
Explainable AI for the Arts

Nick Bryan-Kinns

Queen Mary University of
London
London, UK
n.bryan-kinns@qmul.ac.uk

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

Copyright held by the owner/author(s).
Human-Centered Explainable AI (HCXAI) Workshop at CHI 2023, April 28, 2023, Hamburg, Germany

Abstract

The Arts offer a complex and demanding domain in which to undertake AI and Explainable AI (XAI) research with substantially lower life-critical risks than domains such as health, mental health, medicine, and automotive contexts. However, despite the growth in XAI research there is very little research on XAI for the Arts (XAIxArts) with most XAI research examining task-oriented explanations of AI decisions. This paper outlines a framework for characterising XAI for the Arts and illustrates its use by reviewing 87 AI music generation systems in terms of their explainability. A demo XAI generative music system is introduced which offers meaningful real-time interaction with latent spaces for music generation. The paper concludes by reflecting on the conundrum of how much explanation we should strive for in XAI for the Arts.

Author Keywords

XAI; XAIxArts; AI Arts; generative music.

CCS Concepts

•**Computing methodologies** → **Artificial intelligence**;
•**Human-centered computing** → **Human computer interaction (HCI)**; *Interaction design*; Visualization; •**Applied computing** → **Arts and humanities**;

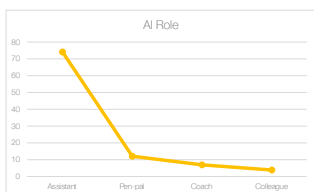


Figure 1: The role of the AI in a survey AI music systems

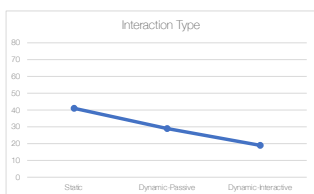


Figure 2: Interaction with the AI in a survey AI music systems



Figure 3: Common ground with the AI in a survey AI music systems

Introduction

Artificial Intelligence (AI) has been used in artistic and creative practice since the birth of computing. However, despite the growth in research on explainable AI (XAI) there is very little research on XAI for the Arts [4] (XAIxArts). Current XAI research predominantly examines explanations of AI decisions in task-oriented and goal-directed situations such as house price prediction (e.g. [11] from HCXAI22). The Arts offer a complex and demanding domain in which to undertake AI and XAI research with substantially lower life-critical risks than current XAI domains such as health, mental health (e.g. [13] from HCXAI22), medicine, and automotive contexts. For example, in a musical performance an AI system must be robust, reliable, and responsive – it is a catastrophic failure if the system needs to be rebooted during a performance in front of an audience of 10,000 or even 10 people – and yet if something does go wrong the results are not life-endangering.

In terms of the HCXAI23 workshop goals and themes, examining XAI for the Arts diversifies the *application areas* for XAI to more creative and aesthetic domains. It expands the *landscape of 'whos' in XAI* to include artists and creative practitioners whose *user goals* are typically open-ended and experiential. Moreover, examining XAI for the Arts needs to account for users who are more focused on expressive and aesthetic values rather than task completion and explanation efficiency as is typically considered in current XAI *evaluation methods*. Expanding XAI to examine XAI for the Arts shares many parallels with the move to third wave Human-Computer Interaction (HCI) [3] and the examination of the experiential properties of HCI.

Papers in previous HCXAI workshops touched on creative and designerly aspects related to XAI, but did not examine the potential of XAI in creative and artistic practice directly.

For example, [8] in HCXAI21 considered the design of XAI rather than XAI for design, and [2] in HCXAI22 examined the creative (unexpected) uses of XAI rather than designing XAI for creative practice per se. The lack of research on XAI for the Arts offers an opportunity to explore a new domain for XAI where an AI's role may range from a tool which creates content overnight to a collaborator which engages in co-creation in-the-moment cf. [14]. In these situations the nature of explanations may be very different to current forms of XAI explanations suggested by [10] and [9].

A Framework for Categorising XAI for the Arts

There are several extremely thorough surveys of XAI systems e.g. [9]. However, these focus on functional tasks and explanations of decision making rather than understanding AI in creative and interactive arts contexts. Instead we developed our own categorisation of XAI for the Arts [4] based on three existing frameworks to capture key features of AI in creative settings: *The role of the AI* – from Assistant to Colleague drawing on [14]; *The possible Interaction with the AI* – from Static (no human input) to Dynamic-interactive (responding to human input) drawing on [6]; and *How much common ground* a user might establish with the AI – from no understanding of the AI (Stage 0) to understanding how a user's input changed the AI output (Stage 3) and what possible responses are (Stage 4) drawing on [5]. In this view the explainability of AI for the Arts is a combination of its role, interaction, and grounding. These three elements are entangled, and real-time interaction is crucial to creating more explainable XAI for the Arts given the aesthetic, exploratory, and subjective nature of artistic endeavour.

The framework was used to analyse XAI for a key form of artistic endeavour – music making – by surveying 87 recent AI music research papers (summarised in [4]). Figures 1, 2, and 3 illustrate the finding that most AI music systems

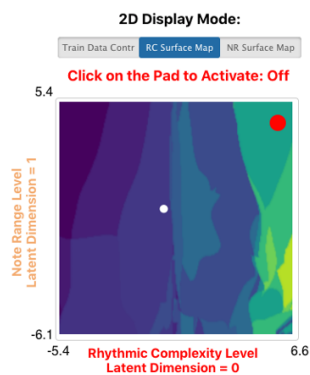


Figure 4: Part of the XAI Generative Music user interface (from [4])

surveyed offered very little explanation of what they do. Indeed, most AI systems took the role of creative tool, taking care of background tasks rather than being a coach or a colleague (fig. 1). It was surprising to find that most systems were static and did not respond directly to human interaction (fig. 2). And, the vast majority of AI Music systems were at stage 1 of grounding (fig. 3) which means that the AI makes a contribution, but the user cannot really discern what the AI did based on their own input. Naturally there were a small number of systems which did exhibit XAI properties. For example, Shimon the robotic marimba player [12] uses a real-time feedback loop within the artwork to make the music generation highly dynamic and interactive and Hyperscore [7] visualises its musical responses to user input allowing users to develop a greater understanding of what the system generated in response to their input.

Case Study: XAI Generative Music

To explore the design of more explainable AI Music generative systems a real-time generative music tool [4] (<https://xai-with-lsr-ui.vercel.app/>) was created based on MeasureVAE [15]. The MeasureVAE system generates a short piece of music similar to a given example piece of music by encoding to and decoding from a 256 dimensional latent space. Latent space regularisation can be employed to force some dimensions to relate to musical metrics, thereby offering some explainability of the otherwise opaque latent space. Figure 4 shows part of the user interface (UI) in which 2 regularised dimensions of the latent space are visualised allowing users to navigate the latent space using musical dimensions of note range and rhythmic complexity, with another part of the UI supporting navigation by note density, and average pitch interval. The visualisations of the latent space in combination with the real-time interaction and musical labelling increase the explainability of the system as: i) the *role* the AI becomes more of a pen-pal or col-

laborator, interactively responding to users as they navigate the space, almost like a duet; ii) the *interaction* is dynamic-interactive, responding in real-time, and can be dynamic-interactive-varying when the musical input is changed; iii) the human-in-the-loop real-time interaction with the system allows for higher levels of *grounding* (Stage 2 or 3) as the AI's latent space is exposed to the user and labelled, and the effect of the user's input is immediately reflected in the latent space visualisation and generated musical output.

To further explore XAI for the Arts the system has been packaged as a plug-in [1] for consumer music making software. It is now being deployed and tested with musicians, embedding the AI model directly into their music making toolchains to explore how XAI for the Arts tools might be used and appropriated in creative practice.

Conclusions

Exploring XAI for the Arts offers a new domain for XAI research which diversifies the set of XAI stakeholders, includes more open-ended user goals, and broadens the range of XAI evaluation criteria. However, the nature of the Arts opens a somewhat philosophical question about what it might mean to understand a creative AI and what kinds of explanations are appropriate in artistic settings. For example, when we co-create music with humans we partly rely on an intuitive understanding of each others' musical intention to mutually engage with each other, but we do not have an in-depth and explicit understanding of why each musician improvises in the way that they do – musicians don't usually ask each other about their rationale and motivation for every single note that they play when they are jamming. As XAI researchers we should question whether to aim for detailed explanations from our creative AI systems at all, or whether conveying the *gist* of what an AI is doing is really what is valuable to strive for in XAI for the Arts.

Acknowledgements

Supported by the Queen Mary University of London Research Enabling Fund. Thanks to the researchers who worked on earlier papers [4][1]: Berker Banar, Corey Ford, Courtney N. Reed, Yixiao Zhang, and Jack Armitage.

REFERENCES

- [1] Berker Banar, Nick Bryan-Kinns, and Simon Colton. 2023. A Tool for Generating Controllable Variations of Musical Themes using Variational Autoencoders with Latent Space Regularisation. In *Proceedings of the 37th AAAI Conference on Artificial Intelligence*.
- [2] Michaela Benk, Raphael Weibel, and Andrea Ferrario. 2022. Creative Uses of AI Systems and their Explanations: A Case Study from Insurance. In *ACM CHI HCXAI22 Workshop*.
- [3] Susanne Bødker. 2015. Third-wave HCI, 10 years later—participation and sharing. *Interactions* 22, 5 (aug 2015), 24–31.
- [4] Nick Bryan-Kinns, Berker Banar, Corey Ford, Courtney N. Reed, Yixiao Zhang, Simon Colton, and Jack Armitage. 2021. Exploring XAI for the Arts: Explaining Latent Space in Generative Music. In *Proceedings of the 1st Workshop on eXplainable AI Approaches for Debugging and Diagnosis, NeurIPS*.
- [5] H Clark and S Brennan. 1992. Grounding in Communication. *Archives* 7, July (1992), 734–738.
- [6] Stroud Cornock and Ernest Edmonds. 1973. The Creative Process Where the Artist Is Amplified or Superseded by the Computer. *Leonardo* 6, 1 (1973), 11–16.
- [7] Morwaread Farbood, Henry Kaufman, and Kevin Jennings. 2007. Composing with Hyperscore: An Intuitive Interface for Visualising Musical Structure. In *International Computer Music Conference (ICMC 2007)*. 111–117.
- [8] Juliana Jansen Ferreira and Mateus de Souza Monteiro. 2021. Designer-User Communication for XAI: An epistemological approach to discuss XAI design. In *ACM CHI HCXAI21 Workshop*.
- [9] Riccardo Guidotti, Anna Monreale, Salvatore Ruggieri, Franco Turini, Dino Pedreschi, and Fosca Giannotti. 2018. A Survey Of Methods For Explaining Black Box Models. (2018).
- [10] David Gunning. 2016. Explainable Artificial Intelligence (XAI). *DARPA/I2O Proposers Day* (Aug 2016).
- [11] Patrick Hemmer, Max Schemmer, Niklas Kühn, Michael Vössing, and Gerhard Satzger. 2022. The Role of Information Asymmetry in Human-AI Decision-Making. In *ACM CHI HCXAI22 Workshop*.
- [12] Guy Hoffman and Gil Weinberg. 2010. Shimon: an interactive improvisational robotic marimba player. In *CHI'10 Extended Abstracts on Human Factors in Computing Systems*. 3097–3102.
- [13] Elena Korshakova and Sang Won Bae. 2022. Towards Human-Centric XAI Chatbots in Mental Health for End-User Experiences. In *ACM CHI HCXAI22 Workshop*.
- [14] Todd Lubart. 2005. How can computers be partners in the creative process: Classification and commentary on the Special Issue. *International Journal of Human-Computer Studies* 63, 4 (2005), 365–369.
- [15] Ashis Pati, Alexander Lerch, and Gaëtan Hadjeres. 2019. Learning to traverse latent spaces for musical score inpainting. *arXiv preprint arXiv:1907.01164* (2019).