PAT 464/564 (Winter 2026)

Generative AI for Music & Audio Creation

Lecture 1: Introduction

Instructor: Hao-Wen Dong

UNIVERSITY OF MICHIGAN

Welcome! Tell Us about Yourself!

- Name
- Pronouns
- Program/year
- What is your **most familiar instrument** (if any)?
- What is your **most familiar programming language** (if any)?
- Have you ever used any **AI Music tools**? Which tools?
- What are you **expecting from this course**?

About Me

- Hao-Wen (**Herman**) Dong
- Pronouns: he/him
- Email: **hwdong**@umich.edu
- Office: **Stearns 131** (15 min walk to the north from Moore)
- Office hours: By appointments (link on the course website)
- Research areas: Generative AI for music, audio, and video

Generative AI for Music & Audio Creation

Generative AI for Music & Audio Creation

Generative AI for **Music & Audio Creation**

What is this course all about?

An introduction to generative AI and its applications to music and audio creation. Topics include **music generation**, **audio synthesis** and **AI-assisted music creation tools**.

Topics

- **Deep Generative Models**
  - RNN, LSTM, transformer, VAE, GAN, diffusion model, latent diffusion model …
- **Symbolic Music Generation**
  - Automatic music composition, arrangement, improvisation …
- **Audio Synthesis**
  - Performance synthesis, text-to-audio synthesis, video-to-audio synthesis …
- **AI-assisted Music Creation Tools**
  - Neural audio effects, auto mixing, interactive tools, audiovisual tools …

Learning Objectives

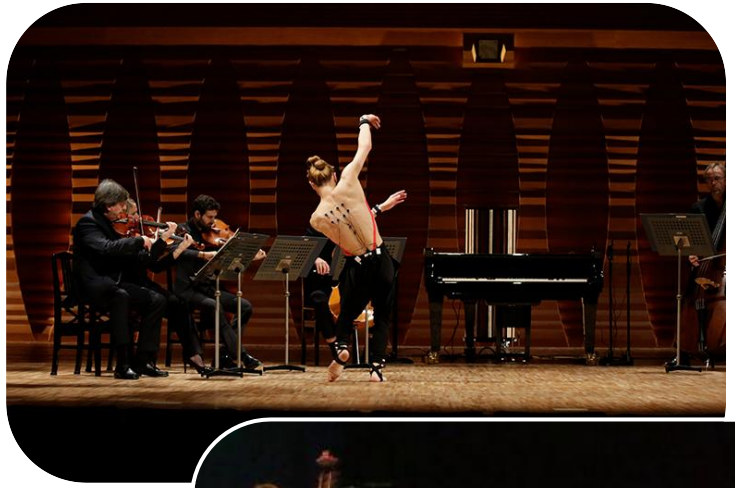
- Gain an **overview** of the field of AI music and audio creation
- Learn **fundamental AI concepts and principles**
- Learn **representative AI tools** for music and audio creation
- Gain **hands-on experience** on creating music using AI tools

Music & Technology Co-evolves

Hildegard Dodel, Public domain, via Wikimedia Commons.
Taken at Hamamatsu Museum of Musical Instruments, August 2019.
yan, CC BY-SA 4.0, via Wikimedia Commons.

Music & AI

(Source: Yamaha)



(Source: Sankei Shimbun)

(Source: Robot Gizmos)

(Source: NBC DFW)

yamaha.com/en/news_release/2018/18013101/
sankei.com/article/20240113-CQCOSQHJWFIYPJJKZDCITRTRVI/
roboticgizmos.com/shimon-musical-robot-deep-learning/
nbcdfw.com/entertainment/the-scene/how-verdigris-ensemble-is-using-ai-to-create-a-new-concert-experience/3366031/

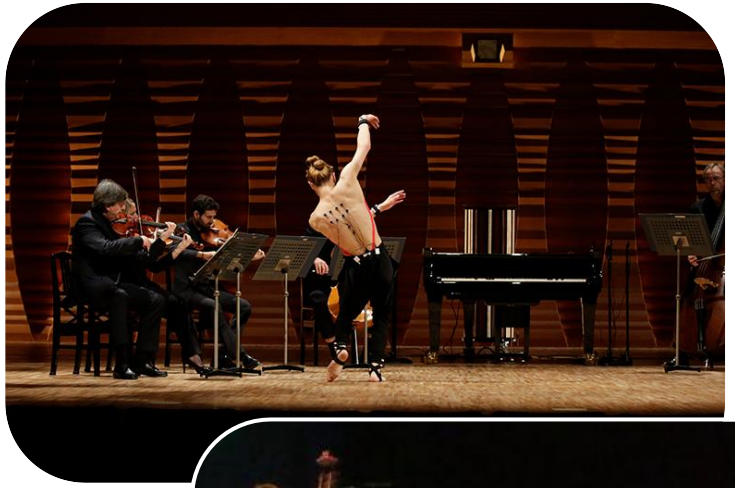
Music & AI

(Source: Yamaha)

yamaha.com/en/news_release/2018/18013101/
Yamaha Global, "Yamaha Artificial Intelligence (AI) Transforms a Dancer into a Pianist - Short Version," *YouTube*, youtu.be/21injmy1wsU, 2018.

Music & AI

(Source: Yamaha)



(Source: Sankei Shimbun)

(Source: Robot Gizmos)

(Source: NBC DFW)

yamaha.com/en/news_release/2018/18013101/
sankei.com/article/20240113-CQCOSQHJWFIYPJJKZDCITRTRVI/
roboticgizmos.com/shimon-musical-robot-deep-learning/
nbcdfw.com/entertainment/the-scene/how-verdigris-ensemble-is-using-ai-to-create-a-new-concert-experience/3366031/

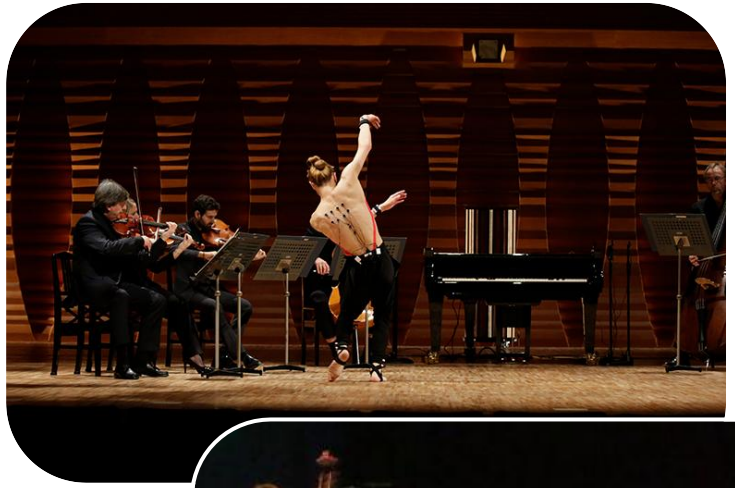
Music & AI

(Source: Robot Gizmos)

roboticgizmos.com/shimon-musical-robot-deep-learning/
Georgia Tech Center for Music Technology, "Shimon with the Aarhus Jazz Orchestra," *YouTube*, youtu.be/ZpTV1-acSU8, 2021.

Music & AI

(Source: Yamaha)



(Source: Sankei Shimbun)

(Source: Robot Gizmos)

(Source: NBC DFW)

yamaha.com/en/news_release/2018/18013101/
sankei.com/article/20240113-CQCOSQHJWFIYPJJKZDCITRTRVI/
roboticgizmos.com/shimon-musical-robot-deep-learning/
nbcdfw.com/entertainment/the-scene/how-verdigris-ensemble-is-using-ai-to-create-a-new-concert-experience/3366031/

Music & AI

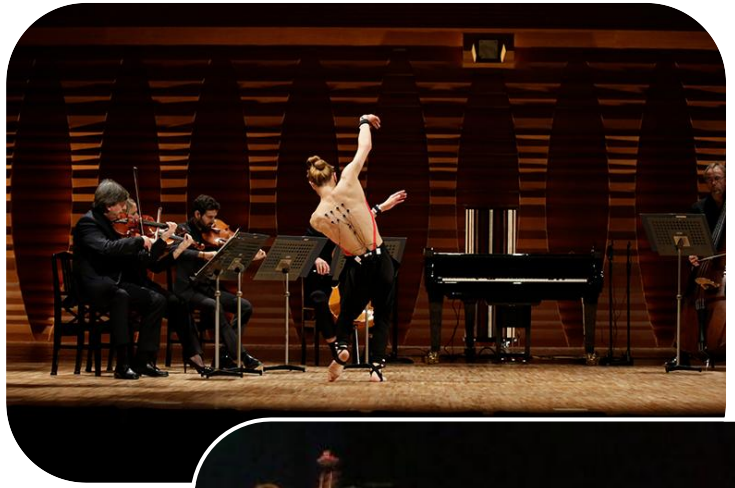
(Source: Sankei Shimbun)

Shlizerman et al., "Audio to Body Dynamics," *Proc. CVPR*, 2018.
yamaha.com/en/news_release/2018/18013101/
sankei.com/article/20240113-CQCOSQHJWFIYPJJKZDCITRTRVI/
roboticgizmos.com/shimon-musical-robot-deep-learning/
nbcdfw.com/entertainment/the-scene/how-verdigris-ensemble-is-using-ai-to-create-a-new-concert-experience/3366031/


第4楽章

Music & AI

(Source: Yamaha)



(Source: Sankei Shimbun)

(Source: Robot Gizmos)

(Source: NBC DFW)

yamaha.com/en/news_release/2018/18013101/
sankei.com/article/20240113-CQCOSQHJWFIYPJJKZDCITRTRVI/
roboticgizmos.com/shimon-musical-robot-deep-learning/
nbcdfw.com/entertainment/the-scene/how-verdigris-ensemble-is-using-ai-to-create-a-new-concert-experience/3366031/

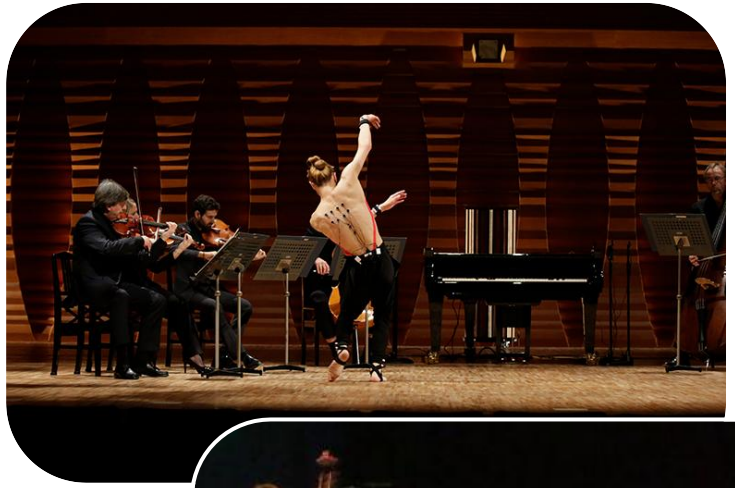
Music & AI

(Source: NBC DFW)

nbcdfw.com/entertainment/the-scene/how-verdigris-ensemble-is-using-ai-to-create-a-new-concert-experience/3366031/

Music & AI

(Source: Yamaha)



(Source: Sankei Shimbun)

(Source: Robot Gizmos)

(Source: NBC DFW)

yamaha.com/en/news_release/2018/18013101/
sankei.com/article/20240113-CQCOSQHJWFIYPJJKZDCITRTRVI/
roboticgizmos.com/shimon-musical-robot-deep-learning/
nbcdfw.com/entertainment/the-scene/how-verdigris-ensemble-is-using-ai-to-create-a-new-concert-experience/3366031/

Generative AI for Music, Audio & Video Creation

(Source: The Denver Post)

(Source: UploadVR)

(Source: Descript)

(Source: Daily Bruin)

(Source: Wikimedia Commons)

Universitaetsmedizin, CC BY-SA 4.0, via Wikimedia Commons
uploadvr.com/iron-man-vr-breaks-free-from-cords-load-screens-on-quest-2/
descript.com/blog/article/what-is-the-best-audio-interface-for-recording-a-podcast
denverpost.com/2019/08/02/colorado-symphony-movie-scores-harry-potter-star-wars/
dailybruin.com/2023/08/04/theater-review-the-musical-les-misrables-offers-stellar-displays-and-impassioned-vocals

Course Logistics

Communications

- **Course website**: Syllabus, schedule, readings, recordings, etc.
- **Email**: Announcements
- **Google Chat**: Q&A

hermandong.com/teaching/pat464_564

Prerequisites

- **Prior coding experience** is recommended!
  - If you've taken any programming course, you should be fine
  - You should be comfortable reading code written by others
- Prior AI/ML experience is recommended!
  - If you've taken any AI/ML course, you should be fine

Assignments (tentative)

- Music & audio processing
- Music instrumentation
- Symbolic music generation
- Audio generation

Assignment Policies

- All assignments must be **completed on your own**
  - You are welcome to exchange ideas with your peers, but this should be in the form of concepts and discussion, not in the form of writing and code.
- **Please turn in your work even if it's incomplete!**
- Due at **11:59pm ET** on the date specified
- Late submissions: **up to a week**, **1 point deducted per day**

Project

- **Open-ended project**
- Group size: 1–2
- **Milestones** (tentative)
  - **Pitch** Mar 18
  - **Presentation** Apr 20
  - **Report** Apr 27
- Deliverables are due at **11:59pm ET** on the date specified
- **No late submissions!** Submit your work early and update it later

Project Topics

- Building a new AI music tool
- Exploring creative & artistic use of AI tools
- Analyzing existing AI music tools

Grading

- All **grading** and **regrade requests** will be handled on Gradescope
- **Programming assignments** (**60%**)
- **Project** (**40%**)
  - Presentation (20%)
  - Report (20%)

Optional Reading

- "*Deep and Shallow: Machine Learning in Music and Audio*" by Shlomo Dubnov and Ross Greer
- "*Fundamentals of Music Processing*" by Meinard Müller (notebooks)
- "*Intelligent Music Production*" by Brecht De Man, Ryan Stables, and Joshua D. Reiss
- Course material for "*Deep Learning for Music Analysis and Generation*" by Yi-Hsuan Yang
- Course material for "*Musical Applications of Machine Learning*" by Juhan Nam
- Course material for "*Machine Learning for Music*" by Julian McAuley

Policies: Attendance

- In-person attendance is strongly encouraged
- Course lectures will be recorded and made available on the course website
- Please **attend in-person** for
  - **Project pitch** (currently scheduled on Mar 18)
  - **Project presentation** (currently scheduled on Apr 20)

Policies: Generative AI Usage

- Feel free to use GenAI tools in your workflow. However, **you must disclose your usage of GenAI services in your write-ups**.
  - For example, U-M GPT, ChatGPT, Gemini, Claude, GitHub Copilot, DALLE, Midjourney, Stable Diffusion, Firefly, etc.
- **You take full responsibility for AI-generated materials as if you had produced them yourself**: ideas should be attributed and facts should be true.

Policies: Academic Integrity

- Plagiarism and cheating violate SMTD's Academic Code of Conduct. **All plagiarism, cheating and other academic misconduct cases will be reported to SMTD's Office of Academic and Student Affairs.**
- **All assignments must be completed on your own**. You are welcome to exchange ideas with your peers, but this should be in the form of concepts and discussion, not in the form of writing and code.
- You must **provide proper citations/references for any external resources** you use in your writing and code.

Performing Arts Technology (PAT)

Minor in Performing Arts Technology (PAT)

myumi.ch/VVADN

Intro (Required)

PAT 100: Music in Technology
PAT 200: Intro to Electronic Music

PAT Theory/Studies (6 credits)

PAT 150: Experiential Music Theory
PAT 205: Intermedia AI Music Practice
PAT 305: Video Game Music
PAT 315: Diversity in Music Technology
PAT 316: NOISE

PAT Practice (3 credits)

PAT 202: Computer Music
PAT 204: Creative Coding
PAT 220: Songwriting Workshop
PAT 280: Sound Reinforcement
PAT 412: Digital Music Ensemble
PAT 413: Electronic Chamber Music

Electives (6 credits)

PAT 421: Advanced Psychoacoustics
PAT 422: Technical Ear Training & Critical Listening
PAT 424: Dialog of the Senses
PAT 431/432: Contemporary Practice in Studio Production I/II
PAT 441: Sound for Film and Games
PAT 443: Immersive Media
PAT 451/452: Interactive Media Design I/II
PAT 454: Digital Fabrication for Acoustics
PAT 461: Performance Systems
PAT 462: Digital Sound Synthesis
PAT 463: Music & AI
PAT 464: Generative AI for Music & Audio Creation
PAT 472: Business of Music

PAT PERFORMING ARTS TECHNOLOGY UNIVERSITY OF MICHIGAN

Questions on the Syllabus?

hermandong.com/teaching/pat464_564

~~The State of the Art~~

Not so long ago in June 2023

Prompt: relaxing and smooth jazz played in a stylish cafe

Prompt: delightful country music with acoustic guitars

Prompt: cinematic and suspenseful orchestral music

huggingface.co/spaces/facebook/MusicGen

Made with Facebook MusicGen (huggingface.co/spaces/facebook/MusicGen)

Not so long ago in July 2023

Video **Runway Gen-2**

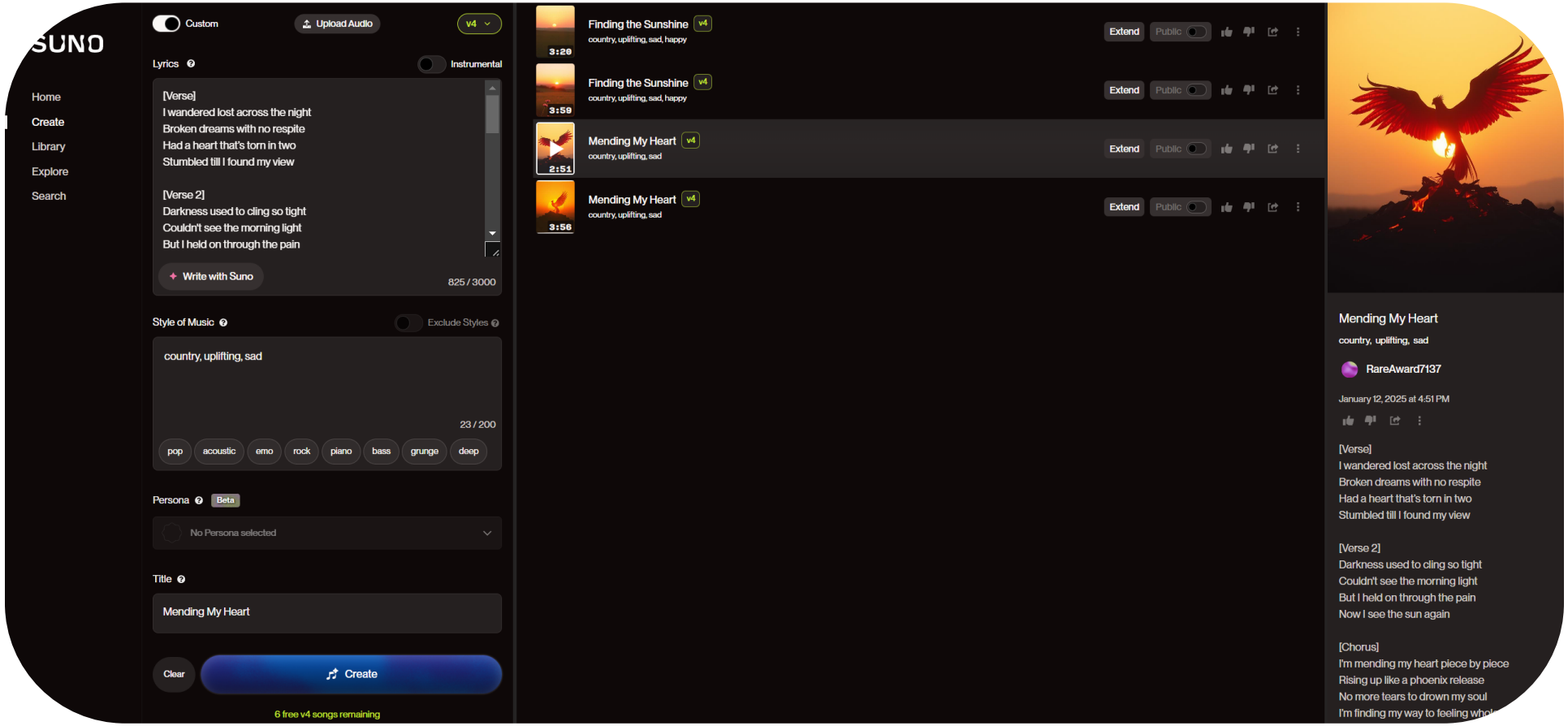
Music **MusicGen**

進捗共有チャンネル, "Runway Gen-2 Sample / Tiger Whiskey," *YouTube*, July 9, 2023.
Kaoru Naito, *Twitter*, https://twitter.com/ka0ru_1620/status/1678313226453520385, July 10, 2023.

The State of the Art

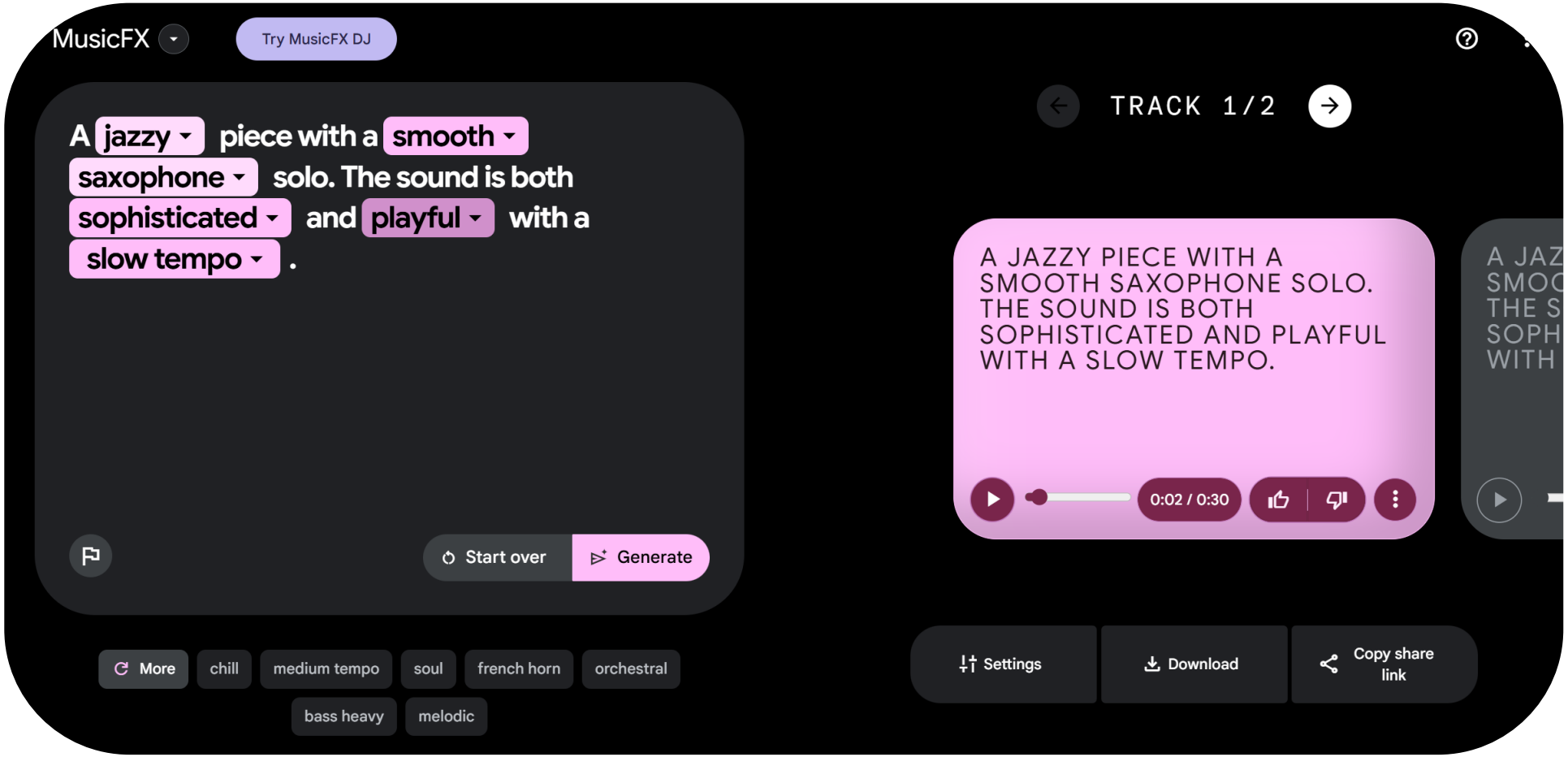
Until when...!?

Suno AI v4 (2024)



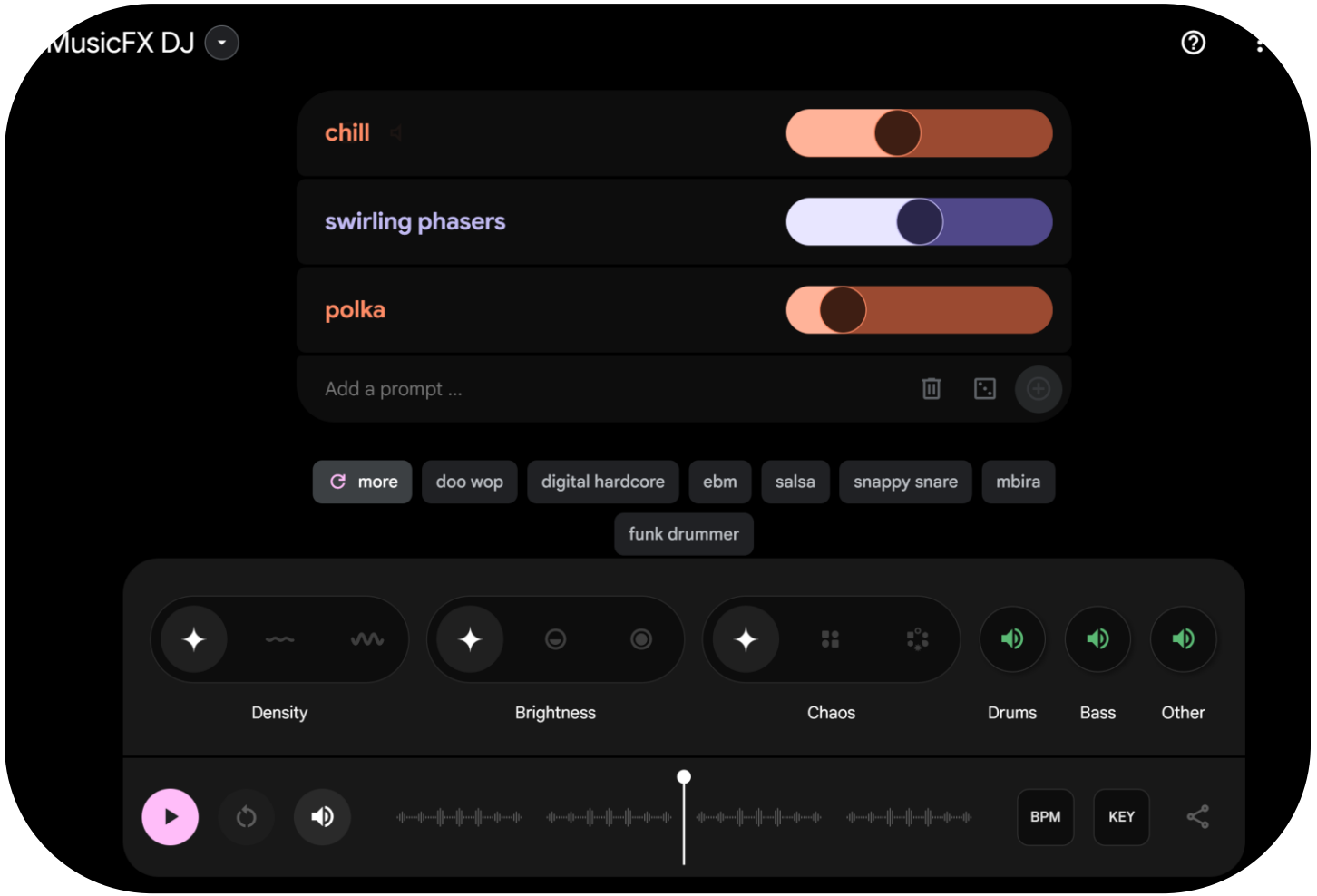
suno.com

Music FX (2024)



aitestkitchen.withgoogle.com/tools/music-fx

Music FX DJ (2024)

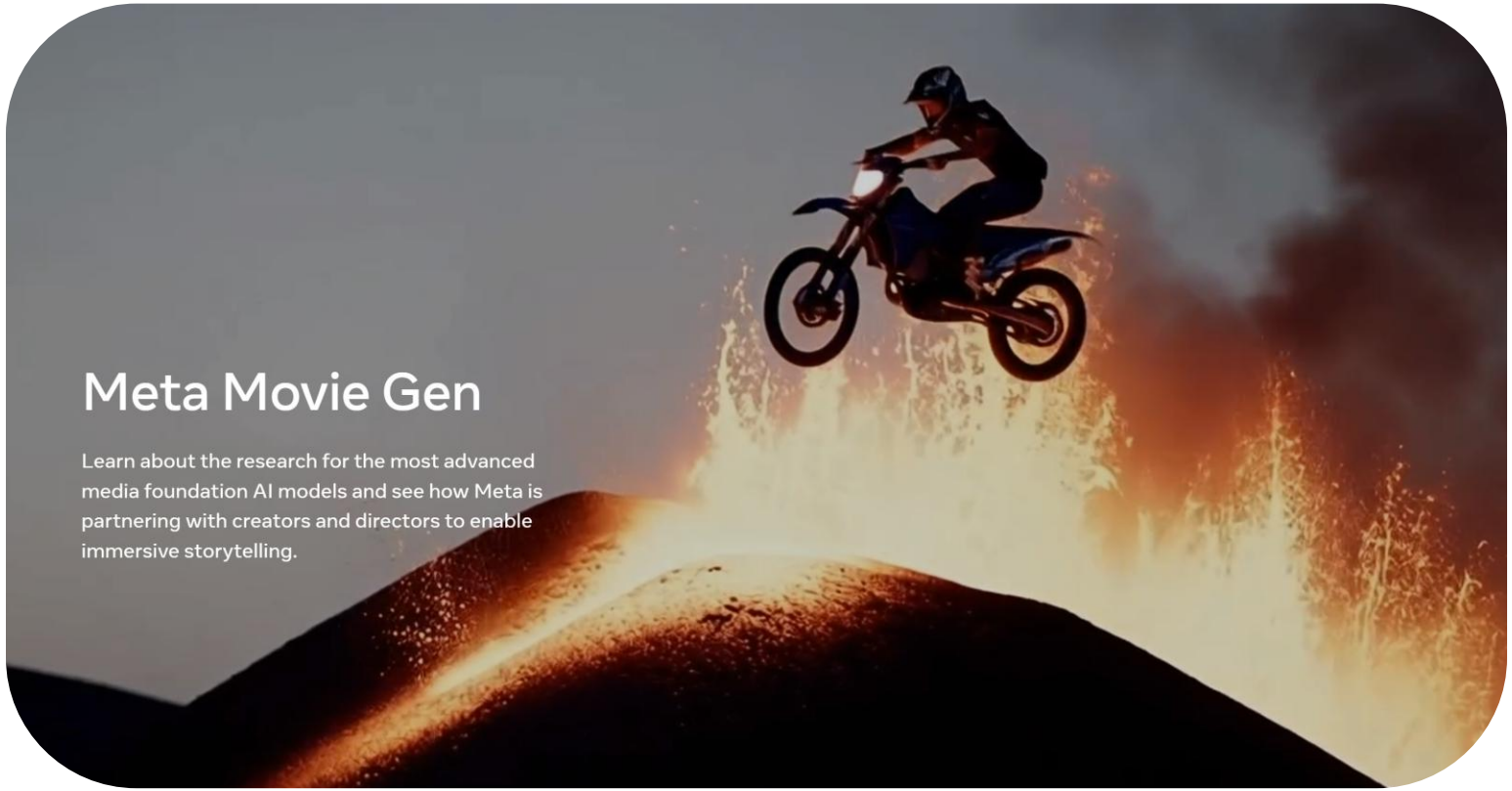


aitestkitchen.withgoogle.com/tools/music-fx-dj

Music FX DJ (2024)

youtube.com/live/IUQW5LgBZvQ

MovieGen (2024)



ai.meta.com/research/movie-gen/

Neural Sound Engine (Saito et al., 2025)

(Source: Saito et al., 2025)

Koichi Saito, Julian Tanke, Christian Simon, Masato Ishii, Kazuki Shimada, Zachary Novack, Zhi Zhong, Akio Hayakawa, Takashi Shibuya, and Yuki Mitsufuji, "SoundReactor: Frame-level Online Video-to-Audio Generation," arXiv preprint arXiv:2510.02110, 2025.

Neural Sound Engine (Saito et al., 2025)

(Source: Saito et al., 2025)

Koichi Saito, Julian Tanke, Christian Simon, Masato Ishii, Kazuki Shimada, Zachary Novack, Zhi Zhong, Akio Hayakawa, Takashi Shibuya, and Yuki Mitsufuji, "SoundReactor: Frame-level Online Video-to-Audio Generation," arXiv preprint arXiv:2510.02110, 2025.
koichi-saito-sony.github.io/soundreactor/

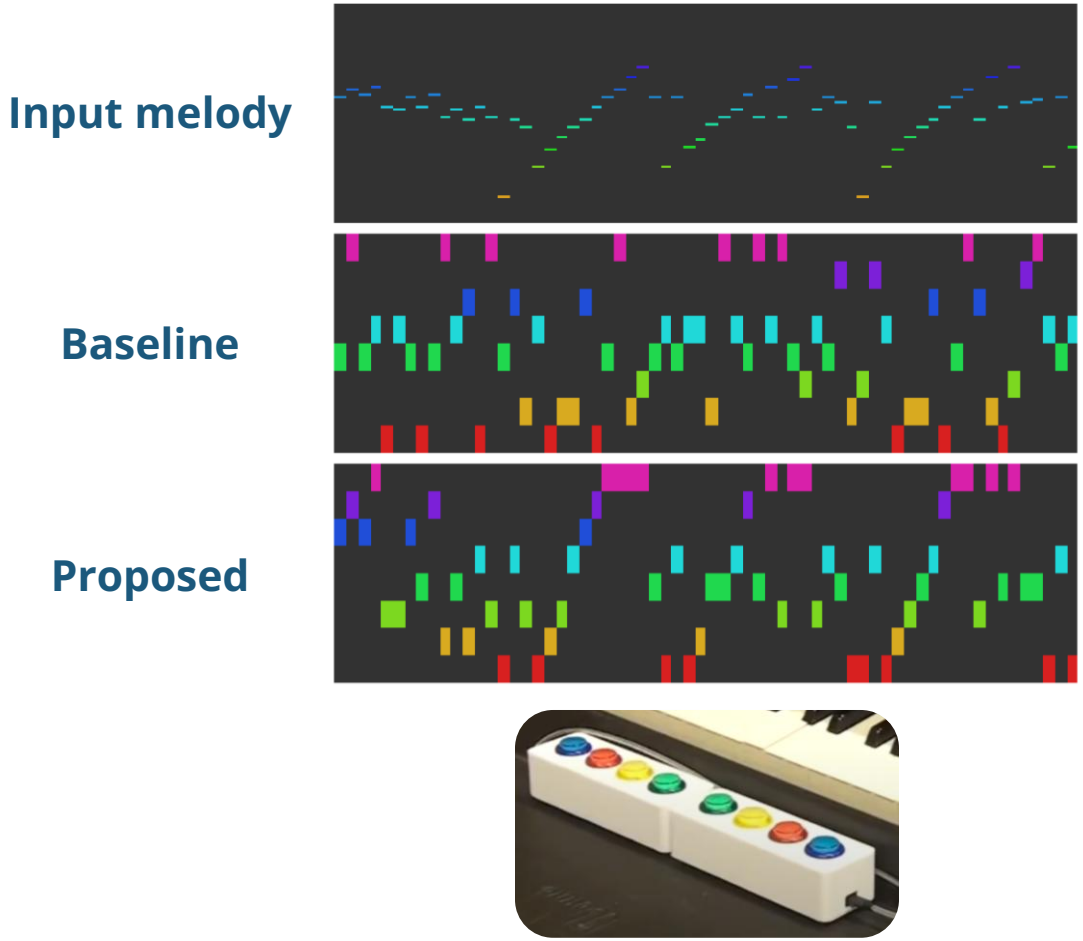
Piano Genie (Donahue et al., 2018)

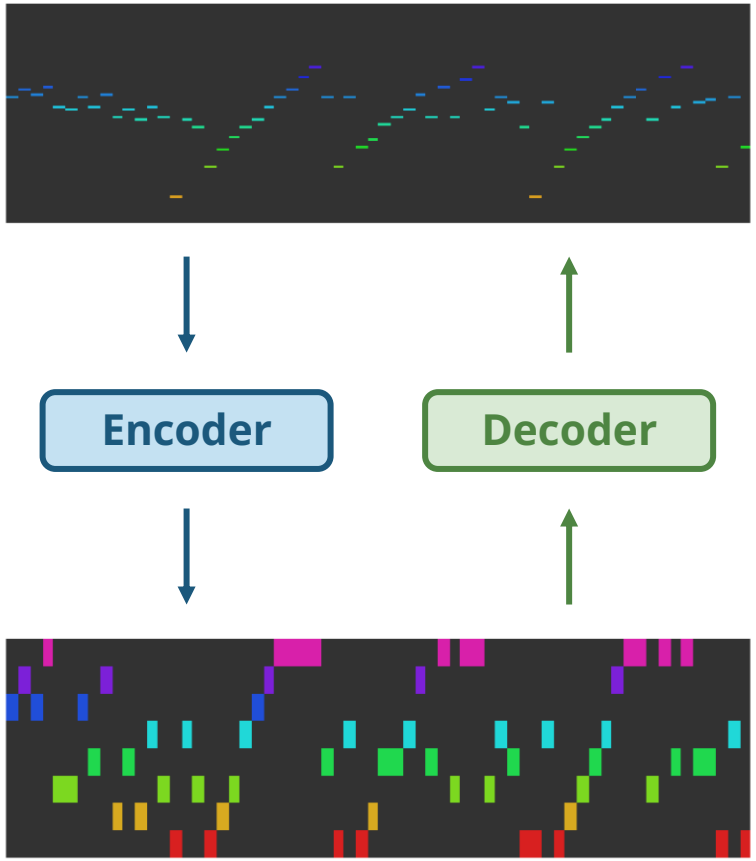
piano-genie.glitch.me/

youtu.be/YRb0XAnUpIk & magenta.tensorflow.org/pianogenie

Chris Donahue, Ian Simon, and Sander Dieleman, "Piano Genie," IUI, 2019.

Piano Genie (Donahue et al., 2018)





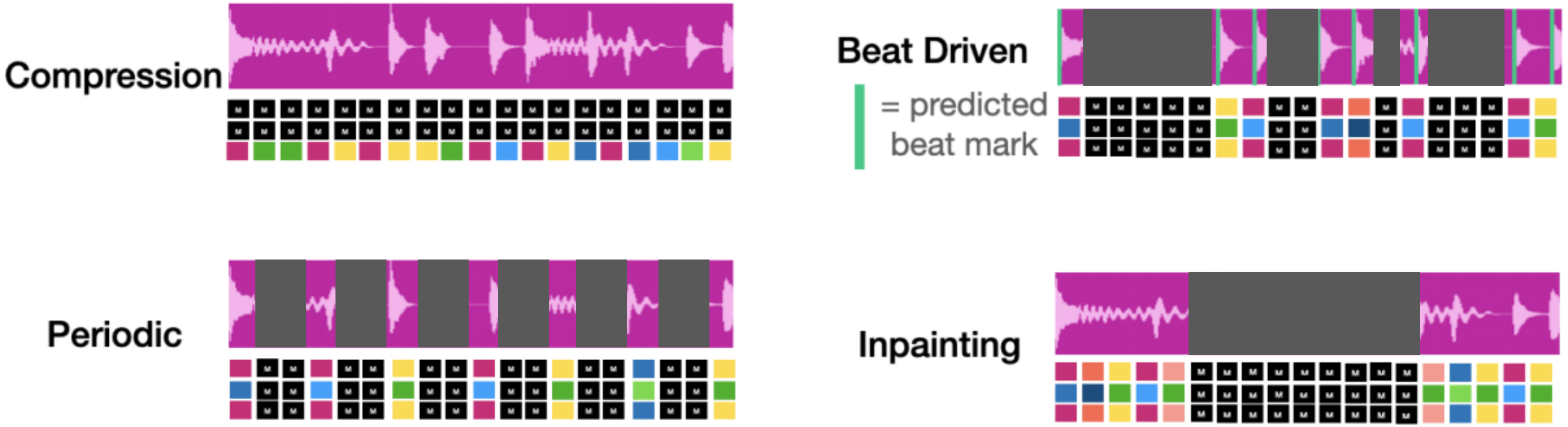
(Source: Donahue et al., 2019)

Chris Donahue, Ian Simon, and Sander Dieleman, "Piano Genie," *IUI*, 2019.

`unloop` (Garcia et al., 2023)

youtu.be/yzBI8Vcjd2s & github.com/hugofloresgarcia/unloop

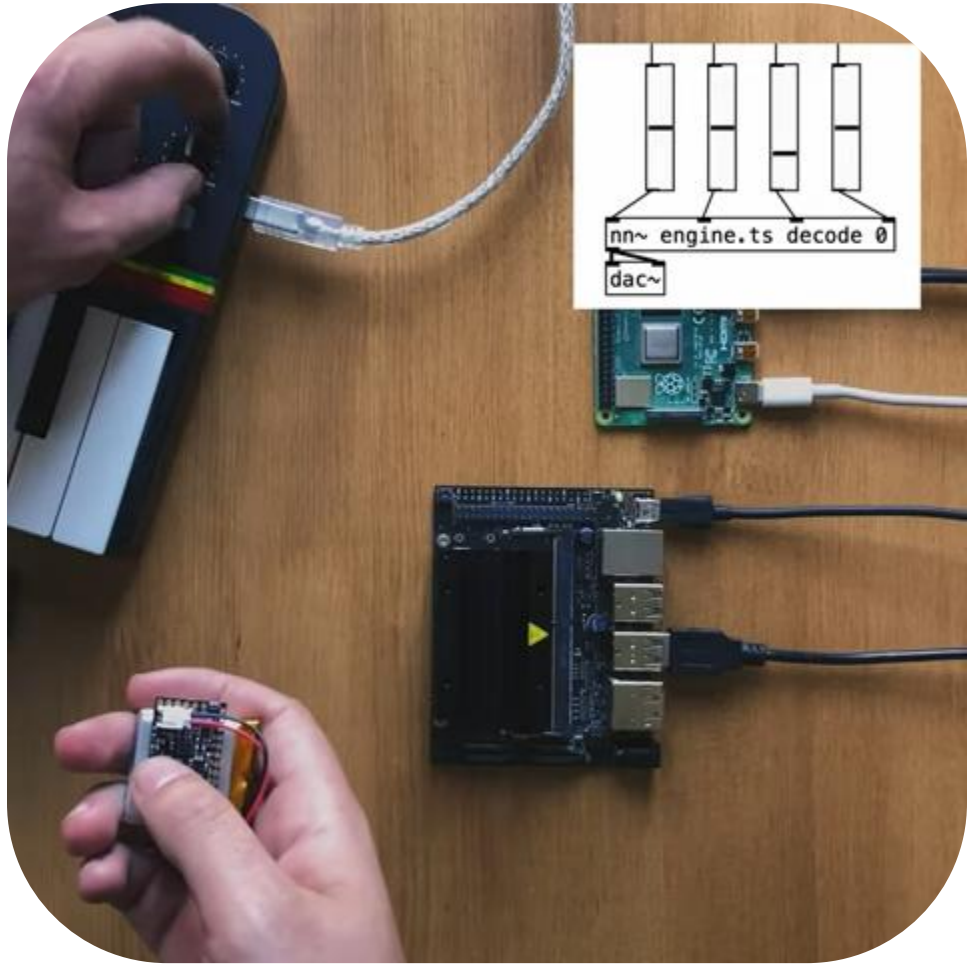
VampNet (Garcia et al., 2023)



(Source: Garcia et al., 2023)

Hugo Flores Garcia, Prem Seetharaman, Rithesh Kumar, and Bryan Pardo, "VampNet: Music Generation via Masked Acoustic Token Modeling," *ISMIR*, 2023.

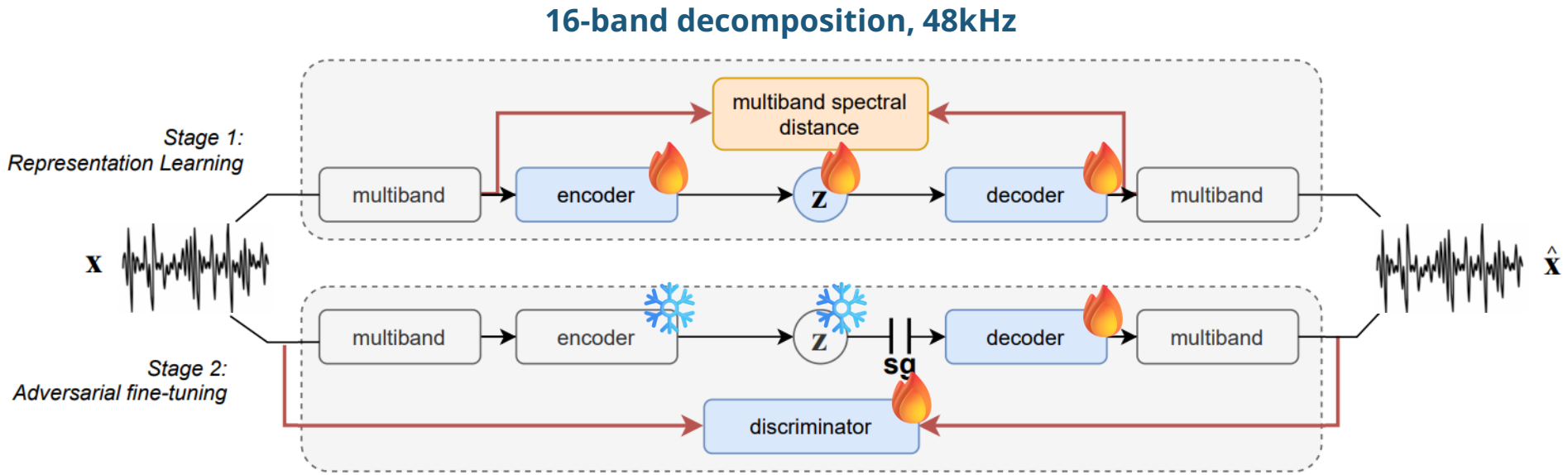
RAVE: Real-time Audio Synthesis (Caillon & Esling, 2022)



youtu.be/jAIRf4nGgYI

Antoine Caillon and Philippe Esling, "RAVE: A variational autoencoder for fast and high-quality neural audio synthesis," *arXiv preprint arXiv:2111.05011*, 2021.

RAVE: Real-time Audio Synthesis (Caillon & Esling, 2022)



Model	CPU synthesis	GPU synthesis
NSynth	18 Hz	57 Hz
SING	304 kHz	9.8 MHz
RAVE (Ours) w/o multiband	38 kHz	3.7 MHz
RAVE (Ours)	**985 kHz**	**11.7 MHz**

Realtime capable on CPUs & GPUs

anonymous84654.github.io/RAVE_anonymous

Antoine Caillon and Philippe Esling, "RAVE: A variational autoencoder for fast and high-quality neural audio synthesis," *arXiv preprint arXiv:2111.05011*, 2021.

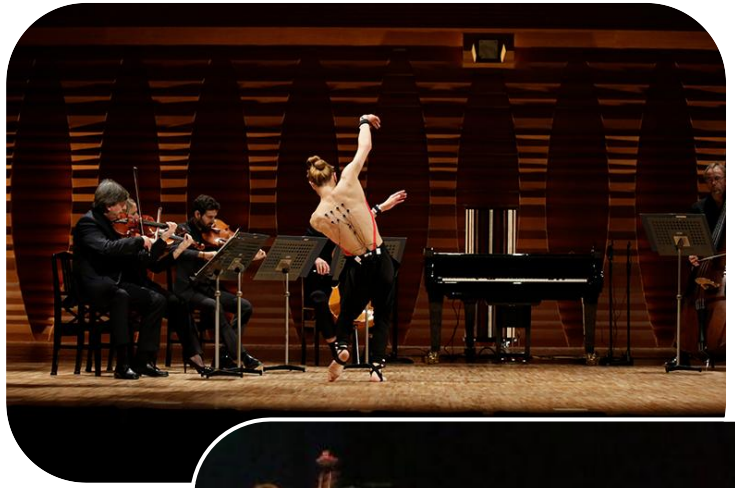
Recap

Music & Technology Co-evolves

Hildegard Dodel, Public domain, via Wikimedia Commons.
Taken at Hamamatsu Museum of Musical Instruments, August 2019.
yan, CC BY-SA 4.0, via Wikimedia Commons.

Music & AI

(Source: Yamaha)



(Source: Sankei Shimbun)

(Source: Robot Gizmos)

(Source: NBC DFW)

yamaha.com/en/news_release/2018/18013101/
sankei.com/article/20240113-CQCOSQHJWFIYPJJKZDCITRTRVI/
roboticgizmos.com/shimon-musical-robot-deep-learning/
nbcdfw.com/entertainment/the-scene/how-verdigris-ensemble-is-using-ai-to-create-a-new-concert-experience/3366031/

Generative AI for Music, Audio & Video Creation

(Source: The Denver Post)

(Source: UploadVR)

(Source: Descript)

(Source: Daily Bruin)

(Source: Wikimedia Commons)

Universitaetsmedizin, CC BY-SA 4.0, via Wikimedia Commons
uploadvr.com/iron-man-vr-breaks-free-from-cords-load-screens-on-quest-2/
descript.com/blog/article/what-is-the-best-audio-interface-for-recording-a-podcast
denverpost.com/2019/08/02/colorado-symphony-movie-scores-harry-potter-star-wars/
dailybruin.com/2023/08/04/theater-review-the-musical-les-misrables-offers-stellar-displays-and-impassioned-vocals

Not so long ago in June 2023

Prompt: relaxing and smooth jazz played in a stylish cafe

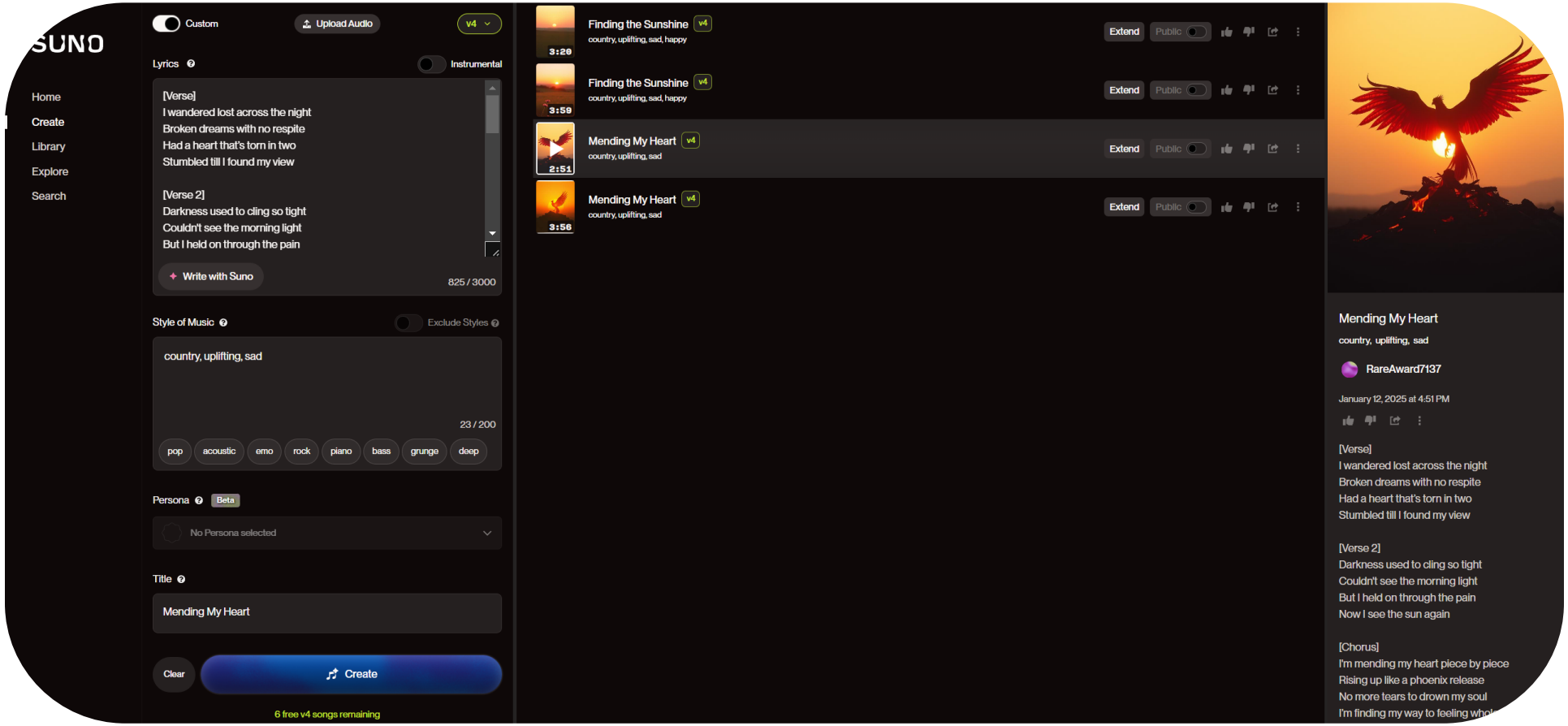
Prompt: delightful country music with acoustic guitars

Prompt: cinematic and suspenseful orchestral music

huggingface.co/spaces/facebook/MusicGen

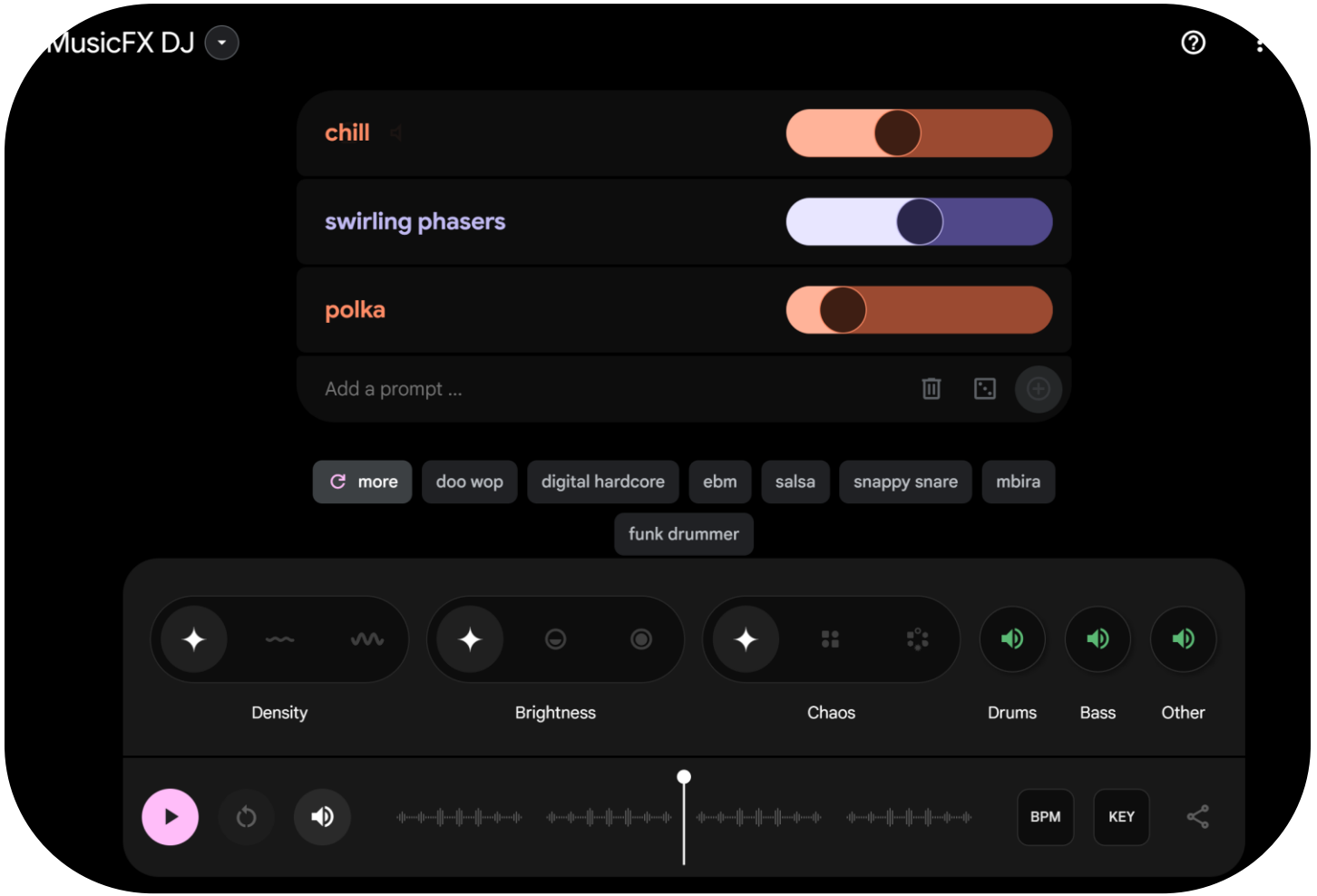
Made with Facebook MusicGen (huggingface.co/spaces/facebook/MusicGen)

Suno AI v4 (2024)



suno.com

Music FX DJ (2024)



aitestkitchen.withgoogle.com/tools/music-fx-dj

Music FX DJ (2024)

youtube.com/live/IUQW5LgBZvQ

Neural Sound Engine (Saito et al., 2025)

(Source: Saito et al., 2025)

Koichi Saito, Julian Tanke, Christian Simon, Masato Ishii, Kazuki Shimada, Zachary Novack, Zhi Zhong, Akio Hayakawa, Takashi Shibuya, and Yuki Mitsufuji, "SoundReactor: Frame-level Online Video-to-Audio Generation," arXiv preprint arXiv:2510.02110, 2025.
koichi-saito-sony.github.io/soundreactor/

Piano Genie (Donahue et al., 2018)

piano-genie.glitch.me/

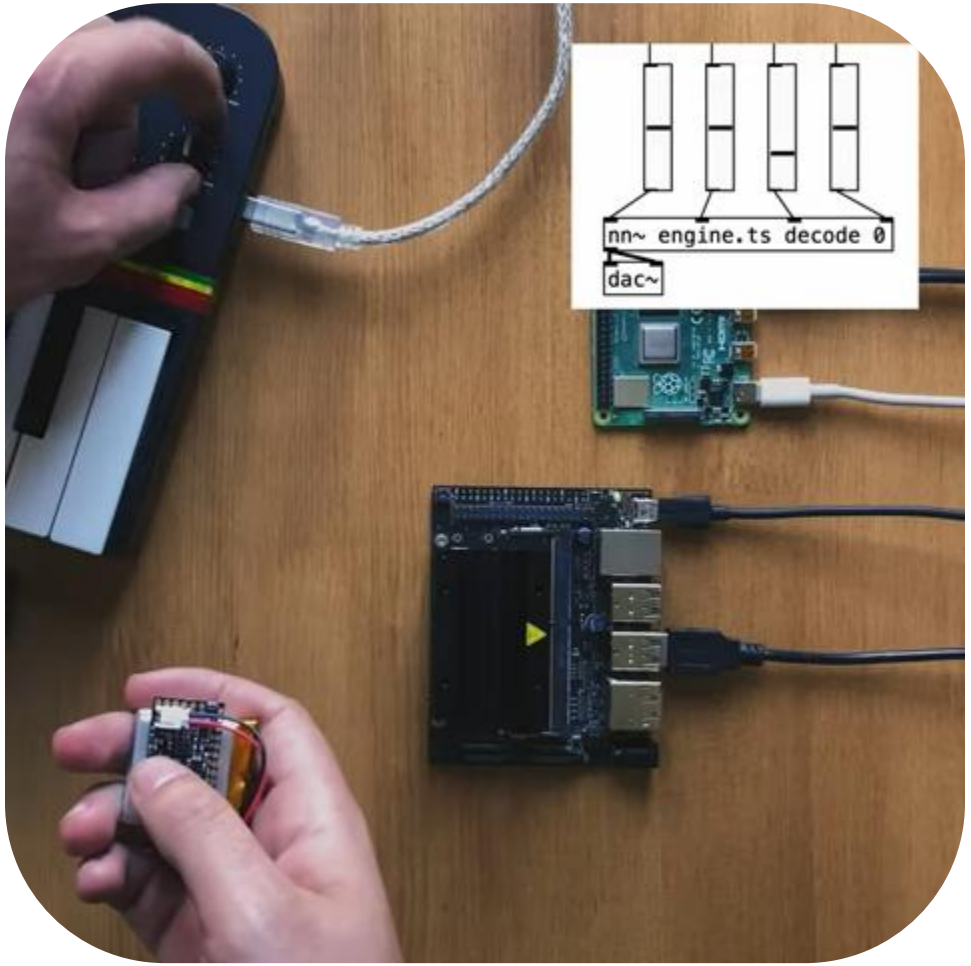
youtu.be/YRb0XAnUpIk & magenta.tensorflow.org/pianogenie

Chris Donahue, Ian Simon, and Sander Dieleman, "Piano Genie," IUI, 2019.

`unloop` (Garcia et al., 2023)

youtu.be/yzBI8Vcjd2s & github.com/hugofloresgarcia/unloop

RAVE: Real-time Audio Synthesis (Caillon & Esling, 2022)

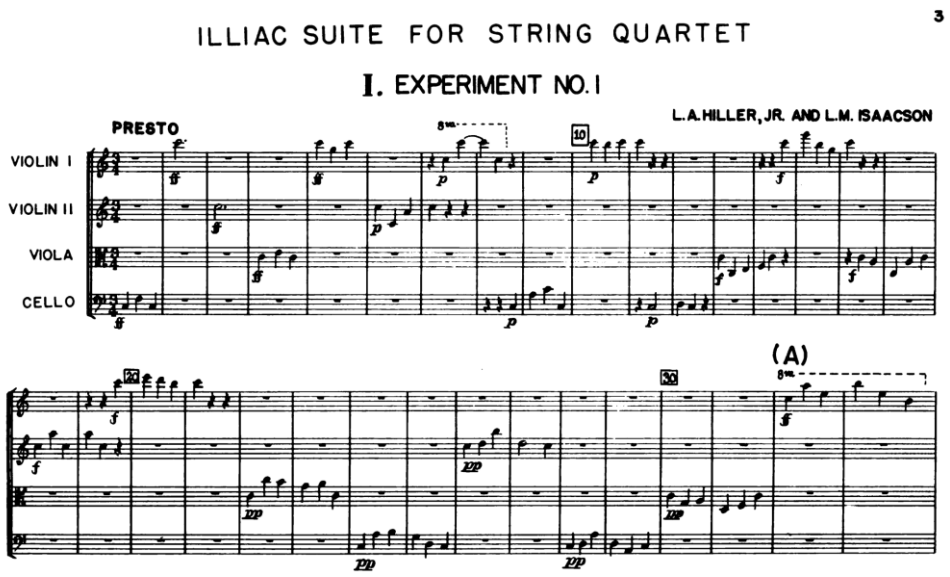


youtu.be/jAIRf4nGgYI

Antoine Caillon and Philippe Esling, "RAVE: A variational autoencoder for fast and high-quality neural audio synthesis," *arXiv preprint arXiv:2111.05011*, 2021.

Next Lecture

AI & Music



(Source: Illinois Distributed Museum)