Human-AI Collaboration Insights from Music Composition

ERIC TRON GIANET, LUIGI DI CARO, and AMON RAPP, University of Turin, Italy

As AI becomes increasingly generative, the role of humans in creative processes is being redefined. We focus on music composition as a case study to explore the nature of human-AI collaboration. We advocate for a qualitative and ethnographic study that delves into the practices and needs of music composers. By understanding the situated nature of music creation, we aim to uncover insights that can inform the design of more effective human-AI collaborative tools that support composers' creative practices instead of replacing them.

1 INTRODUCTION

The ethical, social, and cultural implications of ever deeper integration of Artificial Intelligence (AI) systems into everyday life and into human practices are pressing issues, and increasing recognition exists that current approaches to the development of AI can overshadow human values, priorities, and lived experience. This adumbration of the human has led to numerous appeals towards a Human-Centered AI [14, 21], that is: an AI that enhances and empowers human users; transparently reveals its values, biases, limitations and the ethics behind its algorithms and data collection; and encourages ethical, interactive, and contestable use [2].

AI, which is often built around the logic of automation, is redefining and challenging the concept of the "end user". Notably, recent advances in generative AI (GenAI) have given rise to systems not only capable of performing classification tasks but also with the ability to generate artifacts such as text, images, and more, making them active agents with creative capabilities. This new landscape draws attention towards the need to understand how the role of humans in decision-making, creative, and information processes is being redefined, and how the idea of "end user" might evolve towards the figure of "collaborator" or "co-creator".

Within this broad research agenda, a strand has emerged that posits that, by working together, both AI and humans can perform better and increase their capabilities. Capel and Brereton [2] call this perspective human-AI teaming: *"the purpose of human-AI teaming is to make the best use of both human and AI capabilities, rather than the human simply being called upon to do what the AI cannot yet manage in an AI led project"*. Similar stances have been called *human-computer collaboration* [18], *human-AI co-creation* [20] or *human-AI collaboration* [12]. Besides the assumption that through collaboration humans and AI can perform better, this approach to human-AI interaction follows what can be understood as an "agentistic turn" [12], i.e., the attribution of agency to AI, in a system of distributed agency between humans and AI.

This perspective on agency seems to be in line with Latour's conception of technological agency: in fact, Latour does not limit agency exclusively to humans but also extends it to other non-human actors, such as objects, technologies, and animals [8]. Although Latour's theories emphasize the complexity of the relationships between human beings and the world around them, it is also worth noting that this agentistic turn of AI may be criticized as it could hide the huge amount of hidden labor behind AI systems, performed by human beings (such as data labelers) often originating from the Global South [12]. This "ghost work" [5], often low-paid and sourced from poor countries, contributes to the capabilities of AI but goes unrecognized.

Now, if the interaction with AI really takes place as a collaboration, and if the user becomes a collaborator or co-creator, what effects does this paradigm and role change generate?

To answer this question, it may be useful to explore a specific collaboration domain, where humans and GenAI systems may work together to achieve situated objectives. This could unveil how collaboration and co-creation unfold "in practice", considering not only the "user"'s objectives, needs, and routines, but also the social and cultural context in which the collaboration is performed. With this aim, in the following section we will look at AI research in music, and more precisely in music composition. This provides an illustrative example, where collaboration, authorship, creation, and ownership serve as pivotal themes, thus offering valuable insights into the broader dynamics of collaboration between humans and AI. We will first review recent literature on human-AI collaboration in music composition, and then propose a study aimed at investigating the situated practices [16] of composing music with GenAI systems.

2   AI IN MUSIC COMPOSITION

Since the 1980s, interest in computer-based music composition has grown steadily, with the use of various techniques that can be grouped under the field of Algorithmic Composition (e.g. Markov Models, Generative Grammars, Genetic Algorithms, etc.). Later, the use of neural networks for music composition gained momentum, and the advent of Deep Learning led to the adoption of architectures established in fields such as Computer Vision or Natural Language Processing [6].

Even though early research on interactive musical systems [4, 13, 19] already emphasized the importance of user-centered approaches in helping people with learning and exploration, as well as the necessity of technology that supports composers' creative processes, more recently, Huang et al. [7] further explored the challenges and strategies of co-creating music with AI. They raise issues related to context awareness and control of models through steerable interfaces, suggesting that future model design should follow a study of the practices currently in use and commonly adopted by music composers, thus adapting to them rather than demanding the opposite.

Louie et al. [9] study the co-creation of music through steering tools as well, showing that these instruments increase users' sense of control, trust, and understanding of AI, making them feel more involved in the creative process. They also observe that people may already carry mental models about music composition, namely, established patterns for dealing with problems. These should therefore be studied and mapped to understand recurring practices and ways of thinking and to design intuitive interfaces that adapt naturally to how people are accustomed to tackling and thinking about certain tasks, reducing cognitive effort and increasing the user's sense of agency.

Newman et al. [11] investigate how AI tools concretely influence musical creativity, and they propose a new model for the development of productive and ethical collaborative tools for artists. This model has at its core the division of roles and distribution of control. Common traits of what they identify as positive uses of AI perceived by users are human maintenance of control, agency, intention, and choice, through which creators can carry out an iterative process of generation and evaluation [11].

In this sense, "context" plays an important role in how much control and sense of agency is perceived and required in human-AI collaboration: Louie et al. [9] argue that AI's role should adapt to the creative context and user needs. They highlight that, while in exploratory phases one seeks inspiration, even the most unexpected, and is willing to relinquish control, maintaining it over specific details becomes crucial during the realization stage. To achieve a true human-AI synergy, the system should be sensitive to the different contexts, in order to provide a suitable collaborative role.

Somewhat related, Suh et al. [17] examine how AI can support human-human collaborative and creative practices, revealing how AI can function as a social glue. In short, AI can help alleviate latent tensions typical of creative

collaboration by facilitating the exchange of ideas and group cohesion. They conclude by suggesting that AI should be designed more intentionally to facilitate social collaboration. However, while AI can become a support for musical collaboration, Suh et al. [17] also observed a shift in roles: their study participants reported feeling more like curators or co-producers rather than co-composers, focusing mainly on the evaluation of AI-generated proposals rather than on the direct development of ideas, leading to a weaker creative involvement. The same is also observed by Civit et al. [3] when they came to realize how *"the composer became more of an arranger of different melodies, something like a producer from the 70s rock and roll scene trying to order the wild creativity of some misbehaving rock stars"*, a shift that they instead consider *"very creative, fruitful process, one with an endless flow of new ideas"*. This relates to the adaptive role future AI systems should have, responding to the creative context, user needs, and compositional intention.

However, current research on human-AI collaborative practices in music seems to overlook the complexity of music composition, focusing on measuring the impact of current generative systems and on the strategies adopted by composers to deal with the challenges they pose, or on how best to introduce steering tools into current interfaces. Those who try to bridge AI research on musical generation with the peculiarities of music composition do not engage in the actual field study of compositional practice. For example, Bown [1] discusses the role of AI in music production from a sociological and design perspective. He proposes to consider factors such as cultural contexts and human motivations, in order to create more sophisticated systems capable of learning tastes and identities. However, the social nature of music includes factors not taken into account by current musical AI models: humans create music situated in cultural contexts, motivated by social factors; AI, on the other hand, operates without these social dimensions, producing music only based on algorithms and predictive models, and this risks limiting its impact by not fully reflecting human creative processes [1]. By and large, how the composer's personal motivations and sensibilities, as well as musical genres and sociocultural contexts, may influence human-AI collaboration in the music domain has not yet been fully explored. This research gap highlights the need to study human compositional practices, investigating the personal motivations and needs of music composers, as well as the cultural and social contexts in which music composition takes place, also considering the peculiarities entailed by different musical genres.

2.1 Studying music composition as a situated practice

As evidenced by the literature reviewed above, the need emerges for a deeper understanding of how music is composed, and to design more effective human-AI collaborative tools that are attentive to the individual's subjective experience as well as the social and cultural context in which music is created.

Although efforts such as that of Hernandez-Olivan and Beltrán [6] to model a general scheme of music composition have the value of drawing attention to some "basic music principles", a rigid mapping of such diverse and fluid processes may prove ineffective. Music creation is a vast landscape, equally influenced by ever-evolving genres and stylistic conventions as by individual creative choices, improvisation, and unique social and cultural contexts in which musicians operate.

Therefore, to bridge the research gap, we propose an ethnographic approach of inquiry that emphasizes the situated nature of music creation, taking into account both the musicians' personal motivations (e.g. creative aspirations, musical sensibilities and taste, career) and the sociocultural context in which they operate. Ethnography, here, is an appropriate research method because it allows us to delve both into the social and cultural context in which the music composition takes place and the lived experiences of composers (using e.g., semi-structured interviews). This is also in line with a more general call to engage the social sciences in AI research, and to move beyond the dominant quantitative approach to AI, to understand the social and cultural context in which AI is used [10, 15]. Ethnography helps to see users as actors

that shape contexts, circumstances, meanings, and consequences of technologies, and not as recipients of predefined ones. Specifically, we will try to answer questions like:

- How do sociocultural factors and music genres influence the music composition process?
- How do personal motivations and the sociocultural context interact?
- Does the use of Generative AI tools really constitute a collaboration?
- How does the role of the user change when Generative AI tools are used in music composition?
- How are creative control and authorship negotiated in the collaboration between musicians and AI systems?

To answer these research questions, we propose an ethnographic research approach combining semi-structured interviews and participant observation. We will conduct 18 semi-structured interviews with musicians with experience in music composition in various fields (e.g. classical music, jazz, electronic music, advertising jingles). The interviews will explore topics such as the role of context in compositional work, creative control and authorship, intentions and motivations, creative aspirations, musical sensibilities, social aspects of composition, experiences of working with AI, and perceptions of the role of AI in music composition. Subsequently, we will conduct a round of participant observation of compositional practices. This will allow us to directly observe the lived experience of composition, explore the role of context, and understand how the process unfolds through situated actions. We propose a minimum number of 60 hours of observation of compositional practices.

In sum, this study aims to provide a deeper understanding of the situated nature of music composition, highlighting the role of personal motivations and sociocultural contexts, and how they influence composers' needs and choices. The results of the study will inform the design of a more effective human-AI collaborative system that supports musicians' compositional practices. We believe that this study could shed light not only on how to support human music composers through new design efforts but, more in general, on human-AI collaboration, uncovering insights that may prove valuable in other domains. Preliminary findings and insights derived from the initial interviews and observations will be shared during the workshop.

REFERENCES

[1] Oliver Bown. 2021. Sociocultural and Design Perspectives on AI-Based Music Production: Why Do We Make Music and What Changes If AI Makes It for Us? In *Handbook of Artificial Intelligence for Music*, Eduardo Reck Miranda (Ed.). Springer International Publishing, Cham, 1–20. https://doi.org/10.1007/978-3-030-72116-9_1

[2] Tara Capel and Margot Brereton. 2023. 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, 23. https://doi.org/10.1145/3544548.3580959

[3] Miguel Civit, Javier Civit-Masot, Francisco Cuadrado, and Maria J. Escalona. 2022. A Systematic Review of Artificial Intelligence-Based Music Generation: Scope, Applications, and Future Trends. *Expert Systems with Applications* 209 (Dec. 2022), 118–190. https://doi.org/10.1016/j.eswa.2022.118190

[4] Rebecca Fiebrink, Daniel Trueman, Cameron Britt, Michelle Nagai, Konrad Kaczmarek, Michael Early, MR Daniel, Anne Hege, and Perry Cook. 2010. Toward Understanding Human-Computer Interaction In Composing The Instrument. In *Proceedings of the International Computer Music Conference*. International Computer Music Association, New York, 135–142.

[5] Mary L. Gray and Siddharth Suri. 2019. *Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass*. Houghton Mifflin Harcourt, Boston.

[6] Carlos Hernandez-Olivan and José R. Beltrán. 2023. Music Composition with Deep Learning: A Review. In *Advances in Speech and Music Technology: Computational Aspects and Applications*, Anupam Biswas, Emile Wennekes, Alicja Wieczorkowska, and Rabul Hussain Laskar (Eds.). Springer International Publishing, Cham, 25–50. https://doi.org/10.1007/978-3-031-18444-4_2

[7] Cheng-Zhi Anna Huang, Hendrik Vincent Koops, and Ed Newton-Rex. 2020. 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, 708–716.

[8] Bruno Latour. 2005. *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford University Press, Oxford.

[9] Ryan Louie, Andy Coenen, Cheng Zhi Huang, Michael Terry, and Carrie J. Cai. 2020. 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, 1–13. https://doi.org/10.1145/3313831.3376739

[10] Vidushi Marda and Shivangi Narayan. 2021. On the Importance of Ethnographic Methods in AI Research. *Nature Machine Intelligence* 3, 3 (March 2021), 187–189. https://doi.org/10.1038/s42256-021-00323-0

[11] Michele Newman, Lidia Morris, and Jin Ha Lee. 2023. 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, 80–88.

[12] Advait Sarkar. 2023. 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, 8. https://doi.org/10.1145/3544549.3582735

[13] Hugo Scurto and Frédéric Bevilacqua. 2018. Appropriating Music Computing Practices Through Human-AI Collaboration. In *Journées d'Informatique Musicale (JIM 2018)*. Amiens, France.

[14] Ben Shneiderman. 2022. *Human-Centered AI.* Oxford University Press, Oxford.

[15] Mona Sloane and Emanuel Moss. 2019. AI's Social Sciences Deficit. *Nature Machine Intelligence* 1, 8 (Aug. 2019), 330–331. https://doi.org/10.1038/s42256-019-0084-6

[16] Lucy A. Suchman. 1987. *Plans and Situated Actions: The Problem of Human-Machine Communication.* Vol. 103. Cambridge University Press, USA.

[17] Minhyang (Mia) Suh, Emily Youngblom, Michael Terry, and Carrie J Cai. 2021. 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, 11. https://doi.org/10.1145/3411764.3445219

[18] Loren G. Terveen. 1995. Overview of Human-Computer Collaboration. *Knowledge-Based Systems* 8, 2-3 (April 1995), 67–81. https://doi.org/10.1016/0950-7051(95)98369-H

[19] Yongmeng Wu and Nick Bryan-Kinns. 2017. Supporting Non-Musicians? Creative Engagement with Musical Interfaces. In *Proceedings of the 2017 ACM SIGCHI Conference on Creativity and Cognition (C&C '17)*. Association for Computing Machinery, New York, NY, USA, 275–286. https://doi.org/10.1145/3059454.3059457

[20] Zhuohao Wu, Danwen Ji, Kaiwen Yu, Xianxu Zeng, Dingming Wu, and Mohammad Shidujaman. 2021. AI Creativity and the Human-AI Co-creation Model. In *Human-Computer Interaction. Theory, Methods and Tools*, Masaaki Kurosu (Ed.). Lecture Notes in Computer Science, Vol. 12762. Springer International Publishing, Cham, 171–190. https://doi.org/10.1007/978-3-030-78462-1_13

[21] Wei Xu. 2019. Toward Human-Centered AI: A Perspective from Human-Computer Interaction. *Interactions* 26, 4 (June 2019), 42–46. https://doi.org/10.1145/3328485