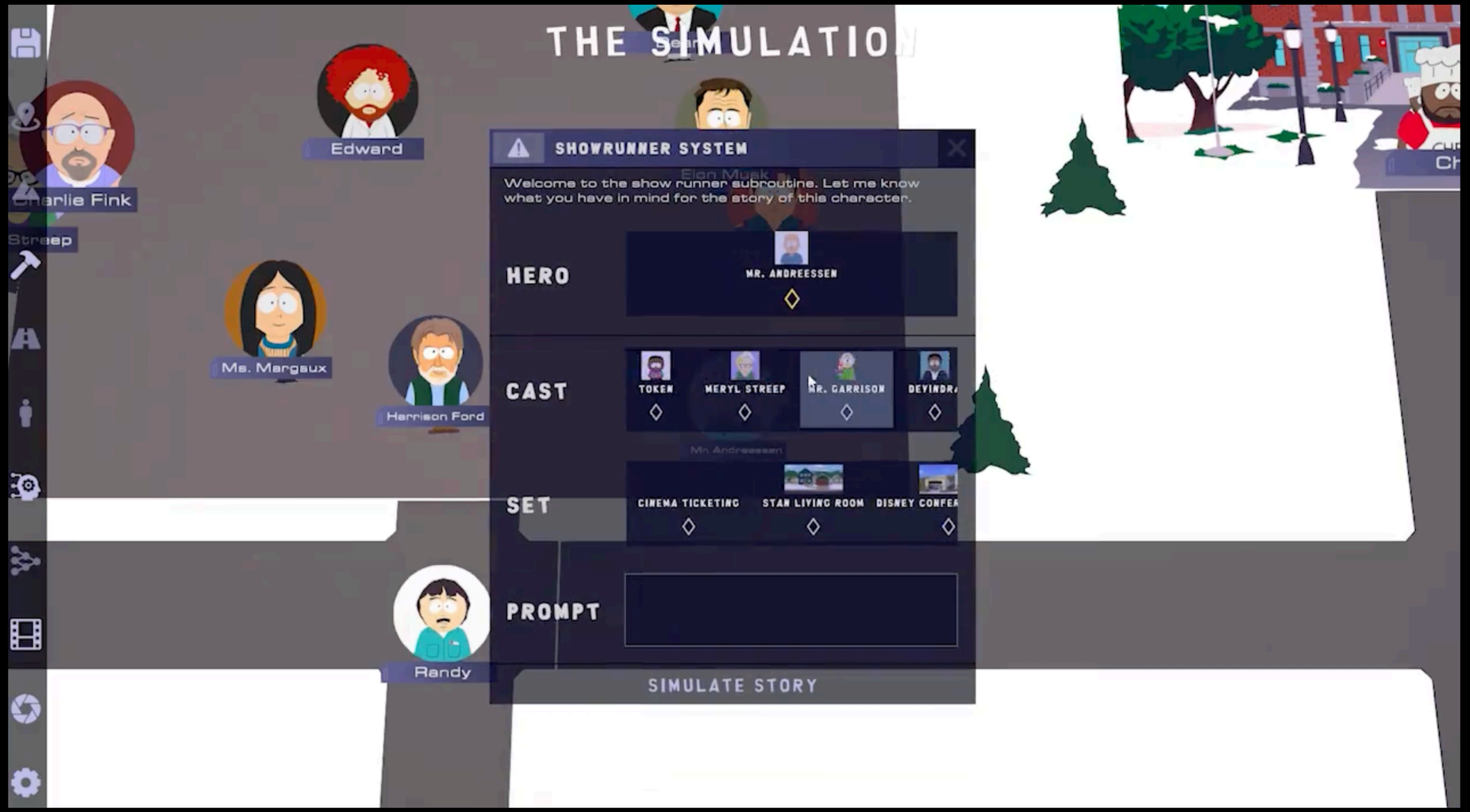
FAQ AIFF 2023 ↗ SUBMIT YOUR FILM

PRESENTING PARTNERS NVIDIA runway

AIFF 2023 Grand Prix

"Generation"

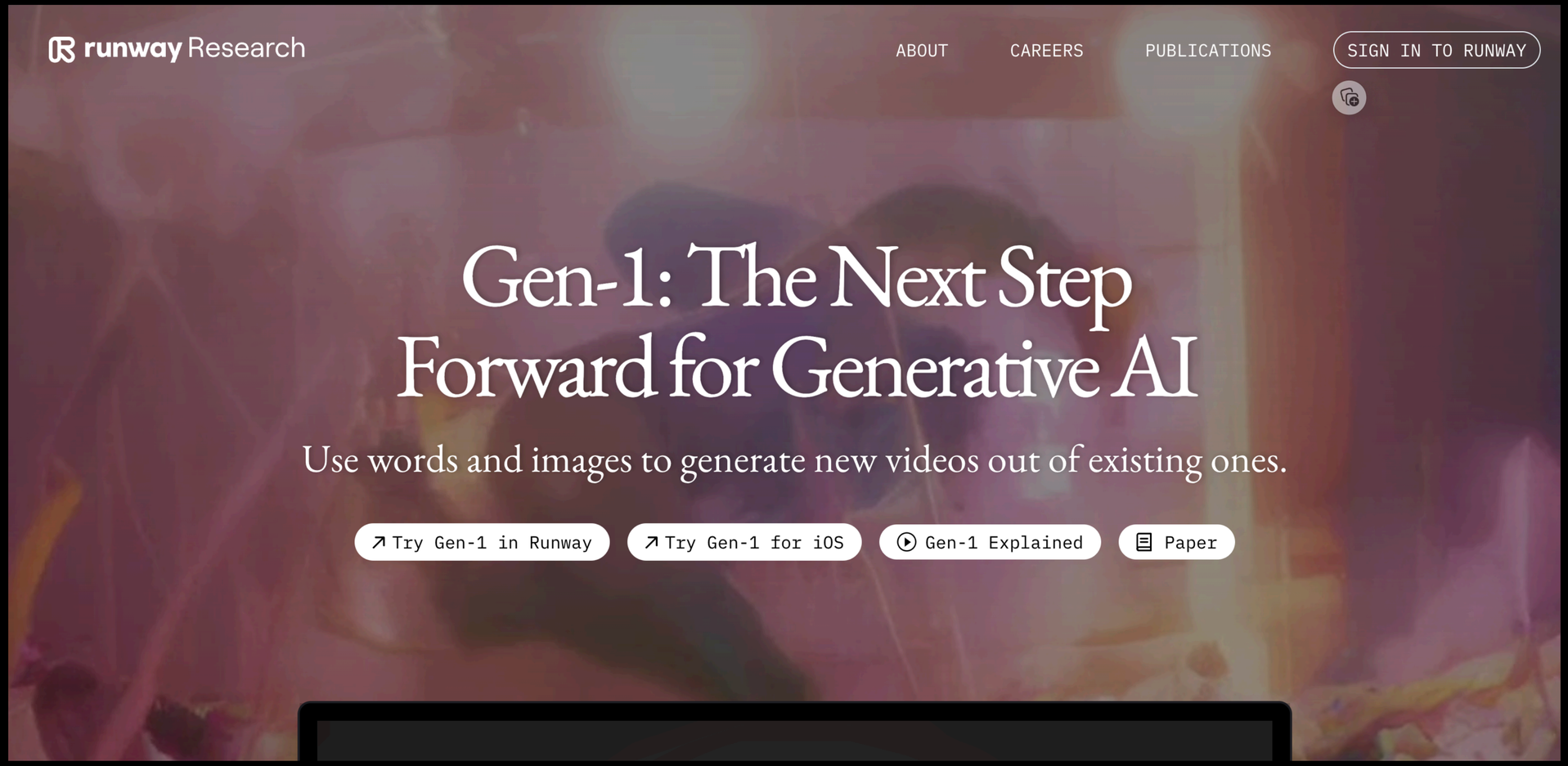
https://aiff.runwayml.com/



"The Simulation" - SouthPark - https://twitter.com/fablesimulation/status/1681352904152850437?lang=en

LAB

TEXT TO VIDEO
IMAGE TO VIDEO
VIDEO TO VIDEO



https://research.runwayml.com/gen1

Computer Science > Computer Vision and Pattern Recognition

[Submitted on 29 May 2022]

CogVideo: Large-scale Pretraining for Text-to-Video Generation via Transformers

Wenyi Hong, Ming Ding, Wendi Zheng, Xinghan Liu, Jie Tang

Large-scale pretrained transformers have created milestones in text (GPT-3) and text-to-image (DALL-E and CogView) generation. Its application to video generation is still facing many challenges: The potential huge computation cost makes the training from scratch unaffordable; The scarcity and weak relevance of text-video datasets hinder the model understanding complex movement semantics. In this work, we present 9B-parameter transformer CogVideo, trained by inheriting a pretrained text-to-image model, CogView2. We also propose multi-frame-rate hierarchical training strategy to better align text and video clips. As (probably) the first open-source large-scale pretrained text-to-video model, CogVideo outperforms all publicly available models at a large margin in machine and human evaluations.

Subjects: **Computer Vision and Pattern Recognition (cs.CV)**; Computation and Language (cs.CL); Machine Learning (cs.LG)
Cite as: arXiv:2205.15868 **[cs.CV]**
(or arXiv:2205.15868v1 **[cs.CV]** for this version)
https://doi.org/10.48550/arXiv.2205.15868

Submission history

From: Ming Ding [view email]
[v1] Sun, 29 May 2022 19:02:15 UTC (37,711 KB)

Bibliographic Tools | Code, Data, Media | Demos | Related Papers | About arXivLabs

https://arxiv.org/abs/2205.15868

Text to Video Samples - CogVideo by Tsinghua University

https://www.youtube.com/watch?v=90-XgkXcP3o

THUDM / CogVideo Public

Notifications | Fork 330 | Star 3.2k

Code | Issues 9 | Pull requests 2 | Actions | Projects | Security | Insights

main | 1 branch | 0 tags | Go to file | Code

wenyihong Update requirements.txt 9000d1b on Jun 9 24 commits

assets	change to png	last year
models	Release code	last year
paper	add paper	last year
scripts	Release code	last year
sr_pipeline	Release code	last year
.gitignore	Release code	last year
Dockerfile	updated docker file and added script to instsall needed package pos...	last year
LICENSE	Create LICENSE	last year
Model_License	Create Model_License	10 months ago
README.md	Update demo url	4 months ago
build_image.sh	added Docker support	last year

About

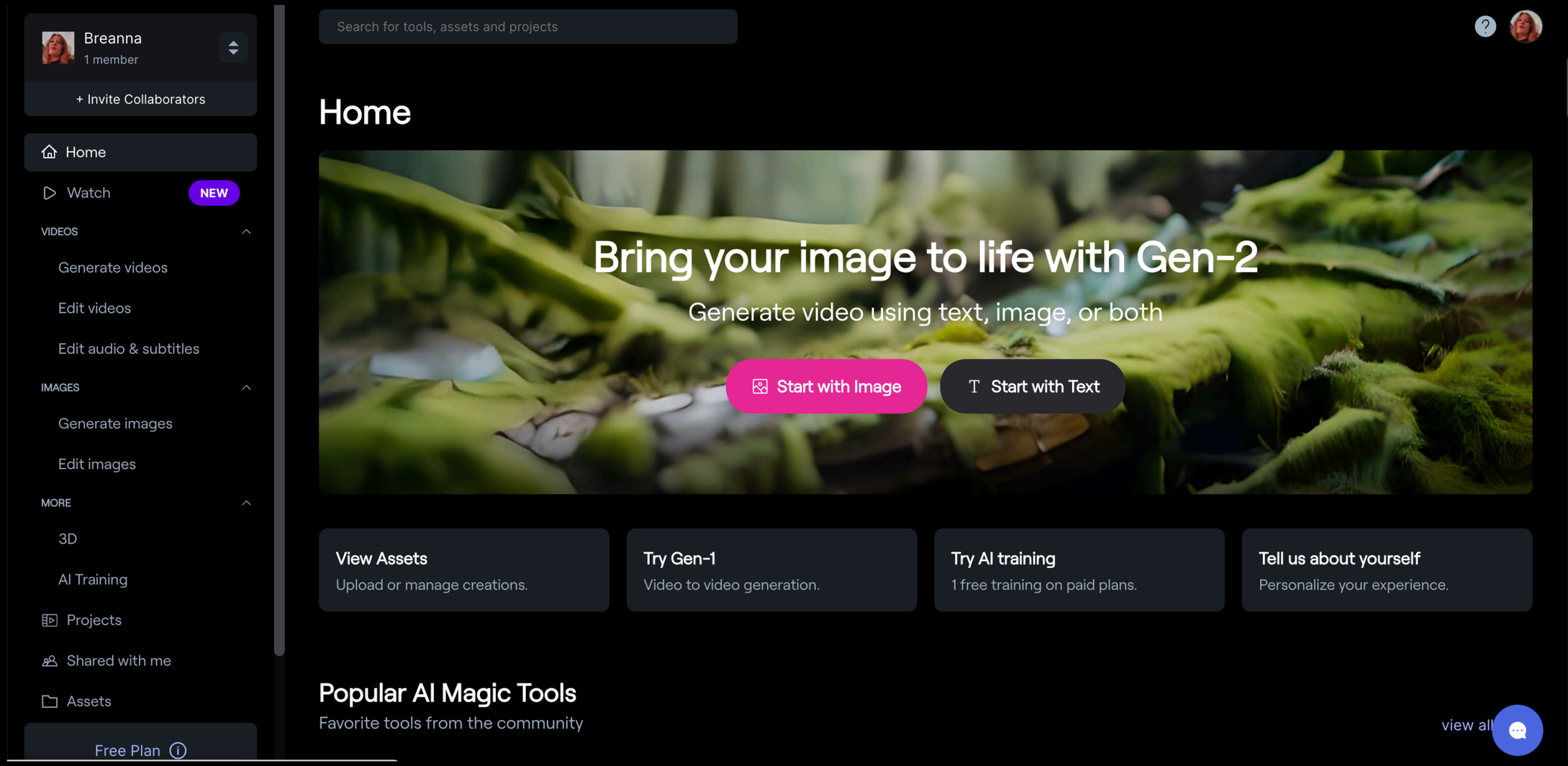
Text-to-video generation. The repo for ICLR2023 paper "CogVideo: Large-scale Pretraining for Text-to-Video Generation via Transformers"

Readme
Apache-2.0 license
Activity
3.2k stars
99 watching
330 forks
Report repository

Releases

No releases published

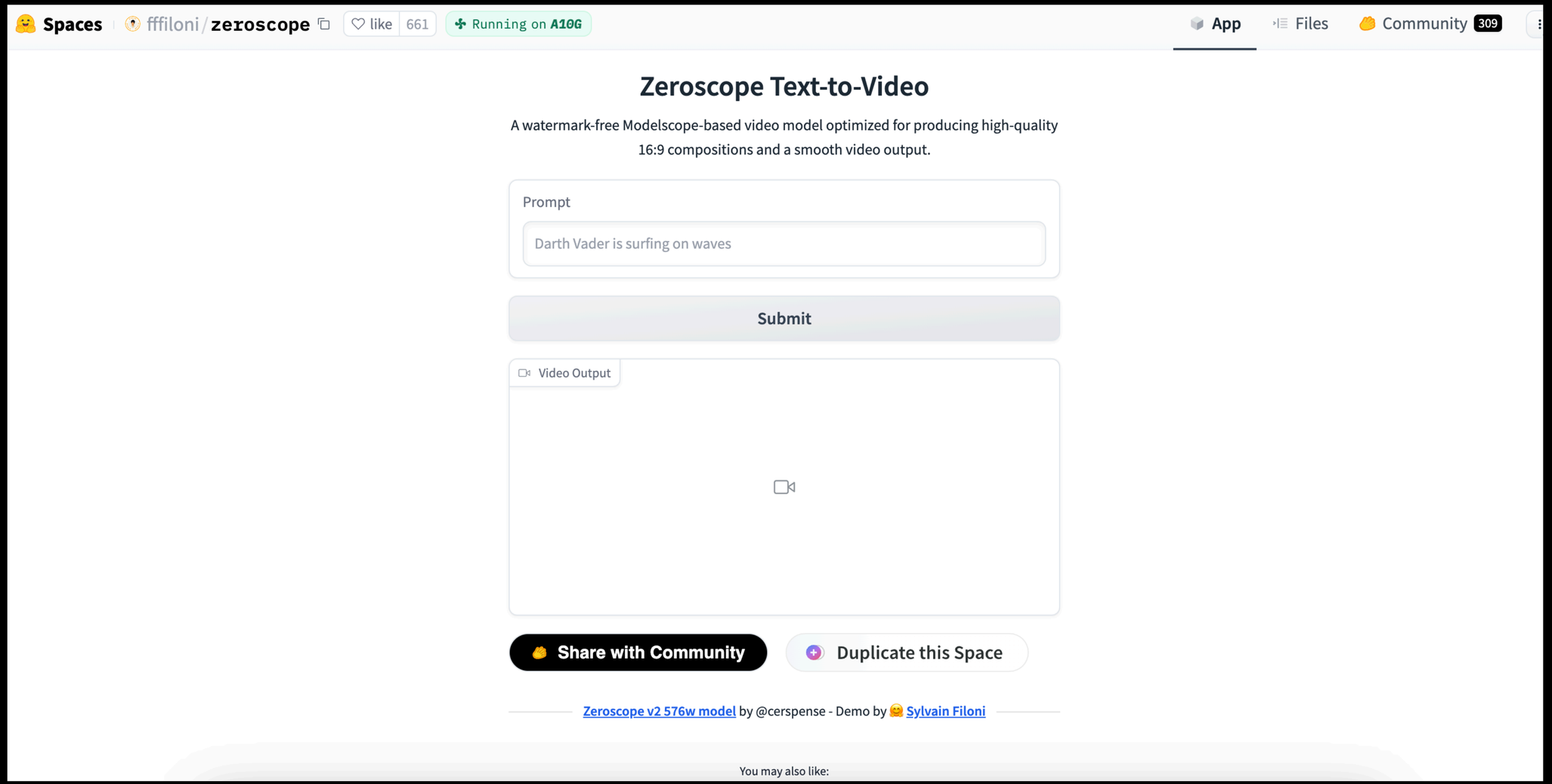
https://github.com/THUDM/CogVideo



https://runwayml.com/ai-magic-tools/gen-2/
https://app.runwayml.com/

“cat is in a magical forest looking around at the surroundings”

"Waves at the beach crashing onto the shore cinematic"



https://huggingface.co/spaces/fffiloni/zeroscope

“Big wave surfer cinematic”

alembics / disco-diffusion Public

Code | Issues 54 | Pull requests 17 | Actions | Projects | Security | Insights

main | 27 branches | 1 tag

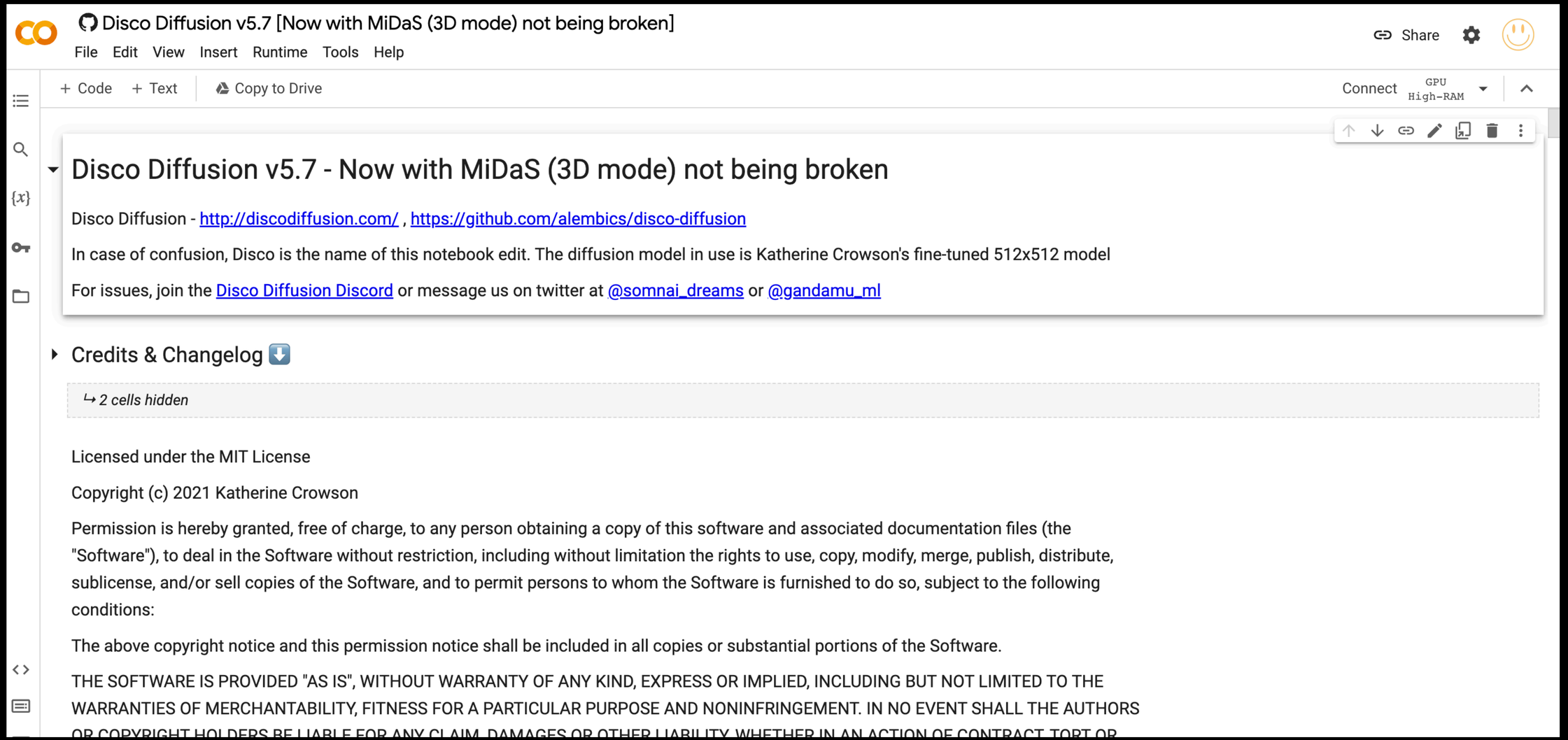
aletts Merge pull request #176 from lmagoncalo/main ... 37eb39b on F

archive	Update to Disco v5
docker	Add mssing libraries
.gitignore	Update .gitignore
Disco_Diffusion.ipynb	Corrected open_clip import error, fixes #175
LICENSE	The first official warp/flow release in official Disco Diffusion -- v5.
README.md	Release notes for the MiDaS v3 fix
disco.py	Release notes for the MiDaS v3 fix
disco_utils.py	Adding disco_utils.py
disco_xform_utils.py	Allow 3D animation to run on GPU's other than GPU0

README.md

https://mgxm.replicate.delivery/e9e07816-6147-4a86-8b27-10058e3b4620/progress.png

https://github.com/alembics/disco-diffusion



https://colab.research.google.com/github/alembics/disco-diffusion/blob/main/Disco_Diffusion.ipynb

deforum.art

Github Discord Twitter X Instagram Youtube Patreon

Animate with AI

Create beautiful animations with the power of generative AI. At Deforum, we're more than just a platform; we're a vibrant, open-source community of innovative developers and talented artists. Dive deep into collaborative projects, share insights, and grow your skills with like-minded individuals. Connect with us on Discord and embark on an unparalleled AI-driven creative journey.

Open-Source

Deforum seamlessly integrates into the Automatic Web UI. Dive into powerful features like video style transfer with Controlnet, Hybrid Video, 2D/3D motion, frame interpolation, and upscaling. Getting started is simple. Search for "Deforum" in the extension tab or download the Deforum Web UI Extension.

Deforum Auto Extension

Deforum extension for the Automatic Web UI, the most feature-rich implementation of Deforum

Deforum Colab

Original implementation of Deforum Stable Diffusion optimized for Google Colab

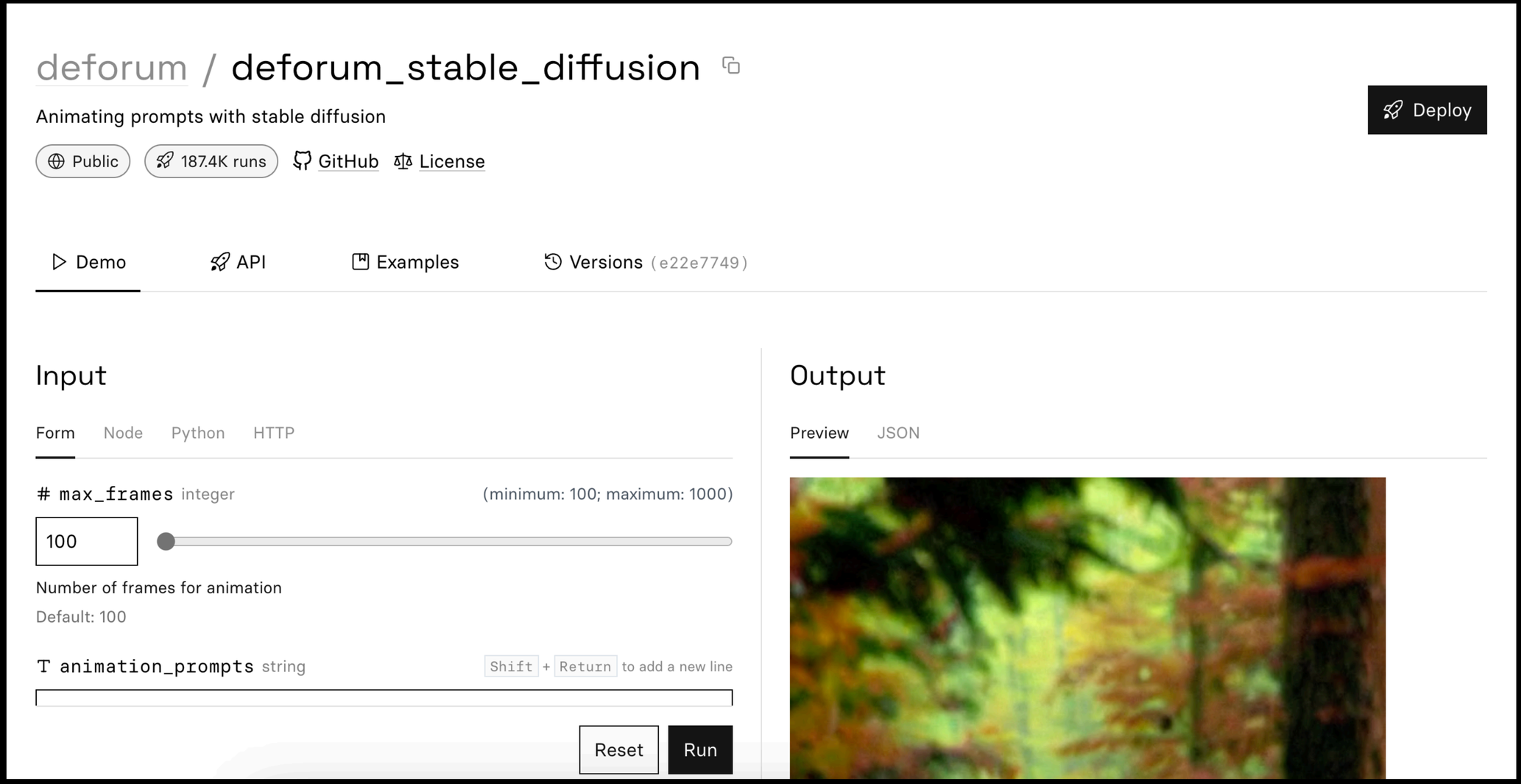
Deforum Parseq

Parseq is a parameter sequencer for the Deforum extension for the UI Deforum Web UI Extension

aiNodes

Simple and easy-to-use Python-based AI image and motion picture generator node engine

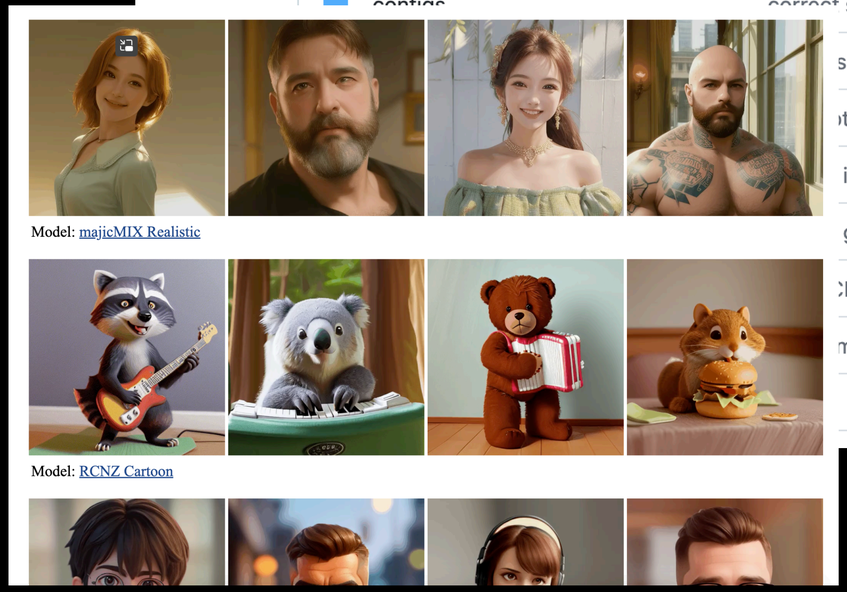
https://deforum.art/



https://replicate.com/deforum/deforum_stable_diffusion?prediction=oqtqckrb5peivw4tngvdhrdwb4

“a beautiful forest in Appalachia”

“a beautiful forest by Asher Brown Durand, trending on Artstation”



https://animatediff.github.io/

guoyww / AnimateDiff

Type / to search

Code | Issues 129 | Pull requests 12 | Actions | Projects | Security | Insights

AnimateDiff Public

Watch 64 | Fork 404 | Star 5.1k

main | 3 branches | 0 tags | Go to file | Add file | Code

AnyiRao add comfyui ae7c52c 2 weeks ago 66 commits

__assets__	add lora figures	last month
animatediff	update infer script	
configs	correct shot name	
shscripts		
tionlora folder		
infer script		
gitignore		
CENSE		
mfyui		

About

Official implementation of AnimateDiff.

animatediff.github.io/

Readme

Zoom In	Zoom Out	Zoom Pan Left	Zoom Pan Right
Tilt Up	Tilt Down	Rolling Anti-Clockwise	Rolling Clockwise

- **[2023/09/10]** New Motion Module release! `mm_sd_v15_v2.ckpt` was trained on larger resolution & batch size, and gains noticeable quality improvements. Check it out at Google Drive / HuggingFace / CivitAI and use it

Packages

https://github.com/guoyww/animatediff/

https://twitter.com/rainisto/status/1712804230401003938

TEXT TO VIDEO: TOOLS IN DEVELOPMENT

https://imagen.research.google/video/
https://aitestkitchen.withgoogle.com/

Input images

Make-A-Video output

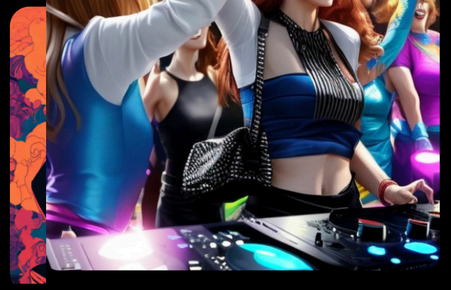
https://makeavideo.studio/

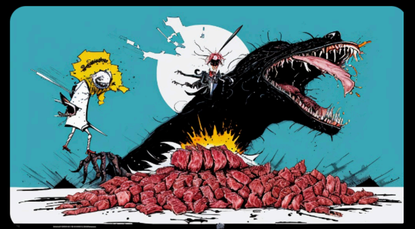
ANIMATE IMAGES

Kaiber Product Pricing Gallery Kaiber Studio About Start free trial Login

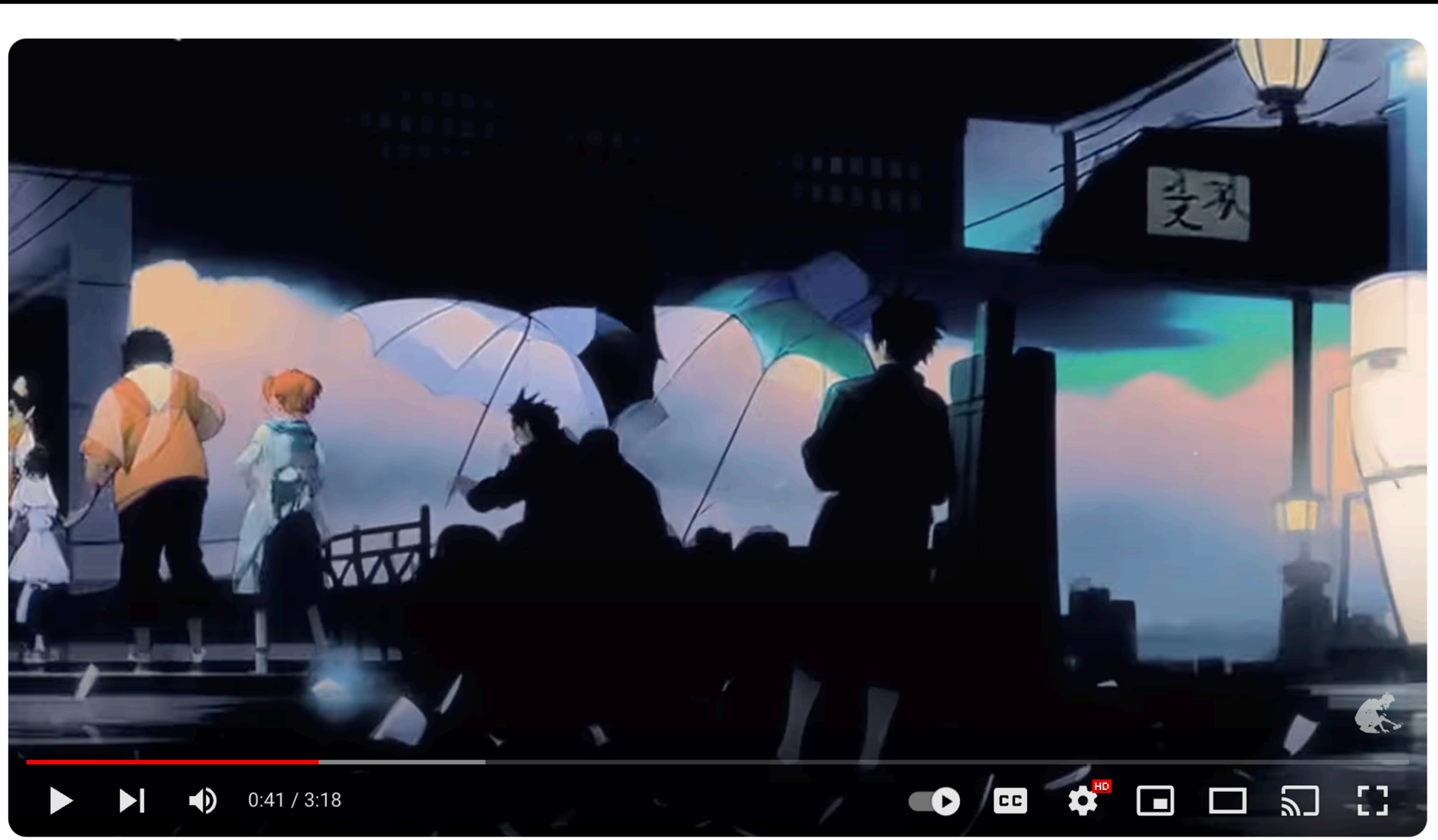
Kaiber Core Features

Made for artists, by artists, we are an AI creative lab on a mission to unlock creativity through powerful and intuitive generative audi





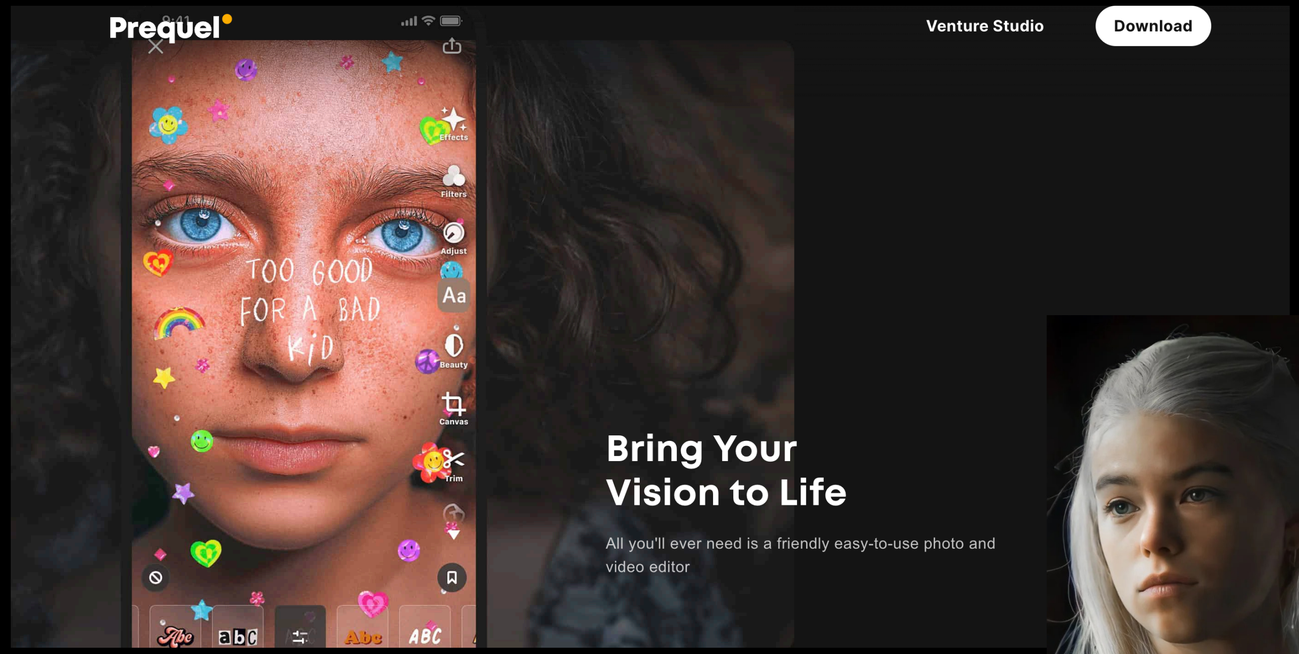
https://kaiber.ai/



Lost [Official Music Video] - Linkin Park

https://www.youtube.com/watch?v=7NK_JOkuSVY

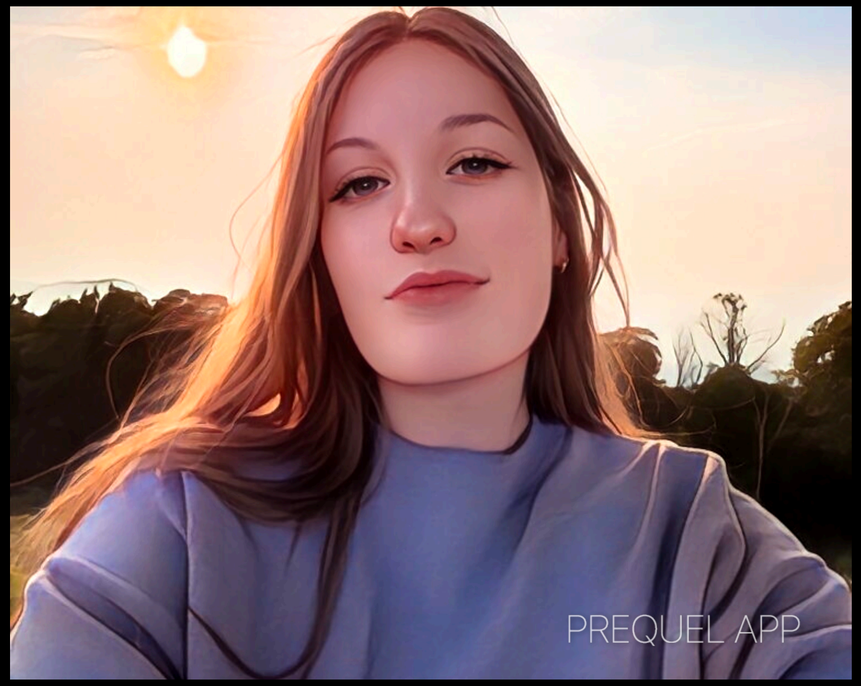
STYLIZE / EDIT VIDEO

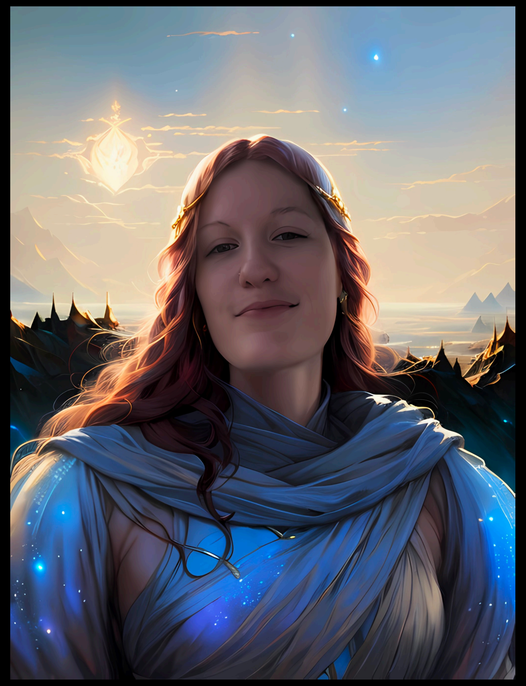


https://prequel.app/

https://preview.redd.it/ov8cclcehrt91.jpg?width=1280&format=pjpg&auto=webp&s=0a506a137742413ade7a8e13051bbf54ce96cc21

Original photo



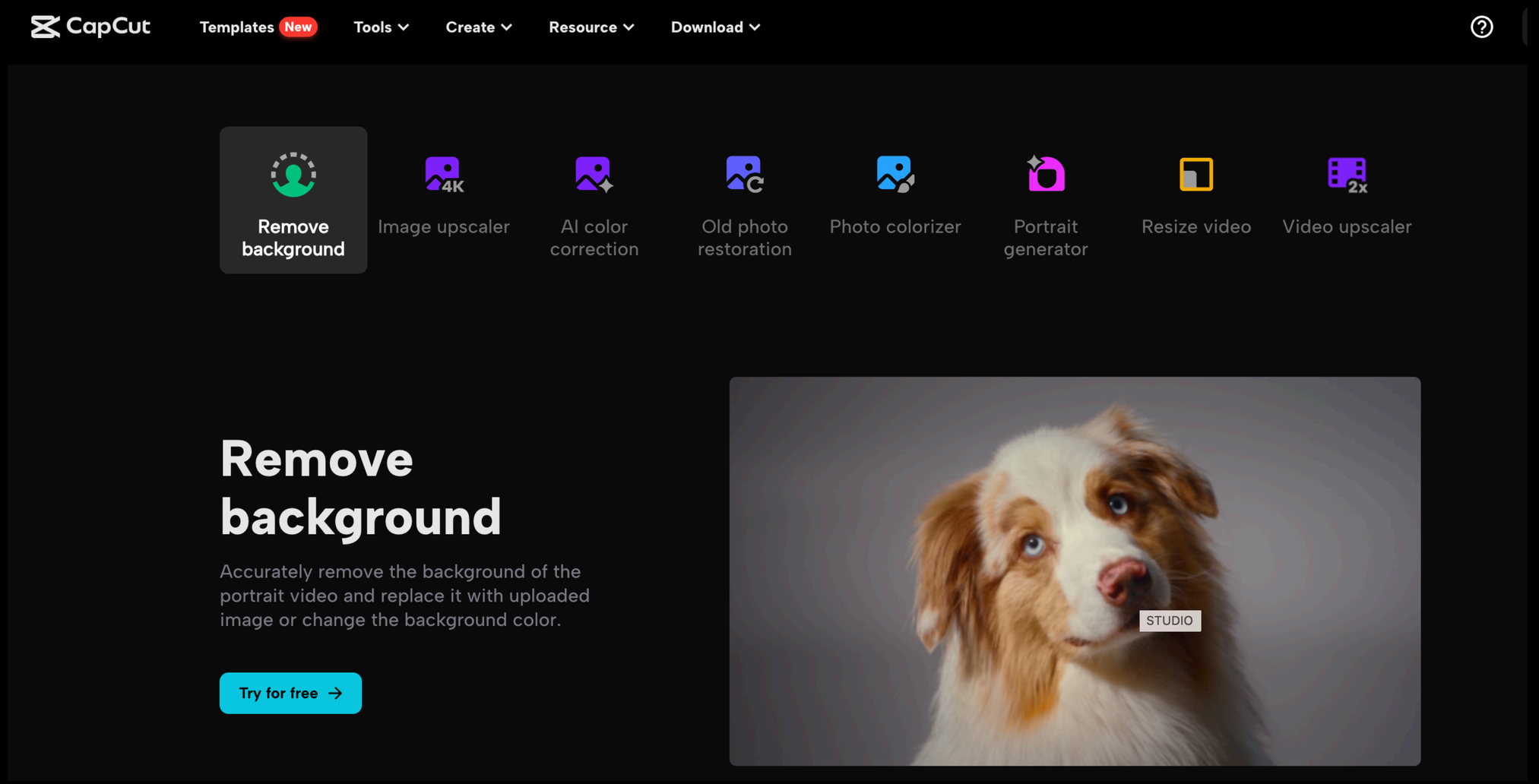


📝 Introduction to ShortGPT

ShortGPT is a powerful framework for automating content creation. It simplifies video creation, footage sourcing, voiceover synthesis, and editing tasks.

- 🎞️ **Automated editing framework**: Streamlines the video creation process with an LLM oriented video editing language.
- ⌨️ **Scripts and Prompts**: Provides ready-to-use scripts and prompts for various LLM automated editing processes.
- 🗣️ **Voiceover / Content Creation**: Supports multiple languages including English 🇺🇸, Spanish 🇪🇸, Arabic 🇦🇪, French 🇫🇷, Polish 🇵🇱, German 🇩🇪, Italian 🇮🇹, Portuguese 🇵🇹, Russian 🇷🇺, Mandarin Chinese 🇨🇳, Japanese 🇯🇵, Hindi 🇮🇳,Korean 🇰🇷, and way over 30 more languages (with EdgeTTS)
- 🔗 **Caption Generation**: Automates the generation of video captions.
- 🌐📹 **Asset Sourcing**: Sources images and video footage from the internet, connecting with the web and Pexels API as necessary.
- 🧠 **Memory and persistency**: Ensures long-term persistency of automated editing variables with TinyDB.

https://github.com/RayVentura/ShortGPT



https://www.capcut.com/

CapCut

+ Create new

Start

Recent drafts

Share history

Templates

Magic tools Beta

Spaces

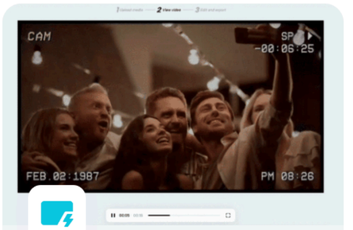
U user9431041984721'...

+ Create new space

New features
Text to Image & Ima...
8

Magic tools

Script to video
Generate scripts to create videos in a few clicks.



AutoCut
Automatically create videos with your media in 1 click.

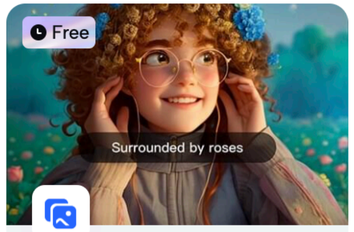
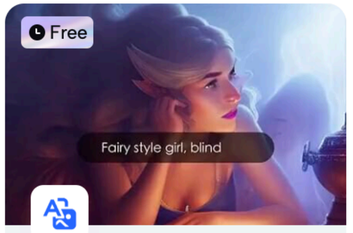
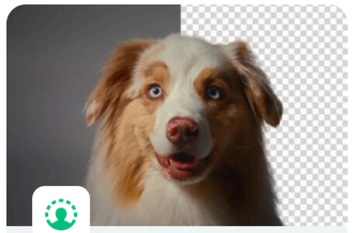


Image to image
Generate images with an image and word prompt.



Text to image
Generate images with word prompts.

Ad script
Generate high-converting TikTok ad scripts.



Remove background
Remove backgrounds in videos and photos.

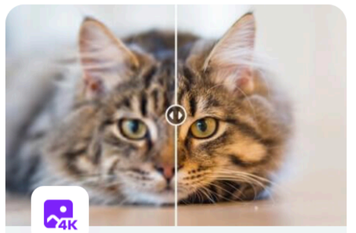


Image upscaler
Upscale images by increasing resolution.

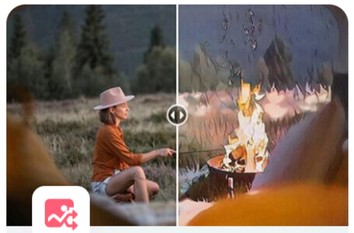
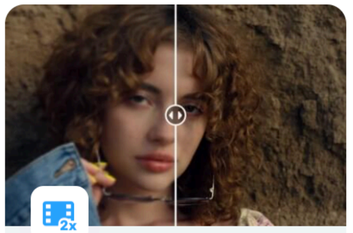
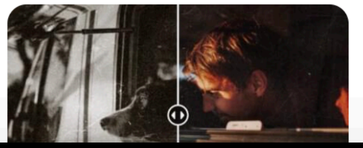


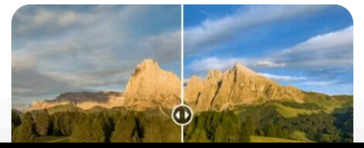
Image style transfer
Convert your images into various styles.

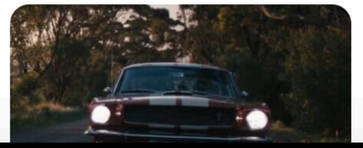


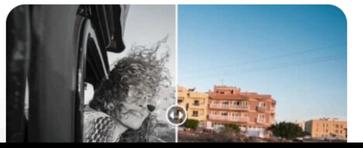
Video upscaler
Increase video resolution to make it clearer without losing it...

Portrait generator
Generate AI portraits in various styles.

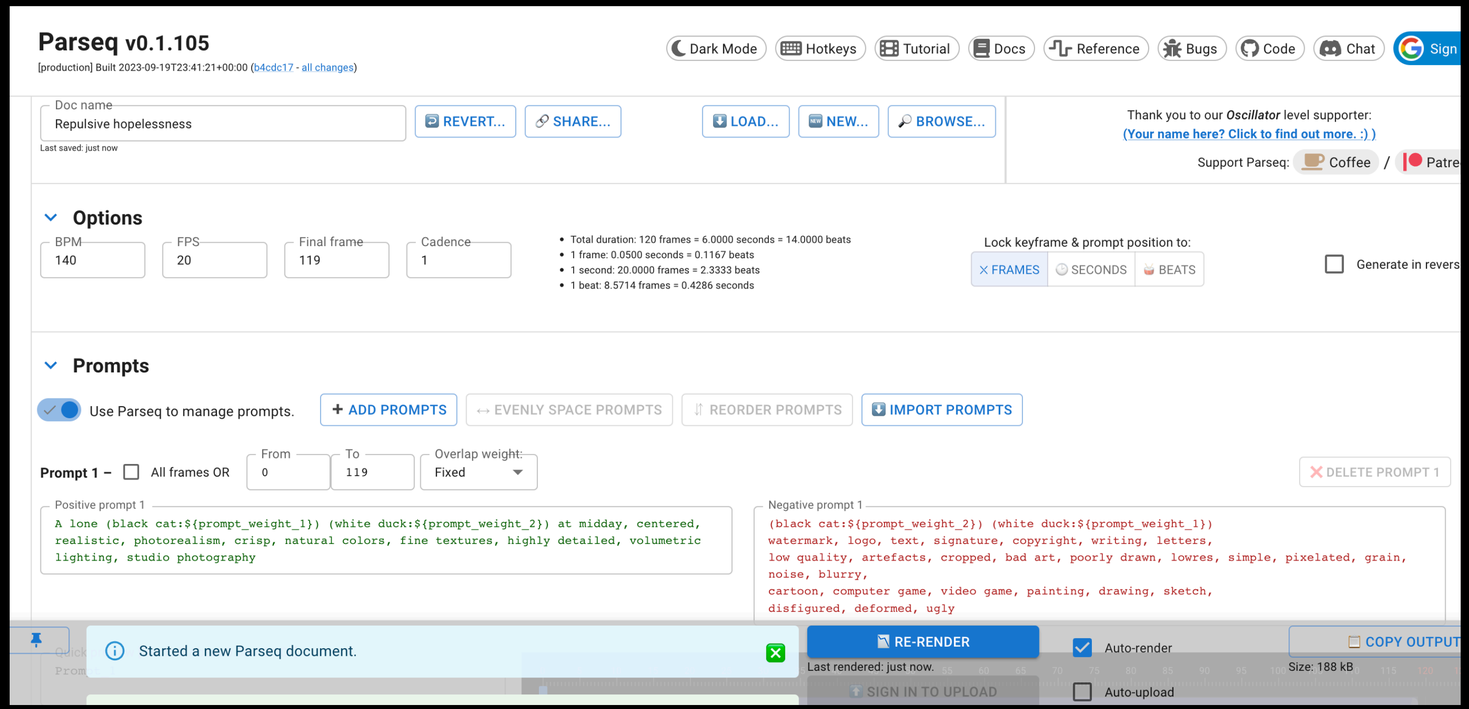




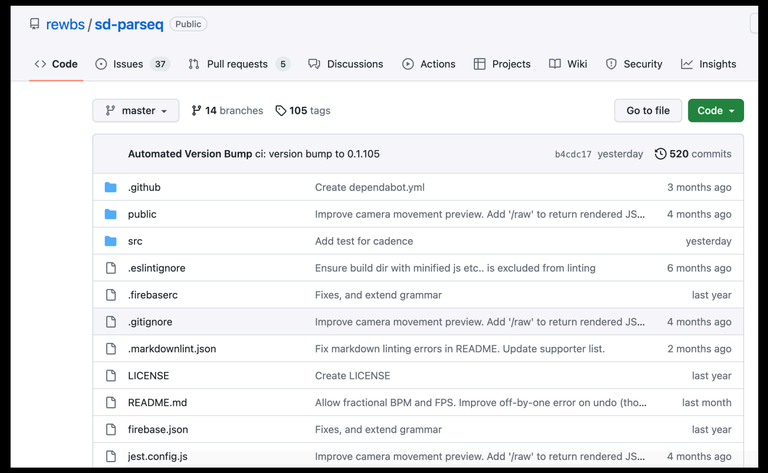




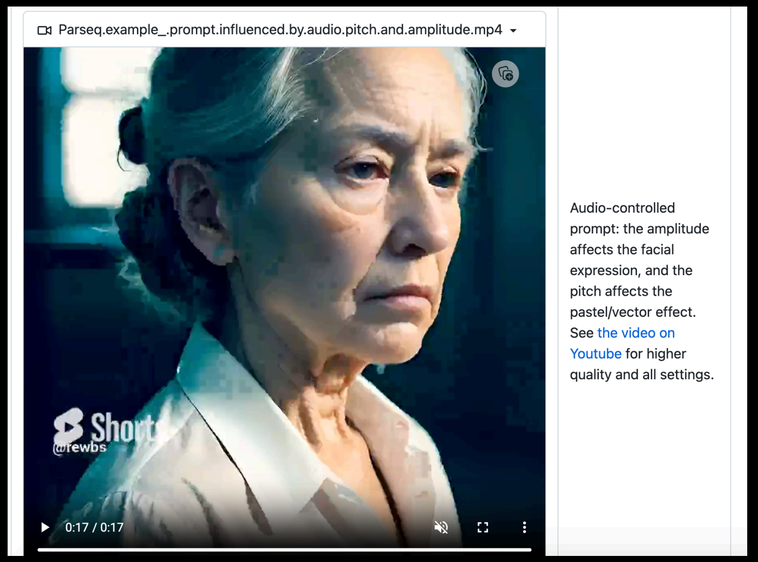
ANIMATE TO MUSIC



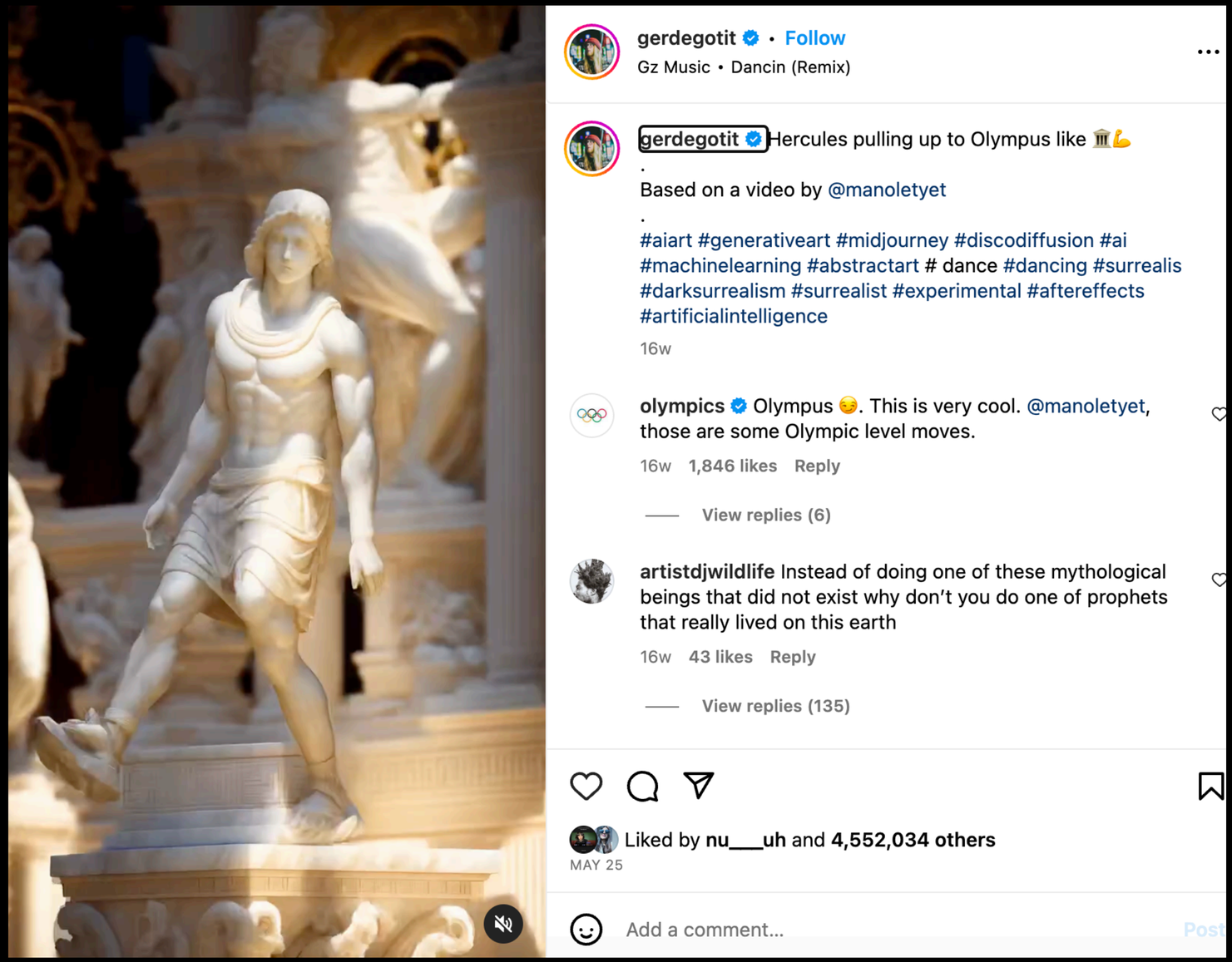
https://sd-parseq.web.app/?docId=doc-1af794e2-d7a2-4954-98af-46af787f5db8



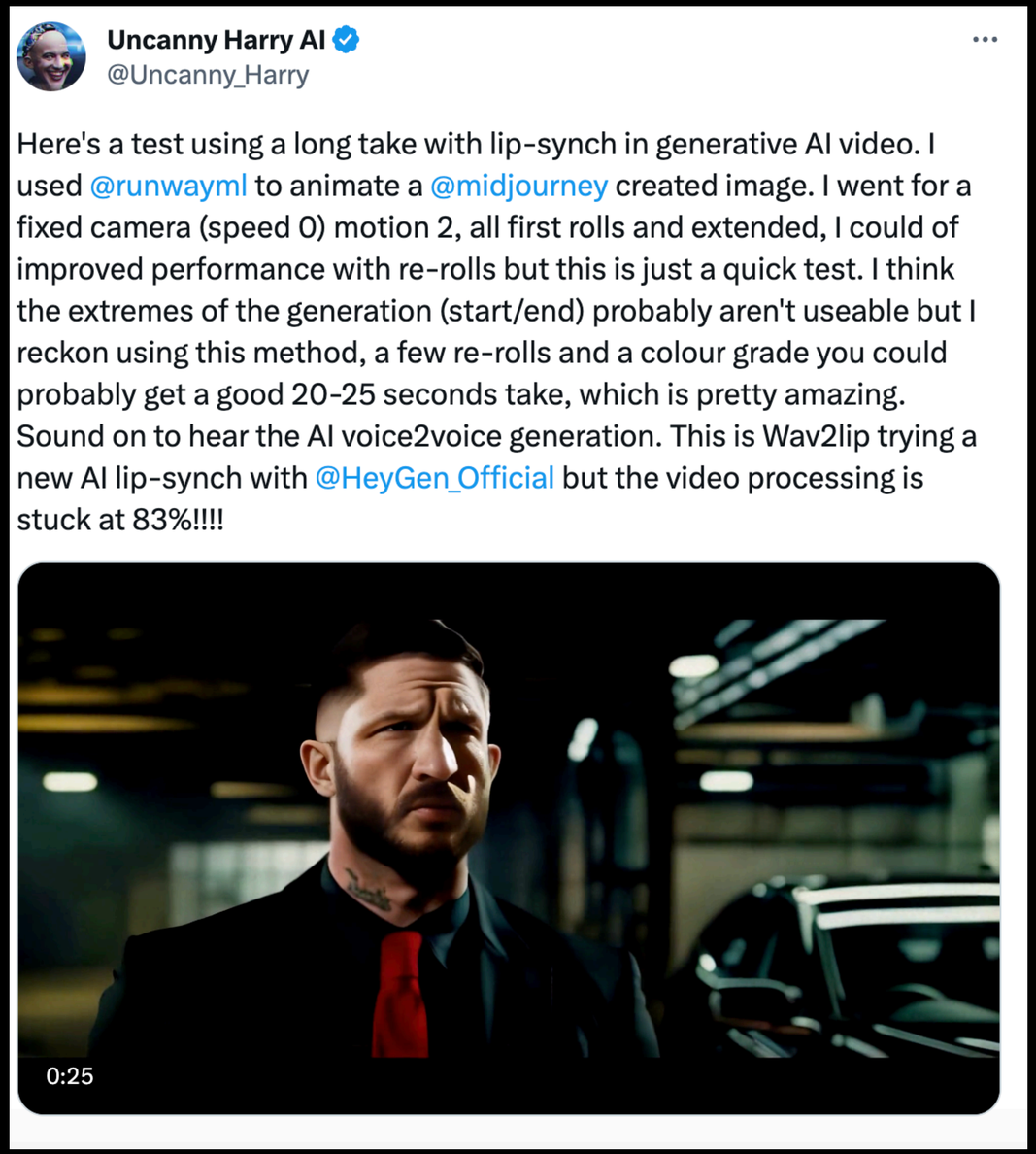
https://github.com/rewbs/sd-parseq#what-is-this



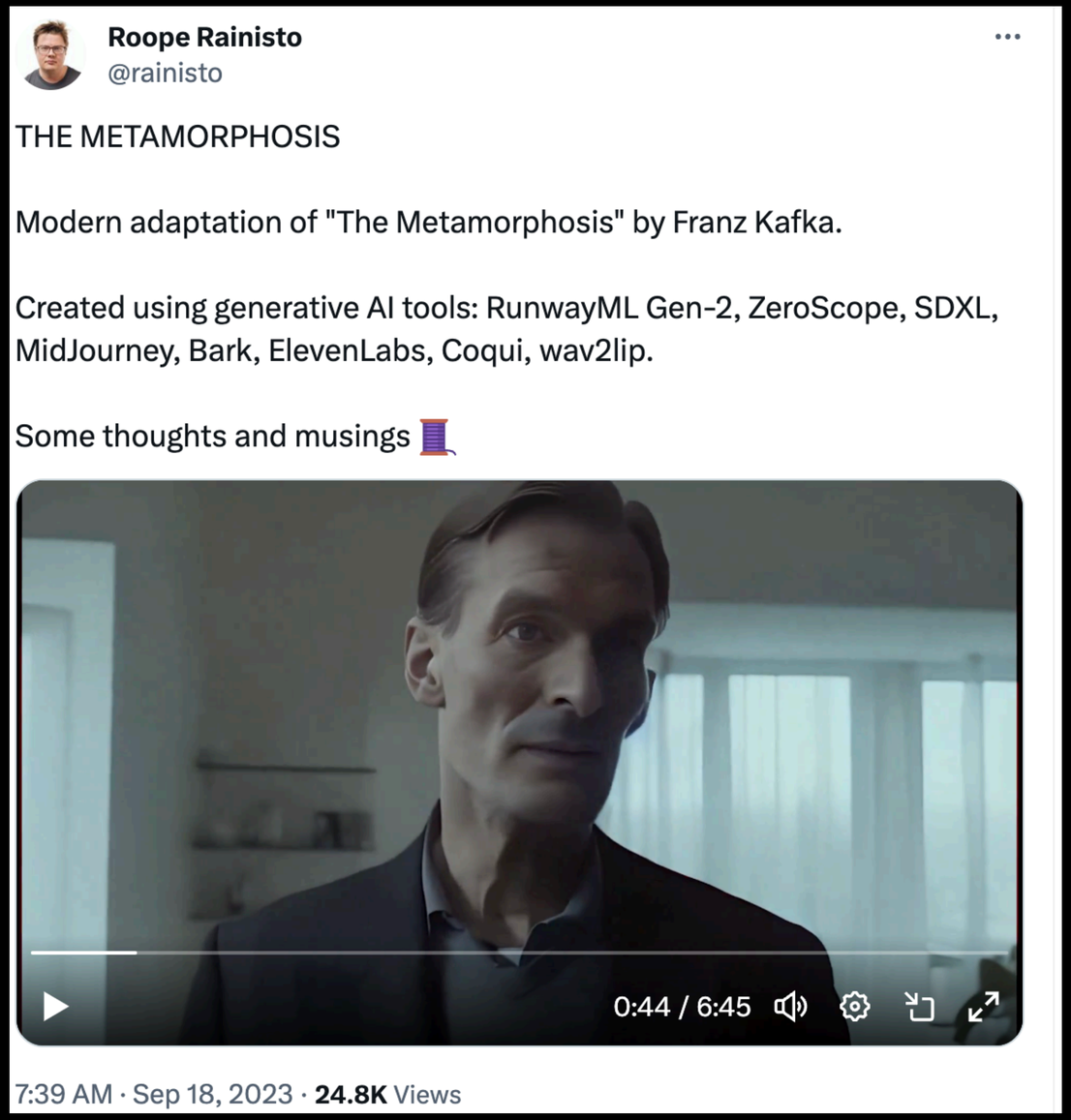
COMBINATORIAL TECHNIQUES / EXAMPLES



https://www.instagram.com/p/CsreGaqNgMz/?hl=en



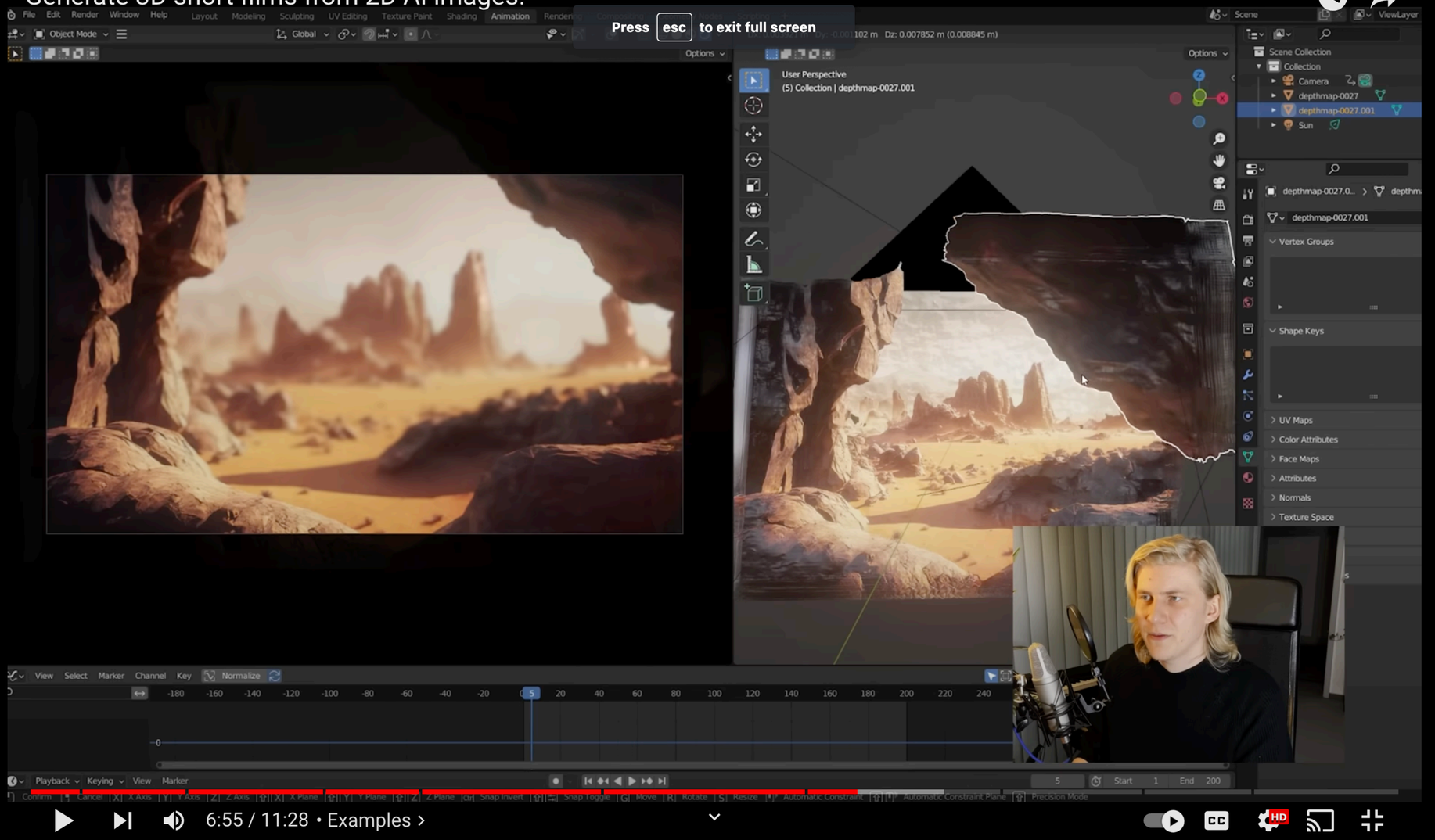
https://twitter.com/uncanny_harry/status/1701547787895107787?s=46&t=MHjSXuuPGHmWLD4It3ZkFA



https://twitter.com/rainisto/status/1703780912436056150?s=46&t=MHjSXuuPGHmWLD4It3ZkFA

https://twitter.com/sekiguchiaimi/status/1705855926148984915?s=46&t=MHjSXuuPGHmWLD4It3ZkFA

Generate 3D short films from 2D AI images!



https://www.youtube.com/watch?v=PFybvdKH8hc

What other composite workflows could you imagine exploring?

ARCHIVAL RESOURCES

Further Approaches

- **Video editing:** These tools use AI to automate certain aspects of the video editing process, such as color correction, audio sync, and effects. Some popular examples include Adobe Premiere Pro, Magisto, and Animoto.
- **Video generation:** These tools use AI to generate videos from scratch, often using voiceovers, text-to-speech, and other techniques to create engaging and informative content. Some popular examples include Lumen5, Wave.video, and Wibbitz.
- **Video recognition:** These tools use computer vision algorithms to detect and track objects within videos. They can be used for a wide range of tasks, including video surveillance, image search, and advertising. Popular examples include Amazon Rekognition, OpenCV, and TensorFlow.
- **Lumen5:** An AI tool that turns blog posts into engaging videos
- **Animoto:** A video creation tool that uses AI to quickly generate polished videos
- **Magisto:** An AI-powered video editing tool that automatically creates professional-looking videos from footage and photos
- **Wibbitz:** A tool that uses AI to create videos by analyzing the text content of articles and turning them into short video summaries
- **VidMob:** An AI-powered video editing platform that matches clients with a team of editors to create professional-quality ads, short videos, and other pieces of video content

HOMEWORK

Homework

- In groups of 2, explore the use of 3+ of the video/animation/moving image generation tools presented in class
- Create videos, artworks, etc.. related to your interests / topic for the course
- This work should be able to be presented as a standalone project, but should ideally contribute towards the development of your final project
- Presentation format is open to student preference, but should be limited to 3-5 minutes