Generative AI for Musicians: Small-Data Prototyping to Design Intelligent Musical Instruments

CHARLES PATRICK MARTIN, The Australian National University, Australia

Music is one of many fields touched by recent innovation in generative artificial intelligence (AI). AI-driven music generation systems have evoked a sense of both promise and threat, artists have adopted AI-generated alter-egos, and the use of music data for AI training has led to legal interventions. Despite the apparent level of activity there are, so far, no mainstream musical instruments that apply generative AI. One of the problems is that music instrument playing, a real-time phenomenon involving complex gestures and embodied interaction, presents difficulties for machine learning methods and involves highly diverse data types. A better approach may be for artists themselves to collect and curate training data and develop their own intelligent musical instruments. In this paper we consider some of the challenges facing artist-led generative AI for music and propose a prototyping approach for intelligent musical instruments based on small-data principles.

Additional Key Words and Phrases: generative AI, small data, human-AI interaction, intelligent musical instruments

1 INTRODUCTION

The use of generative AI for music has been seen as a source of both promise and threat. The idea of generating music fully automatically would seem to inevitably impact artistic careers [3] and the powerful recording industry has started to react to un-authorised use of music for training AI systems [10]. In contrast to the many writing and coding tools using generative AI (e.g., ChatGPT), there are few examples of generative AI tools embedded within music-making tools. AI certainly has a long history in music with early examples appearing in the 1980s [13, e.g.,]. Present AI approaches in music are focussed on producing audio from text [1], assisting in the recording, editing, and mastering processes [2] or generating musical scores [5]. In contrast, the amount of research focussed on how AI can be created and used by a live musician is relatively small. Existing examples of tools for making AI instruments use now-outdated machine-learning models [4] or have command-line interfaces (CLIs) that are not accessible to musicians [9].

Some of the issues facing embedding of generative AI within musical instruments are the real-time nature of musical performance and the embodied interaction that takes place between a musician and their instrument. Musical instruments must be designed to support real-time processing of input gesture to output sound with very tight time constraints [12]. It's not acceptable, as it might be in an image or text generation context, to wait several seconds for a generative AI response. The embodied nature of musical performance [7] relates to the kinds of data that a generative AI system might model. Modeling these data in AI systems is complicated by existing issues of representation in music with notated scores, digital audio recordings in addition to the gestural complexity of live performance and the cultural impact of *really being there*. Although there are standards such as MIDI for data connection between musical instruments, different instruments use different subsets of MIDI's available control channels and some control systems go beyond MIDI's limitations. Efforts have been made to create musical AI systems that can model general gestural information [9], but a single model would not easily map across the wildly different interaction modalities found in musical instruments.

The diversity of available musical interfaces and interaction styles requires a diversity of generative AI approaches. A small-data [14] approach where artists generate, curate, and train generative AI models may be an appropriate way to manage generative AI within music that is both more useful and less threatening than a monolithic approach. However, to undertake small-data generative AI creation, musicians need tools that are more appropriate and guiding examples

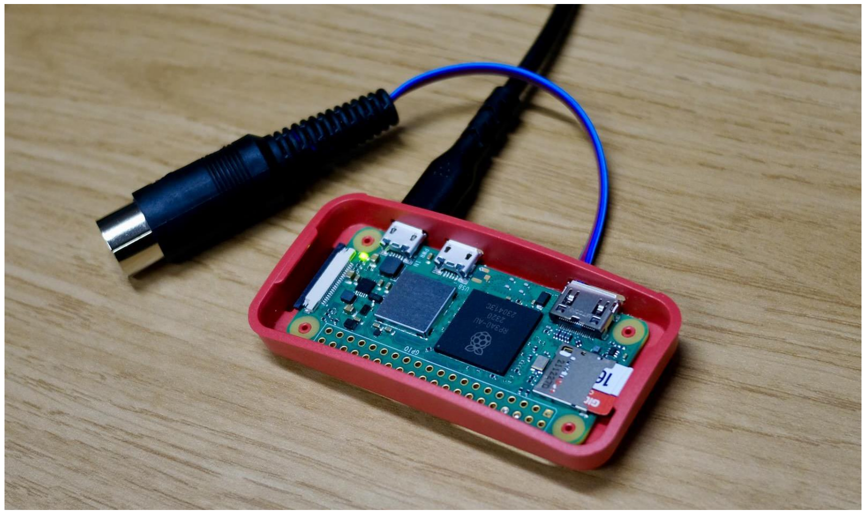

Fig. 1. A Raspberry Pi-based Generative AI music system. This system is intended to interface via the MIDI connector with electronic musical instruments to prototype new intelligent musical instrument interactions. The Raspberry Pi Zero 2 W used here costs 15USD and is capable of running a small deep neural network generating real-time interaction data. In this example system, the generative AI model plays a stream of MIDI messages trained on data of a human turning a single synthesiser control knob.

of what intelligent musical instruments could really be. In the remainder of this workshop paper we discuss what intelligent musical instruments are or could be. We present an example of a Raspberry Pi-based hardware module designed to equip standard electronic musical instruments with generative AI capabilities. Finally, we consider the requirements for enabling artist-centered small-data prototyping that might generate the diversity of intelligent musical instruments needed to match the variation in musical practices.

2 WHAT IS AN INTELLIGENT MUSICAL INSTRUMENT?

There are many ways that musical systems can apply AI technologies for the benefit of musicians and audiences [6]. The use of machine learning can be used to map complex synthesiser data to synthesiser parameters [15] is a well-known example although it is not a generative application. Lewis' *Voyager* system [8] exemplifies the use of generative techniques to create backing accompaniments to an acoustic musician however the accompaniments are framed as separate instruments. The *Continuator* [11] represents what could be termed a true intelligent musical instrument as the machine learning model connected to a digital piano "continues" playing in the style of what the musician has recently played, although the musician and ML model cannot play at the same time. It is notable that both Continuator and Voyager are more than 20 years old and do not apply present AI techniques such as deep neural networks.

Reflecting on these two examples, a definition of an intelligent musical instrument could be:

> An instrument where an AI system generates actions *independently* of a musician's either reacting to the musician's actions or not.

For the purposes of this paper we consider instruments that a human may also play simultaneously. Many acoustic instruments have scope for simultaneous playing of a human and generative AI system (e.g., two humans can play duets on one 88-key keyboard) but this is an area where electronic musical instruments excel. Electronic synthesisers

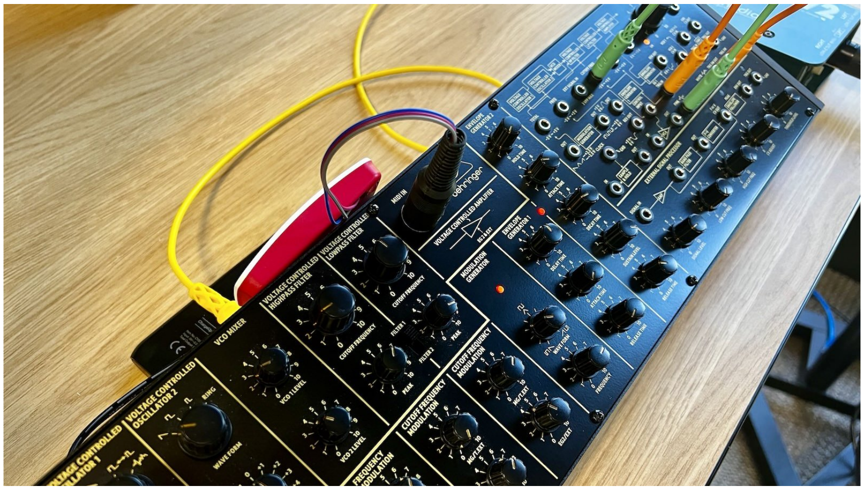

Fig. 2. A commercial desktop synthesiser with the GenAI MIDI plug controlling pitch and rhythm. There are many other interface connections for a musician to express themselves.

and digital musical instruments have extended and extensible interfaces that allow in-depth timbral adjustment, live re-wiring of electronic circuits and simultaneous control over multiple sequences and control sources. Indeed, contemporary electronic music practice often involves building highly customised synthesisers from commercially available modules, interconnected by a tangle of patch cables. Electronic musical instruments would seem to be an ideal context for experimenting with introducing one or multiple generative AI systems to control aspects of the instrument alongside the human.

3 AN EXAMPLE SYSTEM FOR PROTOTYPING INTELLIGENT MUSICAL INSTRUMENTS

As an example of what this could look like we present the GenAI MIDI Plug, shown in Figure 1, a minimal and inexpensive Raspberry Pi-based system for introducing a deep neural network model into a wide variety of commercial electronic musical instruments. This system uses the very widely implemented MIDI standard to interface with a musical instrument. While MIDI is simple, it can represent both musical notes over time and gestural information, such as adjusting control sliders or interacting with sensors, that connects to control parameters. In our system, the Raspberry Pi Zero 2 W runs a mixture-density recurrent neural network adapted from the IMPS system [9] to continually perform an expressive sequence of MIDI notes. The training data for this system consisted of a human performing with one control knob on a synthesiser for around 60 minutes. This minimal and inexpensive (15USD) system is designed for experimentation by electronic musicians in their own systems which frequently include MIDI inputs as well as many other modes for control.

Figure 2 shows how this GenAI MIDI Plug might be deployed in practice with a commercial musical desktop synthesiser which does not include a keyboard. The Raspberry Pi is connected with hook-and-loop tape to the back of the synth and the generative AI model outputs MIDI signals control the pitch and note onsets (gate) of the synthesiser. As Figure 2 shows, this synthesiser has a plethora of other interfaces to occupy a real-time musician including knobs for timbre, amplitude envelope and low frequency oscillators and a patch bay where the synthesiser circuit can be re-wired in real time. The synthesiser itself also includes non-intelligent generative capacity where noise can be used as a source for modifying parameters over time. At present, the genAI MIDI plug acts as an additional low-frequency oscillator although more expressive and less predictable than those typically available on electronic musical instruments. It could be hypothesised that in co-design with electronic musicians, the GenAI MIDI Plug could be more deeply integrated into electronic instruments producing new design knowledge about how intelligent musical instruments might behave.

4 SMALL-DATA PROTOTYPING FOR INTELLIGENT MUSIC

The previous section outlines a way forward for prototyping generative AI within a musical instrument, but not how the models themselves might be prototyped by artists. The concept of small data [14], collected and curated by artists and trained locally on their own computers presents an alternative to the commercial big-data approach to generative AI creation. Within music, this has been established in the context of interactive machine learning for the use-case of mapping sensor data in *Wekinator* [4]. More recently small-data approaches have been proposed for learning gestural models [9] and deploying AI on embedded platforms [12].

Where *Wekinator* provided a graphical interface for collected data on a computer, some of the current systems involve collecting data on a separate device [12], or with a command line program [9] and then running a separate training step. Although these approaches work, they may not appeal to regular musicians for longer term deployment. Additionally, data *curation* after collection is a so-far under explored aspect of this process. Deep learning models are capable of learning from more data than previous machine learning approaches and so artists may wish to collect data over time and, perhaps build different models for different purposes. They may also wish to exclude data introducing undesirable behaviours or experiment with diving back through their artistic history. So far, there are no established systems for these tasks within a musical AI context.

REFERENCES

[1] Jade Copet, Felix Kreuk, Itai Gat, Tal Remez, David Kant, Gabriel Synnaeve, Yossi Adi, and Alexandre Défossez. 2023. Simple and Controllable Music Generation. arXiv:2306.05284 [cs.SD]

[2] Emmanuel Deruty, Maarten Grachten, Stefan Lattner, Javier Nistal, and Cyran Aouameur. 2022. On the Development and Practice of AI Technology for Contemporary Popular Music Production. *Transactions of the International Society for Music Information Retrieval* (Feb 2022). https://doi.org/10.5334/tismir.100

[3] Ben Eltham. 2023. Welcome to 'the robot soundscape': Australia's music industry braces for the rise of music AI. *The Guardian* (sep 2023). https://www.theguardian.com/music/2023/sep/08/ai-music-bigsound-brisbane-ai-dj-spotify-beatles

[4] Rebecca Fiebrink, Dan Trueman, and Perry R Cook. 2009. A Meta-Instrument for Interactive, On-the-fly Machine Learning. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME '09)*. Pittsburgh, PA, USA, 280–285. http://www.nime.org/proceedings/2009/nime2009_280.pdf

[5] Shulei Ji, Xinyu Yang, and Jing Luo. 2023. A Survey on Deep Learning for Symbolic Music Generation: Representations, Algorithms, Evaluations, and Challenges. *ACM Comput. Surv.* 56, 1 (2023). https://doi.org/10.1145/3597493

[6] Théo Jourdan and Baptiste Caramiaux. 2023. Machine Learning for Musical Expression: A Systematic Literature Review. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. 13 pages. http://nime.org/proceedings/2023/nime2023_46.pdf

[7] Marc Leman, Pieter-Jan Maes, Luc Nijs, and Edith Van Dyck. 2018. What Is Embodied Music Cognition? In *Springer Handbook of Systematic Musicology*, Rolf Bader (Ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 747–760. https://doi.org/10.1007/978-3-662-55004-5_34

[8] George E. Lewis. 2000. Too Many Notes: Computers, Complexity and Culture in "Voyager". *Leonardo Music Journal* 10 (2000), 33–39. https://doi.org/10.1162/096112100570585

[9] Charles Patrick Martin and Jim Torresen. 2019. An Interactive Musical Prediction System with Mixture Density Recurrent Neural Networks. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. 260–265. https://doi.org/10.5281/zenodo.3672952

[10] Blake Montgomery. 2023. Music publishers sue Amazon-backed AI company over song lyrics. *The Guardian* (oct 2023). https://www.theguardian.com/technology/2023/oct/19/music-lawsuit-ai-song-lyrics-anthropic

[11] François Pachet. 2003. The Continuator: Musical Interaction With Style. *Journal of New Music Research* 32, 3 (2003), 333–341. https://doi.org/10.1076/jnmr.32.3.333.16861

[12] Teresa Pelinski, Rodrigo Diaz, Adán L. Benito Temprano, and Andrew McPherson. 2023. Pipeline for recording datasets and running neural networks on the Bela embedded hardware platform. In *Proceedings of the International Conference on New Interfaces for Musical Expression*.

[13] Curtis Roads. 1985. Research in Music and Artificial Intelligence. *ACM Comput. Surv.* 17, 2 (jun 1985), 163–190. https://doi.org/10.1145/4468.4469

[14] Gabriel Vigliensoni, Phoenix Perry, and Rebecca Fiebrink. 2022. A small-data mindset for generative AI creative work. In *Proceedings of the CHI 2022 Workshop on Generative AI and HCI*. https://doi.org/10.5281/zenodo.7086327

[15] Federico Ghelli Visi and Atau Tanaka. 2021. Interactive Machine Learning of Musical Gesture. In *Handbook of Artificial Intelligence for Music: Foundations, Advanced Approaches, and Developments for Creativity*. Springer, 771–798. https://doi.org/10.1007/978-3-030-72116-9_27