

Redefining the User in Human-Generative AI Collaboration: Insights from Music Composition

Eric Tron Gianet¹, Luigi Di Caro¹ and Amon Rapp¹

¹University of Turin, Computer Science Department, Torino, Italy

Abstract

The rise of Generative Artificial Intelligence (GenAI) is transforming the role of end users. We investigate human-AI collaboration in music composition as an illustrative example. Existing studies analyze the impact of current AI music models on composers, focusing on challenges and strategies for interaction. However, these studies often neglect the multifaceted nature of music composition, influenced by personal aspirations, social and cultural context, and distinct genre characteristics. We propose an ethnographic approach to explore composers' practices and needs, which can inform the design of human-AI collaborative tools that empower and support them.

Keywords

Human-AI Collaboration, Music Composition, Generative AI, User role

1. Introduction

The growing presence of Artificial Intelligence (AI) in everyday life and human practices raises critical ethical, social, and cultural questions. There is increasing concern that current approaches to AI development may overlook human values and needs. This new landscape has led to calls for a Human-Centered AI (HCAI) [1, 2]: an AI that empowers users, reveals its values, biases, limitations, and the ethics behind its algorithms and data collection, and promotes ethical, interactive, and contestable use [3]. Notably, recent advancements in GenAI have resulted in systems that not only perform classification tasks but can also create artifacts like text and images, making them active agents with creative abilities, thus challenging the traditional concept of the “end user”.

Within HCAI, a focus has emerged that strives towards leveraging the strengths of both humans and AI systems, rather than relying solely on the latter. This strand of research has been called human-AI teaming by Capel and Brereton [3], but various terms have emerged to describe the same collaborative approach: human-computer collaboration [4], human-AI co-creation [5], and human-AI collaboration [6]. Beyond the assumption that, by working together, we can obtain better performance compared to humans or AI systems alone, this collaborative approach represents an “agentistic turn” [6], which means attributing agency to AI within a system of distributed agency. This perspective aligns with Bruno Latour's notion of technological agency, which extends it beyond humans to non-human entities like objects, technologies, and animals

Proceedings of the 8th International Workshop on Cultures of Participation in the Digital Age (CoPDA 2024): Differentiating and Deepening the Concept of "End User" in the Digital Age, June 2024, Arenzano, Italy

✉ erictrngnt@gmail.com (E. Tron Gianet); luigi.dicaro@unito.it (L. Di Caro); amon.rapp@unito.it (A. Rapp)

🆔 0009-0009-3092-7089 (E. Tron Gianet); 0000-0002-7570-637X (L. Di Caro); 0000-0003-3855-9961 (A. Rapp)

© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).



[7]. However, while Latour’s theories highlight the intricate relationships between humans and their surroundings, it is crucial to acknowledge potential critiques of this agentistic turn, as it might obscure the vast amount of “ghost work” [8] that fuels AI, often performed by individuals from the Global South (e.g., data labelers) for low wages [6].

Assuming that interacting with GenAI systems consists of collaboration, and while “collaboration” appears to take place also in interactions with non-generative AIs [9], in this new landscape characterized by GenAI systems that carry out human-like tasks and produce human-like outputs, it is even more necessary to understand how the human role in decision-making, creative, and information processes is being redefined and how “end users” might evolve towards becoming “collaborators” and “co-creators”.

To further examine this, we can focus on a specific domain that could help us understand how collaboration and co-creation occur in practice, taking into account not only the user’s objectives, needs, and motivations, but also the social and cultural context within which humans and GenAI systems may collaborate to achieve situated goals. With this in mind, the following section will explore AI research in music composition as an illustrative example. Collaboration, authorship, creation, and ownership are here central themes and offer valuable insights into the broader dynamics of human-AI collaboration. In the next section we will begin by reviewing recent literature on human-AI music composition. Then, we will propose a study investigating the situated practices [10] of music composition.

2. Music Composition and Artificial Intelligence

Computer-generated music and computer-based music composition have a long history, going back at least to 1957 with the “Illiac Suite”, regarded as the first computer-composed score [11, 12]. Then, from the 80s, there has been a surge in interest, with Markov models, Generative Grammars, and other techniques that can be grouped as Algorithmic Music Composition. Later, the rise of Neural Networks and Deep Learning led to the application of established architectures from fields like Computer Vision and Natural Language Processing to music generation as well [13]. Different authors have provided reviews and taxonomies of AI systems for music generation and composition [14, 13, 12], surveying different architectures and methods, and highlighting current challenges and use cases. In addition, recent studies have then explored the specific challenges and strategies of co-creating music with AI. Here follows a short review.

Louie et al. [15] identified two main challenges in using AI for music generation: information overload and non-deterministic outputs. They then designed “steering tools” for an interactive musical AI system and evidenced that they can enhance the user’s sense of control, trust, and understanding of the AI system, leading to a greater sense of involvement. Moreover, the authors argue that users often rely on their pre-existing mental models of music composition to tackle problems, suggesting that AI systems and interfaces should be designed to adapt to these preconceptions in order to be more intuitive, require less cognitive effort, and ultimately increase user agency. Furthermore, they assert that the AI role in music creation should adapt to the user’s needs and to the creative context. For instance, during the exploratory phase, where the user is searching for unexpected inspiration, relinquishing control is more acceptable. In contrast, during production, maintaining it over specific details becomes critical. In other

words, context plays a vital role in shaping human-AI collaboration dynamics, influencing how control and agency are perceived by the user.

Huang et al. [16] surveyed an AI song contest to investigate the challenges and strategies of co-creating music with AI systems. Their findings emphasize the importance of context awareness and user control and they recommend designing future AI systems to adapt to existing compositional practices rather than imposing new AI-driven workflows.

Newman et al. [17] explored through interviews with composers how current AI tools influence musical creativity. They proposed a model for developing ethical and productive collaborative AI tools, highlighting the importance of user control and clearly defined roles for the AI. Their research suggests that composers value AI use cases where the user maintains control, agency, and choice throughout the creative cycle. The authors recommend that designers of AI tools for creators should consider the expected role of their tools in specific creation processes and make choices that support this, while also recognizing that composers' needs may change over the creative process. On one hand, this echoes what Louie et al. [15] pointed out about adaptability to the creative context, on the other hand, it calls attention to the user in that *"there is still much to do in relation to understanding the exact needs of creative users"* [17].

In their study, Suh et al. [18] examined how AI systems can act as a "social glue" to support human-human collaboration in music composition. According to their findings, AI can promote the exchange of ideas and group cohesion, which can reduce the tensions that often arise during collaboration. The authors thus recommend that AI systems should be intentionally designed to enhance this social collaboration. However, they also observed a potential shift in roles for the users. Specifically, participants reported feeling like curators or co-producers, focusing on evaluating AI-generated material instead of actively developing their own ideas, which can lead to a weaker sense of creative involvement. This finding is consistent with the results of Civit et al. [19] who noted that *"the composer became more of an arranger of different melodies"*, similar to a producer managing a misbehaving band. Although this shift was viewed as a *"very creative, fruitful process"* by Civit et al. [19], it underscores again the need for future AI systems to adapt to the creative context and the changing needs and intentions of the composers.

Despite the increasing interest in human-AI collaboration for music composition, the current research seems to be limited in scope. Much of the focus is on evaluating existing AI systems, the user strategies for navigating the challenges they pose, and integrating steering tools for better control. There seems to be a lack of engagement with field studies on actual compositional practices to bridge the gap between AI music generation and the social and cultural complexities of music composition. While human composers are influenced by cultural context and personal and social motivations, AI systems currently rely solely on algorithmic and predictive logic. This may limit their effectiveness and ability to capture the nuances of human creativity [20]. This research gap points toward the importance of investigating current compositional practices, composers' motivations, artistic sensibilities, the broader cultural and social context influencing their work, and the specific characteristics of various musical genres. By doing so, we can better define how the role of the end user is changing (and could fruitfully change) in these new practices of collaboration between humans and machines.

3. An Ethnographic Study of Music Composition

The literature reviewed so far highlights the need for a more nuanced understanding of music and music composition in order to design human-AI collaborative tools that take into consideration both individual experiences and the socio-cultural context of music creation. While attempts like Hernandez-Olivan and Beltrán [13] to create a generalized model of music composition are valuable for identifying “basic music principles”, imposing a rigid structure on such diverse and fluid processes might be counterproductive. Music creation is shaped by constantly evolving genres, stylistic conventions, individual choices, improvisation, and the unique social and cultural context within which musicians operate.

We then propose an ethnographic approach that can foreground the situated nature of music creation. This approach will consider both the musicians’ personal motivations (e.g., creative aspirations, and career goals) and the socio-cultural context they operate in. Ethnography allows us to delve both into the social and cultural aspects of music creation and the lived experiences of composers. This also aligns with a broader call to integrate social sciences into AI research [21, 22]. By employing ethnography, we look at users as active agents who shape the context, meanings, and consequences of technologies, and not simply as passive recipients. This approach emphasizes the complexity and context-dependent nature of music composition, laying the groundwork for designing human-AI collaborative tools that are more sensitive to the nuances of human creativity and the situated settings in which music is created. Specifically, our research aims to answer the following provisional questions:

- Can the use of Generative AI tools be considered a true collaboration?
- What is the user’s role when utilizing these tools in music composition?
- How do musicians and AI negotiate creative control and authorship during collaboration?

To answer these questions, we propose an ethnographic approach combining semi-structured interviews and participant observation. We will interview 18 musicians with experience in composing diverse genres. Interviews will explore topics like the role of context in composition, creative control and authorship, individual motivations and aspirations, musical sensibilities, social aspects of composing, experiences with AI tools, and their perception of AI’s role and their own role in music creation. After the interviews, we will conduct participant observation (minimum 60 hours) of composers’ practices.

In sum, this study aims to gain a deeper understanding of the situated nature of music composition, considering the role of personal motivations and socio-cultural context in shaping composers’ needs and choices. The findings will inform the design of more effective human-AI collaborative systems that support musicians’ practices. Moreover, we believe this study can contribute valuable insights into human-AI collaboration practices across different domains, allowing us to better define the role of the “end user” in such practices.

3.1. Preliminary Findings

Analysis of the first interviews we carried out has led to some preliminary findings. For instance, the centrality of creative intention in music composition: these intentions guide both compositional strategies and the search and evaluation of creative outcomes. Composers’

choices, like starting with melody, harmony, or timbre, are driven both by an understanding of the intended use (e.g., a commercial, a soundtrack, a song for a personal music project) and an intention to communicate something, reflecting a process of meaning-making where a “*coherent discourse*” is sought. AI systems should support these creative intentions rather than override them. Also, music creation is often a collaborative process, involving bandmates, clients, or sound engineers. Therefore, when it comes to the use of AI systems in music creation, the end user is typically already involved in a collaborative process with other people. This raises a question: should AI function as a tool that enhances this human-human dynamic, or should it be viewed as an additional collaborator? More findings will be shared during the workshop.

References

- [1] W. Xu, Toward human-centered AI: A perspective from human-computer interaction, *Interactions* 26 (2019) 42–46. doi:10.1145/3328485.
- [2] B. Shneiderman, *Human-Centered AI*, Oxford University Press, Oxford, 2022.
- [3] T. Capel, M. Brereton, What is Human-Centered about Human-Centered AI? A Map of the Research Landscape, in: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, CHI '23*, Association for Computing Machinery, Hamburg, Germany, 2023, p. 23. doi:10.1145/3544548.3580959.
- [4] L. G. Terveen, Overview of human-computer collaboration, *Knowledge-Based Systems* 8 (1995) 67–81. doi:10.1016/0950-7051(95)98369-H.
- [5] Z. Wu, D. Ji, K. Yu, X. Zeng, D. Wu, M. Shidujaman, AI Creativity and the Human-AI Co-creation Model, in: M. Kurosu (Ed.), *Human-Computer Interaction. Theory, Methods and Tools*, volume 12762 of *Lecture Notes in Computer Science*, Springer International Publishing, Cham, 2021, pp. 171–190. doi:10.1007/978-3-030-78462-1_13.
- [6] A. Sarkar, Enough With “Human-AI Collaboration”, in: *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems, CHI EA '23*, Association for Computing Machinery, Hamburg, Germany, 2023, p. 8. doi:10.1145/3544549.3582735.
- [7] B. Latour, *Reassembling the Social: An Introduction to Actor-Network-Theory*, Clarendon Lectures in Management Studies, Oxford University Press, Oxford, 2005.
- [8] M. L. Gray, S. Suri, *Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass*, Houghton Mifflin Harcourt, Boston, 2019.
- [9] A. Rapp, A. Boldi, L. Curti, A. Perrucci, R. Simeoni, Collaborating with a Text-Based Chatbot: An Exploration of Real-World Collaboration Strategies Enacted during Human-Chatbot Interactions, in: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, CHI '23*, Association for Computing Machinery, New York, NY, USA, 2023, pp. 1–17. doi:10.1145/3544548.3580995.
- [10] L. A. Suchman, *Plans and Situated Actions: The Problem of Human-Machine Communication*, volume 103, Cambridge University Press, USA, 1987.
- [11] L. A. Hiller, Jr., L. M. Isaacson, Musical composition with a high-speed digital computer, *J. Audio Eng. Soc* 6 (1958) 154–160.
- [12] S. Ji, X. Yang, J. Luo, A Survey on Deep Learning for Symbolic Music Generation: Representations, Algorithms, Evaluations, and Challenges, *ACM Computing Surveys* 56 (2023)

1–39. doi:10.1145/3597493.

- [13] C. Hernandez-Olivan, J. R. Beltrán, Music Composition with Deep Learning: A Review, in: A. Biswas, E. Wennekes, A. Wiczorkowska, R. H. Laskar (Eds.), *Advances in Speech and Music Technology: Computational Aspects and Applications*, Signals and Communication Technology, Springer International Publishing, Cham, 2023, pp. 25–50. doi:10.1007/978-3-031-18444-4_2.
- [14] D. Herremans, C.-H. Chuan, E. Chew, A Functional Taxonomy of Music Generation Systems, *ACM Computing Surveys* 50 (2017) 69:1–69:30. doi:10.1145/3108242.
- [15] R. Louie, A. Coenen, C. Z. Huang, M. Terry, C. J. Cai, Novice-AI Music Co-Creation via AI-Steering Tools for Deep Generative Models, in: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, Association for Computing Machinery, Honolulu, HI, USA, 2020, pp. 1–13. doi:10.1145/3313831.3376739.
- [16] C.-Z. A. Huang, H. V. Koops, E. Newton-Rex, AI Song Contest: Human-AI Co-creation in Songwriting, in: *Proc. of the 21st Int. Society for Music Information Retrieval Conf.*, International Society for Music Information Retrieval, Montréal, Canada, 2020, pp. 708–716.
- [17] M. Newman, L. Morris, J. H. Lee, Human-AI Music Creation: Understanding the Perceptions and Experiences of Music Creators for Ethical and Productive Collaboration, in: *Proc. of the 24th Int. Society for Music Information Retrieval Conf.*, International Society for Music Information Retrieval, Milan, Italy, 2023, pp. 80–88.
- [18] M. M. Suh, E. Youngblom, M. Terry, C. J. Cai, AI as Social Glue: Uncovering the Roles of Deep Generative AI during Social Music Composition, in: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21, ACM, Yokohama, Japan, 2021, p. 11. doi:10.1145/3411764.3445219.
- [19] M. Civit, J. Civit-Masot, F. Cuadrado, M. J. Escalona, A systematic review of artificial intelligence-based music generation: Scope, applications, and future trends, *Expert Systems with Applications* 209 (2022) 118–190. doi:10.1016/j.eswa.2022.118190.
- [20] O. Bown, Sociocultural and Design Perspectives on AI-Based Music Production: Why Do We Make Music and What Changes if AI Makes It for Us?, in: E. R. Miranda (Ed.), *Handbook of Artificial Intelligence for Music*, Springer International Publishing, Cham, 2021, pp. 1–20. doi:10.1007/978-3-030-72116-9_1.
- [21] V. Marda, S. Narayan, On the importance of ethnographic methods in AI research, *Nature Machine Intelligence* 3 (2021) 187–189. doi:10.1038/s42256-021-00323-0.
- [22] M. Sloane, E. Moss, AI's social sciences deficit, *Nature Machine Intelligence* 1 (2019) 330–331. doi:10.1038/s42256-019-0084-6.