



Editorial for TISMIR Special Collection: AI and Musical Creativity

EDITORIAL

BOB L. T. STURM 

ALEXANDRA L. UITDENBOGERD 

HENDRIK VINCENT KOOPS 

ANNA HUANG 

**Author affiliations can be found in the back matter of this article*

]u[ubiquity press

ABSTRACT

This special issue focuses on research developments and critical thought in the domain of artificial intelligence (AI) applied to modeling and creating music. It is motivated by the AI Song Contests of 2020 and 2021, in which the four guest editors adjudicated or participated among many teams from around the world. The 2020 edition had 13 submissions and the 2021 edition had 38. The 2022 edition is now being planned. These unique events provide exciting opportunities for AI music researchers to test the state of the art and push the boundaries of what is possible, within the context of music creation. They portend a future when humans and machines work together as partners in music creation. Maybe “portend” is not the right term, but we must not think that the future of AI and music is only warm and fuzzy. It is important and timely to consider how we, in local and global contexts, can effectively and ethically develop and apply AI in contexts of music creation.

CORRESPONDING AUTHOR:

Bob L. T. Sturm

KTH, SE

bobs@kth.se

KEYWORDS:

Artificial intelligence; music generation; creativity; ethics

TO CITE THIS ARTICLE:

Sturm, B. L. T., Uitdenbogerd, A. L., Koops, H. V., and Huang, A. (2022). Editorial for TISMIR Special Collection: AI and Musical Creativity. *Transactions of the International Society for Music Information Retrieval*, 5(1), 67–70. DOI: <https://doi.org/10.5334/tismir.129>

Nine articles appear in this special collection, contributing technical knowledge, discussions of the practicalities of working and assessing AI applied to music, and reflecting critically on some of its ethical dimensions. Below we provide a brief summary of each article.

In “I Keep Counting: An Experiment in Human/AI Co-creative Songwriting”, Micchi et al. – a team of MIR researchers and amateur musicians – discuss creating their entry to the 2020 AI Song Contest, “*I Keep Counting*” which received fourth place out of the 13 entries overall. The article takes the reader on a detailed and honest tour of the creative process, in which the AI acted as a creative colleague in several aspects of the composition. The AI is seen as suggesting possibilities, which the team balanced with the requirements of the contest and the desire to make a good song. The team trained models that generate song structure, chord progressions, lyrics, and melody (independently). With some curation and random selection, the team settled on specific outputs followed by manual adjustments – some intentional and some accidental. The article reflects on the team’s experience for future iterations of such a contest – and indeed, the team’s 2021 entry won third place overall!

In “We Are Not Groupies... We Are Band Aids’: Assessment Reliability in the AI Song Contest”, Burgoyne and Koops statistically analyze the results of the 2020 AI Song Contest. The ranking of 2020 entries was based on a combination of a popular vote and the assessment of a panel of judges considering multiple criteria. If a team’s popular vote was due more to their effectiveness in getting people to vote for them rather than the quality of their submission, then the AI Song Contest would be a popularity contest instead of being an “international competition exploring the use of AI in the songwriting process” (<https://www.aisongcontest.com/about>). Burgoyne and Koops show that there indeed was a large number of voters exhibiting “groupie-like” behaviors, e.g., giving one team perfect scores and not listening to any other, but involving a jury in the assessment mitigated this impact – at least for the top-tier entries. Burgoyne and Koops go further with suggestions for future AI Song Contests. It will be very interesting to see what their models of voting behavior say for the 2021 edition, where the number of entries expanded from 13 to 38.

In “Evaluating Creativity in Automatic Reactive Accompaniment of Jazz Improvisation”, Ostermann, Vatolkin, and Rudolph take their own “EAR Drummer”, a reactive and generative drum accompaniment system based on evolutionary algorithms, outside of the lab to the test with musicians. Twenty jazz musicians were invited to improvise with and compare between two systems blindly: the reactive EAR Drummer, and iReal Pro which provides a static backing track. Participants were then asked to fill out a specific survey, which shows how a reactive generative system adds to the improvisation experience, and how musicians might have

different needs during a practice session versus a stage performance. This paper provides a compelling example of how to carry out a human-centered evaluation for generative music systems.

In “Evaluating an Analysis-by-Synthesis Model for Jazz Improvisation”, Frieler and Zaddach propose to test hypotheses of musicians’ psychological models using generative models. Studying monophonic jazz improvisation, they hypothesize that the creative process could be modeled by a hierarchical Markov model, by first selecting mid-level musical units, such as an abstract representation of a line or lick, then conditioning on the mid-level unit to generate a series of short melodic units, finally realizing them in a sequence of pitch and rhythm. In the second half of the article, the authors test their hypothesis through a Turing-like listening test, i.e. whether jazz experts and non-experts could distinguish between their model generated solos and solos from human performers. In these experiments, they take a closer look at how “confounding” factors that are added to provide a “fairer” comparison, such as how the generated sequences were rendered (i.e. if expression was manually added) and the level of expertise of the solo performers, can have a large impact on the outcome. The article concludes by offering possible strategies to mitigate some of these challenges and to increase the reliability of similar evaluations.

In “Steerable Music Generation which Satisfies Long-Range Dependency Constraints”, Bodily and Ventura tackle the long-range dependency problem that usually occurs with generative systems for music and/or lyrics. By extending Markov model-based approaches to include constraints learned from existing songs, their approach allows aspects of a generated song to be enforced, such as: repetition; final melodic pitch, duration and chord; stress-based alignment of lyrics and melody; and rhyming conventions. They found that despite satisfying the constraints, there was considerable variety in the resulting generated songs. Their approach is used in their system Pop* (https://www2.cose.isu.edu/~bodipaul/research/pop_star/), which generates “pop, rock, and show tune music using Twitter as an inspiring source from which to compose original music.”

In “Drumroll Please: Modeling Multi-Scale Musical Gestures without Quantizing”, Gillick et al. take a close look at data representations used recently for music generation. More specifically, the authors compare the relative strengths and weaknesses of different approaches to representing expressive percussion data, highlighting opportunities for improvement. The authors propose flexible grid representations. These representations address the inherent problems of modeling expressive percussion patterns with fixed-grid representations. Commonly, in fixed-grid representations, rhythmic events get “lost” when they fall into already occupied grid cells. The flexible grid representation

accounts for this situation by encoding overflow events with additional columns of the grid. When used for music generation, models trained on flexible grids are able to generate music of similar perceptual quality to those using fixed grids, while at the same time incorporating details of the expressive drumming gestures captured by event-based representations.

In “On the Development and Practice of AI Technology for Contemporary Popular Music Production”, Deruty et al. examine how artists of several contemporary popular music styles (e.g., pop, rock, electronica) produce music using a suite of audio-based AI tools developed at Sony Computer Science Laboratories (CSL). Examples of such tools include models that generate drum accompaniment or bass lines given an existing audio track, and tools for mastering tracks. They focus not on the technology itself, but how it is made usable through design, and on its ultimate use by real artists in their personal studio-situated workflows. One interesting observation is that some artists embraced the artifacts or imperfections of AI tools, which are seen to give an identity to the sounds not unlike classic synthesizers. Other artists experimented with using the tools for purposes other than for what they are intended. This article provides a great example of how working with artists can result in AI tools that are relevant, supportive, rewarding and stimulating.

Broader conversations about the ethics of AI applied to music are contributed by two articles. In, “Where Does the Buck Stop? Ethical and Political Issues with AI in Music Creation”, Morreale expands the discussion of AI applied to music from being merely technical, to dimensions that are cultural, political, and economic. He focuses on “commercial AI music” in particular, and calls on the domain of MIR to engage with these dimensions and “self-regulate” in order to reduce adverse harms from the technology it develops and deploys. He writes, “When technological innovation is so firmly situated within and directed by a specific economic system, the argument that describes technology as intrinsically neutral or just maths is a fallacious one.” He argues for the community to acknowledge that some companies do not employ AI in order to generate good music, but to drive down the price of creative labor. Morreale’s article is a strong, timely and ultimately positive call for reflection.

Finally, in “On Creativity, Music’s AI Completeness, and Four Challenges for Artificial Musical Creativity”, Rohrmeier surveys problems inherent to defining “creativity”, in particular its compatibility with computation by formal algorithms, and the use of randomness for overcoming deterministic, and thus uncreative, behavior. He proposes measuring creativity in terms of an output by a system in relation to a formal frame of reference, rather than the procedures that system used in producing its output. Rohrmeier exemplifies this argument with game-playing machines, finding “regions in the search space [involving long-term risky strategies] that have been inaccessible

for [or unthought of by] ... previous [machines] and expert minds.” Although some music practices can be considered as puzzles with solutions, Rohrmeier ultimately argues that music creativity is AI-complete: it requires general intelligence, due to the creative concepts often lying beyond the notes and sounds, and proper evaluation requires models of human cognition, social circumstances, and performance.

This special collection of articles appearing in the transactions of a society historically focused on music information *retrieval* shows just how much the research objectives of the field have diversified – and further motivates interpreting the “R” in ISMIR as “research”. It also seems inevitable that a community interested in extracting information from music data should develop an interest in generating music data from information – at the very least in testing the completeness and accuracy of music models by analysis-by-synthesis. The domain of music generation has been a major focus of communities like the International Computer Music Conference, the International Computational Creativity Conference, Computer Music Modeling and Retrieval, Sound and Music Computing, and the recently formed AI Music Creativity conference. ISMIR has by now clearly joined the fray, and will continue to contribute insightful and technically creative work from its unique and diverse perspectives.

We would like to thank the TISMIR editorial team for their assistance through this process. A special thanks goes to the numerous peer reviewers, for their enthusiasm and diligence in critically engaging with the submitted work, and its revisions. The work of Sturm on this special collection has been supported in part by funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (Grant agreement No. 864189 MUSAiC: Music at the Frontiers of Artificial Creativity and Criticism).

COMPETING INTERESTS

The reviewing of the submission by Burgouyne and Koops (one of the guest editors) was handled by Sturm only, who was not involved in the work. There are no other competing interests.

AUTHOR AFFILIATIONS

Bob L. T. Sturm  orcid.org/0000-0003-2549-6367
KTH, SE

Alexandra L. Uitdenbogerd  orcid.org/0000-0001-9004-8283
RMIT University, AU

Hendrik Vincent Koops  orcid.org/0000-0002-6980-7027
RTL Netherlands, NL

Anna Huang  orcid.org/0000-0002-8440-2378
University of Montreal, CA

TO CITE THIS ARTICLE:

Sturm, B. L. T., Uitdenbogerd, A. L., Koops, H. V., and Huang, A. (2022). Editorial for TISMIR Special Collection: AI and Musical Creativity. *Transactions of the International Society for Music Information Retrieval*, 5(1), 67–70. DOI: <https://doi.org/10.5334/tismir.129>

Submitted: 31 January 2022

Accepted: 11 February 2022

Published: 22 March 2022

COPYRIGHT:

© 2022 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

Transactions of the International Society for Music Information Retrieval is a peer-reviewed open access journal published by Ubiquity Press.

