

Theoretical Reflections on AI-Mediated Jazz Improvisation

Ruixiang Zhu

School of Contemporary Music and Technology, Nanjing University of the Arts, Nanjing, 210013, China

Abstract: The advent of artificial intelligence (AI) technology within music has changed the improvisational performing style and aesthetic underpinning of jazz. Based on the philosophy of technology and the aesthetics of music, this article identifies three levels on which AI intervention has restructured the basic configuration of jazz improvisation: the improvisation subject has expanded from the performer to a human-machine consortium, the creative activity has shifted from the individual's internal expression to distributed generation in the interaction, and the focus of aesthetic attention has shifted from the finished product to the generation process itself. In a sense, then, the question AI poses for jazz improvisation is not whether machines can improvise like humans, but how much definitive meaning the term "improvisation" has in the era of technology. From this perspective, AI is not only a tool or an opponent, but has become an active participant in improvisation. This shift of perspective is important because it encourages the concept of improvisation to expand from the creative behaviour of a single subject to the generation of events during interaction between human and machine, providing a new direction for understanding the relationship between humans and technology in the creation of music.

Keywords: Jazz music; artificial intelligence; improvisation; creative agency; human-computer interaction.

1. Introduction

Since jazz was born, improvisation has been the most distinctive quality. Classical music separates the composition and performance into two separate processes, while jazz improvisation allows conception and expression to coincide. Musicians use their ears and their bodies to make decisions in real time within the limits of harmony and rhythm. "Thinking through action" has long been considered an archetypal example of human creativity. The emergence of AI in the domain of music creation has led to jazz an important test case for examining computational creativity the so-called "creativity".

The relationship between jazz improvisation and AI predates recent advances in deep learning. Generation of jazz melodies using probabilistic models had been explored before the advent of deep learning. Sony's Continuator technology modelled sequences of music using Markov chains and could learn the style of the artist, and respond in real time. This step was not yet on the level of human performance, but it raised the question—which characteristics of improvisation can be formalised, and which cannot? The limits of early systems had not really yet tested the boundaries of improvisation.

Deep learning changed this situation. Transformer-based models are no longer satisfied with reorganising existing data, but can generate cohesive and temporally extended music at a higher level. ImprovNet and BebopNet based on Transformer and Transformer-XL architectures respectively have obtained excellent results in controlled improvisation generation [1]. If the algorithm-generated music is so perceptually close to human improvisation, is there a clear line between "algorithmic production" and "true improvisation"? In controlled experiments, several AI models received listener preference and realism levels close to human performers.

Yet this technical convergence does not mean AI has "solved" improvisation. On the contrary, it exposes a conceptual gap: research on AI and music has focused on

output quality—can AI produce convincing jazz?—neglecting whether the practice itself is fundamentally altered. Improvisation studies, meanwhile, have been slow to engage with AI as a participant rather than a tool. This paper addresses this gap by asking not how well AI performs jazz, but how its presence compels us to rethink what improvisation is and what it may become.

Therefore, the issue for jazz improvisation from AI is not at the technical level but the conceptual level, i.e., if the term "improvisation" still has its referential meaning in the era of AI. The paper studies the technical logic and temporal structure of AI improvisation systems, analyses their impact on the creative process and the attribution of the subject, considers changes in the listening experience and recommends alternate paths of understanding. Specifically, it proceeds in three steps corresponding to the shifts identified in the abstract: Section 2 examines how AI's temporal logic reconfigures the site of improvisational agency (the performer-to-consortium shift); Sections 3-4 trace the reorganization of creative activity from individual expression to distributed, interactive generation (the internal-to-distributed shift); and Sections 5-6 explore the corresponding redirection of aesthetic attention from product to process (the product-to-process shift). The paper explores the question of whether AI improvisation continues to be itself after its defining traits have been algorithmically recreated, as an occasion for aesthetic reflection.

2. The Technical Logic of Algorithmic Improvisation and Its Implicit Temporal Issues

While details may differ, AI improvisation systems share the same underlying architecture: the link between learning and creation. The AI is trained on jazz recordings and then responds in real time to external input. Implicitly, this procedure raises a temporal problem that questions the very idea of improvisation.

Scholars call the temporal framework of traditional jazz improvisation "immediacy"—choices and executions are done at the same instant, with no room for pre-design or post-modification. Each musician's performance simultaneously takes on the twin identities of "being created" and "being presented". This duality also sets improvisation apart from composition and from random play. This is radically different in temporality from the "real-time generation" of AI improvisation systems. From an outside perspective, the AI's "real-time response" might be thought of as an "immediate reaction". Internally, however, the process varies across systems. Some models do little more than sample from probability distributions fixed during training. But interactive systems like Somax2 complicate this picture: they continuously listen to and adapt to the performer's input, updating their internal representations in real time. In both cases, though, the temporal structure differs fundamentally from human immediacy—the AI operates on what we might call "typological time", where each response is an instantiation of learned patterns, rather than the "event time" of a unique, unrepeatable moment. AI's "real time" is "quickly pulling up patterns it has learned" not "choosing without a pattern".

This difference has to do with whether or not we can apply the concept of improvisation to AI [2]. AI is also generating melodies within limits of harmony and rhythm, and facing an open space of options, thereby formally satisfying improvisation. AI can be understood, in formal terms, as a process of making creative choices under constraints. But when we think of improvisation as a particular response to the present auditory, physical, emotional, and micro-interaction with the co-performer, the limits of AI become clear. The "immediacy" it responds to is statistical rather than situational, not the "individual" immediacy determined by the specific performance situation. The former is the "generic" immediacy, the latter the "specific" immediacy. The result is that AI improvisation systems shift improvisational temporality from "event time" to "typological time". The first suggests that the meaning of improvisation is in the non-repetitive moment, whereas the second considers each performance as a reinstantiation of a style type. The architecture of the learning-generation tends to classify improvisation in a natural way. The generated segments are perceived as samples of "jazz styles" rather than as individual musical events with their own intrinsic uniqueness. The audience's deficiency in AI improvisation is typically not a technological problem but a temporal experience anomaly: it sounds like jazz, but without the sense of a "happening" [3]. This temporal reconfiguration is the first and most fundamental shift identified in our analysis: the site of improvisational agency is no longer located solely in the performer, but distributed across the human-AI temporal loop.

3. The Differentiation of Creative Authorship in Improvisation

The temporal disparity between AI improvisation and human improvisation prompts one to ask: Who is actually creating the improvisation? The traditional response is clear: the creator is the performer making choices in the moment. But when AI is a real-time collaborator, the question of authorship becomes unclear.

One might argue that the result of AI belongs to the developers. They wrote the code, selected the data and set the

parameters. But the difficulty is: developers are unable to forecast the real output of the model in a certain performance, and it is difficult to explain their selections afterward. The way large-scale neural networks work is not transparent. Researchers term it "implicit knowledge"—the system has learned a set of operational rules that cannot be reduced to explicit rules through statistical learning. Giving credit for "generative ability" not "specific creations" is really attributing authorship to the developers. The developers have designed a system which is able to improvise, however the solo composition played by the system in a specific performance is not something pre-prepared by the developers. This difference is inconsequential in the conventional creative setting, but becomes a talking point in the AI era.

It is equally hard to attribute the authorship to the AI itself. In the aesthetic tradition, "creation" has always been connected with intentionality: to be an author is not just to create the work but also to have an intention. If authorship is understood through the lens of human phenomenological intentionality—as a conscious, goal-directed act of expression—then current AI systems cannot be said to possess such intentionality. Their output is best understood as the result of optimisation and stochastic sampling, rather than an act of expression in the human sense. If this is also creation, then the boundary of creation will be blurred.

A more plausible perspective is that AI's improvisational creativity is a diffused phenomenon, not unique to the developers or the AI system, but occurring in human-machine interaction [4]. In the interplay between the performer and the AI improvisation, the two parties are constantly responding to each other, and the final music is outside the reach of any party's own expectations. If creativity is understood as "an unpredictable result before interaction", then it is no longer a property of a given object, but an attribute of the interaction relation [5]. Systems like Somax2 demonstrate that the cross input of different modalities is responsible for the dynamics of collective behaviour and also boosts the performer's participation. Such a reconfiguration of creative agency need not result in "the death of the author" but rather can suggest: In the modern era of AI, improvisational creativity is best viewed as an event arising from a particular configuration of interaction, not as the capacity of a particular entity. This redefinition of authorship marks the second shift: creative activity has moved from the individual's internal expression toward a mode of distributed generation that unfolds through human-machine interaction.

4. Subject Position in Human-Computer Interaction and the Transformation of Listening Experience

The specific temporality of AI improvisation is also transforming the role of the performer and the way the audience experiences music. In traditional improvisation the performer is the only productive subject and the audience a passive recipient, establishing a one-way relationship through sound. The coming of AI starts to undo this neat subject-object arrangement.

Improvisation in collaboration with AI leads to a special dual state, from the performer's experience. The general musical direction and aesthetic judgement, when to enter, when to recede, when to change density, these macro decisions are now mostly determined by humans [6]. At the

micro level, however, the performer must accept the "unexpected" information from AI and incorporate these surprises into their improvisation technique. The performer must constantly revise his expectations, re-evaluate the direction of each phrase, and select from the numerous motives given in the materials to weave them into the organic fabric of his own musical story. This condition is similar to human ensemble, but with a critical difference. Human musicians have a reciprocal influence—I can sense your aim and you can sense mine. In the case of AI, the relationship is one-sided; the performer can actively affect the AI, but the AI's response has no knowledge of that intention, it merely produces an auditory event, to which the human must then respond.

In addition, this change in the role of the performer leads to a change in the perception of the audience, an aesthetic attitude that is rarely found in traditional jazz performances. When the audience learns that some of the sound on stage is AI-generated, the original listening framework is no longer enough—what attitude should be used to face these AI-generated music [7]? Do we enjoy it as a "work" or do we perceive it as a "process"? When faced with AI improvisation, audiences typically AI improvisation—they cannot resist using the standards of human improvisation as a yardstick, yet they also vaguely feel that this scale is not quite right. The oscillation itself is a distinct aesthetic experience: the audience not only enjoys the music, but is continually asking "what exactly am I enjoying?" AI improvisation produces not only audio artefacts, but also develops ontological enquiries about the acoustic objects simultaneously.

5. Stylistic Imitation versus Creation: The Aesthetic Position of AI Improvisation

So is AI improvisation just imitating styles to an extreme degree of precision, or is it creating something new? This positioning directly influences the criterion for evaluation.

Imitation is the core of AI improvisation. It learns style grammar from a huge number of jazz recordings and outputs a distribution which statistically resembles the training data—exactly this has led many theorists to conclude that it is mere imitation [8]. But the relationship between imitation and creation in the history of music is not so straightforward. In Baroque fugues, you have imitation counterpoint in which a theme is repeated in several voices. Imitation is a structural tool. In the jazz tradition, the imitation of predecessors is also regarded as the heart of learning improvisation, and musicians typically imitate before creating. This learning method was laid out by Paul Berliner in *Thinking in Jazz*. Could AI follow a similar path to move toward personal style through imitation as humans can?

If we define creativity as "producing unexpected novel outcomes within an existing framework", then AI has done it in some circumstances. ImprovNet is a generative model built on Transformers that gives distinctive outcomes for controlled improvisation creation. Tests of ImprovNet showed that 79% of the participants could properly identify a classical piece that had been changed into a jazz style through improvisations. This unpredictability within the framework is also where the value of human improvisation comes in—listeners know the musician will perform inside the F Blues harmonic framework, but they don't know the exact notes. From this point of view AI improvisation and human

improvisation have in fact comparable structures in form, both yield specific musical outcomes which cannot be foreseen in advance under given constraints.

Nevertheless, this structural closeness does not negate the basic aesthetic distinction. The creativity of human improvisation is not only about making surprises within the frame but also about being able to operate on the frame itself, such as reinterpreting harmony, moving rhythm, regulating timbre, and employing these ways to change the bounds of the frame. This is a higher level of creativity—not content to make choices within the rules, but to change the rules themselves. This kind of higher-order creativity—the capacity to transform the rules rather than merely operate within them—has long been recognised as a distinctive feature of human creativity [9]. By this standard, however, today's AI improvisation systems appear to fall short: they operate within the learned grammar rather than transforming it. They can produce grammatically correct jazz sentences, but they do not appear to actively modify the grammar itself—at least not in ways that current systems reliably demonstrate. It is more accurate to call it "creative imitation": it exhibits some creativity within its imitative framework, but it is still limited by the capabilities of the training data, and it has not yet reached the transformation of this space. But this aesthetic itself poses a more fundamental dilemma—what theoretical perspective should we use to approach AI improvisation? This reframing of the question itself signals the third shift: the focus of aesthetic attention is no longer on the finished musical product, but on the generative process through which human and machine jointly produce meaning.

6. From Ontological Anxiety to Relational Aesthetics: A Theoretical Shift

Why does AI make jazz improvisation perplexing? Because we are utilising the framework that explains human behaviour to comprehend an activity that involves non-human players. The original meaning of the term improvisation refers to human musical actions and presupposes that the creator possesses intentionality, bodily sense and emotional experience. When AI advances to "real-time music generation", this assumption becomes unstable. It is not the technical technique of improvisation that is affected but the validity of the concept itself.

Academics tend to have two kinds of replies. One tends to be conservative, setting the threshold of "true improvisation" very high and insisting that it must involve physical presence, emotional engagement, and creative aim. This approach preserves the essential core of traditional aesthetics but at the expense of defining it too narrowly, eliminating not only AI but perhaps even highly rationalised forms of human improvisation. The other goes to the other extreme, including any sounds made in real time. Nor is this tenable: a concept without boundaries lacks analytical purchase.

However, both these ways have their own limitations—the former narrows the definition too much, possibly removing certain human improvisational forms as well; the latter makes the notion lose its analytical power, and is also difficult to retain [10]. This article takes a different tack. So instead of jumping to conclusions, it's best to see AI as a springboard to revisit the notion of "improvisation". The stress moves from "who improvises" to "what happens in the improvisation"—from the question of the subject to the description of

interaction. The issue is no longer whether AI has the same subjective conditions as humans, but how musical meaning is produced and communicated in human-machine interaction. It may be that this formation of meaning does not require the same planned structure for AI as it does for humans, but merely the ambiguity caused by the interaction itself. Neither humans nor AI musicians can decide about the final conclusion of the relationship. This "unpredictability of interaction" may be the core of improvisational aesthetics in the AI era [11].

The theoretical significance of AI improvisation is not whether we can make a "yes" or "no" judgement, but in compelling us to have to consider a more fundamental question: What are we recognising when we call a piece of music improvisation? The creator's inner state? The timing of creation? The uniqueness of the sound? The audience's feeling of the scene as "ongoing"? AI has neither eliminated the aesthetic value of jazz improvisation, nor has it merely copied it, but has turned a practice formerly taken for granted into a theoretical problem that calls for continuous investigation. The stance of theoretical thinking is questioning itself.

7. Conclusion

The participation of AI in jazz improvisation has put the notion of "improvisation" in a position that needs to be reconsidered. Returning to the three shifts outlined in the abstract—the expansion of the improvisational subject, the turn toward distributed generation, and the reorientation of aesthetic attention—the analysis above has traced each through the technical, creative, and perceptual dimensions of AI-mediated improvisation. When artificial intelligence systems can functionally create music in real time, the original meaning of "improvisation", is placed under pressure. Based on the technical logic and time structure, this study analyses the attribution of creation, the subject's position, and the transformation of listening experience in order, and draws various interconnected judgements. In terms of temporality, AI improvisation redefines improvisation as a type of classified behaviour rather than an event that occurs in a certain instant. Each generation becomes less a response to the immediate situation and more an invocation of stylistic convention". The following aesthetic differences are based on the shift in temporal structure. When attributing invention, the product of AI improvisation cannot be attributed to the developer or the machine. It is expressed in the interplay between humans and machines, and the dispersion of creative agency also redefines the boundaries of "creation". The subject's position makes the performer both the discussion partner and the interpreter of the AI, while the listener oscillates between aesthetic observation and conceptual thought. Aesthetic characterisation: a more adequate name for AI improvisation, on the evidence available, might be "creative imitation", it can develop surprising details within the existing framework, but it has not, in the systems reviewed here, demonstrated the kind of frame-changing creativity that characterises the highest levels of human improvisation. The most critical issue at this moment is not to seek a final

conclusion but to continue the viewpoint change that has been started in the previous text, namely, to move the conversation from "whether artificial intelligence can improvise" to "how improvisation itself is re-understood".

Technology is evolving rapidly and much of the analysis in this paper may be subject to revision in the light of new advances. But this core viewpoint will not be invalidated by technical iterations: AI is an opportunity to re-understand improvisation. The value of AI in jazz improvisation is not whether it can demonstrate the creativity of machines but rather to force us to ask a fundamental question again. The preciousness of improvisation is that it is the way humans make decisions under time pressure, bear the unknown, and coexist with others. It may be more meaningful to see this than to form a conclusion prematurely.

References

- [1] Bhandari, K., Chang, S., Lu, T., Enus, F. R., Bradshaw, L. B., Herremans, D., & Colton, S. (2025, June). ImprovNet: Generating controllable musical improvisations with iterative corruption refinement. In 2025 International Joint Conference on Neural Networks (IJCNN) (pp. 1–10). IEEE.
- [2] Nika, J., Chemillier, M., & Assayag, G. (2017). Improtek: Introducing scenarios into human-computer music improvisation. *Computers in Entertainment (CIE)*, 14(2), 1–27. <https://doi.org/10.1145/3056501>
- [3] Yang, L. C., & Lerch, A. (2020). On the evaluation of generative models in music. *Neural Computing and Applications*, 32(9), 4773–4784. <https://doi.org/10.1007/s00521-019-04049-2>
- [4] Lewis, G. E. (2000). Too many notes: Computers, complexity and culture in voyager. *Leonardo Music Journal*, 10, 33–39. <https://doi.org/10.1162/096112100570585>
- [5] Bown, O. (2021). *Beyond the creative species: Making machines that make art and music*. The MIT Press.
- [6] Benford, S., Amerotti, M., Sturm, B., & Martinez Avila, J. (2024, September). Negotiating autonomy and trust when performing with an AI musician. In *Proceedings of the Second International Symposium on Trustworthy Autonomous Systems* (pp. 1–10).
- [7] Shank, D. B., Stefanik, C., Stuhlsatz, C., Kacirek, K., & Belfi, A. M. (2023). AI composer bias: Listeners like music less when they think it was composed by an AI. *Journal of Experimental Psychology: Applied*, 29(3), 676–688. <https://doi.org/10.1037/xap0000447>
- [8] Berkowitz, A. E. (2024). Artificial intelligence and musicking: A philosophical inquiry. *Music Perception: An Interdisciplinary Journal*, 41(5), 393–412. <https://doi.org/10.1525/mp.2024.41.5.393>
- [9] Colton, S., & Wiggins, G. A. (2012, January). Computational creativity: The final frontier?. In *Ecai* (Vol. 12, pp. 21–26). <https://doi.org/10.3233/978-1-61499-084-0-21>
- [10] Miller, B. (2020). "All of the rules of jazz": Stylistic models and algorithmic creativity in human-computer improvisation. *Music Theory Online*, 26(3).
- [11] Frisk, H. (2020). Aesthetics, interaction and machine improvisation. *Organised Sound*, 25(1), 33–40. <https://doi.org/10.1017/S1355771820000061>