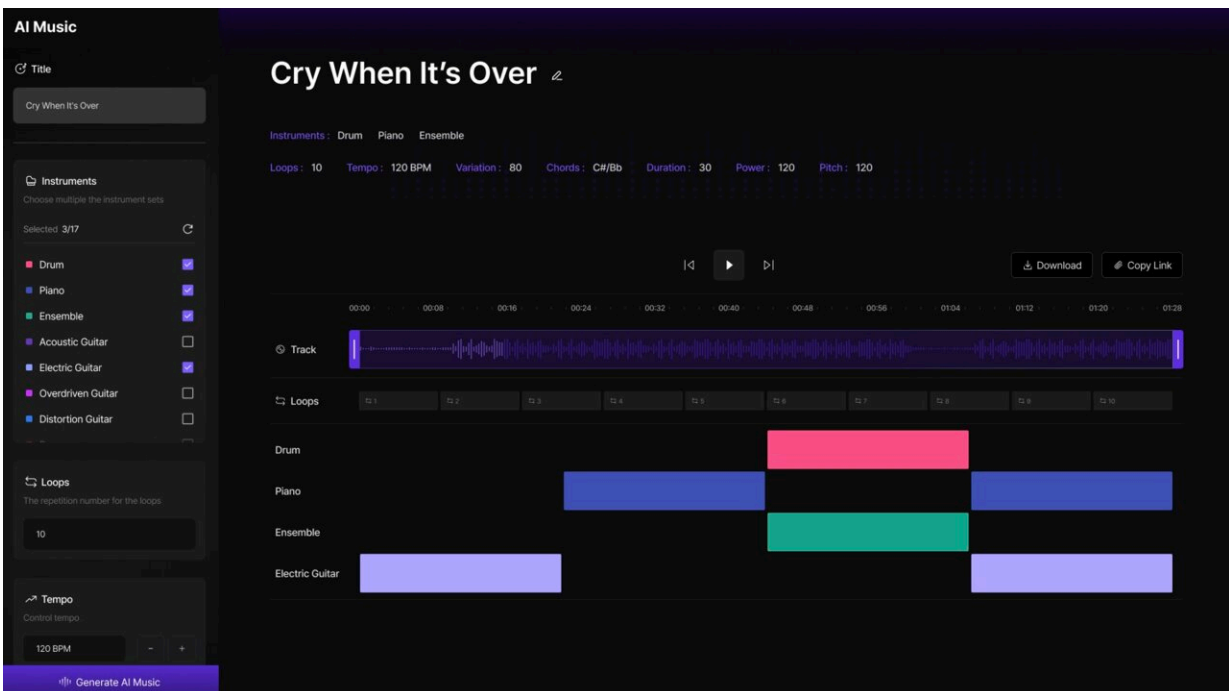


# A new model for symbolic music generation using musical metadata

October 1 2024, by Ingrid Fadelli



Screenshot of the team's system demo showing its user interface. Credit: Han et al.

Artificial intelligence (AI) has opened new interesting opportunities for the music industry, for instance, enabling the development of tools that can automatically generate musical compositions or specific instrument tracks. Yet most existing tools are designed to be used by musicians, composers and music producers, as opposed to non-expert users.

Researchers at LG AI Research recently developed a new interactive system that allows any user to easily translate their ideas into music. This system, outlined in a paper [published](#) on *arXiv* preprint server, combines a decoder-only autoregressive transformer trained on music datasets with an intuitive user interface.

"We introduce the demonstration of symbolic music generation, focusing on providing short musical motifs that serve as the central theme of the narrative," Sangjun Han, Jiwon Ham and their colleagues wrote in their paper. "For the generation, we adopt an autoregressive model which takes musical metadata as inputs and generates 4 bars of multitrack MIDI sequences."

The transformer-based model underpinning the team's symbolic music generation system was trained on two musical datasets, namely the [Lakh MIDI dataset](#) and the [MetaMIDI dataset](#). Collectively, these datasets contain over 400,000 MIDI ([musical instrument](#) digital interface) files, which are data files containing various information about musical tracks (e.g., the notes played, the duration of notes, the speed at which they are played).

To train their model, the team converted each MIDI file into a musical event representation (REMI) file. This specific format encodes MIDI data into tokens representing various music features (e.g., pitch and velocity). REMI files capture the dynamics of music in ways that are particularly favorable for training AI models for music generation.

"During training, we randomly drop tokens from the musical metadata to guarantee flexible control," wrote the researchers. "It provides users with the freedom to select input types while maintaining generative performance, enabling greater flexibility in [music composition](#)."

In addition to developing their transformer-based model for symbolic

music generation, Han, Ham and their colleagues created a simple interface that would make it accessible to both expert and non-expert users. This interface currently consists of a sidebar and a central interactive panel.

In the sidebar, users can specify aspects of the music they want the model to generate, such as what instruments should be playing and the song's tempo. After the model generates a song, they can edit the track in the central panel, for instance, by removing/adding instruments or adjusting the time at which they will start playing music.

"We validate the effectiveness of the strategy through experiments in terms of model capacity, musical fidelity, diversity, and controllability," wrote Han, Ham and their colleagues. "Additionally, we scale up the model and compare it with other music generation models through a subjective test. Our results indicate its superiority in both control and music quality."

The researchers found their model performed significantly well and could reliably generate a maximum of 4 bars of music based on the user's specifications. In their future studies, they could improve their system further by extending the duration of the musical tracks their model can create, broadening the specifications that users can give, and further enhancing the system's user interface.

"Our model, trained to generate 4 bars of music with global control, has limitations in extending music length and controlling bar-level local elements," wrote the researchers. "However, our attempts hold significance in generating high-quality musical themes that can be used as loop."

**More information:** Sangjun Han et al, Flexible Control in Symbolic Music Generation via Musical Metadata, *arXiv* (2024). [DOI:](#)

[10.48550/arxiv.2409.07467](https://arxiv.org/abs/2409.07467)

© 2024 Science X Network

Citation: A new model for symbolic music generation using musical metadata (2024, October 1)  
retrieved 4 October 2024 from

<https://techxplore.com/news/2024-09-music-generation-musical-metadata.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.