A Web Interface for Real-Time Interaction with Machine Learning in Musical Performance

Hongzhe Kevin Zhu
kevin.zhu@anu.edu.au
The Australian National University
Canberra, ACT, Australia

Charles Patrick Martin
charles.martin@anu.edu.au
The Australian National University
Canberra, ACT, Australia

Abstract

Interactive machine learning systems are increasingly enabling new forms of real-time musical collaboration between human performers and artificial intelligence, transforming how musicians access and manipulate performance data. However, many existing tools function as black boxes, obscuring the internal logic of the generative models and limiting the performer's sense of agency and trust. This research addresses this challenge by developing a transparent, web-based interface for an interactive machine learning system which visualises the state of mixture density recurrent neural networks (MDRNNs) in real-time. We present the design and evaluation of this system, detailing how its architecture and two-way visualisation strategy support multiple interaction paradigms. An exploratory user study suggests that while the conceptual complexity of model training presents a learning curve for novices, the system's real-time visual feedback can support user trust and enable expressive co-creation. We conclude that exposing the data provenance and generative processes of AI systems can help transform them from opaque tools into intelligible musical partners. This work contributes a framework for human-centered AI interface design in the NIME community, suggesting that transparency is an important design factor for creative agency in intelligent instruments.

Keywords

Human-AI collaboration, interactive machine learning, web-based musical interfaces, user-centered design, real-time performance

1 Introduction

The integration of deep learning into real-time musical systems offers new possibilities for control and improvisation. In particular, mixture density recurrent neural networks (MDRNNs) and recurrent neural networks (RNNs) have proven effective for modelling continuous, temporal musical interaction from performance control data, enabling predictive and improvisational behaviours in interactive instruments [16, 24–26]. However, while the generative capabilities of these models are well-established, their integration into performance workflows remains technically demanding. Current approaches often present the machine learning component as a black box, obscuring the relationship between a performer's input and the system's output, and making it difficult to develop reliable expectations during live interaction [9, 13, 20].

In this paper, we present a human-centered system designed to facilitate transparent, bidirectional interaction between a musician and a machine learning model. Building upon an existing musical AI system that enables call-and-response interaction with a machine-learning model, we introduce a novel interface architecture that exposes the internal state of the generative model in real time. Rather than treating the performer merely as a data source, our system visualises high-dimensional control signals and model predictions with low latency, creating a more transparent environment where the AI's decision-making process is visible and intelligible. This transparency is critical for building trust during live performance, allowing musicians to anticipate system responses and engage in genuine creative dialogue [18, 26].

With this background, this work is motivated by the question: What effects does visualising the interactive machine learning process have on musicians' experience of performing with an intelligent musical instrument?

Our design approach addresses the diversity of the NIME community by supporting users across varying levels of technical and musical expertise. Through a rigorous design process grounded in Human-Computer Interaction (HCI) principles, we developed a progressive interface that scaffolds complexity. The system supports novices through guided exploration and immediate feedback, while simultaneously offering expert users deep control over model hyperparameters and system configuration. This multi-level accessibility ensures that the tool can serve as both an entry point for hobbyists and a robust instrument for professional researchers and performers [6, 11].

The system integrates the entire machine learning lifecycle into the creative workflow. Effective use of musical AI requires an iterative loop of recording, curation, and retraining. By unifying data logging, dataset management, and model deployment within the system, we enable musicians to trace the provenance of their AI partners, understanding exactly which performance data shaped a model's behaviour.

We evaluate these contributions through an exploratory user study combining usability metrics with qualitative thematic analysis, suggesting that transparent feedback and integrated data management can support user engagement and trust. Our contributions could lead to broader participation and inclusion in the design and performance of intelligent musical instruments.

2 Background

2.1 Interactive Machine Learning for Music

Interactive machine learning (IML) systems enable users to iteratively collect data, train models, and experience updated behaviour within a single workflow, making model development experientially accessible [12, 14]. In music, this supports rapid exploration of mappings and performance styles, but it also increases the need for interfaces that render concepts such as data quality, coverage, and generalisation legible to performers rather than remaining abstract [17].

NIME '26, June 23–26, 2026, London, UK

The interactive musical predictive system (IMPSY[1]) [26] exemplifies IML for music performance. It uses mixture density recurrent neural networks (MDRNNs) to predict continuous control streams (e.g., sensor values with explicit time deltas) rather than discrete MIDI notes, modelling inputs with multi-modal probabilistic outputs grounded in recurrent sequence learning [15]. IMPSY is focused on a call-and-response interaction paradigm, where both a human musician and the MDRNN model can control an instrument with the MDRNN listening to human interactions and responding when the human stops playing. In the scenario examined here, a performer plays short gestural phrases on a MIDI controller, IMPSY records these interactions as training data, and the trained MDRNN later generates call-and-response continuations when the performer pauses, similar to the interaction style in the Continuator [29]. IMPSY provides core functionality for logging interactions, constructing datasets, training personalised models, and auditioning predictions in real time. In this paper, IMPSY serves as the foundational system and interaction model; our work builds on this prior platform by designing a web-based interface layer intended to improve workflow continuity and support more interpretable, performer-facing feedback during use.

2.2 Music, Machine Learning Interfaces, and Workflow Continuity

Machine-learning musical instruments often unfold as an iterative loop: musicians record performance data, curate or segment it into datasets, train or select models, and then return to performance. When these stages are distributed across disconnected views and representations, systems can make it difficult for users to understand how changes in one phase (e.g., data selection or training settings) manifest as changes in another (e.g., real-time system behaviour) [16, 24–26]. This fragmentation weakens attribution: musicians may struggle to identify whether an unexpected output is caused by their input, the training data used, or an implicit model/configuration change, reducing confidence and interrupting creative flow.

IMPSY concretises this challenge because it exposes multiple necessary capabilities (e.g., logging, dataset construction, model training/selection, and configuration), but still offers limited feedback that helps performers interpret system behaviour during creative use. While manageable for technically proficient users, coordinating a UI client and model server as separate processes adds configuration steps that may interrupt creative flow, particularly for musicians with limited technical backgrounds. [26].

As a contrasting baseline from general-purpose IML, Teachable Machine foregrounds workflow continuity by keeping example capture, training, and testing within an approachable browser-based pipeline; however, it primarily targets static classification tasks and does not address continuous, low-latency musical control interaction [14].

Related music systems provide partial interface solutions without fully closing this end-to-end loop. NexusUI offers reusable control widgets that speed the construction of browser instruments, but it typically leaves the surrounding workflow (data management, training iteration, and reflective feedback) to the designer [31]. Webmapper makes mapping relationships legible and editable, yet it primarily supports configuration and routing rather than sustained performer-facing learning over time [33]. Algorithmic Power Ballads supports human–AI improvisation via interface-level steering of a generative model, but provides a comparatively narrow feedback loop around the learning process itself and limited visibility into how training data and model updates shape behaviour [22].

2.3 Musical Visualisation and Interpretability in IML

Music information visualisation provides established representational strategies for making temporal structure and musical detail inspectable. Piano-roll views in digital audio workstations map pitch against time, revealing timing patterns and harmonic density at a glance [1, 2, 23], while spectrograms visualise frequency energy over time to support inspection of timbre [23]. Beyond these foundations, augmented scores and enriched notations link performance to representation through interactive annotations, and pedagogical systems use immediate visual feedback to guide practice and improve retention and execution accuracy [28, 34]. Across these contexts, prior work also cautions that visual aids must complement rather than replace listening, particularly when interfaces become central during music-making [19].

For musician-facing IML, visualisation additionally functions as a transparency mechanism: it can connect live performer action to what the system records, learns from, and predicts, thereby improving attribution and trust during iterative use. However, many web-based musical interfaces and IML tools still address only fragments of the creative loop (e.g., control widgets, mapping configuration, or limited AI “steering”) rather than integrating capture, curation, training, and performance into a single interpretable workflow [14, 31, 33]. This motivates instruments that treat the interface itself as part of the musical instrument [8, 11], using visualisation not only to show musical signals but also to expose the learning process and its consequences in real time, so musicians can relate system behaviour to their actions and data.

3 System Design

3.1 Interface Design and Visualisation

The interface design follows the double diamond framework [10], evolving from initial persona definition (novice to expert) to high-fidelity prototyping. Following an initial persona-based design process spanning novice hobbyists to expert creators, we adopted a philosophy of progressive disclosure. The goal was to demystify the MDRNN without overwhelming users or compromising musical accessibility.

The user journey is structured around three views supporting the machine learning lifecycle (Figure 1).

Project and Data Management. To address the opacity often associated with neural networks, the system exposes the lineage of the AI model. The Project Dashboard abstracts the underlying file system, allowing users to organise performance logs and model checkpoints. Crucially, this view enables collaborative provenance, where users can curate datasets from previous sessions and explicitly link them to new training jobs. This transforms the training process from a hidden backend task into a visible, auditable component of the creative workflow.

Training Configuration. The Training View (Figure 2) displays live training metrics, allowing users to monitor model convergence as it occurs. Rather than a static progress indicator, the interface visualises training loss and convergence metrics in real time. This pedagogical scaffolding helps users understand the

[1]https://github.com/cpmpercussion/impsy

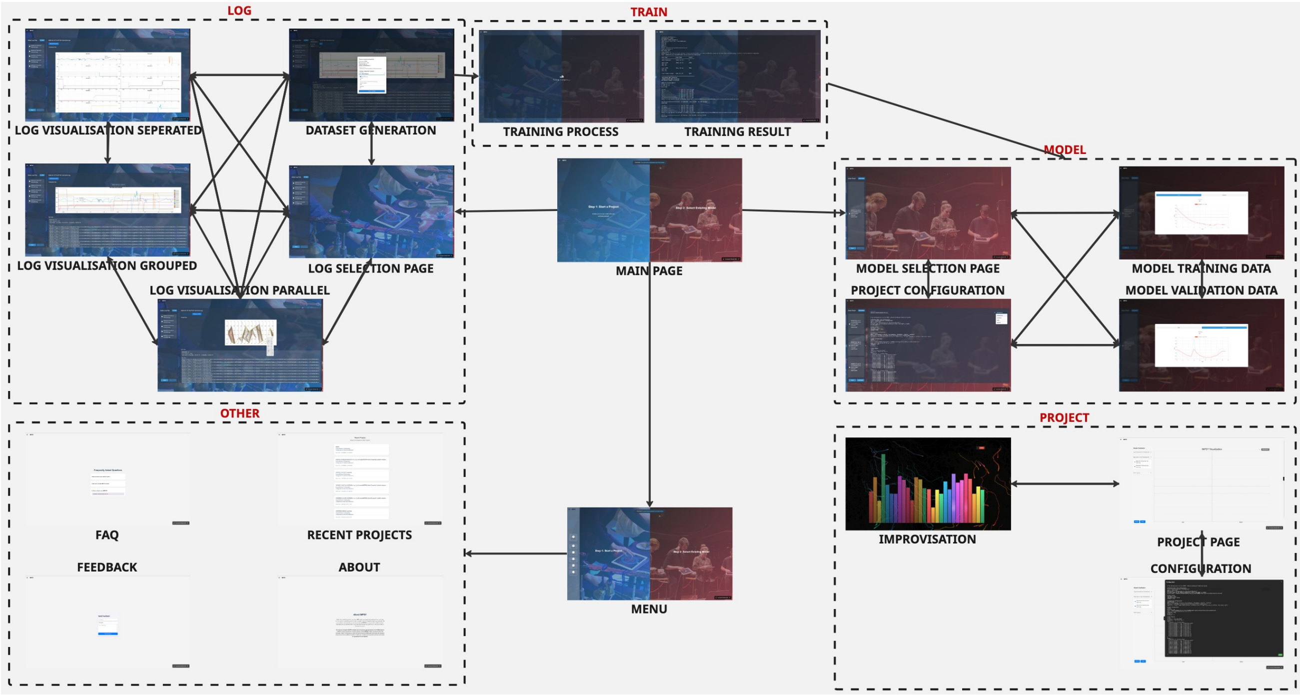

Figure 1: Interaction flow showing navigation between improvisation, log management, model training, project configuration, and supplementary information.

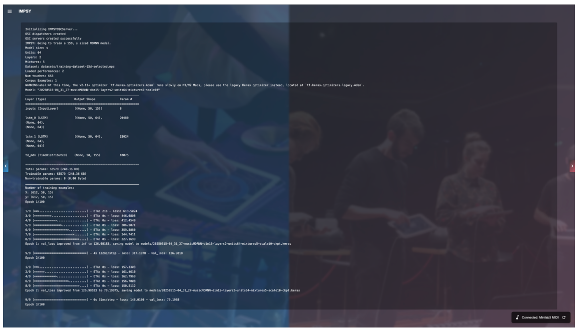

Figure 2: Training View showing model parameters and live training feedback.

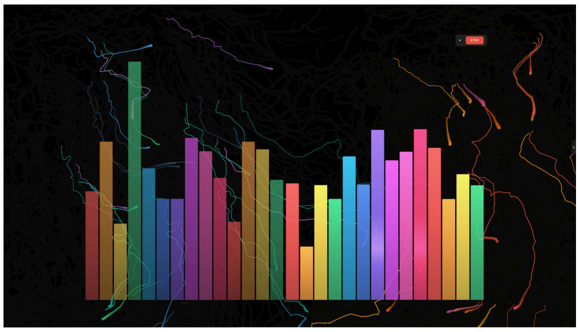

Figure 3: Improvisation Page with I/O Visualisation.

relationship between their dataset and the resulting model performance, bridging the gap between musical intent and algorithmic outcome.

Performance and Real-Time Visualisation. The central contribution of the interface is the Performance View, which implements a dual-path visualisation strategy to support agency during improvisation.

First, to support precise control, the `InputVis` and `OutputVis` components utilise D3.js to render instantaneous control values as dynamic bar graphs. This allows for analytical verification of the musical interaction, ensuring that specific inputs are correctly mapped and received by the system.

Second, the `BackgroundVis` component renders an immersive, canvas-based particle system (Figure 3). Using flow field dynamics, this system responds to musical features such as note density and amplitude. Human inputs are rendered in cool tones (blue and teal), while the AI's generative predictions appear in warm tones (red and orange). This colour scheme enables performers to instantly distinguish their own contributions from the machine's stochastic responses. By rendering the MDRNN model's output as a tangible visual element, the interface shifts the user's role from a passive operator to an active partner in the generative loop.

3.2 Real-Time Data Flow

As illustrated in Figure 4, the architecture bifurcates data streams to ensure that visual feedback remains synchronised with auditory output, which is a critical requirement for perceiving causality in digital instruments.

Upon reception by the backend MIDI server, each incoming hardware MIDI message is processed through a single callback. This callback logs the interaction to a timestamped file on disk, recording raw material for future dataset curation and model training, and enqueues the data into the interface input queue for MDRNN processing. Separately, when the OSC server architecture is active, an intermediary OSC bridge receives device

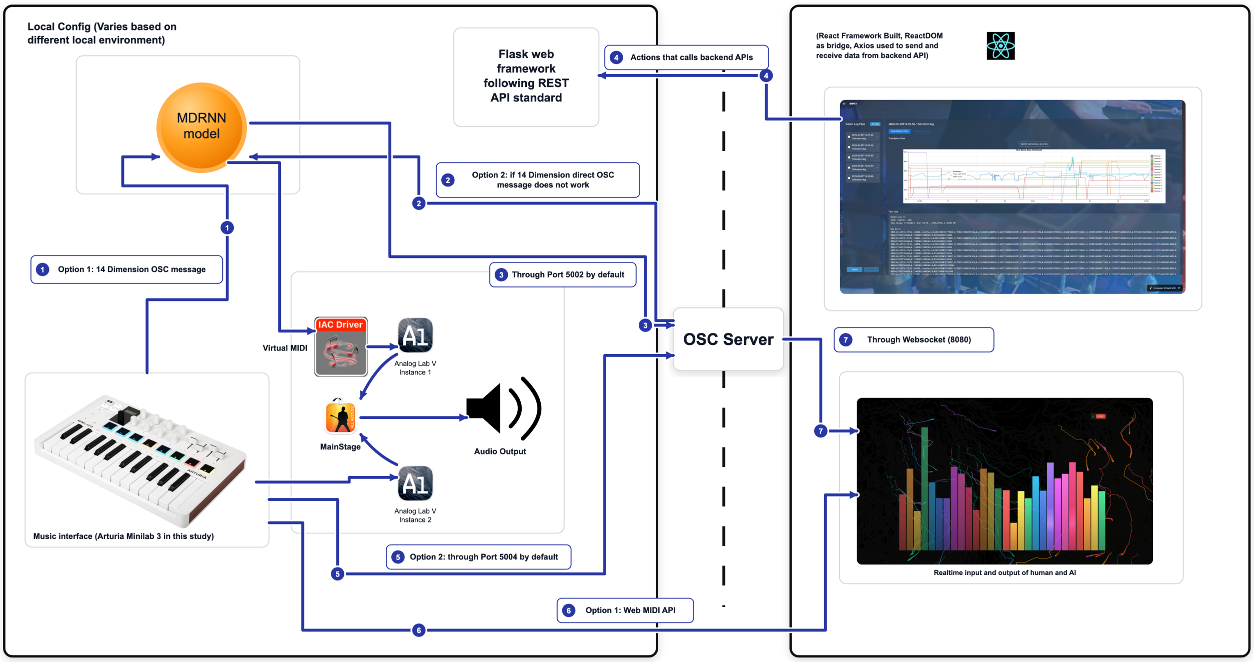

Figure 4: Data flow diagram illustrating the interaction between hardware MIDI devices, the IMPSY backend core, and the frontend visualisation components.

messages and forwards them to the frontend over WebSocket for output visualisation. For input visualisation, the frontend independently accesses the hardware MIDI controller directly via the browser's Web MIDI API, rather than relying on data routed back through the backend. This avoids coupling the input display to the backend's processing pipeline and ensures the frontend can visualise performer input regardless of which backend IO configuration is active.

The generative loop monitors the input queue and feeds data to the MDRNN model. Model outputs are buffered and consumed by a dedicated playback thread that schedules events according to the model's predicted time deltas. These predictions are converted back into standard MIDI messages for synthesis and simultaneously transmitted to the frontend's particle system. This architecture ensures that the AI-associated visual elements appear on screen at the exact moment the corresponding sound is triggered, reinforcing the perception of the AI as a co-present musical entity.

3.3 Backend Architecture

The system implements a decoupled client-server architecture composed of two primary layers: a Python-based backend encapsulating the machine learning logic, and a React-based frontend handling interaction.

The backend (Python 3.11) is organised into modular components to support extensibility. The core generative logic implements mixture density recurrent neural networks using Keras and TensorFlow. Separate modules manage interaction state machines and multi-protocol input and output, orchestrated by a central interaction server. To prevent the time-delayed playback of MDRNN outputs from blocking the main prediction and IO loop, the playback scheduling executes as a separate thread, decoupling the output timing from the inference and input-handling cycle.

Communication employs a polyglot strategy: raw MIDI handles hardware control signals with high precision; WebSockets provide persistent bidirectional state synchronisation for the UI; and standard HTTP/REST endpoints manage non-real-time configuration tasks. This hybrid approach allows the system to deliver the rich user experience of a modern web application without sacrificing the timing stability required for musical performance.

A demo video is provided as supplementary material.

4 Evaluation

4.1 Study Design

For the evaluation study, an Arturia Minilab 3 keyboard was connected via USB to a laptop running the IMPSY web interface. This compact setup approximated a home-studio environment while minimising technical barriers for participants (Figure 5).

Our approach is guided by standard HCI usability frameworks (e.g., think-aloud protocols and post-task interviews) often used in creative musical interface research, such as those demonstrated in [21] and [32]. In these studies, users perform specified musical tasks with a prototype, while researchers observe their workflow and gather qualitative feedback to identify interface strengths and weaknesses.

Further, the user study has been designed following examples from NIME-based evaluations of interactive musical systems [e.g., 30], where short improvisational tasks and post-session interviews offer both quantitative and qualitative insights on usability. The study involved interviews with individuals [21]

Figure 5: Study Setup

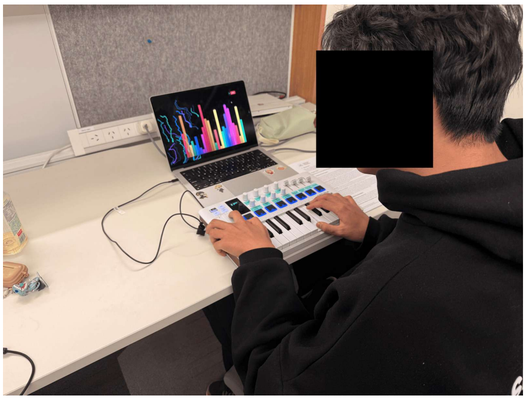

Figure 6: Participant Interacting with the system during Improvisation

and thematic analysis of qualitative data [3, 4, 27], providing a clear framework for interpreting user experiences with the interface.

We recruited five participants from a university campus, with different levels of musical proficiency and technical knowledge (Table 1). Due to the study's exploratory nature, a small sample size of five participants was chosen and we focussed on interpreting qualitative data. While the most common sample size at CHI is 12, studies with small qualitative samples remain within the documented range of accepted practice in HCI research [7].

Table 1: Participant Profiles

ID	Musical Background	Technical Background
P1	Hobbyist (Moderate)	Minimal technical exp.
P2	Music Student (Expert)	Limited computing knowledge
P3	Hobbyist (Little knowledge)	No ML experience
P4	Hobbyist (Limited knowledge)	Studying Machine Learning
P5	Hobbyist (No experience)	Limited computing knowledge

Each session lasted approximately 40 minutes and followed a structured protocol designed to guide users through the full machine learning lifecycle, rather than just performance. The sessions followed this procedure:

(1) **Tutorial & Familiarisation:** A brief tutorial to introduce participants to the web application, its controls, data visualisations, and AI-driven feedback, as well as the electronic music controller. We take cues from prior musical HCI evaluations in which an initial guided session helps users understand basic functionality [21].
(2) **Live Musical Interaction Task:** Participants engaged in short improvisations on the keyboard controller with the system acting as a real-time predictive partner, inspired by best practices from creative AI user studies [32]. Participants used our system to create a small AI model and interact with that model. Participants were observed to understand the interface's feedback, to transition between user-driven and AI-driven performance, and to identify any usability obstacles.
(3) **Survey & Semi-Structured Interview:** After the improvisation, participants completed a survey (adapted from standard usability scales) and a semi-structured interview. The interview focused on eliciting feedback regarding interface design and its impact on creativity, following approaches used in prior musical interface research [30].

4.2 Quantitative Results

The System Usability Scale (SUS) [5] yielded a mean score of **62.50** (SD = 11.9), within the "Marginal" to "OK" range. While slightly below the industry average of 68, given the small sample size and single-session format, this score should be treated as indicative rather than representative, and read alongside the qualitative findings. This interpretation is supported by P4, the only participant with ML experience, who scored substantially higher at 81.25, suggesting that prior domain knowledge may have reduced perceived interface friction.

Table 2: System Usability Scale (SUS) Results

ID	Score	Interpretation
P1	56.25	OK
P2	64.58	OK
P3	58.33	OK
P4	81.25	Excellent
P5	52.08	OK
Mean	**62.50**	**Marginal / OK**

4.2.1 Analysis of Usability Facets. A breakdown of the SUS item scores reveals a distinct dichotomy between the interface's *performance* features and its *configuration* features.

- **Visual Clarity (High):** Question 11 ("Real-time Feedback Clarity") received the highest consensus (Mean 4.0/5), with participants confirming that the particle systems and bar graphs made the AI's status immediately intelligible.
- **Learnability (Low):** Question 10 ("I needed to learn a lot of things before I could get going") indicated friction (Mean 2.6/5 inverted), suggesting that while the *interface* was clean, the *concept* of training an MDRNN remained cognitively demanding for P1 and P5.

P4 (Studying Machine Learning) rated the system as "Excellent" (81.25), substantially higher than P5 (52.08). This contrast may be consistent with the wider variability observed in SUS items related to complexity, need for support, and learning curve. P4's familiarity with concepts such as model training, data selection, and system state may have reduced the perceived cognitive load of IMPSY's end-to-end pipeline, allowing them to treat configuration, logging, and model control as expected parts of an interactive ML workflow rather than as sources of friction. Conversely, P5 may have had to infer these abstractions from the interface alone, which could have amplified uncertainty around whether the AI was "learning correctly" and how their actions shaped the resulting models, contributing to lower confidence (Q9) and a more modest overall SUS score. However, this comparison involves individual cases rather than evidence of a general trend. We therefore treat it as a hypothesis for future study: users with prior ML familiarity may require less scaffolding, while novice users may benefit from clearer explanations of training, data provenance, and model state.

4.3 Qualitative Findings

Thematic analysis [3, 4] of the semi-structured interviews provided critical context for the usability scores, revealing that while the learning curve was steep, the depth of engagement was high. We identified four key themes that explain how transparency facilitates co-creativity.

Theme 1: Empowerment through Flexibility. Participants described the system as enabling a new mode of creative discovery that felt distinct from traditional composition. P1 characterised the session as a "brand new experience," while P3 highlighted the enjoyment of "watching the colours and how they respond." This sense of empowerment was closely linked to the *Call-Response* mode, which fostered a sense of partnership rather than mere control. P5 remarked, "I felt connected to the music... I got good feedback," indicating that the system's ability to "listen" and "reply" created a feedback loop that validated their musical agency.

Theme 2: Trust-Building through Transparent Feedback. Transparency emerged as a prominent factor for user engagement. The real-time visualisations allowed participants to verify the AI's contributions, bridging the trust gap often associated with "black box" generative systems. P4 noted, "The visualisations are very helpful... I understand what's happening," and P2 specifically pointed to the retrospective tools, stating, "The logs are the clearest part." By visually clarifying the human (cold coloured particles) and MDRNN (warm coloured particles) actions on a shared canvas, the interface allowed performers to instantly distinguish their agency from the machine's. This operational transparency demystified the AI's behaviour, transforming it from an erratic generator into a verifiable instrument.

Theme 3: Positive Experience via Engaging Design. Despite the moderate usability scores, the interface's aesthetic and functional design was frequently praised for lowering the barrier to entry. P2 (an expert music software user) commented, "The interface is beautifully made... Very clean," contrasting it favourably with the cluttered, text-heavy interfaces often found in research prototypes. P4 observed that the workflow was "pretty clear and straightforward," taking only "five minutes to figure out." This suggests that the moderate SUS scores may reflect both the conceptual demands of ML training and remaining interface learnability friction, rather than either factor alone. The React-based interface leveraged familiar web UI conventions to reduce the "intimidation factor" of the underlying neural network.

Theme 4: Emotional Resonance and Immersion. Perhaps most importantly, users reported high levels of affective engagement that transcended functional utility. P1 described the interaction as "immersive," and P5 expressed a strong desire to "show an audience how AI works with my creation." This finding is significant for NIME design: it demonstrates that an interface can be challenging to learn (low learnability) yet still highly successful as an expressive instrument (high engagement). The disconnect between the "Marginal" usability scores and the high emotional investment suggests that for creative AI tools, the primary value proposition is not efficiency, but the ability to foster a compelling, transparent dialogue with the machine.

Observational Insights. Direct observations during the study sessions complemented the interview data, revealing interaction patterns that illuminated users' real-time engagement with the system. A consistent finding was participants' high reliance on the real-time visualisations to guide their improvisation, with all five users frequently glancing at the particle canvas to assess the AI's responsiveness before adjusting their input—indicating that the dual-path visualisation strategy successfully anchored their musical decisions. Differences emerged in model training approaches: technically proficient users (e.g., P4) methodically curated inputs over multiple iterations, while novices (e.g., P5) adopted exploratory styles, generating diverse but unstructured data that influenced subsequent AI behaviour. Non-verbal cues further underscored theme cohesion—focused concentration during performance views contrasted with brief hesitations at config controls, affirming the engaging design's role in immersion while highlighting minor learnability friction. These observations support the themes of empowerment and transparency, demonstrating how visual feedback fostered intuitive co-creativity across varying expertise levels.

5 Discussion & Conclusion

This exploratory study examined how visualising an interactive machine learning process shaped musicians' experiences of performing with an intelligent instrument. Our findings suggest that transparency appears to be one important factor in supporting creative agency. By rendering the performer-facing state of the MDRNN and the provenance of its training data as real-time visual elements, the interface helped make the otherwise opaque behaviour of the neural network more intelligible as a musical partner. Qualitative feedback suggested that this visualisation allowed performers to distinguish their own controls from the machine's responses, supporting a sense of dialogue where the AI could be viewed as an active agent rather than a passive tool.

However, the quantitative results highlight the challenge of evaluating such systems in short-term studies. The moderate System Usability Scale (SUS) score (62.5) may reflect a conflation between the usability of the interface and the musical quality of the AI output. The study protocol, limited to a single 40-minute session, meant that participants generated relatively small training datasets. Since MDRNNs require sufficient data density to learn robust stylistic patterns, the resulting models often lacked the depth required for satisfying personalisation. This may have contributed to a mismatch in expectations, particularly for participants with limited musical backgrounds: they anticipated structured "melodies," but often received chaotic control inputs due to the sparse training data. Rather than providing a definitive explanation of the SUS results, this interpretation suggests a design tension between interface clarity, model behaviour, and users' expectations of musical output.

Despite these constraints, the core contribution of this work remains the evidence from this study that exposing data provenance and generative processes can support co-creative engagement among participants, suggesting that it merits further investigation as a design principle for intelligent musical instruments. Even when the musical output was imperfect, the transparent visualisation allowed users to develop some understanding of *why*, shifting the user's role from a passive consumer to an active curator of the system's learning. Directly addressing the RQ, visualisation of the interactive ML process appeared to support performers' experience by contributing to trust, agency, and immersion, and by rendering the MDRNN more verifiable as a musical partner rather than an opaque tool. This aligns with the NIME 2026 theme of *Communities* by bridging the gap between technical logic and musical intuition, allowing diverse stakeholders to engage with AI as a shared creative medium.

Future work must address the data scarcity issue through longitudinal studies, allowing musicians to curate extensive personal datasets over weeks or months. Ultimately, we conclude that while interface transparency cannot compensate for insufficient training data, it can help musicians perceive, trust, and gradually learn to shape intelligent instruments.

6 Ethical Standards

This research applies a small-data machine learning approach emphasising sustainability. The ethical aspects of this study were approved by the ANU Human Research Ethics Committee (Protocol H/2025/0134). Participants were fully informed of the implications of their engagement.

References

[1] Ableton AG. 2025. Ableton Live. https://www.ableton.com/en/live/ Digital audio workstation software.

[2] Apple Inc. 2025. Logic Pro. https://www.apple.com/logic-pro/ Digital audio workstation software.

[3] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3 (01 2006), 77–101. https://doi.org/10.1191/1478088706qp063oa

[4] Virginia Braun and Victoria Clarke. 2019. Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health* 11 (06 2019), 1–9. https://doi.org/10.1080/2159676X.2019.1628806

[5] John Brooke. 2013. SUS: a retrospective. *J. Usability Studies* 8, 2 (Feb. 2013), 29–40.

[6] Dom Brown, Chris Nash, and Tom Mitchell. 2017. A User Experience Review of Music Interaction Evaluations. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Aalborg University Copenhagen, Copenhagen, Denmark, 370–375. https://doi.org/10.5281/zenodo.1176286

[7] Kelly Caine. 2016. Local Standards for Sample Size at CHI. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 981–992. https://doi.org/10.1145/2858036.2858498

[8] John M. Carroll. 2003. *HCI Models, Theories, and Frameworks: Toward a Multidisciplinary Science*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

[9] A. Chatzimparmpas, R. M. Martins, I. Jusufi, K. Kucher, F. Rossi, and A. Kerren. 2020. The State of the Art in Enhancing Trust in Machine Learning Models with the Use of Visualizations. *Computer Graphics Forum* 39, 3 (June 2020), 713–756. https://doi.org/10.1111/cgf.14034

[10] Design Council. 2005. The 'Double Diamond' Design Process Model. https://www.designcouncil.org.uk/news-opinion/design-process-what-double-diamond. Accessed: 2025-05-20.

[11] Alan Dix, Janet E. Finlay, Gregory D. Abowd, and Russell Beale. 2004. *Human-Computer Interaction* (3rd ed.). Pearson Prentice Hall.

[12] Rebecca Fiebrink and Perry Cook. 2010. The Wekinator: A System for Real-time, Interactive Machine Learning in Music. *Proceedings of The Eleventh International Society for Music Information Retrieval Conference (ISMIR 2010)* (01 2010).

[13] Rebecca Fiebrink and Laetitia Sonami. 2020. Reflections on Eight Years of Instrument Creation with Machine Learning. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*. https://www.nime.org/proceedings/2020/nime2020_paper45.pdf

[14] Google Creative Lab. 2025. Teachable Machine. https://teachablemachine.withgoogle.com/. Accessed: 14 March 2025.

[15] Alex Graves. 2014. Generating Sequences With Recurrent Neural Networks. arXiv:1308.0850 [cs.NE] https://arxiv.org/abs/1308.0850

[16] Lamtharn Hantrakul. 2018. GestureRNN: A neural gesture system for the Roli Lightpad Block. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, Thomas Martin Luke Dahl, Douglas Bowman (Ed.). Virginia Tech, Blacksburg, Virginia, USA, 132–137. https://doi.org/10.5281/zenodo.1302703

[17] Marc Hassenzahl, Sarah Diefenbach, and Anja Göritz. 2010. Needs, affect, and interactive products - Facets of user experience. *Interact. Comput.* 22, 5 (Sept. 2010), 353–362. https://doi.org/10.1016/j.intcom.2010.04.002

[18] Jeffrey Heer and Ben Shneiderman. 2012. Interactive dynamics for visual analysis. *Commun. ACM* 55, 4 (April 2012), 45–54. https://doi.org/10.1145/2133806.2133821

[19] Eric J. Isaacson. 2005. What You See Is What You Get: on Visualizing Music. In *International Society for Music Information Retrieval Conference*. https://api.semanticscholar.org/CorpusID:14453331

[20] 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*, Miguel Ortiz and Adnan Marquez-Borbon (Eds.). Mexico City, Mexico, Article 46, 13 pages. https://doi.org/10.5281/zenodo.11189198

[21] Chris Kiefer, Nick Collins, and Geraldine Fitzpatrick. 2008. HCI Methodology for Evaluating Musical Controllers: A Case Study. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*. 87–90. https://www.nime.org/proceedings/2008/nime2008_087.pdf

[22] Shelly Knotts. 2021. Algorithmic Power Ballads. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Shanghai, China, Article 36. https://doi.org/10.21428/92fbeb44.548cca2b

[23] Hugo B. Lima, Carlos G. R. Dos Santos, and Bianchi S. Meiguins. 2021. A Survey of Music Visualization Techniques. *ACM Comput. Surv.* 54, 7, Article 143 (July 2021), 29 pages. https://doi.org/10.1145/3461835

[24] Charles P. Martin, Kai Olav Ellefsen, and Jim Torresen. 2017. Deep Models for Ensemble Touch-Screen Improvisation. In *Proceedings of the 12th International Audio Mostly Conference on Augmented and Participatory Sound and Music Experiences* (London, United Kingdom) *(AM '17)*. Association for Computing Machinery, New York, NY, USA, Article 4, 4 pages. https://doi.org/10.1145/3123514.3123556

[25] Charles Patrick Martin, Kyrre Glette, Tønnes Frostad Nygaard, and Jim Torresen. 2020. Understanding Musical Predictions With an Embodied Interface for Musical Machine Learning. *Frontiers in Artificial Intelligence* Volume 3 - 2020 (2020). https://doi.org/10.3389/frai.2020.00006

[26] 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*, Marcelo Queiroz and Anna Xambó Sedó (Eds.). Porto Alegre, Brazil, 260–265. https://doi.org/10.5281/zenodo.3672952

[27] Muhammad Naeem, Wilson Ozuem, Kerry Howell, and Silvia Ranfagni. 2023. A Step-by-Step Process of Thematic Analysis to Develop a Conceptual Model in Qualitative Research. *International Journal of Qualitative Methods* 22 (2023), 16094069231205789. https://doi.org/10.1177/16094069231205789

[28] University of St. Thomas Playful Learning Lab. 2018. Code+Chords Project. https://www.playfullearninglab.org/code-chords.

[29] 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

[30] Atau Tanaka, Rebecca Fiebrink, and Adam M. Parkinson. 2012. A Qualitative Evaluation of Augmented Human-Human Interaction in a Music Ensemble. In *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*. 144–147. https://www.nime.org/proceedings/2012/nime2012_108.pdf

[31] Benjamin Taylor, Jesse Allison, William Conlin, Yemin Oh, and Daniel Holmes. 2014. Simplified Expressive Mobile Development with NexusUI, NexusUp, and NexusDrop. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Goldsmiths, University of London, London, United Kingdom, 257–262. https://doi.org/10.5281/zenodo.1178951

[32] Renaud Bougueng Tchemeube, Jeffrey Ens, Cale Plut, Philippe Pasquier, Maryam Saf, Yvan Grabit, and Jean-Baptiste Rolland. 2023. Evaluating Human-AI Interaction via Usability, User Experience and Acceptance Measures for MMM-C: A Creative AI System for Music Composition. In *Proceedings of the 32nd International Joint Conference on Artificial Intelligence (IJCAI)*. 5769–5778. https://www.ijcai.org/proceedings/2023/0640.pdf

[33] Johnty Wang, Joseph Malloch, Stephen Sinclair, Jonathan Wilansky, Aaron Krajeski, and Marcelo M. Wanderley. 2019. Webmapper: A Tool for Visualizing and Manipulating Mappings in Digital Musical Instruments. In *Proceedings of the 14th International Symposium on Computer Music Multidisciplinary Research (CMMR 2019)*. Marseille, France.

[34] Xiao Xiao, Basheer Tome, and Hiroshi Ishii. 2014. Andante: Walking Figures on the Piano Keyboard to Visualize Musical Motion. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Goldsmiths, University of London, London, United Kingdom, 629–632. https://doi.org/10

.5281/zenodo.1178987