CULTURAL AUTOMATION WITH MACHINE LEARNING

SESSION 05: GENERATIVE AUDIO

Parag K. Mital
UCLA DMA

CULTURAL CONTEXT + HISTORY OF GENERATIVE AUDIO

Still-Life with Chair Caning, 1912 by Pablo Picasso

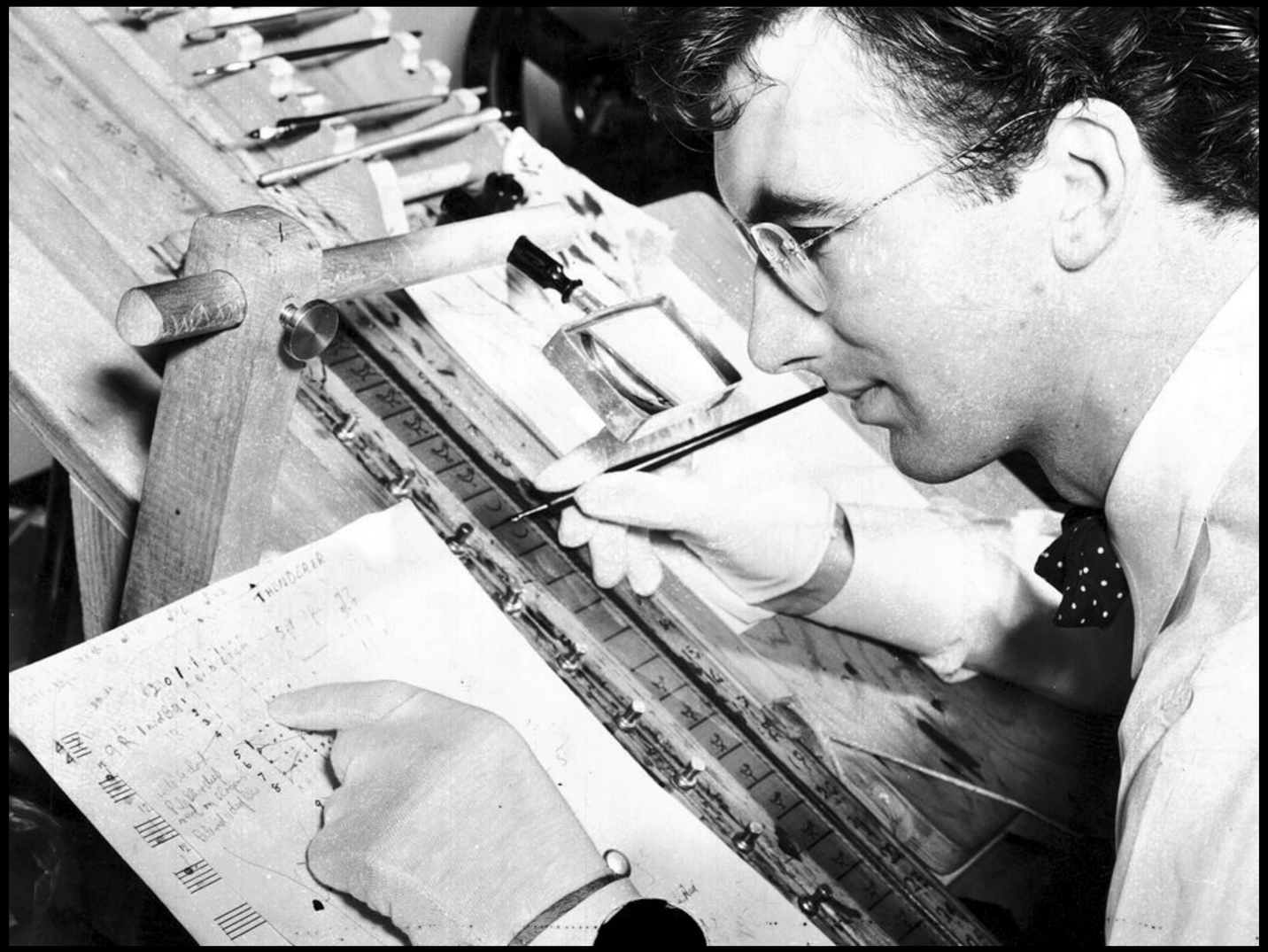
Walter Ruttmann: Weekend (1930)

https://www.youtube.com/watch?v=v3cEAMh3eng

Hans Richter

"The Art of Noise"

Pierre Schaeffer - "Musique concrète"

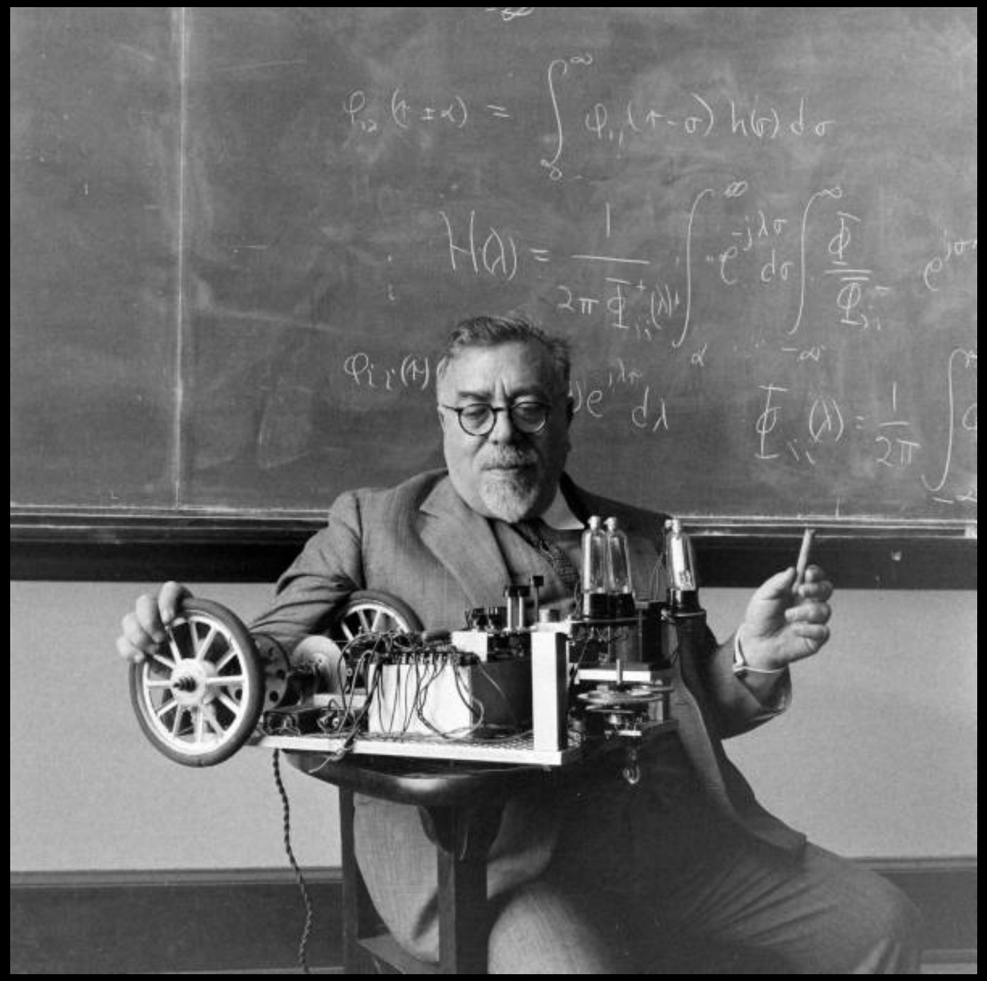


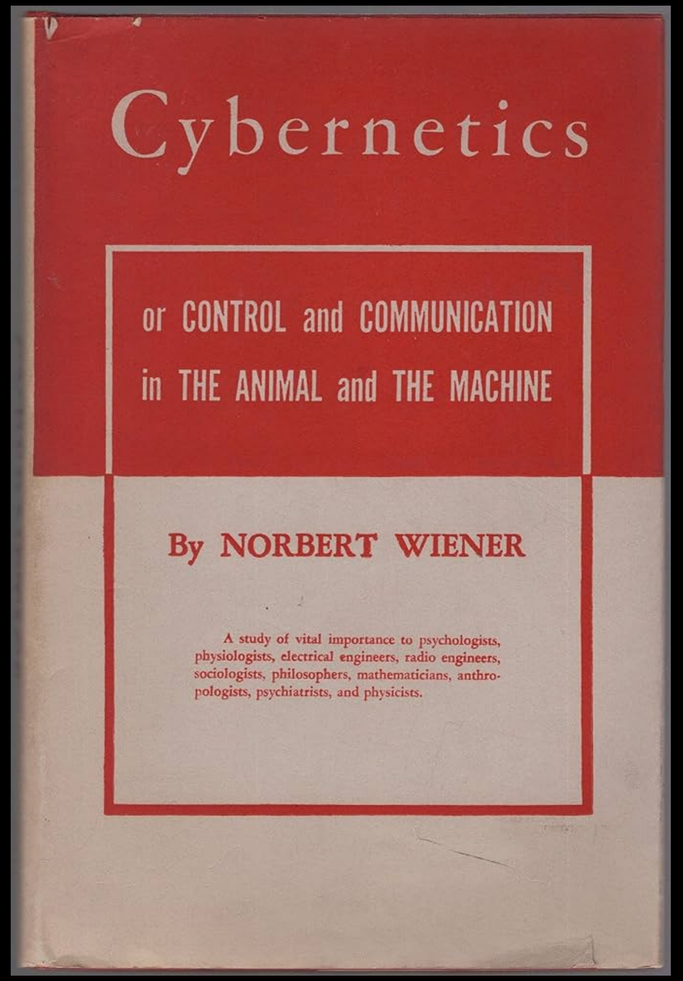
Norman McClaren - visual music

Oskar Fischinger - visual music

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

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



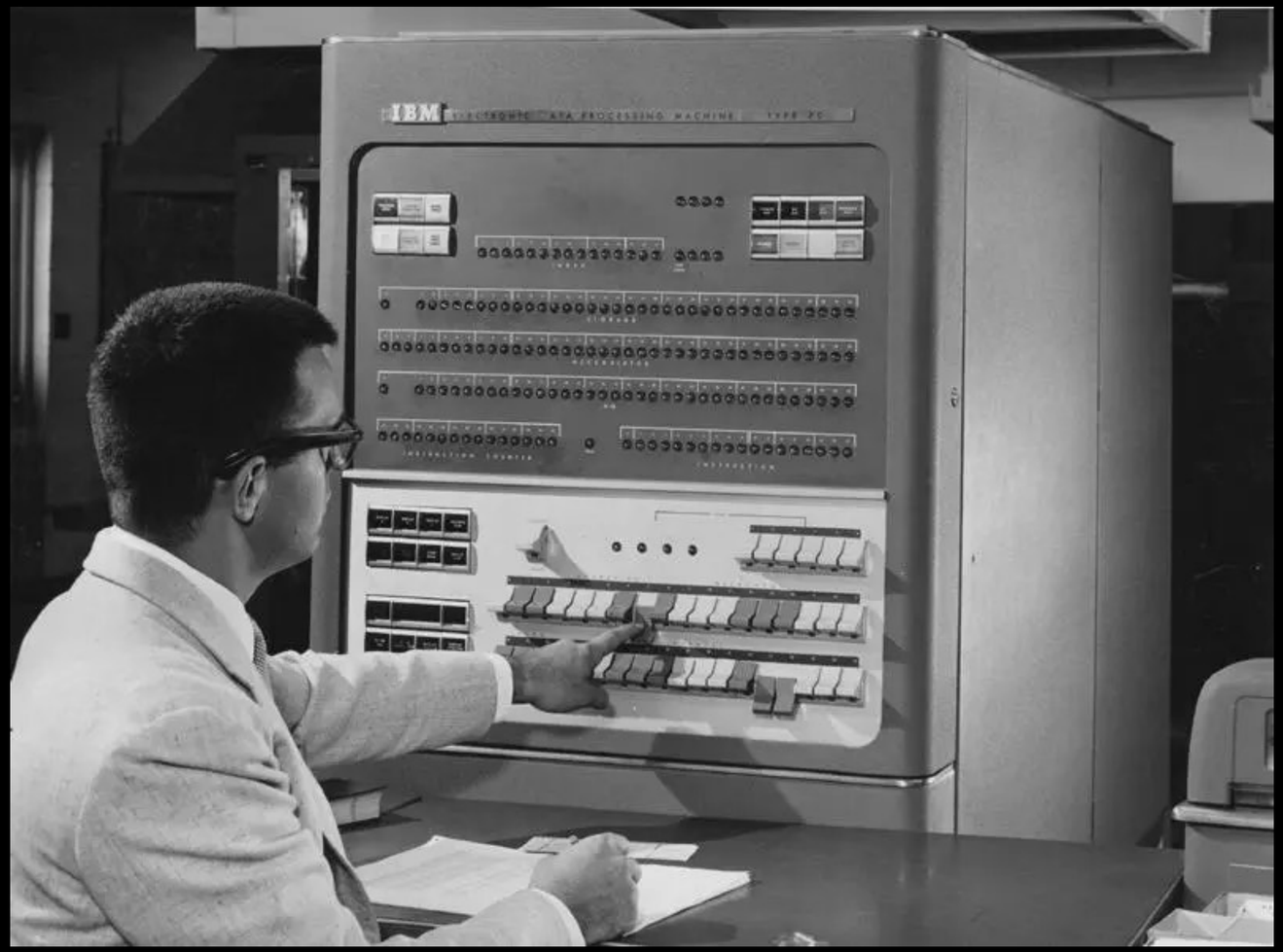


Norbert Wiener - cybernetics research and his 1948 book "Cybernetics: Or Control and Communication in the Animal and the Machine"

https://www.bl.uk/collection-items/earliest-known-recording-of-computer-generated-music#:~:text=In%201951%2C%20a%20BBC%20outside,recording%20of%20computer%2Dgenerated%20music

https://en.wikipedia.org/wiki/Illiac_Suite

https://www.youtube.com/watch?v=fojKZ1ymZlo

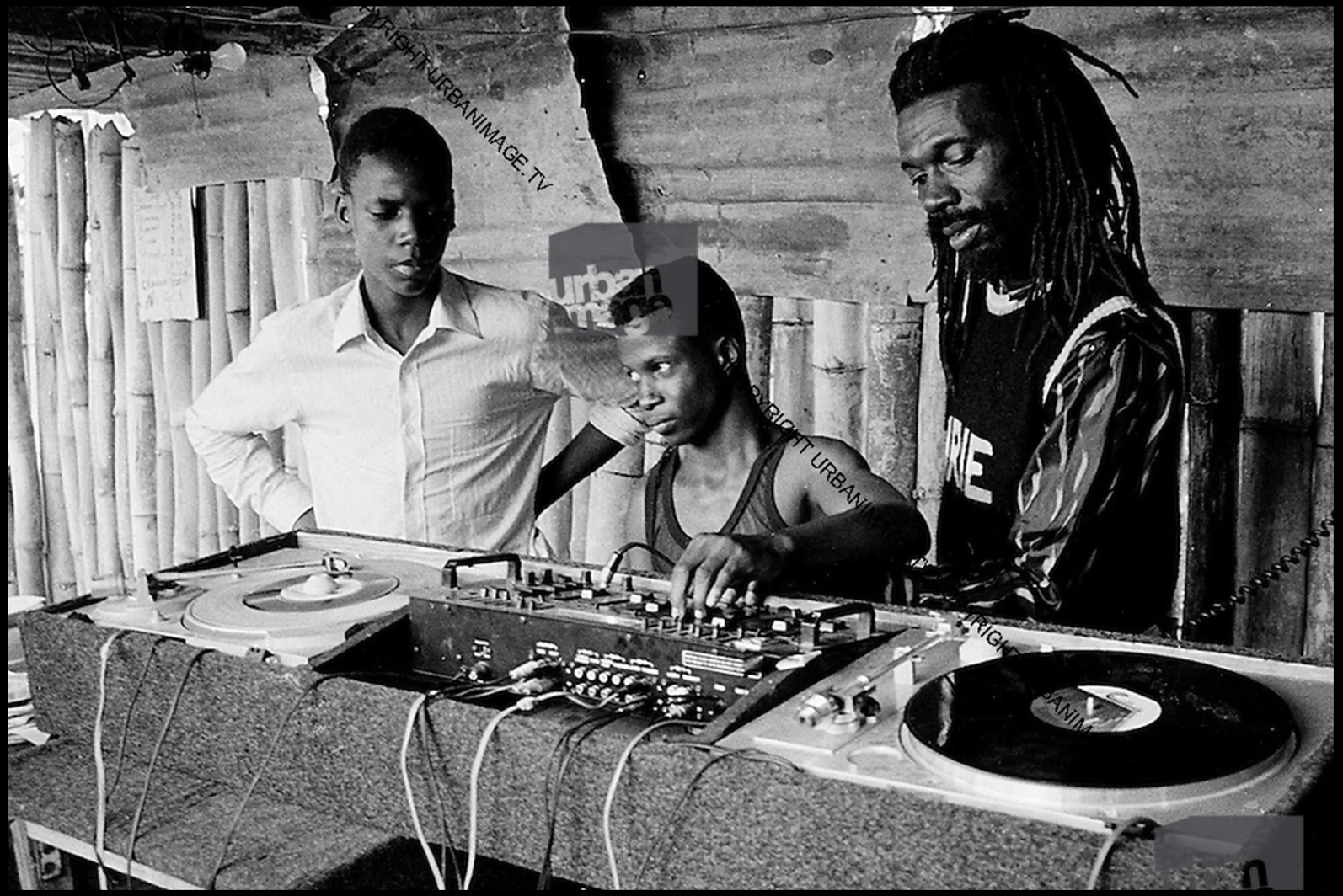


https://www.musicradar.com/news/tech/a-brief-history-of-computer-music-177299

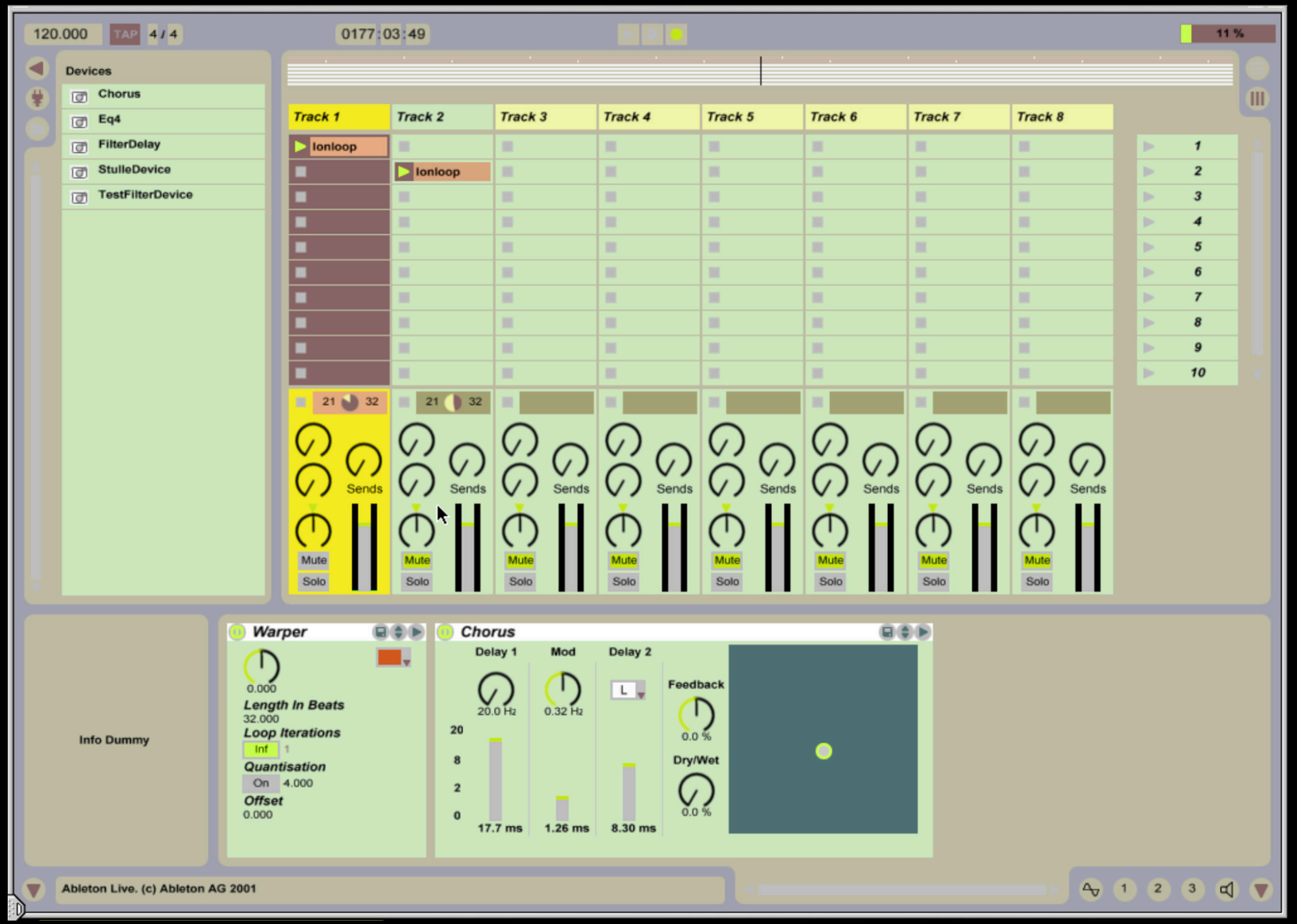
Daphne Oram - visual music (1960's)

https://www.daphneoram.org/oramicsmachine/

David Tudor



Dj Kool Herc



Ableton Live

The Sugarhill Gang - Rapper's Delight (Official Video)

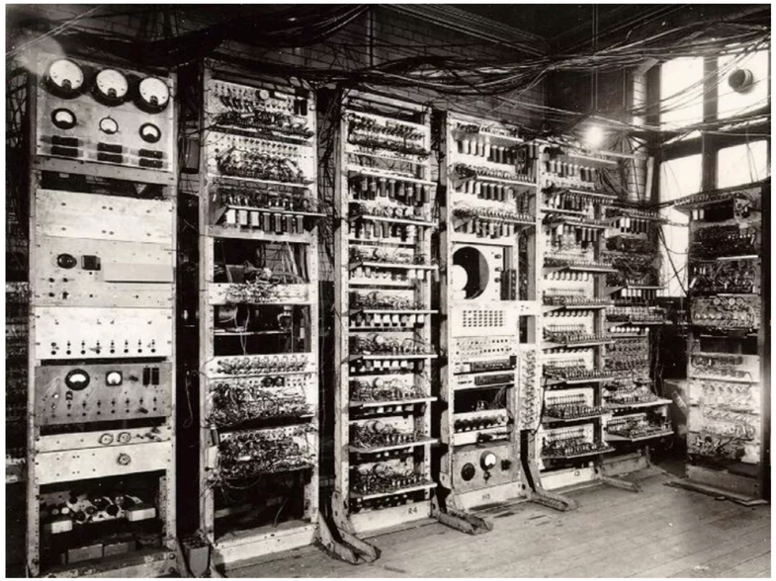
https://www.youtube.com/watch?v=mcCK99wHrk0

News

A brief history of computer music

By Computer Music Specials published October 13, 2008

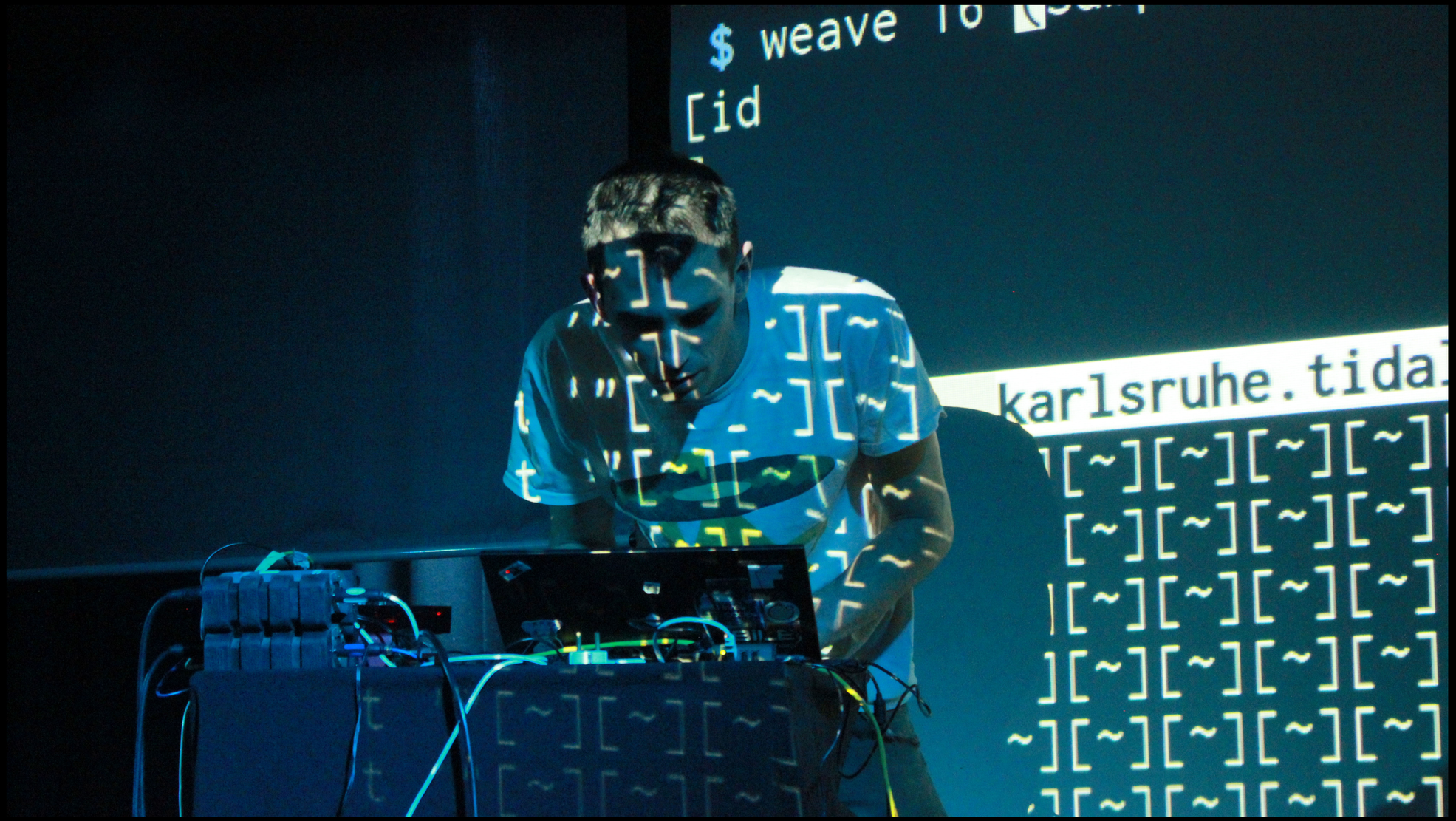
From Baa Baa Black Sheep to Ableton Live

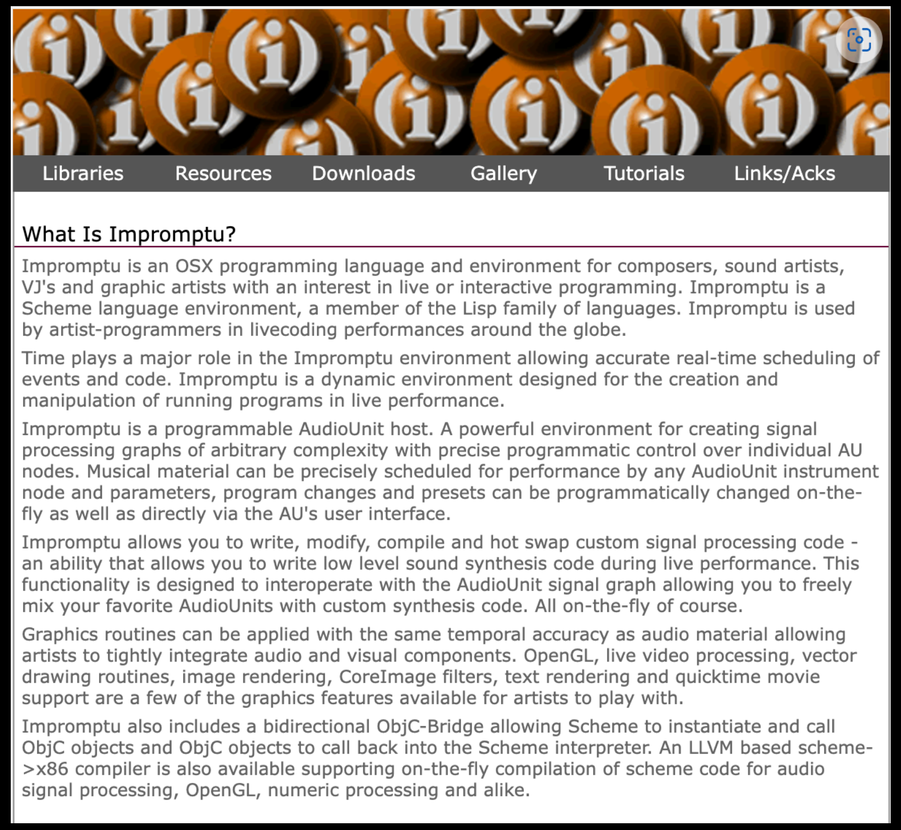


'The Baby' was the forerunner to the Ferranti Mark 1 that was used to make the first computer music recording.

History of Computer Music:
https://www.musicradar.com/news/tech/a-brief-history-of-computer-music-177299

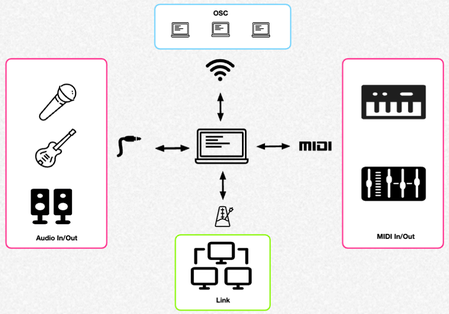
rjdj
https://en.wikipedia.org/wiki/RjDj


$ weave
karlsruhe.tida



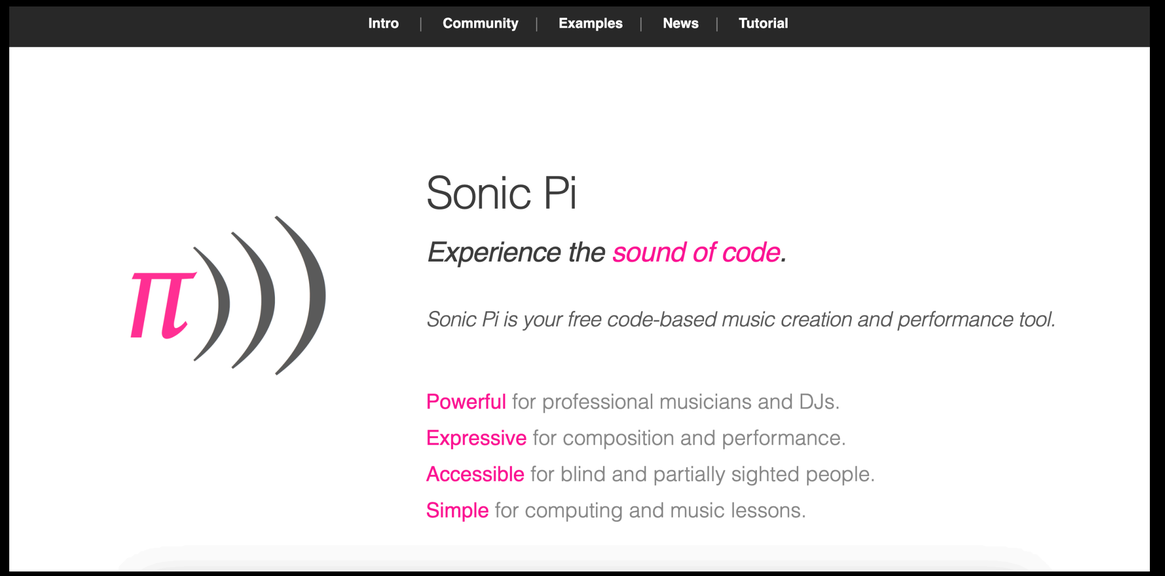
http://impromptu.moso.com.au/

Live Code Everything

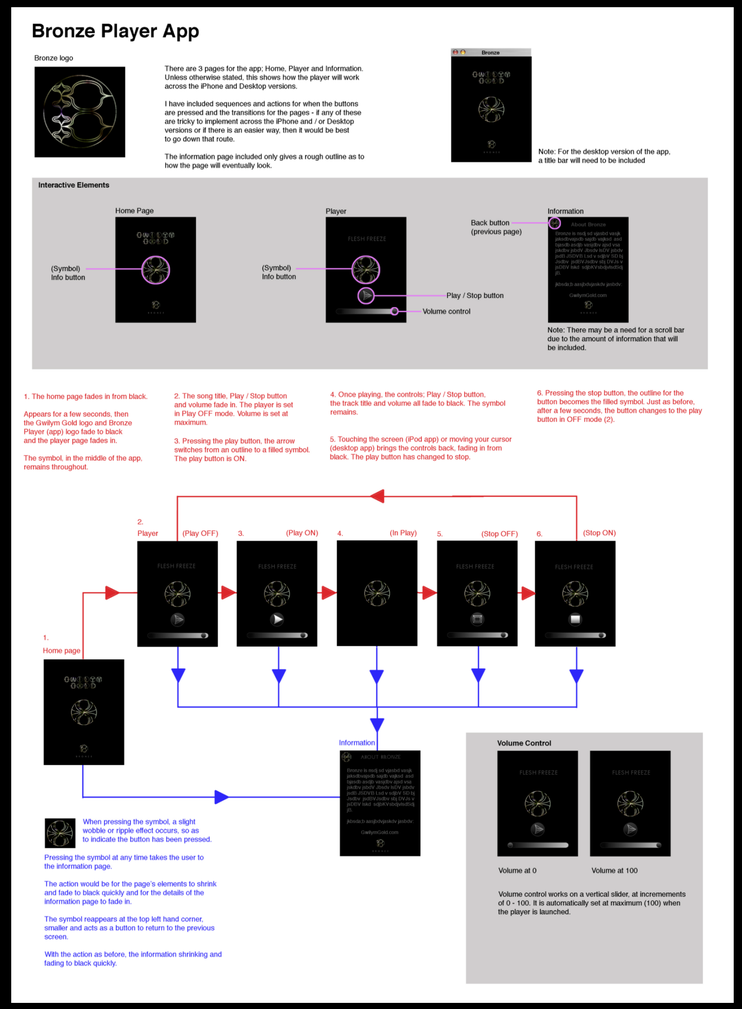


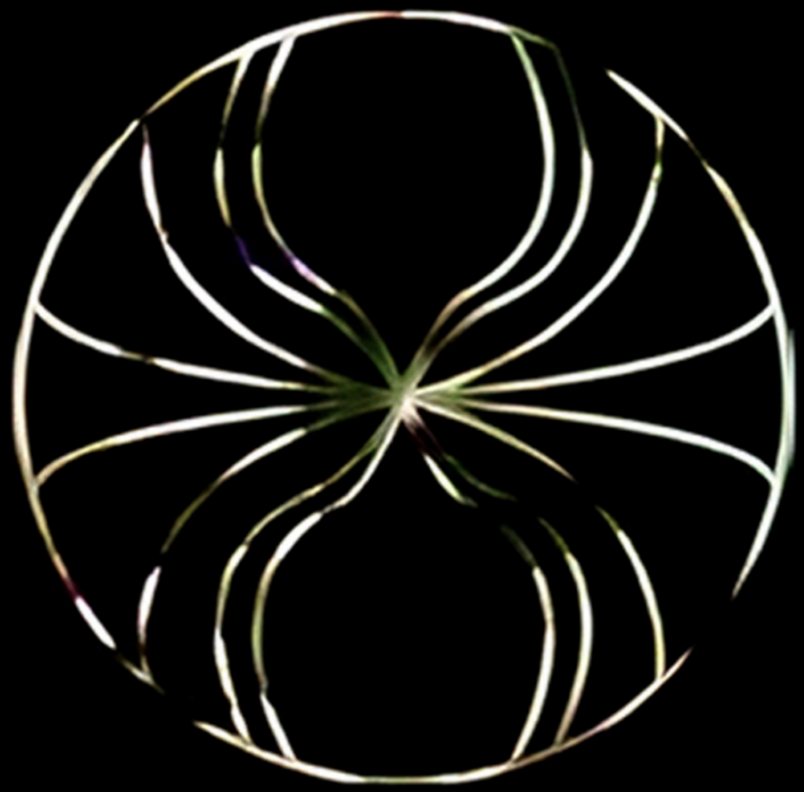
Sonic Pi lets you use simple code to turn your computer into a fully networked live coding music studio:

Multi Channel Audio In/Out
Well-timed MIDI In/Out
Well-timed OSC (Open Sound Control) In/Out
Ableton's Link network metronome built-in



https://sonic-pi.net/





Bronze Format

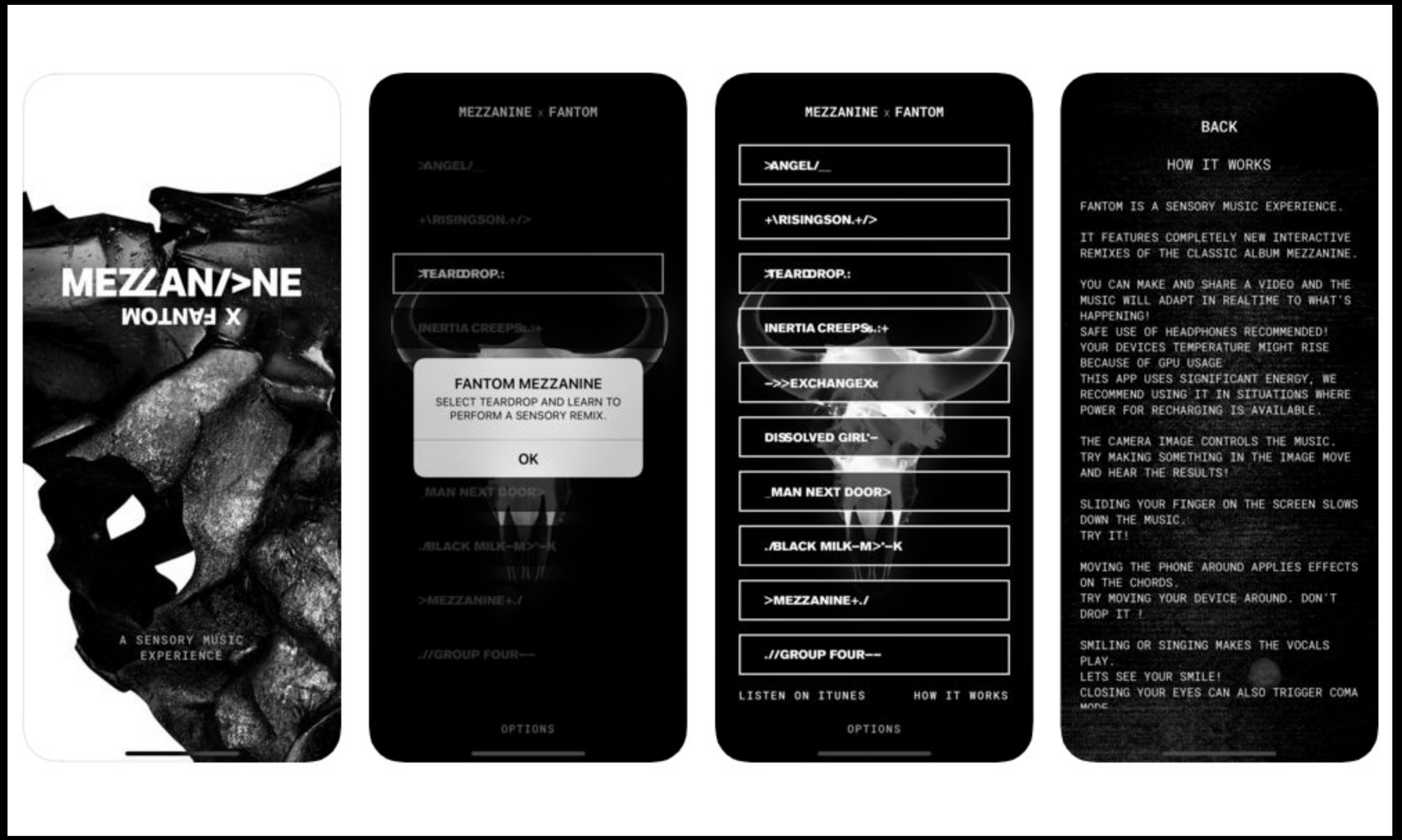
https://pkmital.com/home/works/bronze-format/

Memory Mosaic
https://pkmital.com/home/works/memory-mosaic/

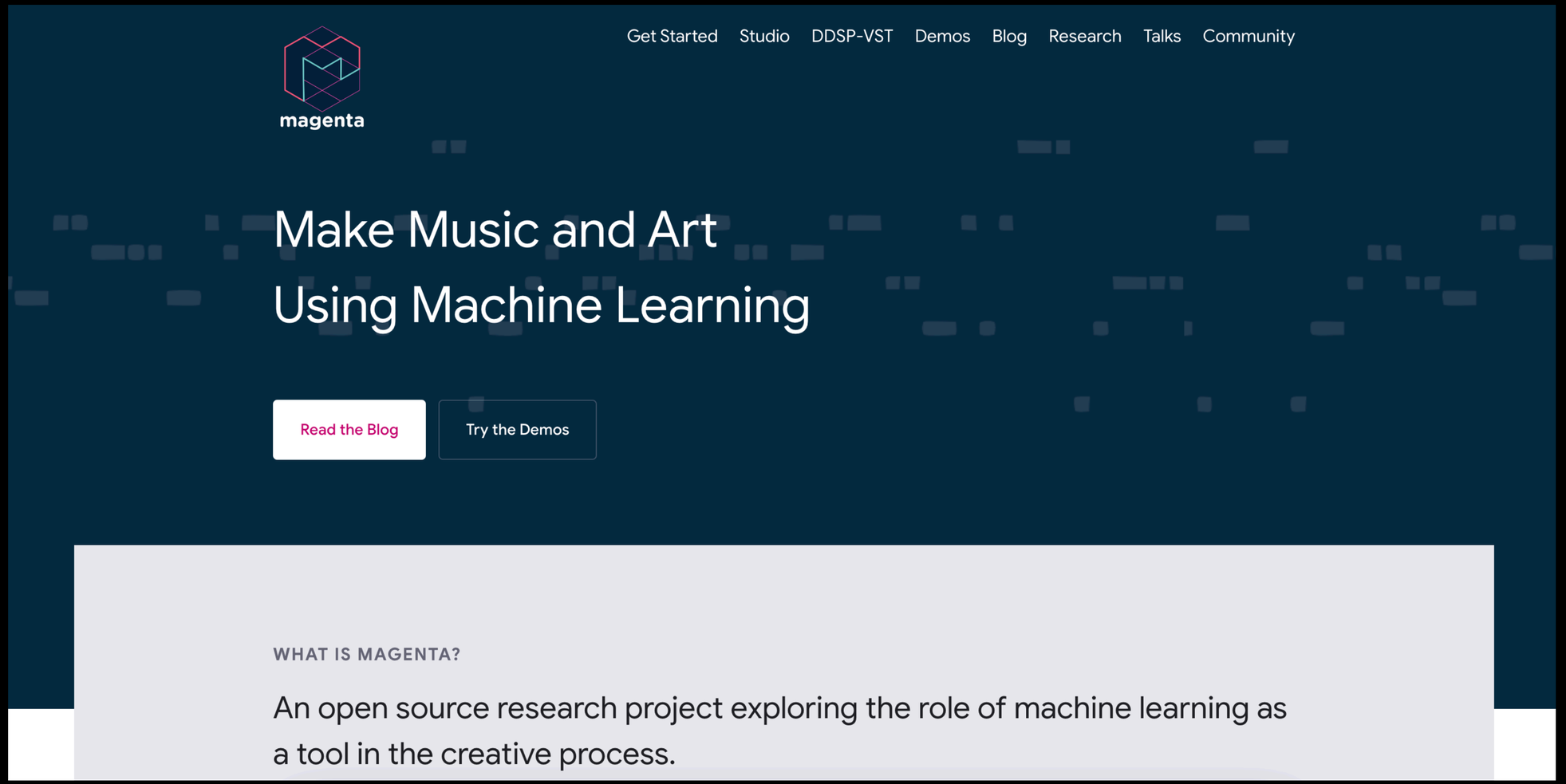
https://vimeo.com/111858776

https://keirvine.com/living-symphonies/

YouTube Smash Up
https://pkmital.com/home/works/youtube-smash-up/



Fantom
https://www.thirdspace.co/fantom



Magenta
https://magenta.tensorflow.org/

MusicVAE: Creating a palette for musical scores with machine learning.

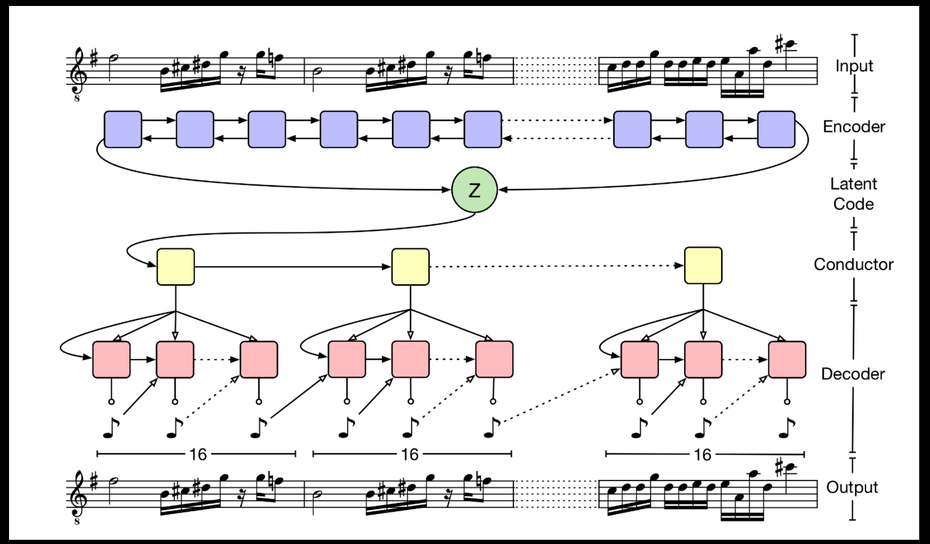
Mar 15, 2018

Adam Roberts adarob ada_rob
Jesse Engel jesseengel jesseengel
Colin Raffel craffel colinraffel
Ian Simon iansimon iansimon
Curtis Hawthorne cghawthorne fjord41

When a painter creates a work of art, she first blends and explores color options on an artist's palette before applying them to the canvas. This process is a creative act in its own right and has a profound effect on the final work.

Musicians and composers have mostly lacked a similar device for exploring and mixing musical ideas, but we are hoping to change that. Below we introduce **MusicVAE**, a machine learning model that lets us create palettes for blending and exploring musical scores.

As an example, listen to this gradual blending of 2 different melodies, A and B. We'll explain how this morph was achieved throughout the post.



Music VAE
https://magenta.tensorflow.org/music-vae

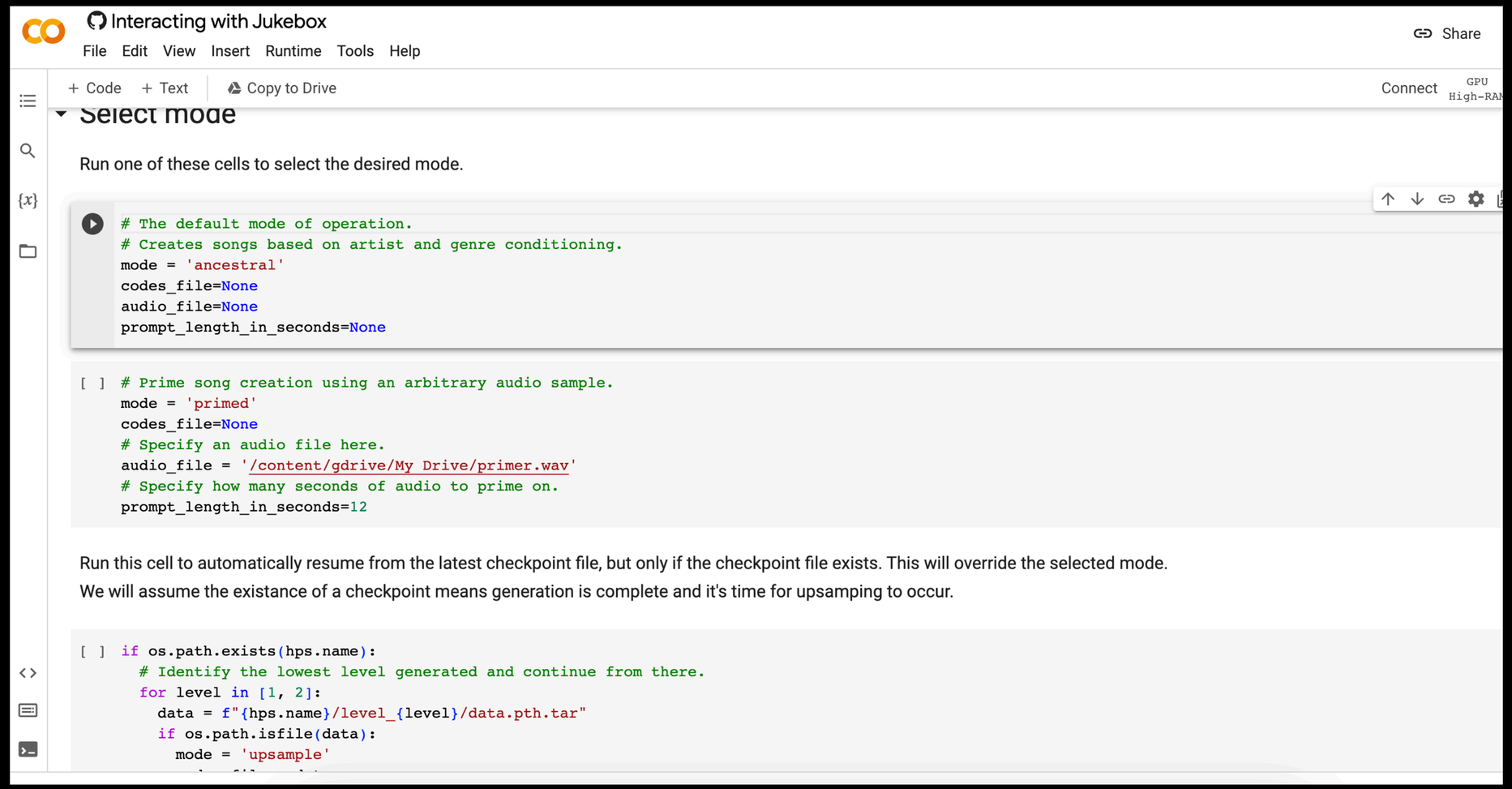
Dadabots
https://dadabots.com/

34.5TB

WDCH Dreams

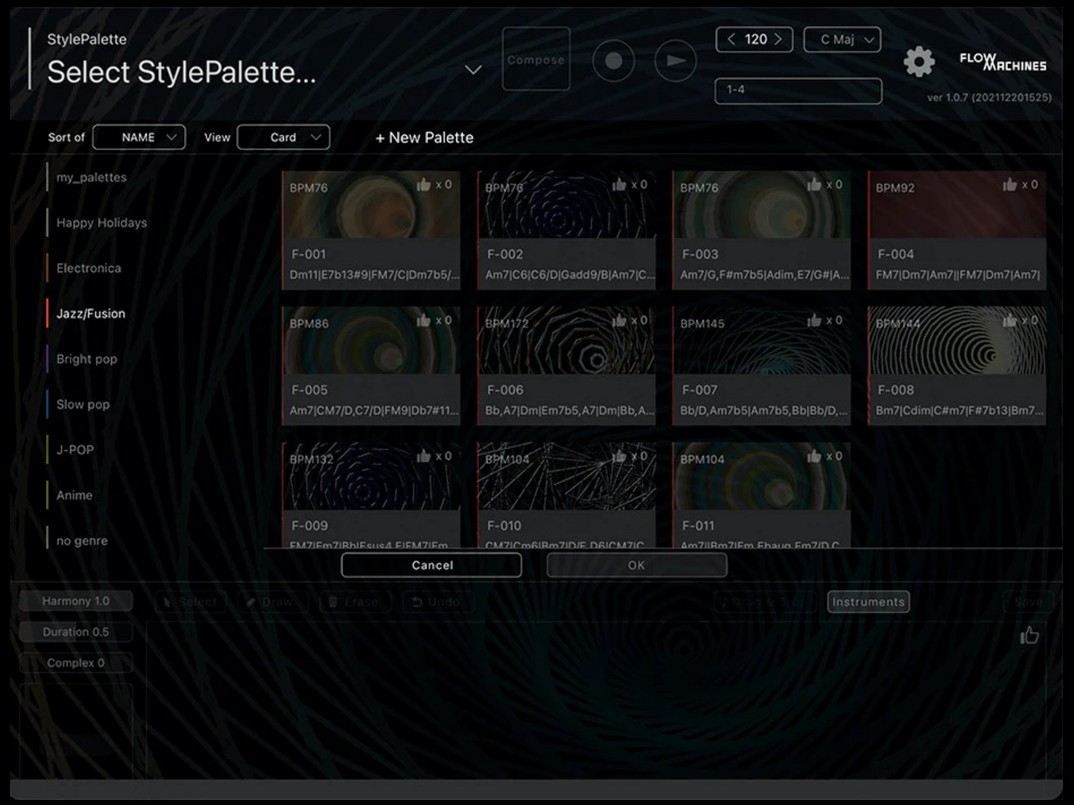
https://pkmital.com/home/works/walt-disney-concert-hall/

OpenAI MuseNet
https://openai.com/research/musenet



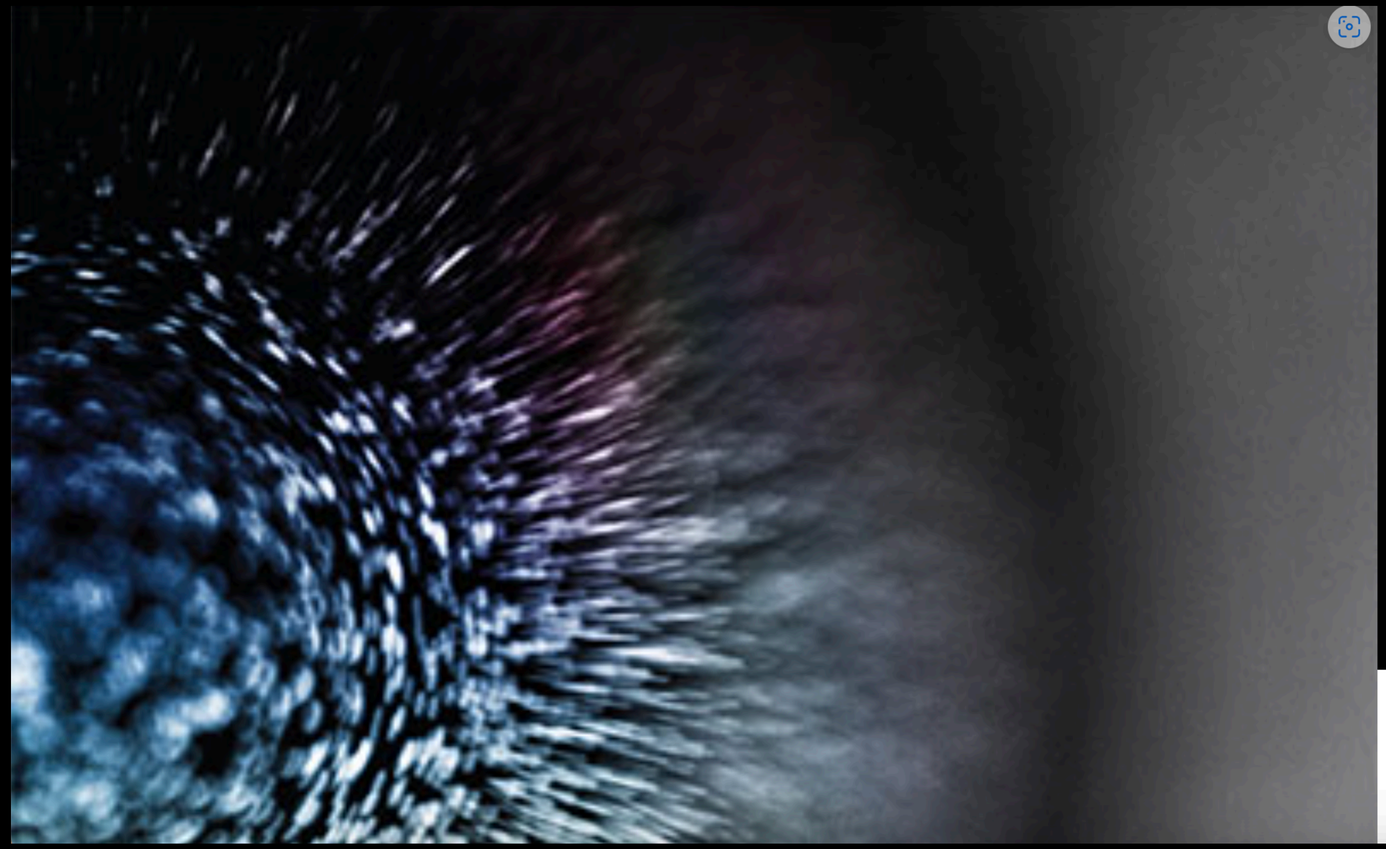
OpenAI Jukebox

https://colab.research.google.com/github/SMarioMan/jukebox/blob/master/jukebox/Interacting_with_Jukebox.ipynb#scrollTo=VVOQ3egdj65y



Sony Flow Machines

https://www.sony.com/en/SonyInfo/design/stories/flow-machines/

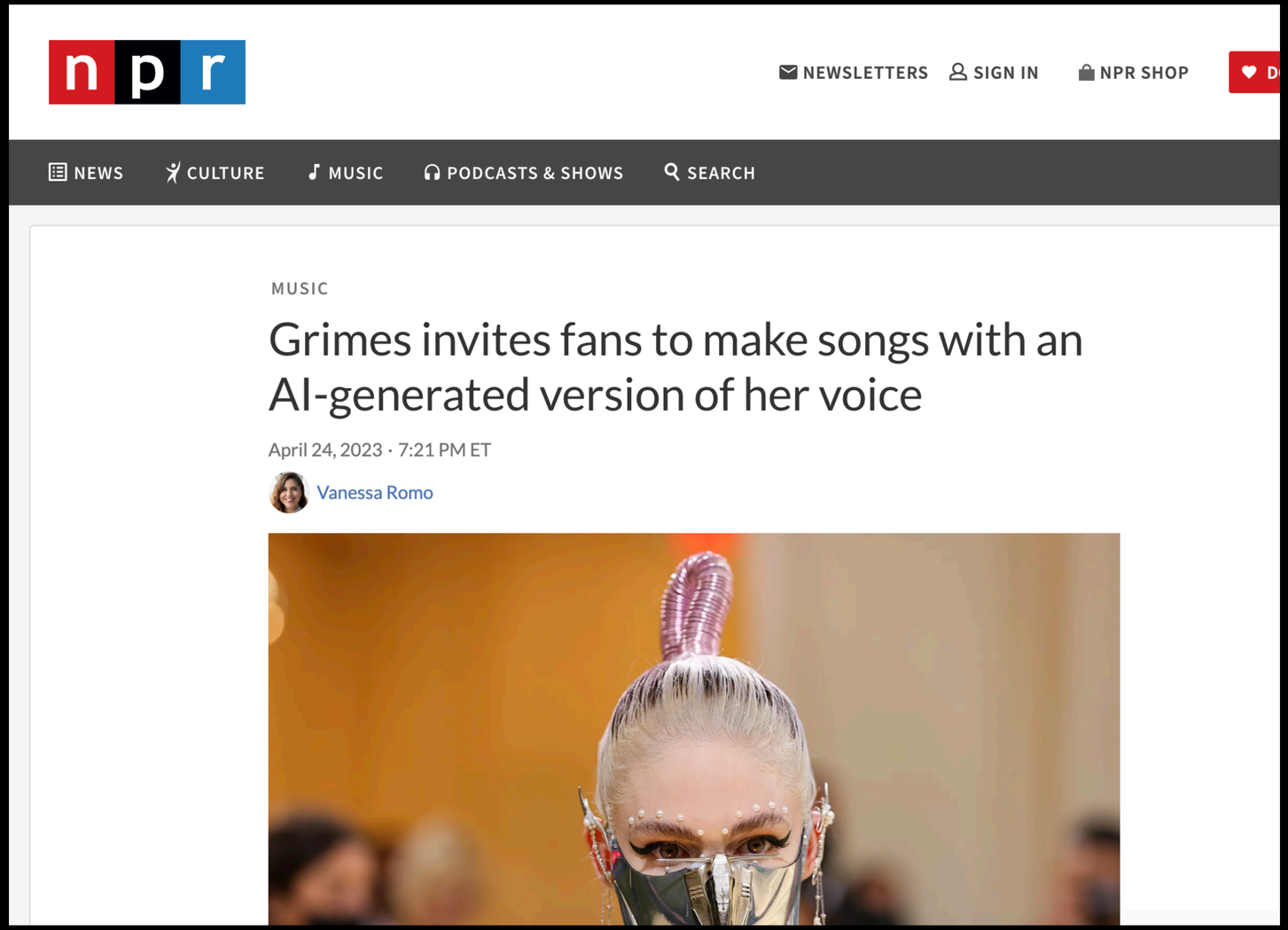


October 8, 2019

EMPTYSET, exploring architecture, algorithms & the material reality of sound

Emptyset
https://clotmag.com/interviews/emptyset-from-the-algorithmic-to-the-architectural-filling-the-spaces-between-our-ears

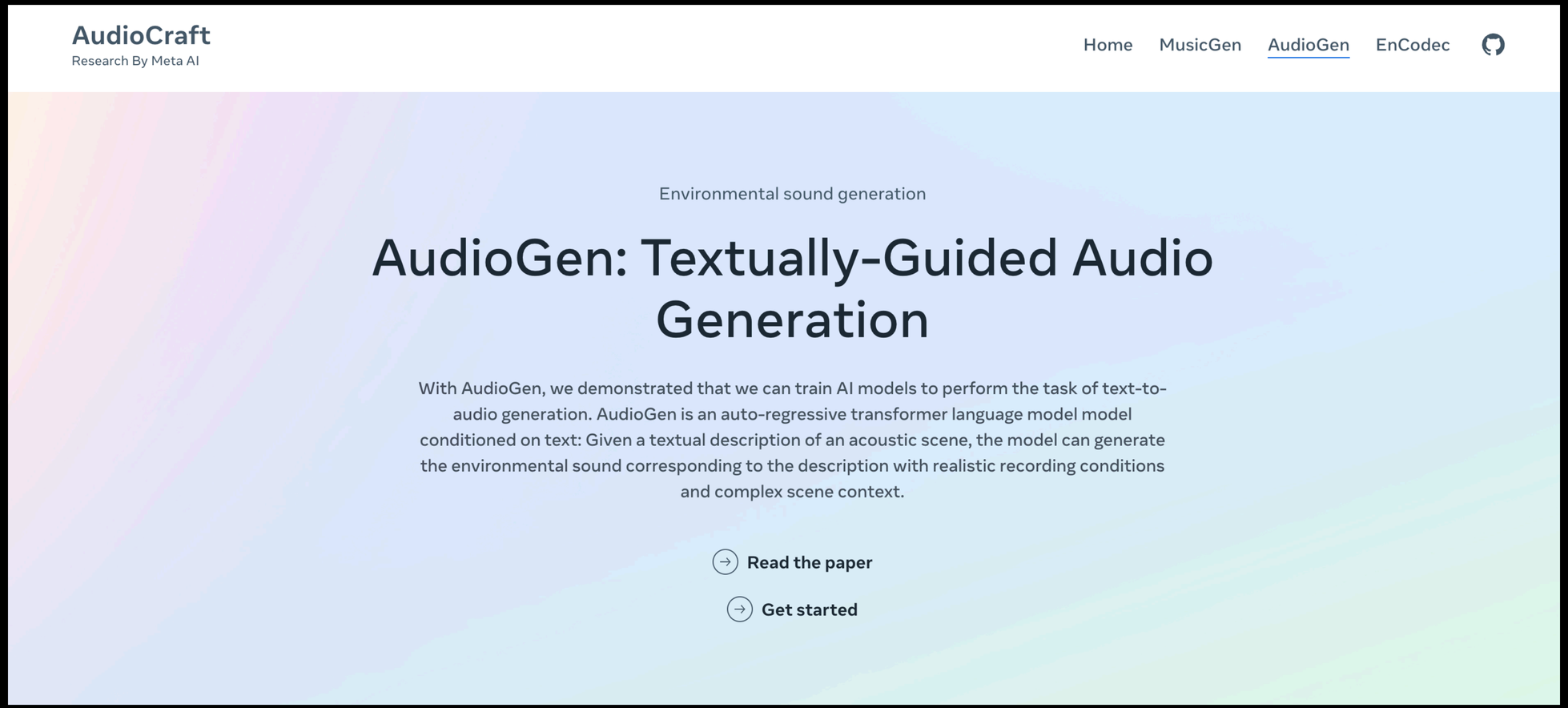
Moises Hexorcismos
https://cdm.link/2023/07/watch-hexorcismos-ai-live/



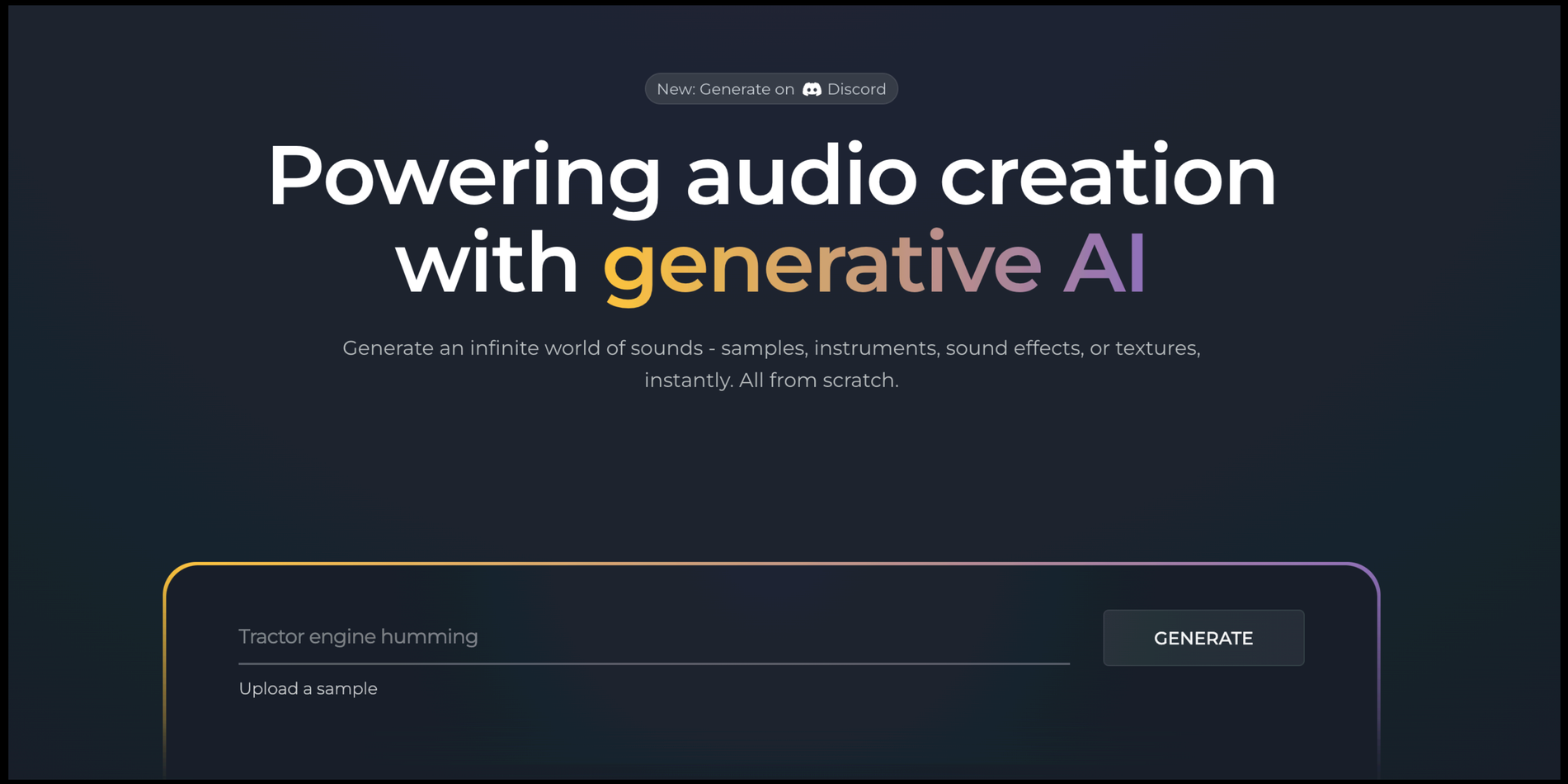
Grimes - AI generated music
https://www.npr.org/2023/04/24/1171738670/grimes-ai-songs-voice

LAB

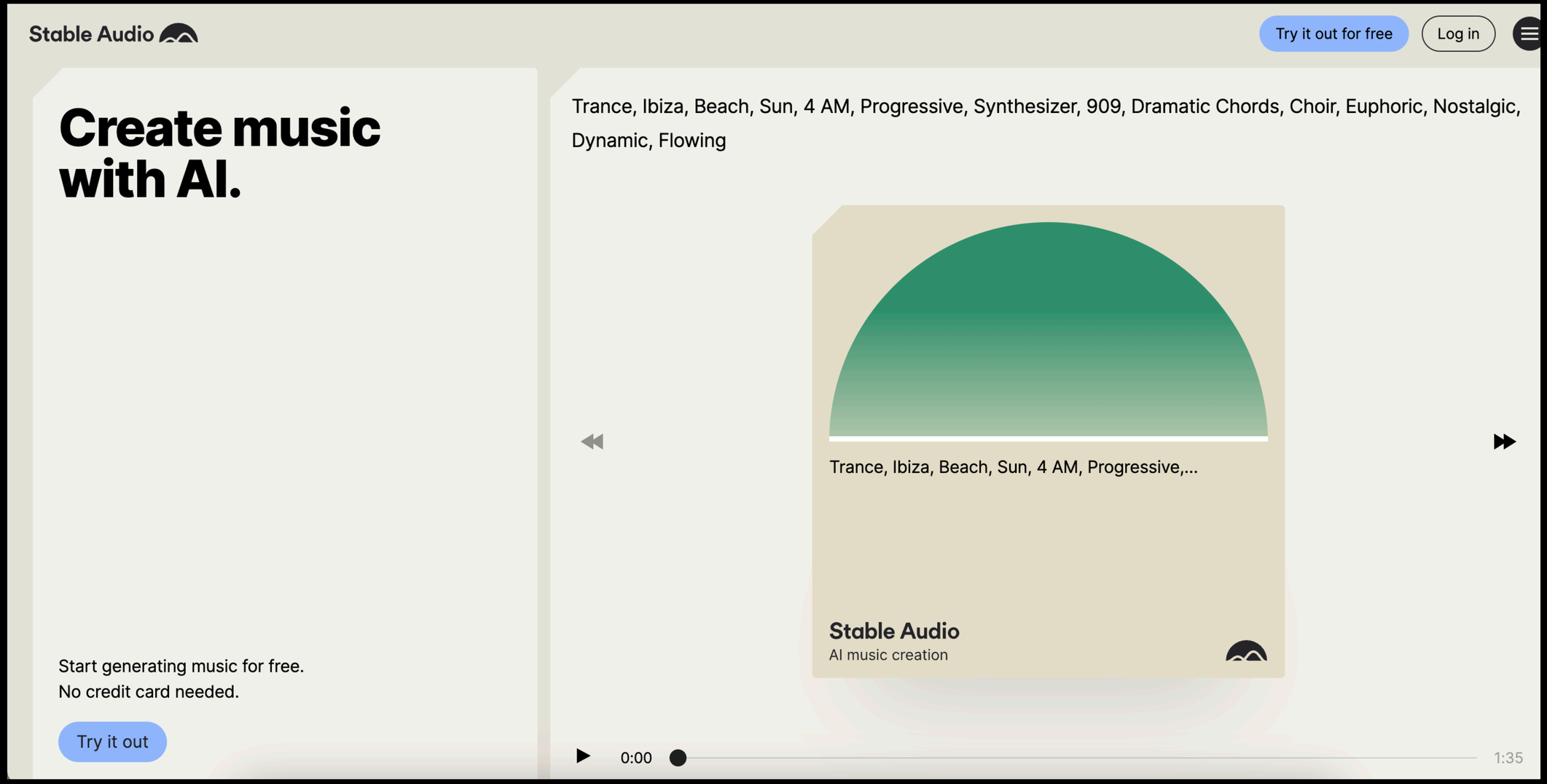
GENERATE MUSIC, SOUND, FX, AND FOLEY



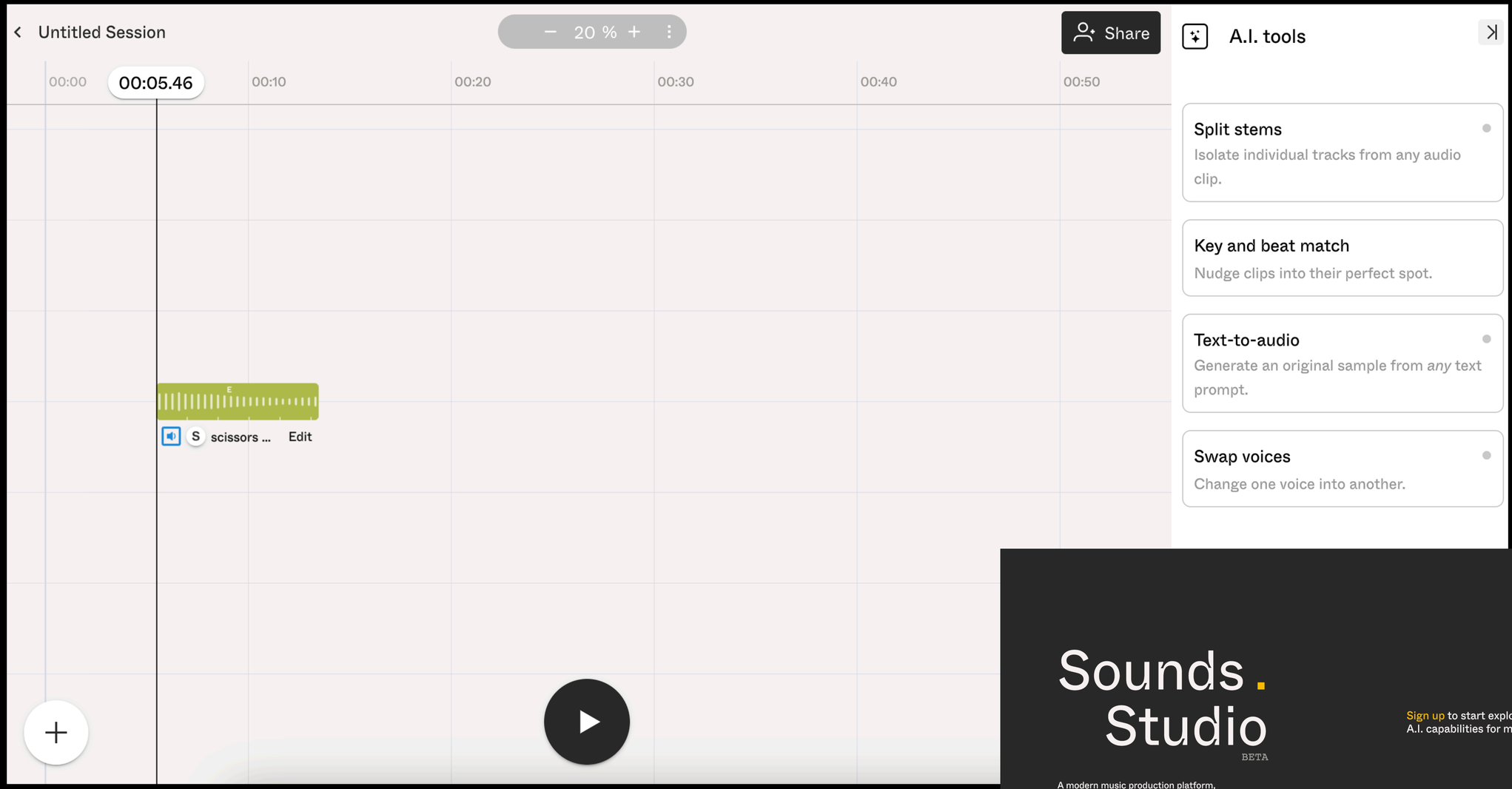
https://audiocraft.metademolab.com/audiogen.html

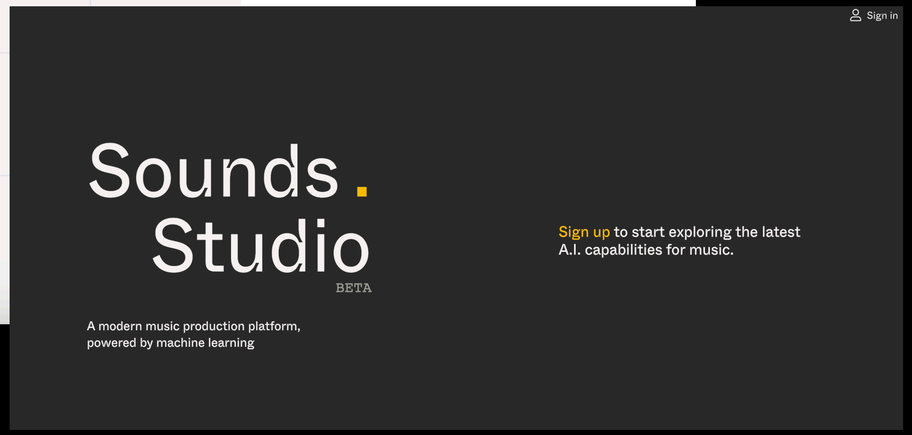


https://www.audiogen.co/

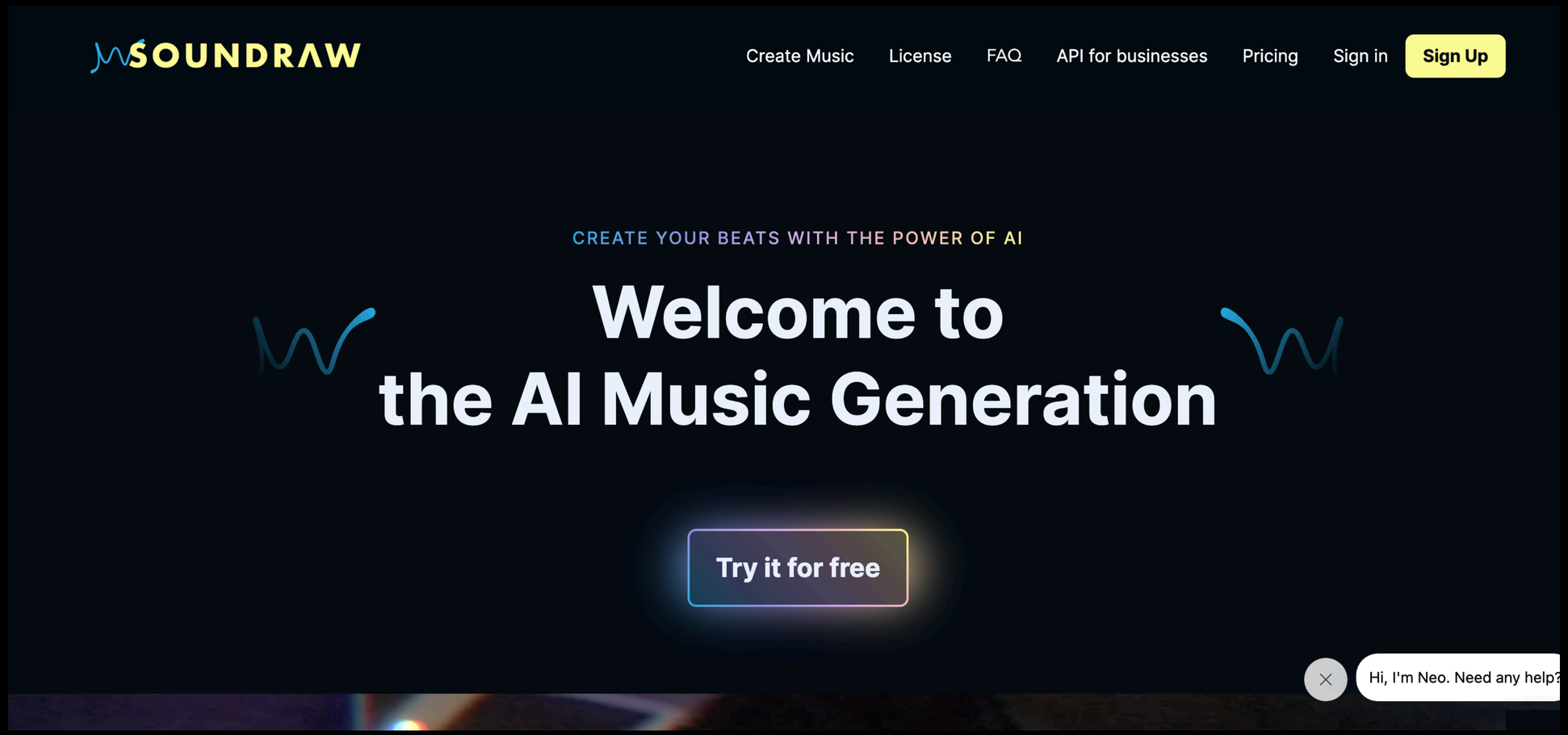


https://stableaudio.com/

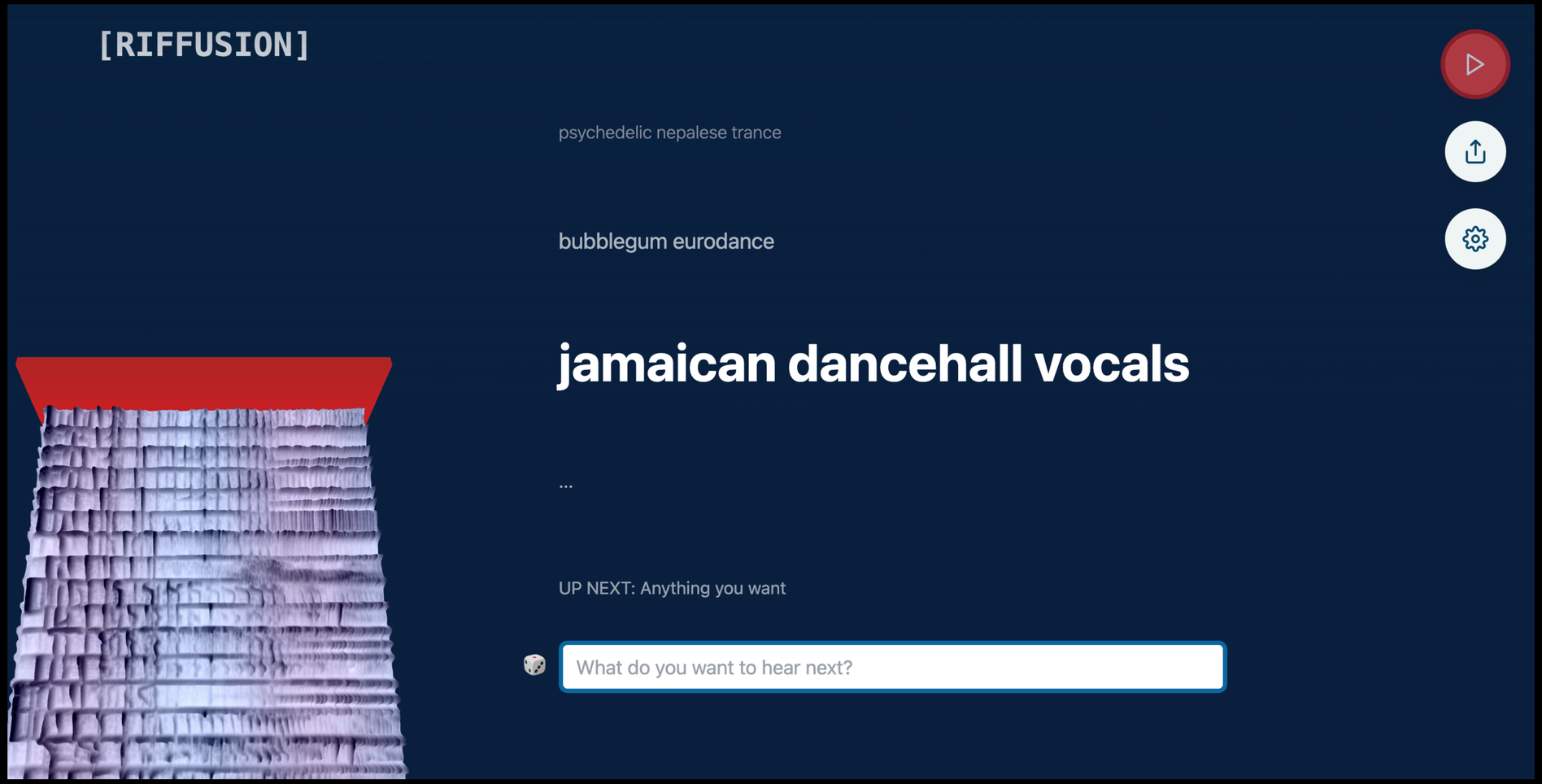




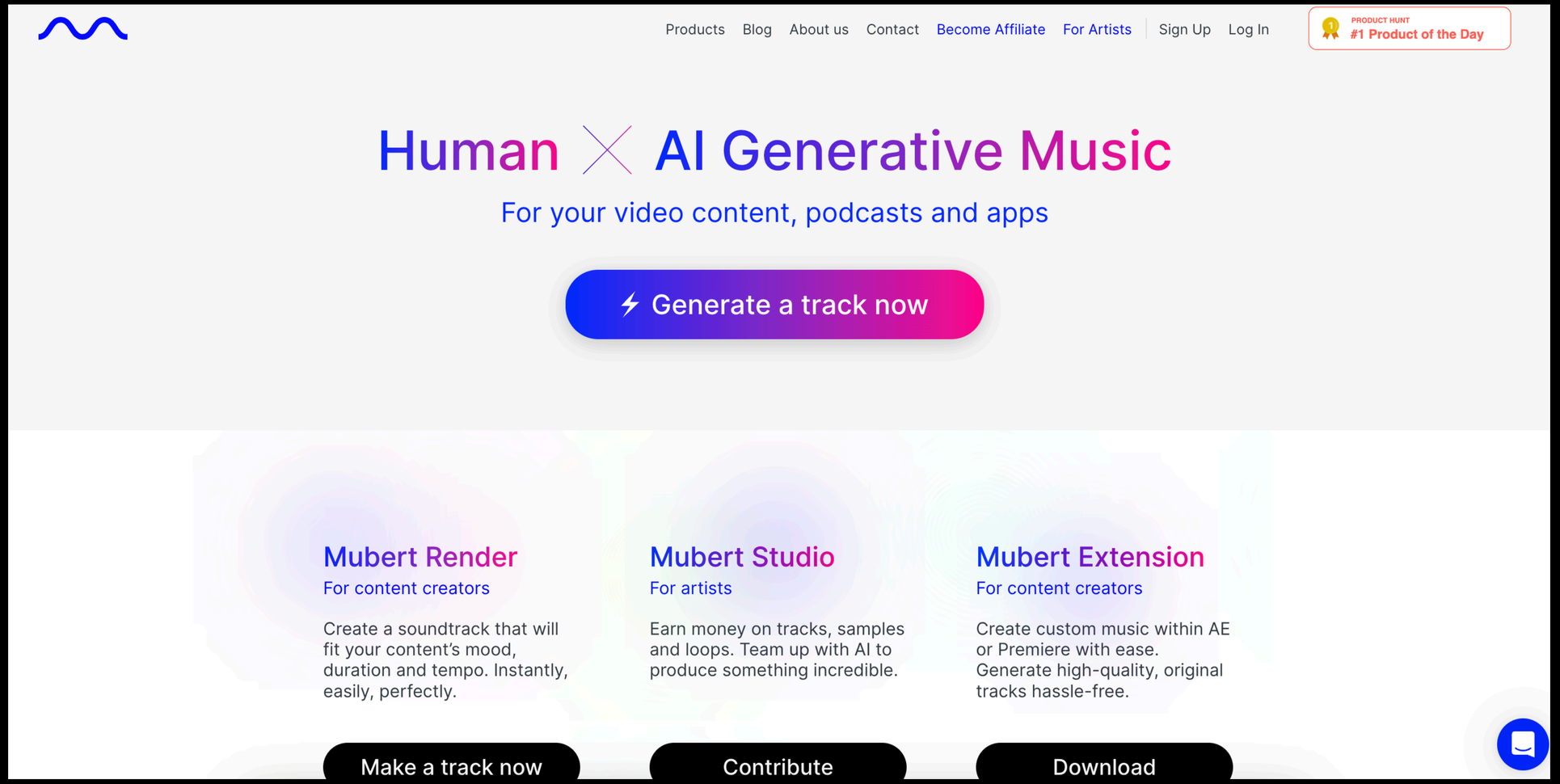
https://sounds.studio/



https://soundraw.io/



https://www.riffusion.com/



https://mubert.com/

AudioLDM: Text-to-Audio Generation with Latent Diffusion Models

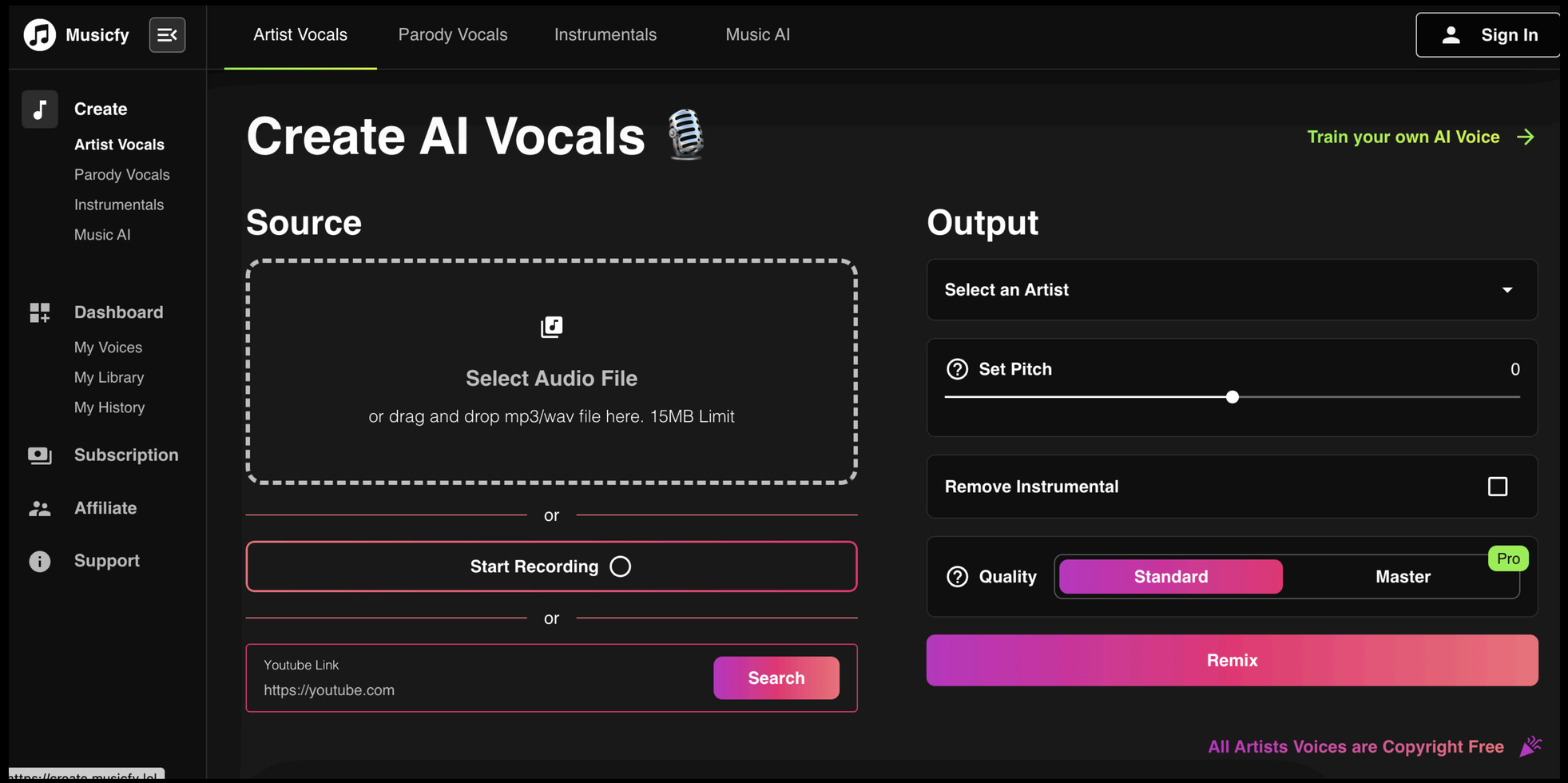
Haohe Liu[*1], Zehua Chen[*2], Yi Yuan[1], Xinhao Mei[1], Xubo Liu[1]
Danilo Mandic[2], Wenwu Wang[1], Mark D. Plumley[1]
[1]CVSSP, University of Surrey, Guildford, UK
[2]Department of EEE, Imperial College London, London, UK
[*]Equal Contribution

[Paper on ArXiv] [Code on GitHub] [Hugging Face Space]

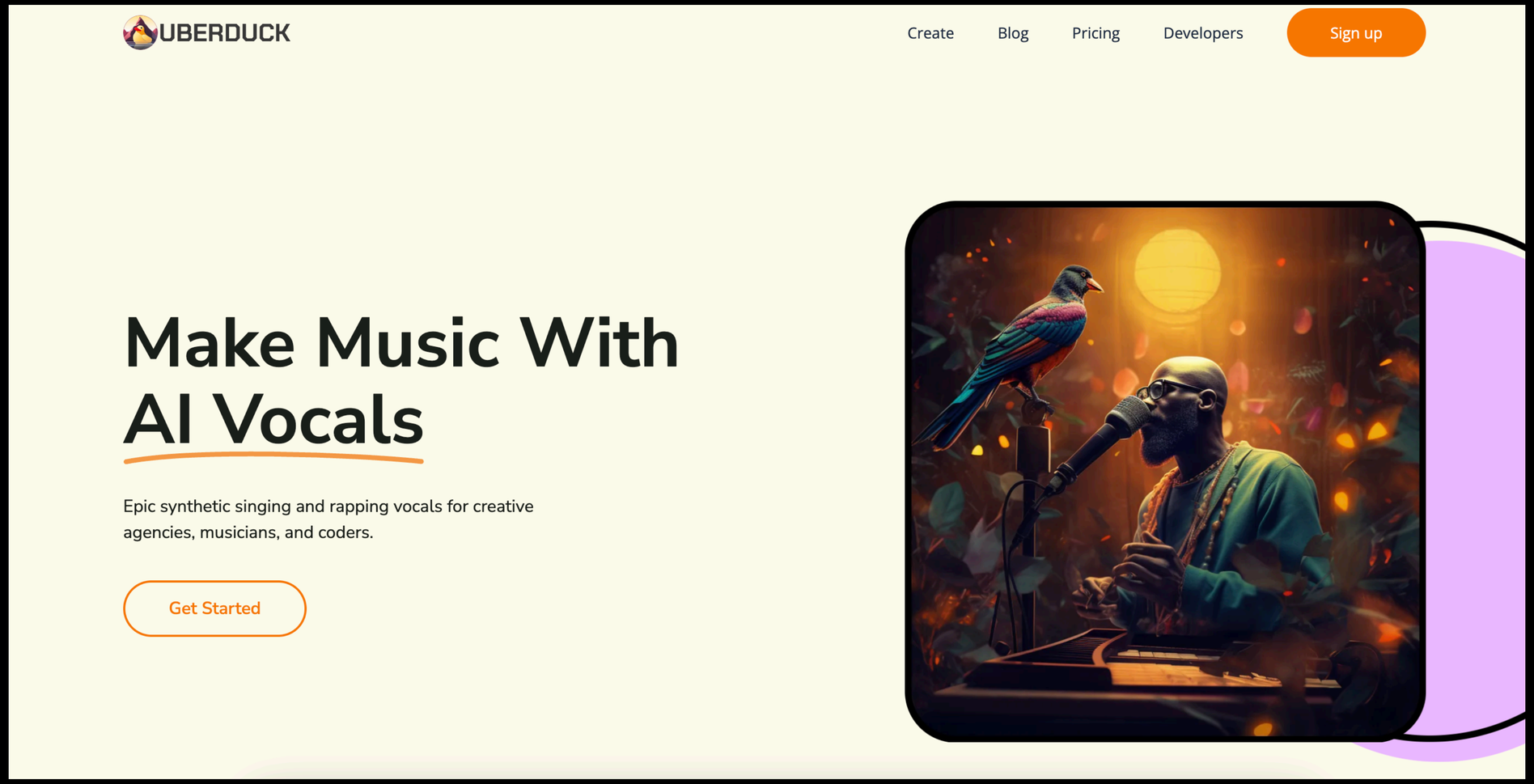
Abstract

Text-to-audio (TTA) system has recently gained attention for its ability to synthesize general audio based on text descriptions. However, previous studies in TTA have limited generation quality with high computational costs. In this study, we propose AudioLDM, a TTA system that is built on a latent space to learn the continuous audio representations from contrastive language-audio pretraining (CLAP) latents. The

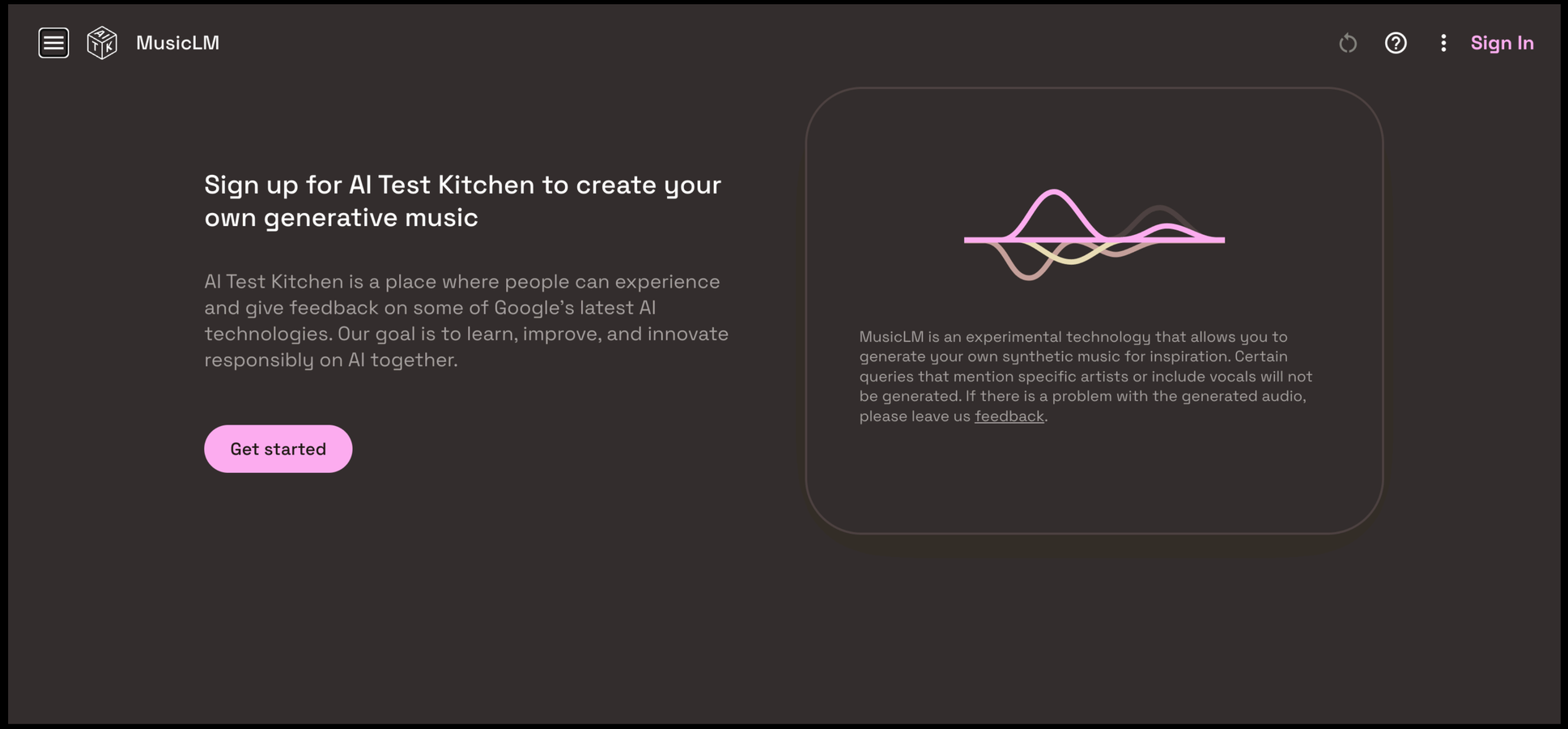
https://audioldm.github.io/



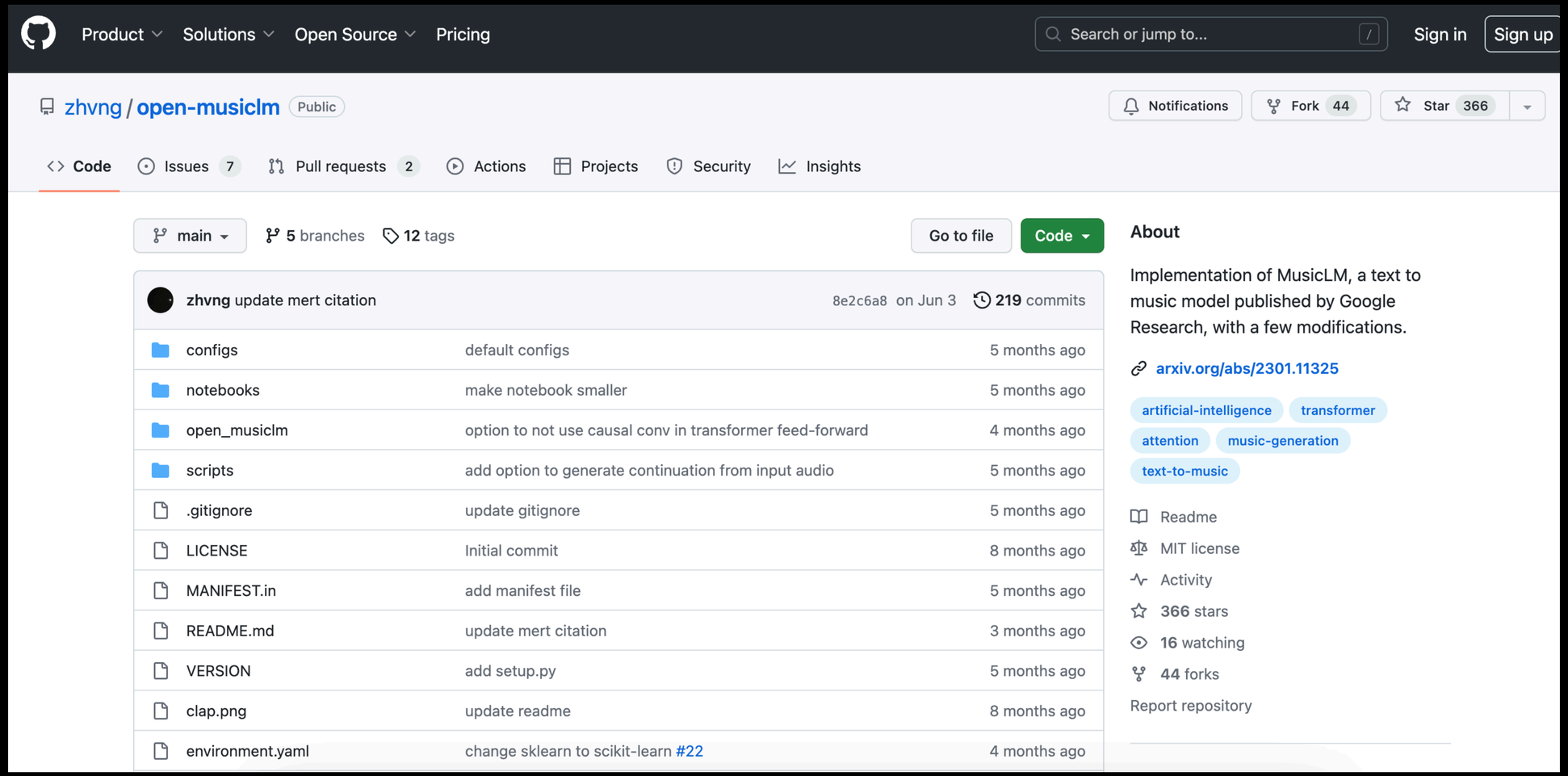
https://create.musicfy.lol/



https://www.uberduck.ai/



https://aitestkitchen.withgoogle.com/experiments/music-lm



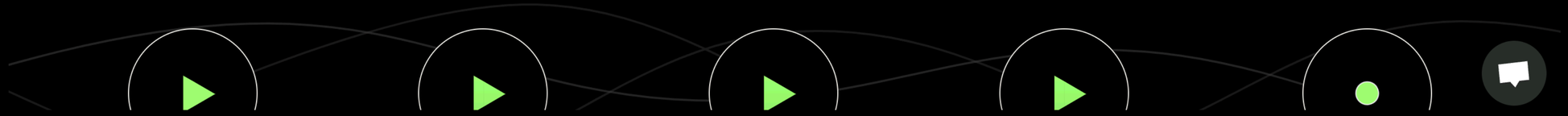
https://github.com/zhvng/open-musiclm

GENERATE VOICES

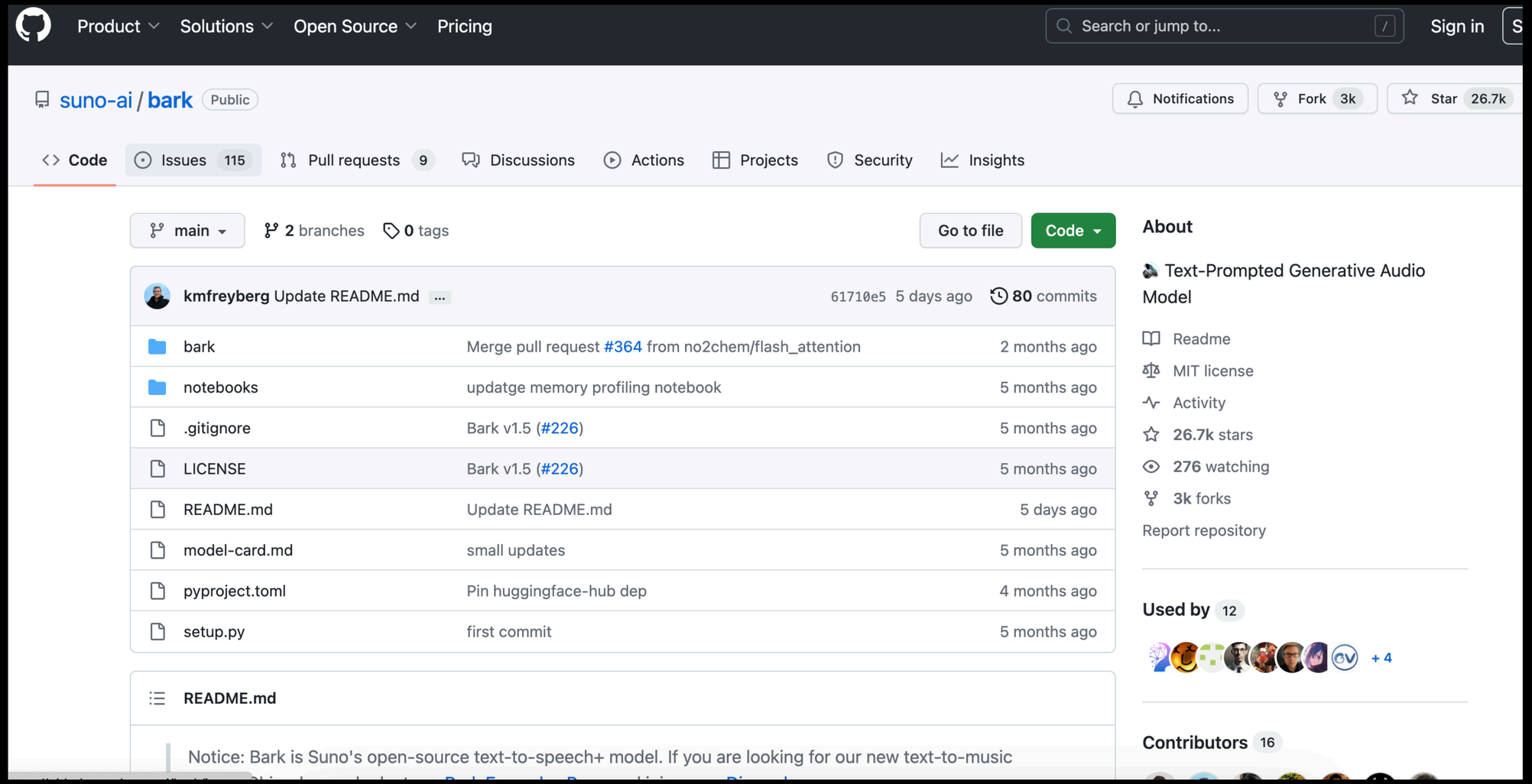
PlayHT Products Use Cases Resources Pricing Log in Try for Free

AI VOICE GENERATOR WITH REALISTIC TEXT TO SPEECH

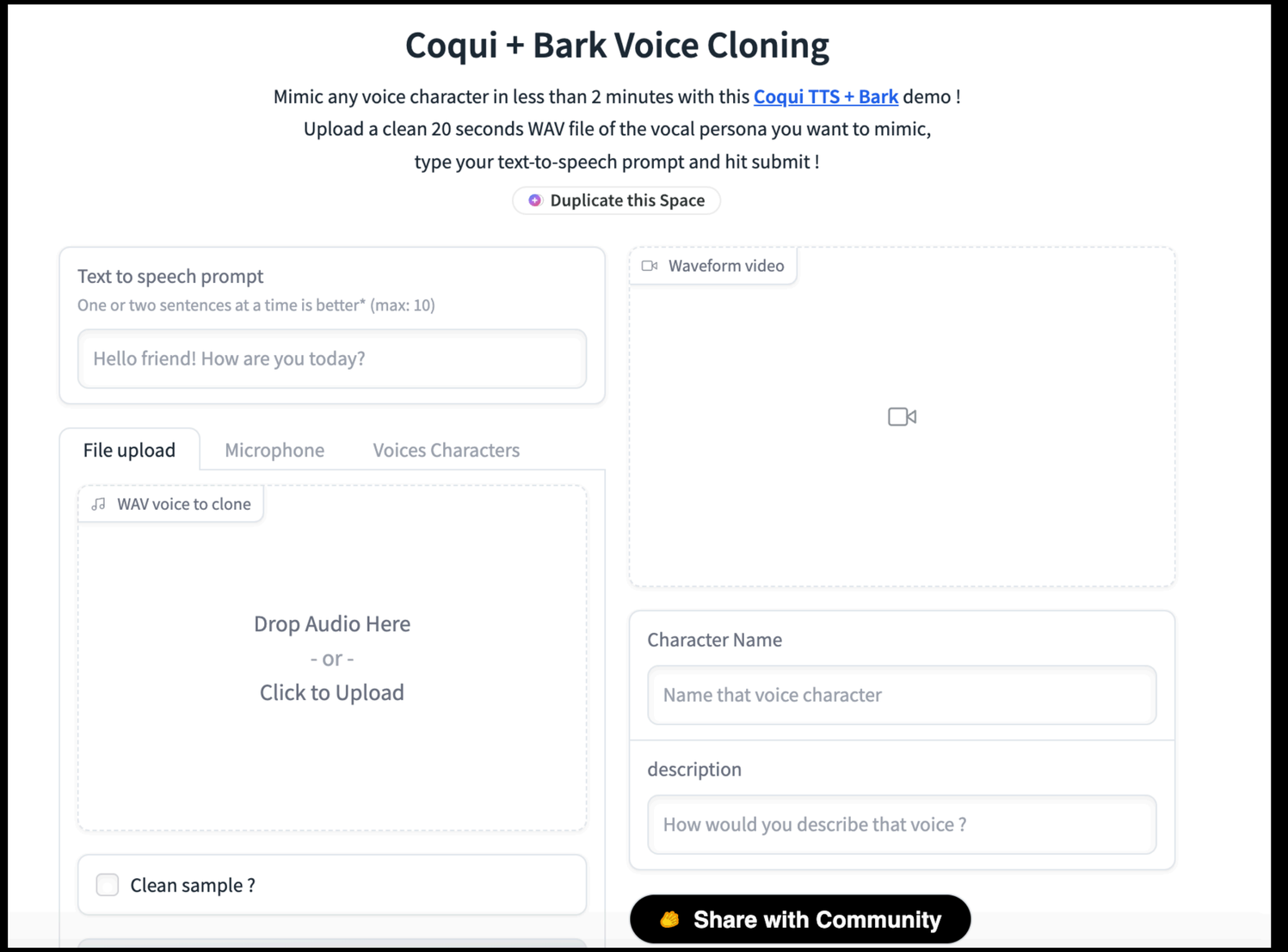
Generate AI Voices, Indistinguishable from Humans



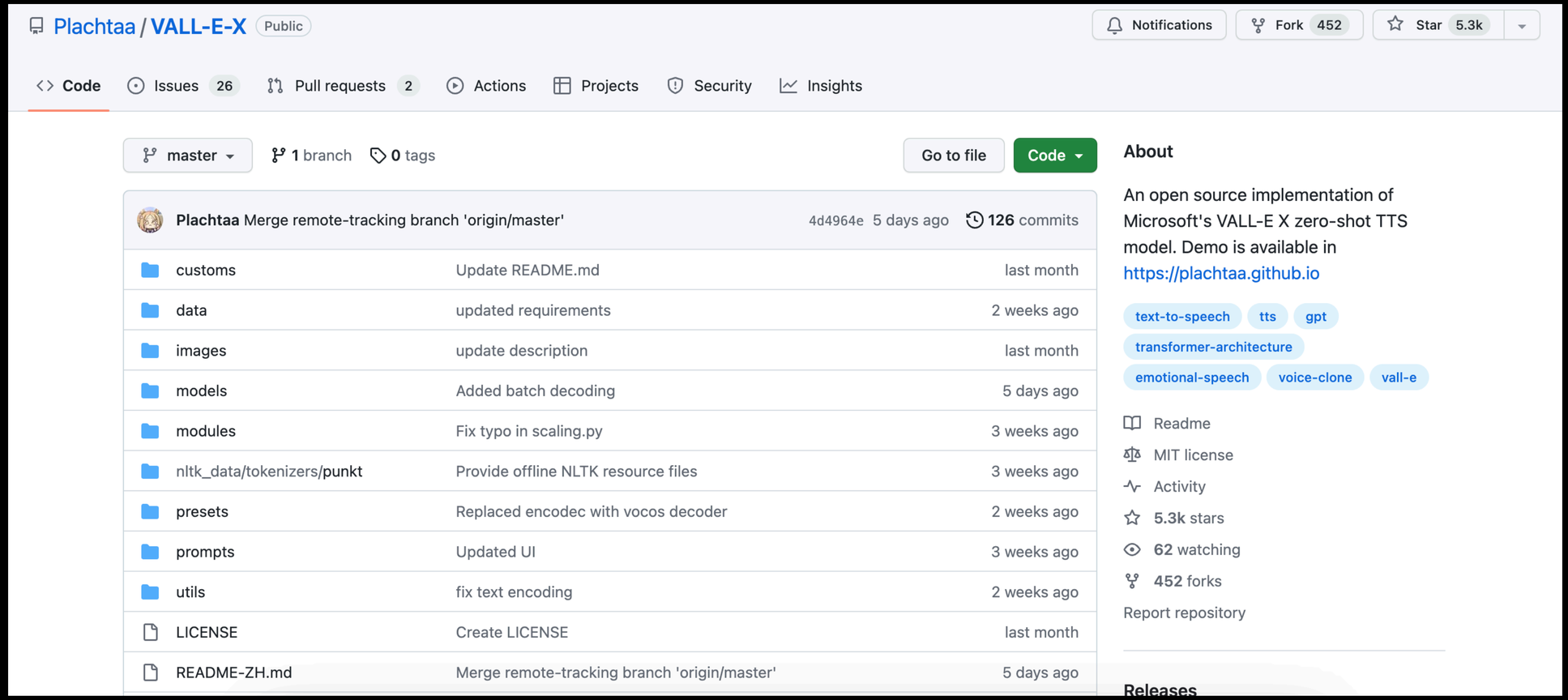
https://play.ht/



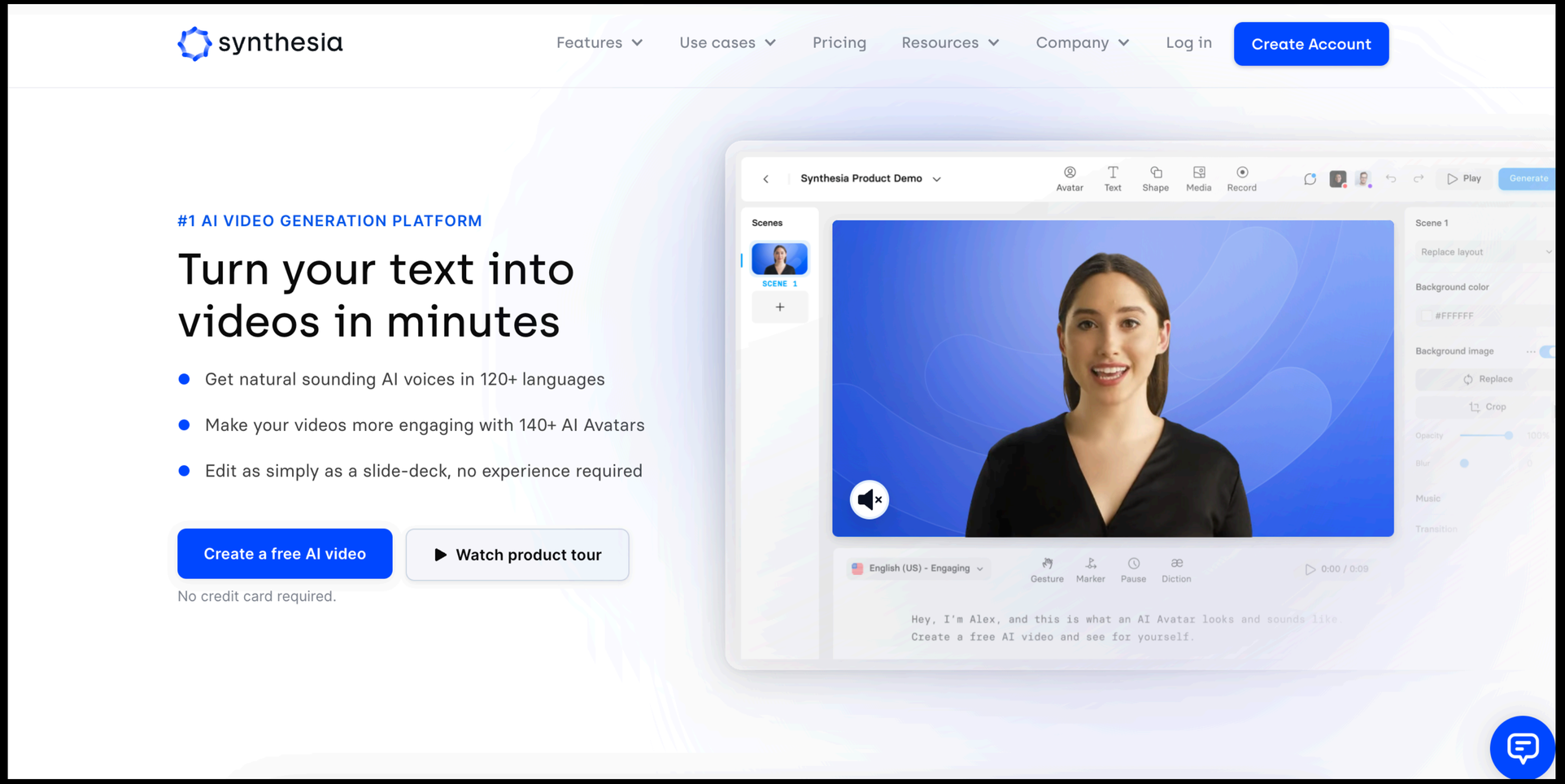
https://github.com/suno-ai/bark



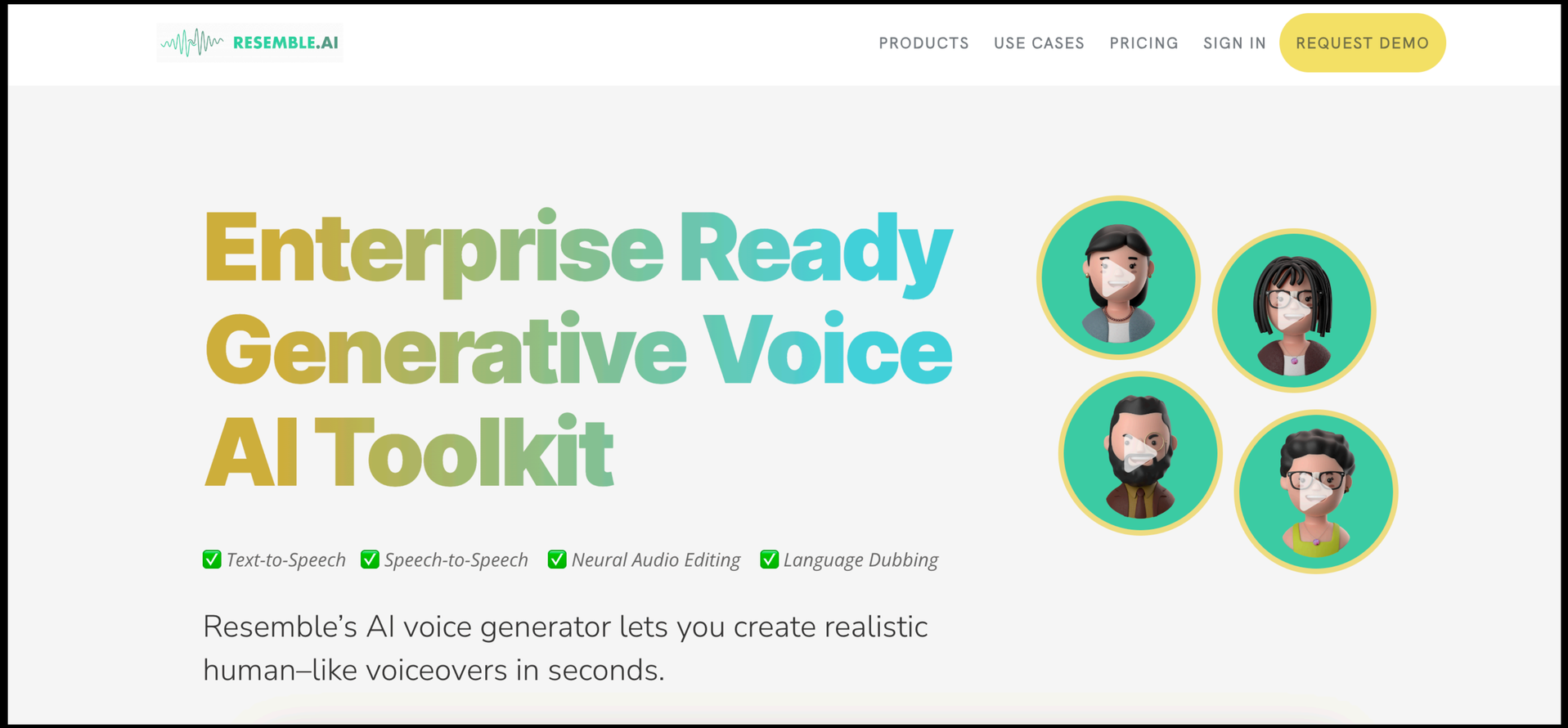
https://huggingface.co/spaces/fffiloni/instant-TTS-Bark-cloning



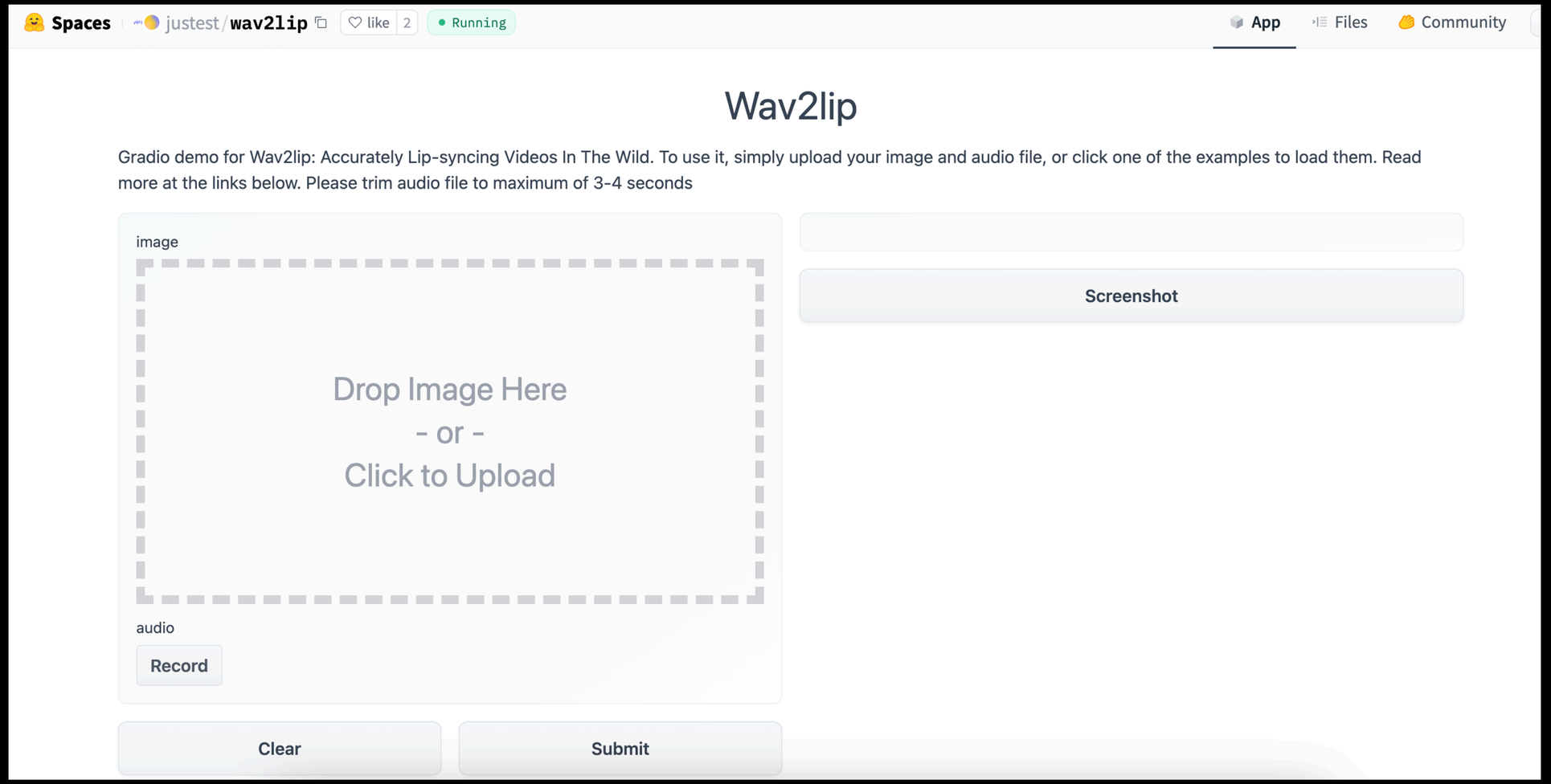
https://github.com/Plachtaa/VALL-E-X



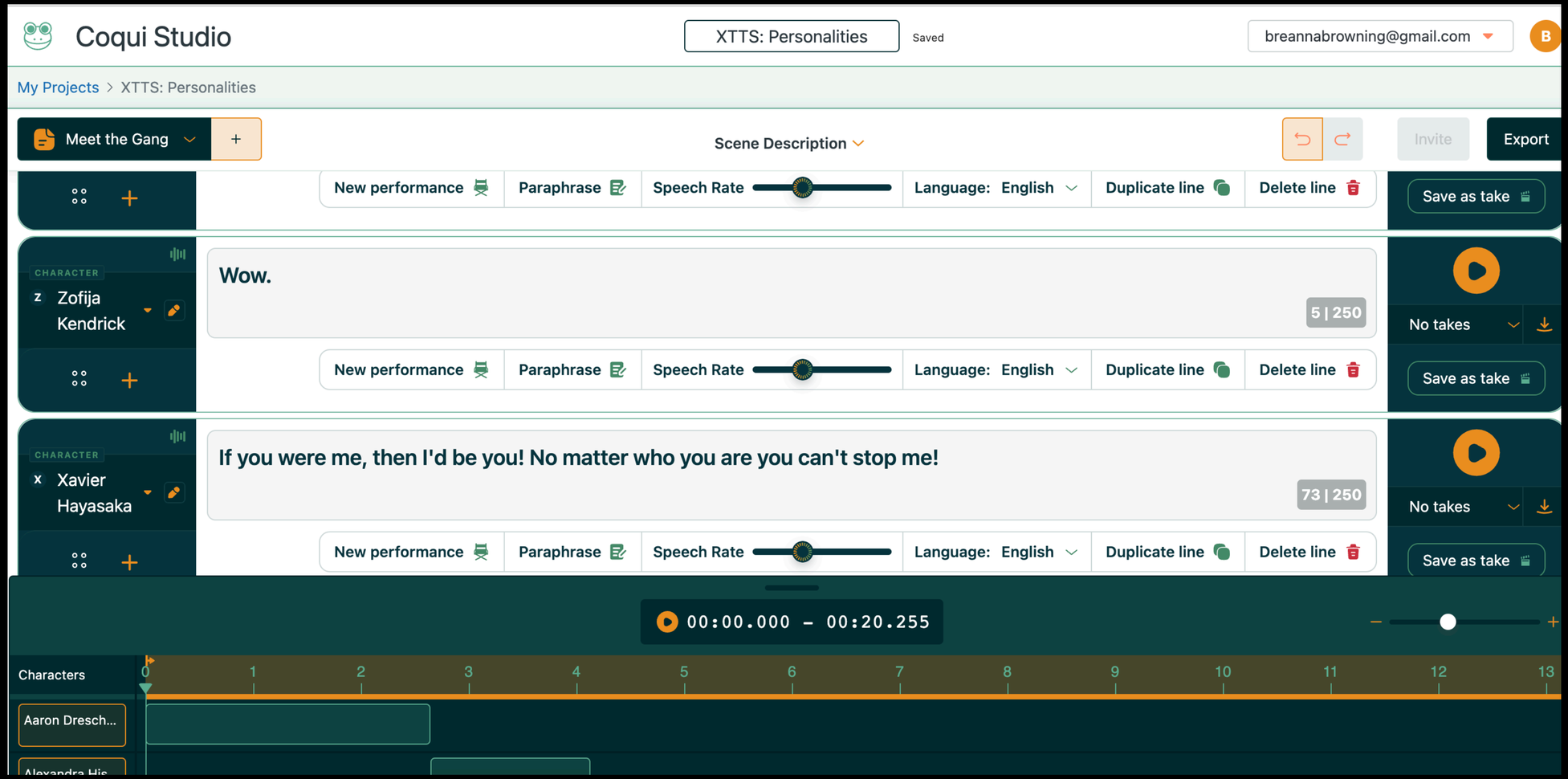
https://www.synthesia.io/



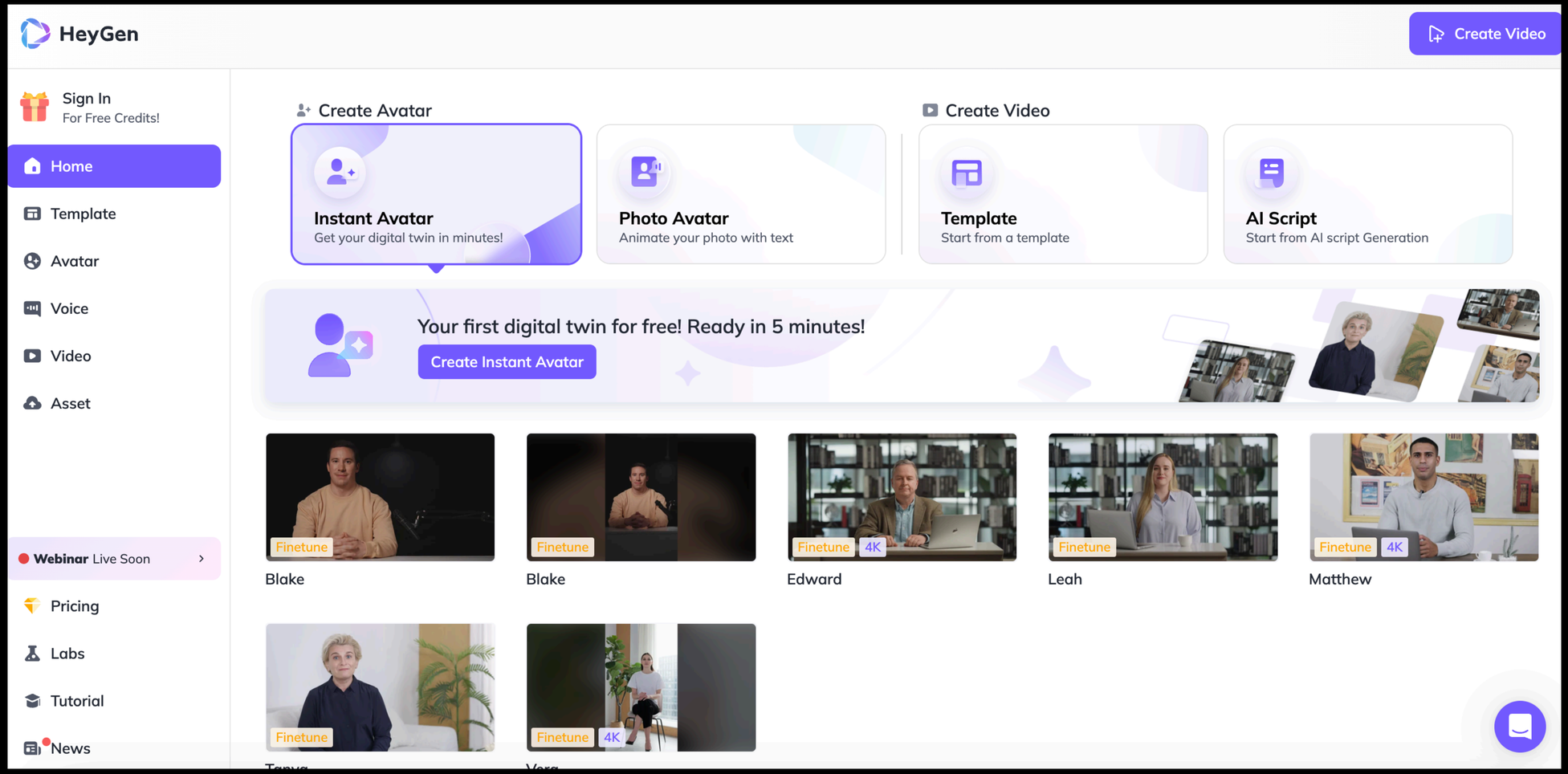
https://www.resemble.ai/



https://huggingface.co/spaces/justest/wav2lip



https://coqui.ai/



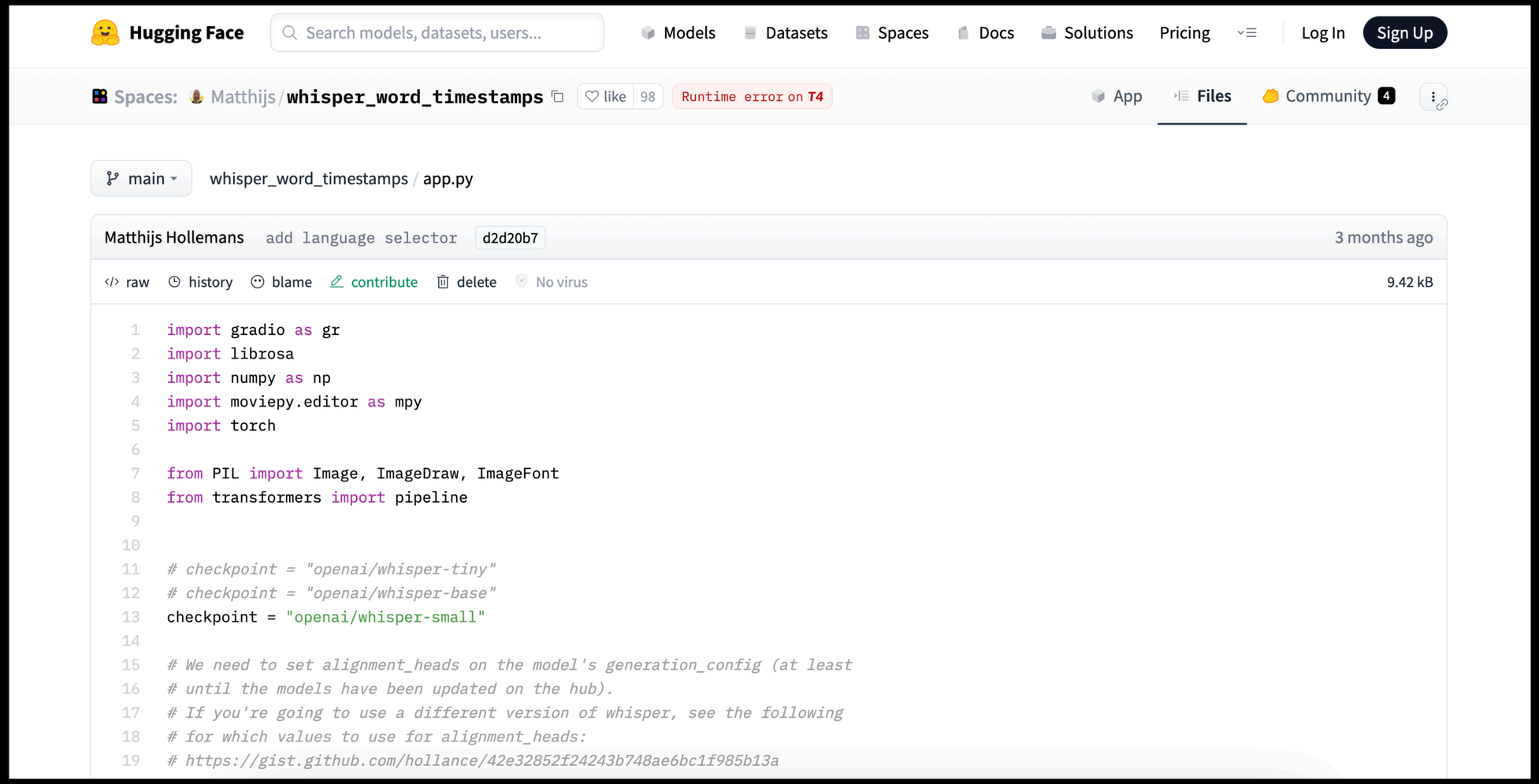
https://app.heygen.com/guest/home

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

Other models for voice generation

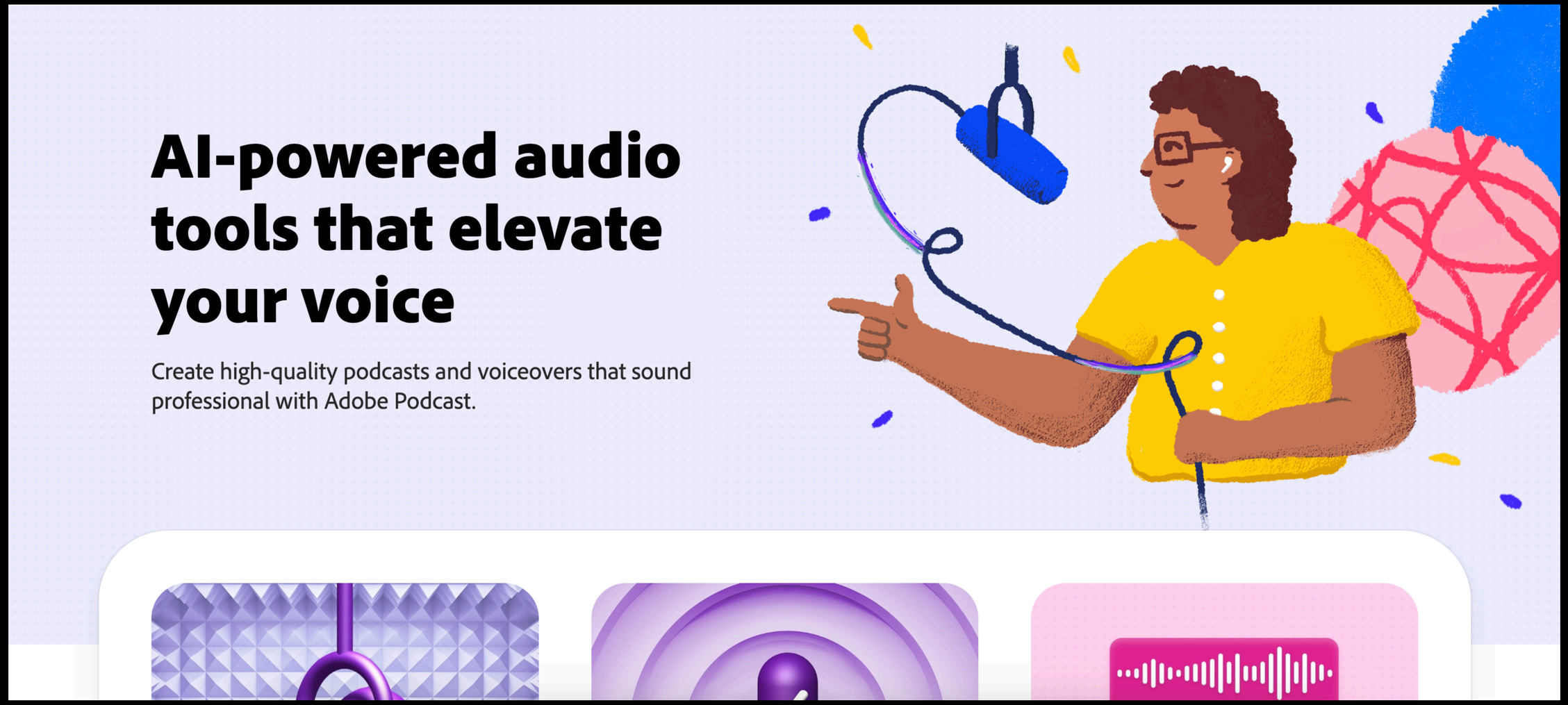
- https://github.com/MycroftAI/mimic3
- https://openreview.net/forum?id=__czv_gqDQt
- https://github.com/heatz123/naturalspeech
- https://github.com/legekka/GanyuTTS
- https://github.com/rendchevi/nix-tts
- https://github.com/mush42/piper-nvda

SEGMENT AUDIO

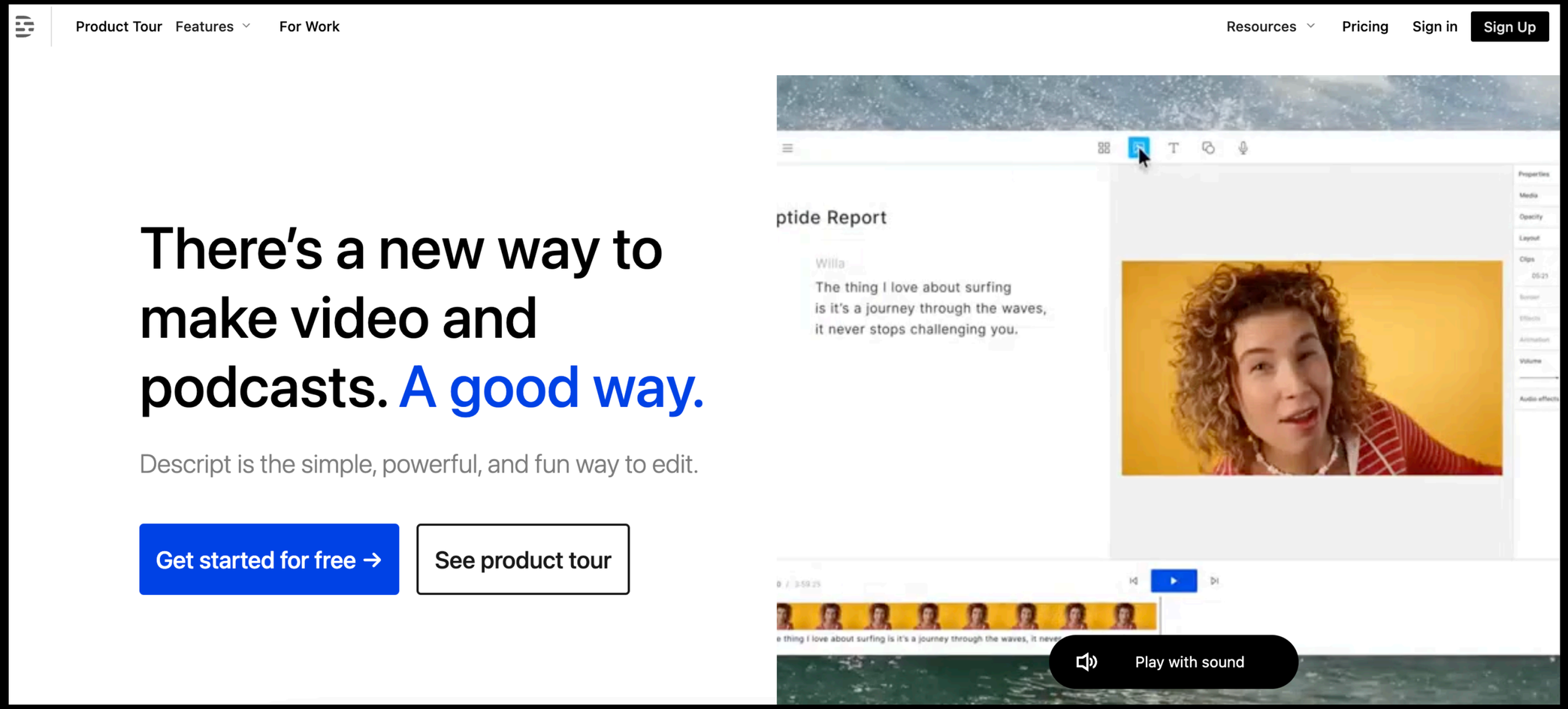


https://huggingface.co/spaces/Matthijs/whisper_word_timestamps/blob/main/app.py

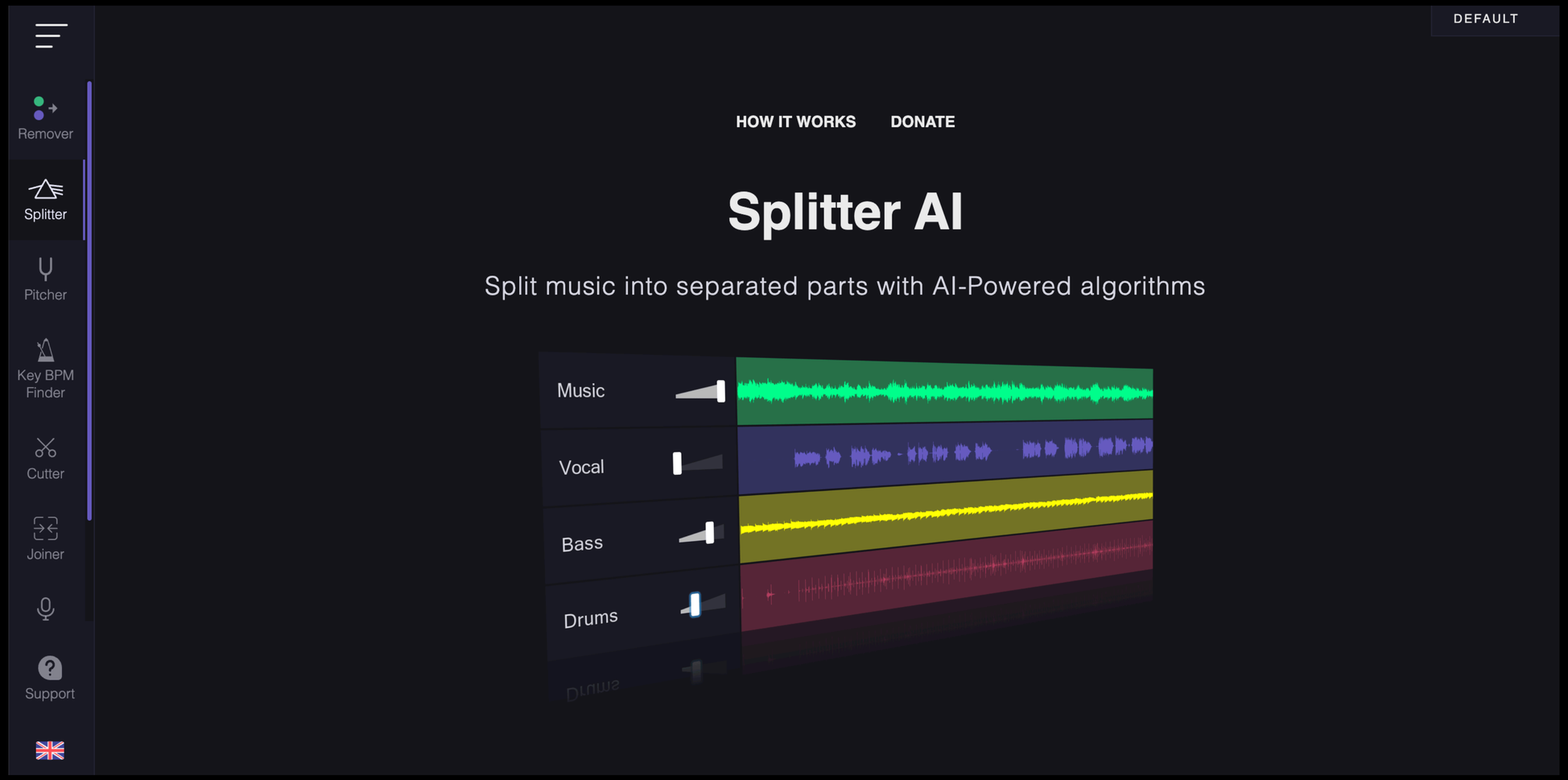
CLEAN VOICES



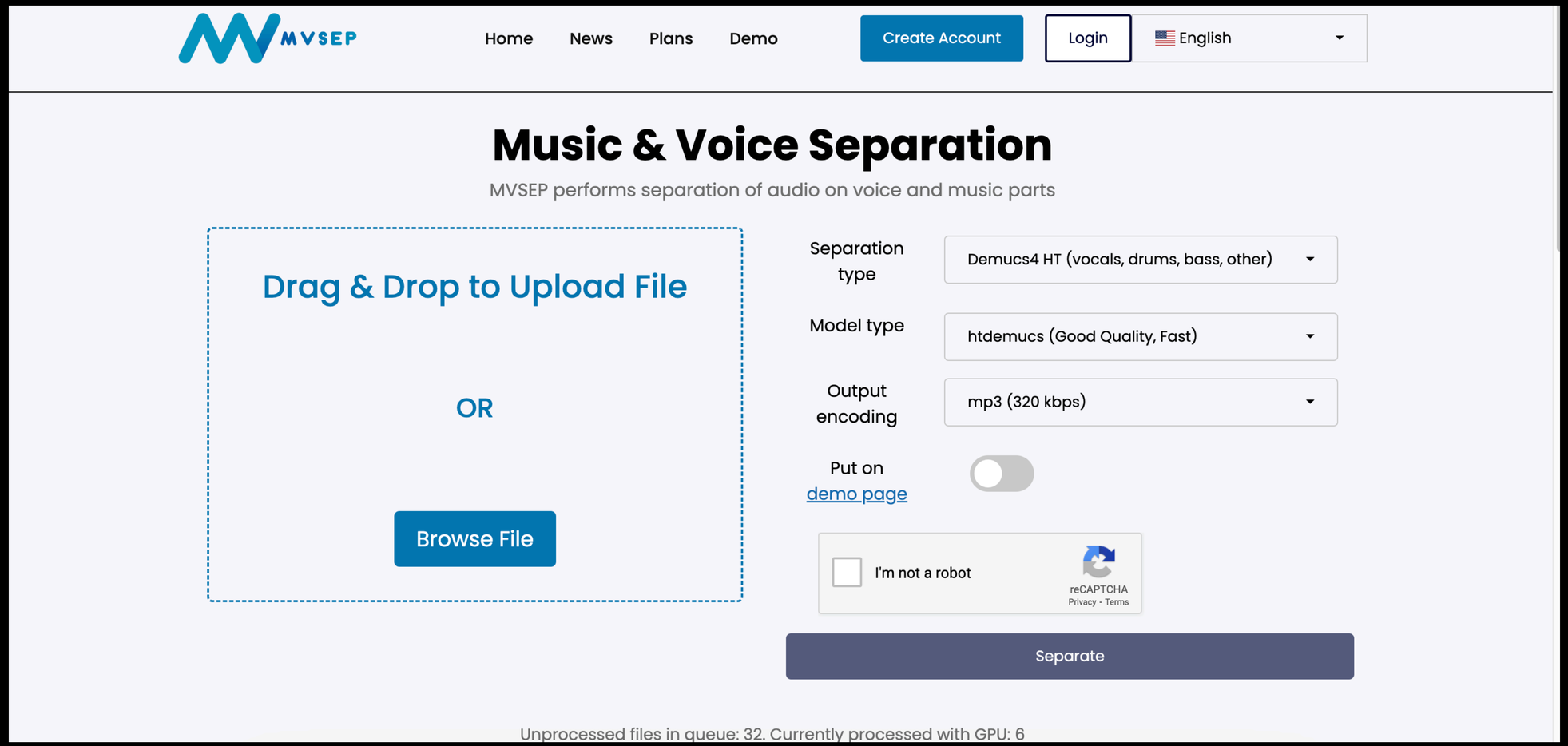
https://podcast.adobe.com/



https://www.descript.com/



https://vocalremover.org/splitter-ai



https://mvsep.com/en

ARCHIVE

Archival resources / audio generation models

- audio style transfer: https://colab.research.google.com/drive/18_DckMGM-fsSthIqZI9sOAlY6I3plDLD#scrollTo=aEH0RfPbasnP
- voice style transfer: https://colab.research.google.com/drive/1tacJzTEJW8Y42kBu-NiYQdVHuOOYbNZ6
- voice style transfer with unpaired collections: https://colab.research.google.com/drive/16t0TVX1D6y941d4uFL32e2eNxeDmB-Xi#scrollTo=V00rptcdKSbq
- speech to text to speech: https://colab.research.google.com/drive/100X8F1uaul6rk1BRbpDZxtLWRIe7CF7r#scrollTo=p-Bi7A11-cll
- speech to text from youtube: https://colab.research.google.com/github/tugstugi/dl-colab-notebooks/blob/master/notebooks/MozillaDeepSpeech.ipynb#scrollTo=fZwM0GtmA7mX
- source separation (complex to stems, including voice): https://colab.research.google.com/github/deezer/spleeter/blob/master/spleeter.ipynb#scrollTo=f8Brdfh6mzEz
- source separation (complex to stems, including voice): https://colab.research.google.com/drive/1mijF0zGWxN-KaxTnd0q6hayAlrID5fEQ
- openai jukebox: https://colab.research.google.com/github/SMarioMan/jukebox/blob/master/jukebox/Interacting_with_Jukebox.ipynb#scrollTo=65aR2OZxmfzq

Further approaches

- **AIVA:** An AI-powered software that can compose classical music in a range of styles
- **Amper Music:** An AI-powered tool that allows users to create custom music tracks by selecting parameters like mood, style, and pace
- **Sonix:** An AI transcription tool that automatically transcribes audio recordings into written text
- **Descript:** A tool that uses AI to produce accurate transcripts that you can edit by editing the text, which in turn edits the audio

Resources

- Cycling 74
- Pure Data
- Bob Sturm
- Magenta
- GRUV
- Nick Collins
- Ali Momeni Max MSP
- Rebecca Fiebrink Wekinator and Kadenze Course
- NIME
- Neurips Creativity Workshop
- ProcGen
- ISEA
- MAH

HOMEWORK

Homework

- No lab / student presentations due to class cancellation Wednesday October 25
- Students should explore audio generation tools independently, considering their application in final projects