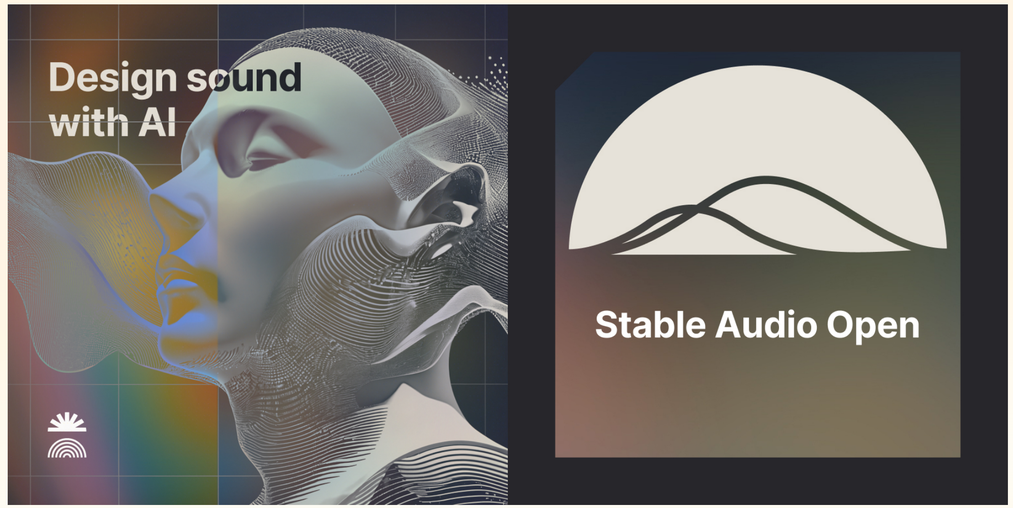
STABLE AUDIO OPEN

STABILITY AI

STABLE AUDIO OPEN

- Open-weights text-conditioned diffusion model and autoencoder for sound & music.
- ‘Production-quality’ i.e. 44.1kHz stereo
- Trained on CC-0 and CC-BY data from Freesound and the Free Music Archive.
- Base model concentrates on short form generation of sound effects, musical samples up to 47s i.e. building blocks in the production process not final results.



WHY IS IT INTERESTING TO THIS COMMUNITY?

- Competitive with state-of-the-art in terms of audio quality and prompt coherence
- Very permissive license (even for small commercial operations)
- Documented dataset including attribution
- Great starting point for further work
- Active community developing new variations + techniques

POSSIBILITIES

- Use audio-input functionality to transform existing sample content
- Fine-tune Stable Audio Open base model on your own small datasets (LoRA, DoRA etc)
- Build other types of generative model based on the autoencoder.
- Experiment with new types of control.
- Use for sample production in a larger AI pipeline.

RESOURCES

- Model weights on Huggingface (https://huggingface.co/stabilityai/stable-audio-open-1.0)
- Code in Stable Audio Tools (https://github.com/Stability-AI/stable-audio-tools)
- Discussion / guidance on the Harmonai Discord server (https://discord.gg/TmqzxSeav4)
- Great livestream showing how to fine-tune, by *Lyraaaa* (https://www.youtube.com/watch?v=ex4OBD_lrds)