Robust Personalized Human-AI Collaboration with *SmartLooper*

Sageev Oore
Dalhousie University
Halifax, Canada
osageev@gmail.com

Finlay Miller
Dalhousie University
Halifax, Canada

Chandramouli Sastry*
Amazon
Toronto, Canada

Sri Harsha Dumpala
Dalhousie University
Halifax, Canada

Marvin F. da Silva
Dalhousie University
Halifax, Canada

Daniel Oore
Aix-Marseille University
Marseilles, France

Scott C. Lowe
Vector Institute
Toronto, Canada

https://osageev.github.io/projects/looper/

1 Short Abstract

This demonstration can be presented either as a live improvised performance or as a recorded video of one. Like most music, it is better listened to rather than read about. We developed an AI "smart looping" system for collaborative human-AI real-time performance. A dataset is created by the artist who will be using the system. In performance, the artist begins by improvising a short musical fragment—say, a bass groove—and the system starts riffing on it in their style, based on the data they provided. Sometimes it repeats a phrase exactly; other times it makes small changes—shifting the rhythm, altering harmonic progressions—but not too much at once. The artist plays over this evolving loop, adding new lines or textures as the system continues to develop its own variations. The result is a fluid, co-created improvisation. It has been demonstrated in multiple live performances.

2 Extended Abstract

2.1 Intuitive Overview

Using the System. From the user's perspective, the main steps in this interactive demo consist of the following:

- **Launching the system** with options (set the tempo, preferences, etc).
- **Recording a short musical seed, in MIDI,** that will form the basis of the upcoming loop/groove.
- **Letting the system find a MIDI-based groove** based on that seed.
- **Playing a duet with the system** as it continues to loop and very slowly change the groove, thus creating a extended, dynamic, collaborative musical interaction.

Dataset. Crucially, in order for this to be possible, the user must first have created a dataset of relevant material. This dataset consists of numerous short loopable segments. Short means roughly 4-16 beats, and loopable means that they can be concatenated and repeated to form longer looping

*Work done while at Dalhousie University.

39th Conference on Neural Information Processing Systems (NeurIPS 2025) Workshop: Artificial Intelligence for Music: Where Creativity Meets Computation.

sequences. This dataset will form the basis of the grooves and loops that the system will later draw from. This means that all of the musical building blocks *played* by the system are derived directly and exclusively from the material created by the artist using the system.

Underlying Process. How are these building blocks connected together? In otherwords, given a set of loopable segments, how does this system decide which segment to play next? At the heart of this stochastic process is a mechanism for measuring a musical "distance" between any two segments. For this we do rely on a larger dataset; the distance between musical segments is computed using the cosine distance between their respective embeddings, as obtained from a pre-trained diffusion model [6, 5]. We have explored a variety of techniques for subsequently using these distances, along with other information (i.e. weak labels, other heuristics to react to what the human is playing, etc), to give rise to different algorithms for choosing the next segment. More details on this system are provided in a corresponding report [13].

Personalized Data In a musician's voice: *"The generative aspect of the system is based exclusively on a dataset I created from my own playing, so it reflects my rhythmic feel, harmonic tendencies, and stylistic voice. In particular, I created a set of musical MIDI recordings with this system in mind, and so it reflects the kind of rhythmic and harmonic tendencies that I chose, as an artist, to have present when performing with it. And so, in performance, it feels like a duet with a version of myself: I begin by improvising a short musical fragment—perhaps a bass groove or chord progression—and the system starts riffing on it in a style based on the training data I created. Sometimes it will repeat the phrase exactly; other times it makes small shifts in rhythm or harmonic role, but not so much that it loses the musical thread. I thus set this in motion at the beginning of a piece, and then I play over this evolving loop, responding to the systems steady yet slowly-shifting presence."*

Robust Interaction. The various elements of this system (dataset, looping framework, distance computation in embedding space, stochastic process to sequence loop segments) work together to give rise to a system that is remarkably robust during real-time interaction. For example, the distance metric helps prevent jarring transitions and ensure smooth ones during the looping. The fact that the next looping segment is being determined while the current loop is still playing (and long before it finishes playing) means that, from the user's point of view, there appears to be zero latency. Altogether, this robustness is a crucial design criterion for any kind of live performance, and for the musician, it provides an effective balance between predictability and surprise. This allows for truly fun and spontaneous interactions (see the demonstration video).

Demonstration Notes. The work can be presented as a live improvisation or as a recorded video performance. The musician's role is to provide an initial spark and then, together with the system, they co-create the arc of the piece; the system sustains and transforms the seed material. It does not try to predict the musician's next move—it simply develops their own material in ways that feel both familiar and surprising.

While the accompanying video shows performance clips in which the musician was playing a Yamaha Disklavier (i.e. a MIDI-enabled grand piano), they have also performed with this system using an electric keyboard.

Related Work. Numerous interesting and effective MIDI-generation systems have been proposed in recent years [8, 2, 16, 1, 4, 10, 12, 7, 3, 19, 9, 15, 14, 11, 17, 18], ranging from iterative co-composition and editing to interactive systems for live performance. Our system uses a single user-recorded seed to initiate a looping process grounded in a personalized corpus. The AI selects phrase-level segments–also recorded by the user–that are musically similar to the seed, and continues evolving the loop by traversing this corpus in real time, allowing the output to shift gradually while remaining musically coherent.

Because the personalized corpus reflects groove-based playing, the system tends to evolve rhythmically coherent, harmonically shifting textures that prioritize feel and flow–but this framework could be applied just as well to other musical idioms depending on the source data. In contrast to systems focused on audio continuation or harmonic-grid accompaniment, SmartLooper is intended to support a fluid, groove-centered interaction.

References

[1] Mason Bretan and Gil Weinberg. Chronicles of a robotic musical companion. In *Proceedings of the International Conference on New Interfaces for Musical Expression*.

[2] Mason Bretan, Sageev Oore, Jesse Engel, Douglas Eck, and Larry Heck. Deep music: Towards musical dialogue. *Proceedings of the AAAI Conference on Artificial Intelligence*, 31(1), Feb 2017. doi:10.1609/aaai.v31i1.10544.

[3] Antoine Caillon, Brian McWilliams, Cassie Tarakajian, Ian Simon, Ilaria Manco, Jesse Engel, Noah Constant, Pen Li, Timo I. Denk, Alberto Lalama, Andrea Agostinelli, Anna Huang, Ethan Manilow, George Brower, Hakan Erdogan, Heidi Lei, Itai Rolnick, Ivan Grishchenko, Manu Orsini, Matej Kastelic, Mauricio Zuluaga, Mauro Verzetti, Michael Dooley, Ondrej Skopek, Rafael Ferrer, Zalán Borsos, Äaron van den Oord, Douglas Eck, Eli Collins, Jason Baldridge, Tom Hume, Chris Donahue, Kehang Han, and Adam Roberts. Live music models. *arXiv:2508.04651*, 2025. doi:10.48550/arXiv.2508.04651.

[4] Jason d'Eon, Sri Harsha Dumpla, Chandramouli Shama Sastry, Daniel Oore, and Sageev Oore. Musical speech: a transformer-based composition tool. In *NeurIPS 2020 Competition and Demonstration Track*, pages 253–274. PMLR. URL https://proceedings.mlr.press/v133/d-eon21a.

[5] Josh Gardner, Ian Simon, Ethan Manilow, Curtis Hawthorne, and Jesse Engel. MT3: Multi-task multitrack music transcription, 2022. URL https://arxiv.org/abs/2111.03017.

[6] Curtis Hawthorne, Ian Simon, Adam Roberts, Neil Zeghidour, Josh Gardner, Ethan Manilow, and Jesse Engel. Multi-instrument music synthesis with spectrogram diffusion, 2022. URL https://arxiv.org/abs/2206.05408.

[7] Cheng-Zhi Anna Huang, Ashish Vaswani, Jakob Uszkoreit, Noam Shazeer, Curtis Hawthorne, Andrew M Dai, Matthew D Hoffman, and Douglas Eck. Music transformer: Generating music with long-term structure. *arXiv preprint arXiv:1809.04281*, 2018. doi:10.48550/arXiv.1809.04281.

[8] Yewon Kim, Sung-Ju Lee, and Chris Donahue. Amuse: Human-AI collaborative songwriting with multimodal inspirations, 2025. URL https://arxiv.org/abs/2412.18940.

[9] Jing Luo, Xinyu Yang, and Dorien Herremans. Bandcondinet: Parallel transformers-based conditional popular music generation with multi-view features. *Expert Systems with Applications*, 299:130059, March 2026. ISSN 0957-4174. doi:10.1016/j.eswa.2025.130059. URL http://dx.doi.org/10.1016/j.eswa.2025.130059.

[10] Google Magenta. Studio. Ableton Live Plugin, 2025. URL https://magenta.withgoogle.com/studio.

[11] Marco Marchini, François Pachet, and Benoît Carré. Rethinking Reflexive Looper for structured pop music. *NIME*, 2017.

[12] Nicholas Meade, Nicholas Barreyre, Scott C. Lowe, and Sageev Oore. Exploring conditioning for generative music systems with human-interpretable controls. In *International Conference on Computational Creativity (ICCC)*, 2019. doi:10.48550/arXiv.1907.04352. URL https://computationalcreativity.net/iccc2019/papers/iccc19-paper-57.pdf.

[13] Sageev Oore, Finlay Miller, Chandramouli Shama Sastry, Sri Harsha Dumpala, Marvin F da Silva, Daniel Oore, and Scott C Lowe. A loopy framework and tool for real-time human-AI music collaboration. In *NeurIPS AI for Music Workshop (paper track)*.

[14] François Pachet, Pierre Roy, Julian Moreira, and Mark d'Inverno. Reflexive loopers for solo musical improvisation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2205–2208, Paris France, April 2013. ACM. ISBN 978-1-4503-1899-0. doi:10.1145/2470654.2481303.

[15] Jordi Pons, Zack Zukowsi, Julian D. Parker, CJ Carr, Josiah Taylor, and Zach Evans. Music and artificial intelligence: Artistic trends, 2025. URL https://arxiv.org/abs/2508.11694v1.

[16] Adam Roberts, Jesse Engel, Curtis Hawthorne, Ian Simon, Elliot Waite, Sageev Oore, Natasha Jaques, Cinjon Resnick, and Douglas Eck. Interactive musical improvisation with Magenta. In *Advances in Neural Information Processing Systems*, 2016.

[17] Victor Shepardson and Thor Magnusson. The living looper: Rethinking the musical loop as a machine action-perception loop. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. URL http://nime.org/proceedings/2023/nime2023_32.pdf.

[18] Victor Shepardson, Halla Steinunn Stefánsdóttir, and Thor Magnusson. Evolving the Living Looper: Artistic Research, Online Learning and Tentacle Pendula. *NIME*, 2025.

[19] John Thickstun, David Hall, Chris Donahue, and Percy Liang. Anticipatory music transformer, 2024. URL https://arxiv.org/abs/2306.08620.

A System Overview

Figure 1 shows a user's perspective of playing with SmartLooper.

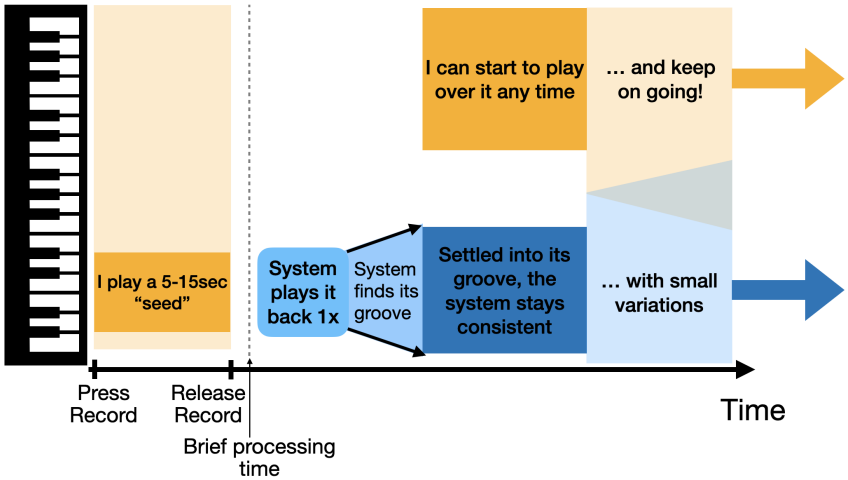

Figure 1: **The user's view of the system.** Orange: user playing; blue: AI system playing. To start the process, the user records a 5–15 second seed, typically in the bass through middle pitches of the keyboard. The system plays back the seed, and then has a short warm-up phase during which it transitions to find its groove. Once it settles in, it stays robust and feels musically solid as it gradually shifts and evolves its groove. While we show the user and AI system mainly playing in upper and lower regions of the piano respectively (after the seed), there is no hard threshold between the two; the player is free to play in all areas of the piano, and the system may meander up or down the keyboard (though this is uncommon). The user can see on the computer screen the next segment queued from the looper before it is played, and plays over top of this continuous groove as they wish. All of these phases (including the user playing across the full keyboard) are shown in the demonstration video.