

Original Paper

The Reconstruction of Artistic Subjectivity and Its Ethical Issues in AI-Based Music Composition

Chen Ni^{1*}

¹ School of Humanities, Tongji University, Shanghai, China

* Corresponding author

Received: September 6, 2026 Accepted: September 29, 2026 Online Published: October 21, 2026

doi:10.22158/asir.v10n3p46

URL: <http://dx.doi.org/10.22158/asir.v10n3p46>

Abstract

Artificial intelligence has altered the material and conceptual conditions of music-making. Composition is no longer imagined solely as a human practice grounded in intention and lived experience; it increasingly unfolds through interactions among users, datasets, models, and platforms. This study considers how AI-based composition reworks artistic subjectivity and with what ethical consequences. It first outlines the technical foundations of AI music generation and then examines shifts in creative agency, aesthetic judgment, and the relation between human creators and algorithms. The argument is that AI does not simply replace human creativity. It helps produce a distributed human–AI subject of creation. As a result, artistic production moves from experience-based creation toward a hybrid process in which human experience, aesthetic judgment, and algorithmic generation all play a role. The same movement raises ethical questions about authorship, responsibility, creativity, data dependence, and aesthetic homogenization. AI-based music composition is thus better understood as a reconstruction of artistic subjectivity and human–technology relations than as a self-contained technical innovation in music production.

Keywords

artificial intelligence, AI-based music composition, artistic subjectivity, human-AI collaboration, algorithmic generation, ethics

1. Introduction

AI has recast both the technological conditions and the organizational forms of artistic production. Deep learning, large datasets, generative models, and multimodal systems now allow AI to enter music-making directly. Early rule-based and probabilistic experiments have been superseded by Transformer-based and diffusion-based models. AI systems can now produce melody, harmony, rhythm, arrangement, and even

relatively complete pieces from text prompts, musical materials, or user preferences. These capacities broaden the technical scope of composition while unsettling received assumptions about what it means to compose. Composition need no longer begin with a notated idea, a performed gesture, or a single authorial intention; it may begin with a prompt, a corpus, a parameter space, or a process of iterative selection.

The rise of AI composition poses a question that exceeds technical capability: who or what may be counted as the artistic subject of an AI-generated work? In conventional accounts, the human artist remains central. The artist supplies intention, converts experience into form, exercises taste, and bears responsibility for the work. Musical creation is thus understood as an activity rooted in experience, intentionality, aesthetic judgment, and cultural context. AI complicates this picture. Models learn from large musical corpora, detect statistical patterns, generate materials, and optimize outputs according to objectives or feedback. They do not possess human experience or consciousness, yet they perform functions once assigned to certain stages of composition. They therefore occupy an unstable position: neither authors in the traditional sense nor mere instruments.

This transformation is not adequately described as the substitution of machines for human creators. In practice, AI composition unfolds within an assemblage of users, artists, algorithms, datasets, computational infrastructures, and platforms. The process typically involves a user specifying an objective, a system generating candidate outputs, the user selecting, modifying, or rejecting those outputs, and subsequent iterations being shaped by feedback. Production thus acquires a recursive and distributed character; it can no longer be reduced to a linear sequence under the sovereign control of a single maker. The salient theoretical question is therefore not whether an AI system can qualify as a “genuine artist,” but how artistic subjectivity is reconstituted once algorithmic systems operate as active participants in production. Agency is dispersed across human and non-human actors, while intention is mediated by interfaces, default parameters, and optimisation objectives. What is required, accordingly, is a relational rather than author-centred account of creation.

AI composition also unsettles the conventional nexus between experience and creativity. Traditional creation is embedded in lived experience, affect, cultural memory, and individual judgment; the singularity of a work is commonly explained by reference to the creator’s biography and perspective. Algorithmic generation, by contrast, proceeds through computation over extensive datasets and through the extraction, representation, and recombination of patterns. Its generative logic is not experiential in the phenomenological sense. This distinction, however, should not be overstated. Training corpora are themselves sedimented from human musical practice and cultural history; algorithmic outputs consequently bear traces of accumulated experience even as they are produced through statistical operations. The relation between experience and computation is therefore better understood as one of mediation and entanglement rather than pure opposition. Massive data are themselves sedimented from human musical practice and cultural history. Algorithmic generation can therefore be understood, at least partly, as a technologically mediated reproduction and recombination of human musical experience. The

outcome is neither a clean displacement of human experience by machine production nor a simple opposition between the two; it is a hybrid formation in which experience and computation interpenetrate. Experience is drawn into the system through dataset assembly, annotation, curation, and evaluative response; computation then recasts it as latent representations and probabilistic outputs. Creativity is thus dispersed across lived experience and statistical inference, with neither serving as its sole or sovereign source.

Aesthetic questions follow. Conventional evaluation privileges structural unity, formal coherence, emotional expression, intention, and the link between work and author. AI-generated music can satisfy certain formal criteria while also yielding discontinuities, stylistic mixtures, and unexpected structures characteristic of probabilistic generation. These traits prompt debate over whether algorithmic irregularity is a technical flaw or a potential aesthetic value. More fundamentally, different computational methods embody different assumptions about music—as symbolic language, statistical distribution, structured dataset, or acoustic signal. The technological reorganisation of musical production therefore also reconfigures the conception of the musical object. Evaluation may need to address not only the finished work but also the generative process, user–system interaction, and practices of curation and selection.

Ethically, the reconstruction of artistic subjectivity raises further problems. If a work emerges from human intention, algorithmic generation, training data, and platform infrastructure, traditional authorship becomes difficult to maintain in any straightforward form. Responsibility likewise becomes complicated when the final output cannot be traced wholly to one human creator. AI composition relies heavily on training data derived from existing works, raising issues of data use, copyright, creative appropriation, and the reproduction of established aesthetic patterns. Large-scale generative systems may also foster aesthetic homogenisation when models repeatedly reproduce dominant patterns in their training data. The ethics of AI composition therefore cannot be limited to the final output; it must address the entire socio-technical arrangement through which music is produced, circulated, and evaluated. Questions of consent, compensation, transparency, and cultural heritage therefore become central. The transformation of artistic subjectivity is inseparable from new ethical relations among creators, users, algorithms, datasets, and platforms.

Existing studies of AI-generated music have concentrated on technical matters such as representation, architecture, generation methods, datasets, and evaluation metrics. These studies are indispensable, but technical description alone cannot explain changes in the structure of artistic creation. Philosophical and ethical discussions of AI art have addressed creativity, authorship, agency, and machine participation, yet they often need closer engagement with the concrete computational mechanisms of AI music. An interdisciplinary mode of inquiry is therefore indispensable—one capable of linking the technical architecture of AI composition to the reconfiguration of artistic subjectivity and its ethical ramifications. It is against this backdrop that the present study examines the reconstruction of artistic subjectivity in AI-based music composition. The discussion proceeds first through the technical realisation of AI

composition, encompassing music representation, algorithmic generation, and contemporary generative models. It then turns to the transformation of artistic subjectivity as manifested in artistic realisation and aesthetic evaluation, before addressing the displacement of the individual artistic subject by a human–AI collaborative subject. The final section considers the ethical stakes of this displacement, with particular attention to authorship, responsibility, creativity, data dependence, and aesthetic homogenisation. The argument advanced here is that AI-based music composition ought not to be construed as a simple antithesis between human creation and machine generation. It is more adequately understood as a reconfiguration of the relations among human experience, algorithmic computation, aesthetic judgment, and technological systems—a reconfiguration that issues in a more distributed structure of artistic subjectivity.

2. Technical Implementation of AI-Based Music Composition

Music AI is an interdisciplinary formation situated at the intersection of artificial intelligence and musicology. Its scope includes music understanding, algorithmic composition, lyric–music alignment, vocal synthesis, orchestration, and interactive performance. Algorithmic composition constitutes one of its principal branches. Methods developed to date include Markov chains, musical grammars, artificial neural networks, genetic algorithms, and expert systems.

As AI has matured, its application to music and art has become increasingly sophisticated. AI composition now occupies a central position within Music AI systems. By modelling musical structure and style computationally, such systems acquire a degree of generative capacity. This development reorients the technical trajectories of composition and simultaneously reopens questions concerning artistic agency and musical creativity. Technically, AI composition has shifted from early rule-based and probabilistic procedures toward data-driven paradigms grounded in deep learning and large-scale multimodal models. Multiple technical paths now coexist.

AI composition, however, cannot be reduced to a technical problem. It also carries artistic and ethical stakes. The artistic status of algorithmically produced music, together with the criteria by which it is evaluated, remains contested. The collection, training, and deployment of large musical datasets have likewise generated disputes over copyright, privacy, and the equitable distribution of benefits.

2.1 Music Representation

Earlier algorithmic composition generally relied on explicitly defined musical rules to establish a formal framework, which was then refined through optimisation and iteration. More recent approaches assemble and preprocess large musical datasets, segment them into computationally tractable representations, train models, and subsequently generate musical outputs.

Music representation is a foundational condition that determines a model's functional boundaries. Different representational schemes affect how deeply a model can capture musical structure, and they also embody assumptions about the ontology of music and the logic of computational generation. Three approaches are particularly important: symbolic representation, audio-level representation, and hybrid or

multimodal representation.

(1) Symbolic Representation

Symbolic representation remains among the most mature and widely used approaches in AI composition. Its central premise is to treat music as a sequence of discrete symbols, thereby recasting generation as a sequence-modelling problem analogous in certain respects to natural language processing. Typical forms include MIDI, note sequences, and event-based representations. MIDI is a standard event-based representation describing pitch, duration, velocity, and onset time. It is structurally clear, relatively low-dimensional, and compatible with sequence models such as RNNs and Transformers. However, MIDI cannot adequately represent timbre, performance nuance, or expressive detail; its account of the perceptual dimensions of actual audio remains limited. Note-sequence representations encode music linearly according to temporal order. They are simple and well suited to sequence modelling, but they struggle with complex polyphony, and their temporal granularity is limited. Event-based representations refine the units of modelling by decomposing pitch, onset, offset, velocity, and other properties into finer events. This permits more precise modelling of temporal structure and polyphonic relations, and it is especially suitable for Transformer architectures. Overall, symbolic representations are explicit, relatively interpretable, and computationally manageable. They are well suited to learning musical grammar and formal structure. Their major limitation is that timbre, spatial characteristics, and other perceptual dimensions are heavily compressed.

(2) Audio-Level Representation

Audio-level representation operates directly on raw sound signals and emphasises the perceptual and physical properties of music. It mainly includes waveform and spectrogram representations. Waveform representation uses raw audio as input and preserves acoustic information most completely. Its high dimensionality, however, entails substantial computational cost, and high-level musical structure is difficult to model directly. Spectral representations generally use the short-time Fourier transform or Mel spectrograms to map audio into a two-dimensional time–frequency representation, with time and frequency as the two principal dimensions. Such representations are somewhat more compatible with aspects of human auditory perception and reduce the difficulty of computational modelling. Still, they suffer from implicit structural information and insufficient semantic representation. In model implementation, audio generation commonly employs generative adversarial networks, variational autoencoders, and diffusion models. GANs improve generation quality through adversarial training but are prone to instability. VAEs enable style interpolation through latent-space representation but may preserve fine details poorly. Diffusion models generate audio through iterative denoising and have become an important route to high-quality audio generation because of their strong distribution-modelling capacity and generation quality.

(3) Hybrid and Multimodal Representation

To negotiate the divide between symbolic and audio-level representations, recent scholarship has gravitated toward hybrid and multimodal designs. The purpose is to bring structural articulation and

acoustic expressiveness into a single modelling framework, rather than treating them as separate domains. Among the strategies adopted are MIDI-to-audio rendering, audio-to-symbolic transcription, and joint-training arrangements that bind the two modalities together. Multimodal representation extends this programme by admitting text, images, performance gestures, and related sources, thereby widening the semantic horizon available to generation. These advances, however, confront formidable obstacles: cross-modal alignment remains unstable, training costs are high, and semantic consistency across modalities is frequently incomplete. Such difficulties suggest that the integration of symbolic and acoustic domains is not merely an engineering task but a conceptual problem concerning how musical meaning is distributed across media.

2.2 Model Implementation

The evolution of model implementation in AI composition may be periodised into four broad phases. The first, extending roughly from the 1950s to the 1990s, was dominated by rule-based and symbolic systems, with expert systems as the most representative form. Composition was governed by explicitly encoded musical knowledge, while computational generation remained comparatively limited in scope and flexibility.

The second phase, spanning the 1990s to the 2010s, witnessed the ascendancy of statistical learning and probabilistic modelling. Evolutionary algorithms and neural networks became prominent, enabling systems to infer patterns from data rather than rely exclusively on hand-crafted rules. These approaches improved adaptability, yet continued to struggle with long-range structural coherence and stylistic control. The third phase, roughly 2010 to 2020, was marked by the widespread adoption of deep-learning models. Recurrent neural networks, variational autoencoders, genetic algorithms, and related architectures were applied to composition. Each displayed distinct strengths and limitations with respect to controllability, long-term structural modelling, and style transfer.

The current phase is characterised by large-scale models and multimodal generation. Music generation is no longer confined to symbolic or acoustic representation; it increasingly incorporates text, video, and other modalities. This paradigm remains in development and continues to reshape both the technical and conceptual horizons of AI composition.

Current mainstream approaches primarily involve deep learning, diffusion models, and multimodal modeling. These approaches generally offer stronger long-range structural modeling, higher-quality outputs, and relatively greater controllability, though they still show limitations in emotional modeling, evaluation methodology, and multimodal integration. At the model level, recurrent neural networks capture temporal dependencies through sequence modeling and can learn compositional patterns without explicitly encoded rules. They nevertheless face significant limits in modeling long-range structure and are relatively inefficient to train. Transformer models use self-attention to capture global dependencies and have performed strongly in long-sequence music generation. Their disadvantages include high computational complexity, substantial dependence on large datasets, and comparatively limited controllability. Generative adversarial networks use adversarial training between generators and

discriminators to improve diversity and realism. They are applicable to multi-track music modeling, but training is often unstable and susceptible to mode collapse. Variational autoencoders enable continuous generation and style interpolation through latent-space mappings, yet they may struggle to preserve high-fidelity details. Diffusion models generate music through iterative denoising and demonstrate strong distribution-modeling capabilities and high generation quality. Their limitations include relatively slow inference and high computational cost.

Data-driven composition centres on deep-learning architectures that infer musical structure and style from large corpora, thereby reducing reliance on explicitly encoded rules. Such models—spanning recurrent networks, Transformers, and diffusion architectures—learn latent representations of musical regularities and can generate outputs of considerable complexity. Their principal affordances lie in capturing nuanced stylistic features and producing high-dimensional musical material. Yet they remain vulnerable to limited controllability, pronounced dependence on data quality, and biases embedded in training corpora. Because their generative logic is statistical rather than rule-governed, long-range structural coherence is often difficult to secure, and outputs may exhibit mode collapse or stylistic inconsistency.

Rule-based composition relies on music theory and manually encoded constraints. It affords relatively strong controllability and structural consistency, since generation proceeds within an explicitly defined formal framework. Its limitations are equally evident: creative flexibility is constrained, scalability is limited, and the labour of encoding comprehensive rules is considerable. Such systems tend to reproduce established musical grammars rather than venture beyond them.

Genetic algorithms simulate biological evolution through selection, crossover, and mutation to optimise musical segments. They provide exploratory capacity and can discover unforeseen combinations, but they tend to converge slowly, depend heavily on fitness-function design, and may lack stability across generations. Their outputs are often uneven, and the absence of an explicit structural grammar can result in incoherence.

Markov chains model note sequences through state-transition probabilities and can generate music at relatively low computational cost. They are simple, interpretable, and efficient for local dependencies. However, their capacity to represent long-term musical structure is limited; they lack hierarchical memory and struggle to sustain thematic development over extended passages.

Musical grammar approaches formalise musical structure through rule systems, often drawing on generative or hierarchical grammars. They can provide relatively strong structural coherence and syntactic consistency, but comprehensive grammars are complex to construct and difficult to scale. Their rigidity may also impede stylistic flexibility and adaptive generation.

Reinforcement learning formulates composition as a sequential decision-making problem and uses reward functions to guide generation. It can incorporate aesthetic objectives into optimisation and adapt to interactive feedback. Yet its performance depends heavily on reward design; poorly specified rewards may produce policy biases, reward-hacking behaviours, or degenerate outputs. It is also sample-

inefficient and difficult to stabilise. In conclusion, different algorithms reveal a tension among generative capacity, structural control, and emotional expression.

Table 1. Comparison of AI-Based Music Composition Models

Algorithmic Model	Representation		Compositional Tasks	Main Advantages	Main Limitations	Representative Models
Rule-based systems	Music-theoretical rules; symbolic representation		Harmony generation, counterpoint, automatic arrangement	Strong interpretability; structural stability	Limited creativity; low complexity	Expert Systems; rule engines
Genetic algorithms	Symbolic representation		Melody generation, harmony optimization	Structural exploration; incorporation of user feedback	Slow convergence; difficult fitness-function design	Interactive Genetic Composition; EvoComposer
LSTM	MIDI; sequences	note	Melody generation	Stable performance; suitable for melodic generation	Low training efficiency	LSTM Blues Improvisation
RNN	MIDI; sequences	note	Melody generation; monophonic composition	Suitable for sequence-generation tasks	Severe vanishing-gradient problems; limited long-range modeling	Multi-instrument RNN
GAN	Piano Roll; MIDI; audio		Multi-track generation; accompaniment; style generation	High perceptual naturalness; capable of learning musical distributions	Training instability; mode collapse; weak controllability	MidiNet; MuseGAN; Transformer-GAN

2.3 Evaluation of AI-Generated Music

AI composition rests on three conceptual premises. One treats music as language and focuses on symbolic manipulation. A further approach construes music as acoustic waveform, concentrating on raw signal processing. A third construes music as data, generating outputs by estimating the probability of particular data configurations.

Evaluation of AI-generated music differs from that of image or text generation in that it must simultaneously address structural coherence, perceptual quality, and aesthetic judgment. No single metric can adequately capture the quality of AI-generated music. Pitch-distribution measures, including Kullback–Leibler divergence, assess whether generated pitches statistically resemble those in real musical corpora. Such measures can indicate whether a model has acquired basic musical grammar, but they cannot adequately capture global structure or affective expression. Rhythmic-consistency measures gauge the stability and regularity of temporal organisation, including beat alignment and inter-onset intervals; their capacity to capture complex rhythm and dynamic nuance remains limited. Diversity measures quantify variation across generated samples and can help detect or prevent mode collapse. Excessive diversity, however, may produce structural incoherence or stylistic instability. Objective metrics are computationally tractable and comparable across systems. Their fundamental limitation is that they measure statistical resemblance rather than musical quality as such.

Because music possesses a pronounced aesthetic dimension, subjective evaluation remains indispensable. Listening tests are among the most common methods: participants rate or compare generated works, often through Likert-scale ratings or Turing-test-like procedures in which listeners judge whether a piece was produced by a human or a machine. Listening tests directly capture human perceptual responses and thus approximate real-world reception more closely. They are, however, relatively costly, require sufficient participants, and are susceptible to individual preference, which may limit experimental stability and reproducibility. Expert evaluation involves composers, musicologists, or professional musicians assessing generated works along dimensions such as structural coherence, harmonic plausibility, stylistic consistency, and creativity. Compared with general listeners, experts bring greater disciplinary expertise and can identify complex musical structures more deeply. Expert judgment nonetheless remains highly subjective, and substantial disagreement may arise, making inter-rater consistency difficult to guarantee. A robust evaluative framework therefore typically combines objective metrics with human judgment rather than relying exclusively on either.

Taken together, these technological approaches to AI composition embody distinct generative logics. Data-driven methods foreground the learning of statistical distributions; rule-based methods foreground structural control; reinforcement learning attempts to reconcile these dimensions through goal-directed optimisation. Such differences are not confined to technical implementation. They also imply divergent conceptions of music's ontology—as symbolic sequence, probability distribution, or physical signal. The coexistence of these paradigms accounts in part for the complexity of AI composition and furnishes the basis for subsequent analysis of artistic evaluation and ethics. Algorithms also vary markedly in

generative capacity, controllability, and dependence on data, and these technical characteristics may in turn configure corresponding structures of ethical risk.

It is necessary to specify what “generation” denotes in this context. Despite the widespread use of the term “music generation,” the underlying computational operations do not necessarily constitute generation in the conventional ontological sense. Traditional algorithmic models yield musical outputs through probabilistic computation; later systems iteratively optimise candidate solutions relative to predefined objectives; deep-learning models learn statistical patterns from large corpora, after which particular outputs emerge probabilistically from the trained model. Such processes involve variation, stochasticity, and differential output quality. Several mechanisms may accordingly be identified as central to AI composition: probabilistic computation, iterative optimisation toward selected solutions, algorithmic emergence, and acoustic denoising or reconstruction. None corresponds straightforwardly to “creation” in the traditional humanistic sense. “Music generation” is therefore best treated as a convenient technical term for the production of musical outputs through computational processes, rather than as an ontological claim about the nature of the activity.

3. Reconstruction of Artistic Subjectivity in AI-Based Music Composition

AI composition has not only transformed the technical means of music production; it has also altered the structure of subjectivity in artistic production. In traditional accounts, artistic creation is grounded in human subjectivity. Through accumulated experience, emotional expression, and aesthetic judgment, the artist transforms an inner world into an external work. With AI, musical creation gradually shifts from a single human-centered subject structure toward a composite structure in which human agents and intelligent algorithmic systems jointly participate.

The central issue raised by AI composition is therefore not simply whether machines can create art. It concerns how the boundaries of the artistic subject change when algorithms participate in musical production, how creativity should be reconceptualized, and what new subject relations emerge between human beings and technological systems.

3.1 Artistic Realization in AI-Based Music Composition

AI composition is not merely a technical mechanism for producing musical outputs; it is also a form of artistic practice with a distinct aesthetic orientation. An adequate account of the artistic character of AI composition cannot remain confined to algorithmic or engineering description. It must be situated within the broader concerns of philosophical aesthetics. The artistic realisation of AI composition may be analysed along two distinct registers: the experiential and the transcendental.

At the experiential register, the realisation of art depends substantially on sensory experience and the mechanisms of aesthetic pleasure. AI composition models learn from scores, audio corpora, and user behavioural feedback, and they produce works exhibiting formal properties such as rhythmic organisation, melodic contour, and harmonic pattern. Such works do not issue directly from a human composer’s intuitive experience or affective expression; they arise instead from statistical fitting and

structural recombination of pre-existing musical idioms. Even so, they can function as perceptible sonic objects capable of eliciting aesthetic response within auditory experience. At this register, therefore, AI composition possesses a degree of artistic validity.

Relative to traditional composition, AI composition exhibits a pronounced de-subjectivation. Its creative process depends more heavily on data-driven models and parameter optimisation than on individual experience, emotional expression, or aesthetic intuition. This structural shift means that artistic realisation often requires human intervention to complete what may be termed an output-consolidation process: generated materials are selected, edited, and further processed so as to conform more closely to particular aesthetic norms or practical contexts. Contemporary AI composition thus remains embedded within a hybrid structure of human-machine collaboration.

Transcendental aesthetics, particularly in its Kantian formulation, furnishes a significant critique of purely empiricist aesthetics. Kant rejects the reduction of beauty to pleasure or subjective gratification, arguing instead that aesthetic judgment arises from the free and harmonious interplay of sensibility and understanding. Beauty is neither mere sensory stimulus nor purely rational judgment, but a distinctive mode of internal coordination within the subject. From this vantage, the artistic value of an AI-composed work does not hinge solely on whether it reproduces the human process of artistic creation. The pertinent question is whether the work can occasion such structural unity within aesthetic experience.

Considered further from the experiential dimension, AI composition tends toward what may be regarded as optimal musical structures through the aggregation and modelling of heterogeneous data. Its generative logic consists essentially in recombining and optimising extant musical distributions. The decisive question, however, concerns the mechanism by which outputs are selected and consolidated. Generated materials generally require functional filtering or post-processing before they satisfy particular aesthetic criteria. AI composition is therefore not a wholly autonomous system of aesthetic production; it more closely resembles a recursive cycle of production, selection, and reproduction. Nonetheless, AI-composed music can still elicit affective responses and aesthetic experience through its sonic form. Although its computational mechanism operates primarily through simulation and statistical fitting rather than inspiration in the traditional sense, its status as an object of experience sustains the possibility of aesthetic reception. Under certain conditions, imperfections or deviations in algorithmic output may themselves acquire aesthetic significance once reinterpreted within an artistic frame.

AI composition does not negate the transcendental structure of art. It reorganises artistic production at the technological level, shifting creation from direct human authorship toward the participation of technical systems in forming and restructuring production. Artistic subjectivity thereby becomes distributed, hybrid, and systemic rather than exclusively individual. Historically, art and technology were already intertwined in the Greek notion of *techne*; they should not be treated as wholly separate domains. The entry of AI composition into artistic production is thus not necessarily an external intrusion into the essence of art. It may be understood as a historical rearticulation, under new technological conditions, of the original unity between *techne* and artistic practice.

3.2 Aesthetic Evaluation of AI-Based Music Composition

Traditional art theory tends to assess the aesthetic value of a musical work by reference to integrity and completeness: whether structural organisation, expressive content, and semantic orientation achieve internal unity. This framework presupposes an integrated conception of the true, the good, and the beautiful, and it construes the artwork as an entity oriented toward holistic completion and formal perfection. Classical works are accordingly treated as paradigmatic aesthetic objects whose value resides in the coherence of structural logic and the consistency of emotional expression.

Within this evaluative horizon, AI-composed music can still be accommodated when it attains relative stability in rhythm, harmony, and style. Yet because its production depends heavily on data-driven modelling and computational optimisation, it frequently exhibits weaknesses in the logical continuity of emotional expression, the coherence of long-range structure, and the development of thematic material. Such limitations generate uncertainty when AI-generated works are judged by traditional standards of artistic integrity. The issue is not simply technical immaturity; it also reflects a tension between the temporality of statistical generation and the teleological logic presupposed by classical aesthetics.

Contemporary art theory, by contrast, foregrounds incompleteness, contingency, and resistance to established norms. Art is no longer necessarily conceived as a closed structure oriented toward perfection but as an open process marked by continual generation, rupture, and recombination. Artistic value is therefore not reducible to formal unity. At the level of data processing, AI composition decomposes music into discrete symbolic units—pitch, duration, rhythmic pattern—and reconstructs larger structures through probabilistic recombination or mappings within vector spaces. This procedure carries an inherent logic of de-totalisation, so generated outputs may exhibit structural discontinuity. Non-normative passages, abrupt stylistic transitions, and formal ruptures should not be regarded merely as technical defects. Under certain aesthetic conditions, they may function as new elements of artistic experience, disclosing the generative process or reconfiguring habitual modes of listening.

AI composition can therefore produce works acceptable under traditional aesthetic criteria while also generating aesthetic objects characterised by meaningful deviation within contemporary artistic contexts. Such deviation is not simply departure from norms; it may open possibilities for reorganising perceptual experience outside established normative systems, thereby establishing a new pathway of aesthetic experience. The evaluative question thus shifts from whether a work conforms to a given standard to how it reframes the conditions under which aesthetic judgment itself operates.

Different technological approaches to AI composition embody different understandings of the nature of music. Approaches that treat music as a symbolic system emphasise linguistic and grammatical properties. Approaches that treat music as a probability distribution emphasise statistical characteristics and data-driven generative logic. Approaches that treat music as an acoustic signal emphasise physical properties and continuous modes of expression. These philosophical assumptions influence not only technological implementation but also the basic criteria through which musical works are aesthetically evaluated and, consequently, the ways artistic meaning is produced. Aesthetic evaluation of AI-based music composition

must therefore remain reflexive about the musical ontology presupposed by each technical paradigm.

3.3 From the Artistic Subject to a Human–AI Collaborative Subject

Traditional theories of art assign subjectivity a human-centred shape. The maker is presumed to command the origin of ideas, the design of form, and the conferral of meaning; the work is taken as an externalisation of an inner world. Intelligent composition unsettles this arrangement.

Data becomes a new ground of artistic production. Algorithms do not compose out of biographical experience; they derive patterns from existing repertoires. Musical creation thus loses its exclusive anchorage in the individual and becomes entangled with vast data sets and repeated model adjustment. AI composition algorithms act within the creative process rather than standing outside it. Architecture, training, and objective functions all leave traces on musical structure. The algorithm remains external to human consciousness, yet its technical logic is woven into the genesis of works and steers compositional outcomes.

Viewed operationally, algorithms are neither authors in the customary sense nor neutral instruments. Instruments and recording media principally transmit human intention. AI algorithms intervene in the constitution of musical structure through learning, recognition, and generation. They distil regularities from training material and produce melodic, rhythmic, and harmonic arrangements, thereby bending the course of composition. The making of a work therefore escapes the linear formula “intention → execution → object.” It assumes a recursive shape: human goal-setting, algorithmic production of alternatives, human selection and revision, system feedback and recalibration. The creative subject is displaced from the solitary author toward a composite of human and machine.

This does not abolish human agency; it redistributes it. Humans retain aesthetic judgment, value choice, and meaning attribution. Algorithms handle pattern exploration, possibility expansion, and formal generation. The relation is complementary: humans furnish purpose, experience, and cultural context; algorithms furnish computation, data, and an enlarged field of possibility. Creation becomes technologically mediated collaboration. New ethical difficulties follow. The boundaries of creative subjectivity blur, and authorship, copyright, and responsibility resist conventional assignment. Algorithms are not value-neutral. Training corpora, model design, and platform protocols shape outputs and condition artistic practice through infrastructural arrangements. Human–AI collaboration is thus not mere co-working; it reorganises subject relations under algorithmic logic.

AI composition thereby moves artistic subjectivity from an individual to a relational register. Creation is no longer solely the expression of interior consciousness. It becomes an emergent dynamic arising from interactions among humans, algorithms, data, and platform environments. The mode of production changes, and with it the philosophical questions of creativity, subjectivity, and artistic value return in altered form.

3.4 From Experiential Creation to Algorithmic Generation

Traditional creation is rooted in subjective experience. Artists convert life experience, affect, cultural memory, and aesthetic judgment into expressive works. Creation is understood as heavily dependent on

experience; its centre lies in the artist's perception, interpretation, and expression of the world. The work's singularity often derives from the irreplaceable experience and individual expression of its maker. AI composition alters this logic. Algorithmic generation does not proceed from the creator's experience. It relies on large musical corpora, model training, and computational rules to learn existing structures, stylistic patterns, and formal regularities, then to emit new musical outputs. Creation shifts from an experience-driven expression toward a data-driven generation. Algorithms lack human affect. Through computational analysis of relations in data, they model melody, rhythm, and harmony, producing a mode of creative activity unlike traditional art.

Epistemologically, the generation of artistic knowledge changes. Traditional creation depends on the subject's perception and understanding; knowledge arises from accumulated experience and aesthetic judgment. Algorithmic generation builds its representation of musical regularities through statistical analysis and pattern recognition. Algorithms do not necessarily understand music; they establish computational relations among musical elements. Creativity in AI composition no longer appears only as human construction of meaning; it increasingly appears as exploration of possibilities encoded in data relations. Algorithmic generation does not sever art from human experience. The datasets derive from historical human musical practice and carry cultural contexts, aesthetic standards, and value orientations. It can be read as a technologically mediated reworking of human experience: historical experience, compositional practice, and aesthetic structures are encoded into computable form and recombined through models into new outputs.

New ethical questions arise. Algorithmic generation widens creative possibility, allowing creators to exceed the limits of individual experience and explore a broader musical space. Yet dependence on existing data patterns may confine creation within historically sedimented structures, reproducing and reinforcing established styles. It carries both inventive potential and the risk of aesthetic homogenisation and weakened creativity.

AI composition should not be framed as substituting algorithms for experience. It moves creation from purely experiential production toward a hybrid of experience and computation. The question for future art is not only who creates or how a work is generated, but how human creativity, aesthetic judgment, and value choice can be preserved amid algorithmic intervention, so that technological generation becomes an extension of human artistic capacity.

4. Conclusion

AI composition has transformed how music is produced and understood. This study has argued that the change exceeds the arrival of new algorithms or generative technologies. It recasts the relations among creator, work, technical system, and process. AI composition is therefore not merely a technical innovation in music production; it is a reconstruction of artistic subjectivity. On the technical plane, AI music systems have moved from rule-based and probabilistic methods toward deep learning, large-scale data modelling, and generative models. Music may be represented symbolically, statistically, or

acoustically and processed by different computational mechanisms. What is called “music generation” is better described as a compound of probabilistic computation, pattern recognition, iterative optimisation, and signal reconstruction. This technical shift supplies the material basis for changes in artistic practice. At the level of artistic subjectivity, AI does not simply replace the human creator. It enters an increasingly distributed creative process. Humans set goals, exercise aesthetic judgment, select outputs, and assign meaning; AI systems process data, detect patterns, generate alternatives, and enlarge the space of possible musical forms. Artistic production moves from an individual subject-centred model toward human–AI collaboration. The creative subject becomes relational and distributed, encompassing human creators, intelligent systems, datasets, algorithms, and platforms. The relation between experience and creativity also changes. Traditional creation is bound to individual experience, emotion, cultural memory, and aesthetic intuition. Algorithmic generation depends on data, computation, and statistical pattern recognition. The two need not be opposed. As AI systems are trained on corpora sedimented from human musical practice, algorithmic generation does not sever itself from accumulated experience but rearticulates that experience computationally. AI composition is therefore best understood as a hybrid formation in which lived experience, aesthetic judgment, and computational generation mutually condition one another. Experience enters the system through data selection, annotation, curation, and evaluative feedback; computation then recasts it as latent structure and probabilistic output. The relation is not one of substitution but of mediation and entanglement.

The reconstruction of artistic subjectivity is inseparable from an ethical reconfiguration. AI participation unsettles conventional ascriptions of authorship and responsibility, while reliance on training data implicates copyright, consent, compensation, and the reproduction of existing musical materials. The statistical replication of dominant patterns within large corpora may foster aesthetic homogenisation and contract the range of expressive diversity. These are not merely defects of individual outputs; they are structural features of the socio-technical arrangement through which AI music is produced. They cannot be addressed by output evaluation alone. They require sustained examination of the relations among creators, algorithms, data, platforms, and audiences.

The significance of AI composition therefore lies less in whether machines can compose independently than in how artistic subjectivity is reorganised once computational systems become participants in production. AI need not eliminate human creativity; it alters the conditions under which creativity is exercised. Future inquiry should accordingly address the distribution of creative agency, the allocation of responsibility, the governance of training data, and the preservation of human aesthetic judgment within increasingly automated creative systems. These questions are indispensable to any comprehensive account of the artistic and ethical stakes of AI-based music composition.

References

- Alaeddine, M., & Tannoury, A. (2021). Artificial intelligence in music composition. In *IFIP International Conference on Artificial Intelligence Applications and Innovations* (pp. 387-397). Cham: Springer.

- https://doi.org/10.1007/978-3-030-79150-6_31
- Borghuis, V., Angioloni, L., Brusci, L., et al. (2020). Pattern-based music generation with Wasserstein autoencoders and PRC descriptions. In *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence* (pp. 5225-5227). California: IJCAI. <https://doi.org/10.24963/ijcai.2020/751>
- Briot, J. P., & Pachet, F. (2020). Deep learning for music generation: Challenges and directions. *Neural Computing and Applications*, 32(4), 981-993. <https://doi.org/10.1007/s00521-018-3813-6>
- Chen, X. (2019). Ethical system of artificial intelligence: Infrastructure and key issues. *CAAI Transactions on Intelligent Systems*, 14(4), 605-610.
- Chen, Y., Huang, L., & Gou, T. (2024). *Applications and advances of artificial intelligence in music generation: A review*. arXiv preprint arXiv:2409.03715.
- Dash, A., & Agres, K. (2024). AI-based affective music generation systems: A review of methods and challenges. *ACM Computing Surveys*, 56(11), 1-34. <https://doi.org/10.1145/3672554>
- De Prisco, R., Zaccagnino, G., & Zaccagnino, R. (2020). EvoComposer: An evolutionary algorithm for 4-voice music compositions. *Evolutionary Computation*, 28(3), 489-530. https://doi.org/10.1162/evco_a_00265
- Dong, H. W., Hsiao, W. Y., Yang, L. C., et al. (2018). MuseGAN: Multi-track sequential generative adversarial networks for symbolic music generation and accompaniment. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 32(1). <https://doi.org/10.1609/aaai.v32i1.11312>
- Eck, D., & Schmidhuber, J. (2002). Finding temporal structure in music: Blues improvisation with LSTM recurrent networks. In *Proceedings of the 12th IEEE Workshop on Neural Networks for Signal Processing* (pp. 747-756). Piscataway: IEEE. <https://doi.org/10.1109/NNSP.2002.1030094>
- Guan, F., Yu, C., & Yang, S. (2019). A GAN model with self-attention mechanism to generate multi-instruments symbolic music. In *2019 International Joint Conference on Neural Networks (IJCNN)* (pp. 1-6). Piscataway: IEEE. <https://doi.org/10.1109/IJCNN.2019.8852291>
- Guo, H., & Wang, L. (2017). Design of intelligent music composition system based on interactive genetic algorithm. *Microcomputer Applications*, 33(2), 10-13, 34.
- Herremans, D., Chuan, C., & Chew, E. (2017). A functional taxonomy of music generation systems. *ACM Computing Surveys*, 50(5), 1-30. <https://doi.org/10.1145/3108242>
- Huang, C. Z. A. (2019). *Deep learning for music composition: Generation, recommendation and control*. Doctoral dissertation. Cambridge: Harvard University.
- Hung, Y. N., Chiang, I., Chen, Y. A., et al. (2019). *Musical composition style transfer via disentangled timbre representations*. arXiv preprint arXiv:1905.13567. <https://doi.org/10.24963/ijcai.2019/652>
- Jiao, J., Wang, L., & Kang, Q. (2019). Research on emotional music generation methods. *Microcomputer Applications*, 35(2), 52-55.
- Jiang, N., Jin, S., Duan, Z., et al. (2020). RL-Duet: Online music accompaniment generation using deep reinforcement learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(1),

- 710-718. <https://doi.org/10.1609/aaai.v34i01.5413>
- Lopez-Rincon, O., Starostenko, O., & Ayala-San Martín, G. (2018). Algorithmic music composition based on artificial intelligence: A survey. In *2018 International Conference on Electronics, Communications and Computers* (pp. 187-193). Piscataway: IEEE. <https://doi.org/10.1109/CONIELECOMP.2018.8327197>
- Miranda, E. R. (1995). An artificial intelligence approach to sound design. *Computer Music Journal*, 19(2), 59-75. <https://doi.org/10.1109/CONIELECOMP.2018.8327197>
- Mitra, R., & Zualkernan, I. (2025). Music generation using deep learning and generative AI: A systematic review. *IEEE Access*, 13, 18079-18106. <https://doi.org/10.1109/ACCESS.2025.3531798>
- Muhamed, A., Li, L., Shi, X., et al. (2021). Symbolic music generation with Transformer-GANs. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(1), 408-417. <https://doi.org/10.1609/aaai.v35i1.16117>
- Muñoz, E., Cadenas, J. M., Ong, Y. S., et al. (2014). Memetic music composition. *IEEE Transactions on Evolutionary Computation*, 20(1), 1-15. <https://doi.org/10.1109/TEVC.2014.2366871>
- Mycka, J., & Mańdziuk, J. (2025). Artificial intelligence in music: Recent trends and challenges. *Neural Computing and Applications*, 37(2), 801-839. <https://doi.org/10.1007/s00521-024-10555-x>
- Mycka, J., & Mańdziuk, J. (2026). Classification of piano performers with deep learning models. *Journal of Computational Science*, 102804. <https://doi.org/10.1016/j.jocs.2026.102804>
- Rawls, J. (1988). *A theory of justice* (H. He et al., Trans.). Beijing: China Social Sciences Press.
- Roberts, A., Engel, J., Raffel, C., et al. (2018). A hierarchical latent vector model for learning long-term structure in music. In *Proceedings of the International Conference on Machine Learning* (pp. 4364-4373). New York: PMLR.
- Samuel, D., & Pilát, M. (2019). Composing multi-instrumental music with recurrent neural networks. In *2019 International Joint Conference on Neural Networks (IJCNN)* (pp. 1-8). Piscataway: IEEE. <https://doi.org/10.1109/IJCNN.2019.8852430>
- Tan, X., & Li, X. (2021). A tutorial on AI music composition. In *Proceedings of the 29th ACM International Conference on Multimedia* (pp. 5678-5680). New York: ACM Press. <https://doi.org/10.1145/3474085.3478875>
- Wu, J., Liu, X., Hu, X., et al. (2020). PopMNet: Generating structured pop music melodies using neural networks. *Artificial Intelligence*, 286, 103303. <https://doi.org/10.1016/j.artint.2020.103303>
- Yang, Q. (2020). An analysis of ethical principle of AI: Based on the difficult problem of AI. *Philosophical Analysis*, 11(2), 137-150.
- Yu, B., Lu, P., Wang, R., et al. (2022). Museformer: Transformer with fine- and coarse-grained attention for music generation. *Advances in Neural Information Processing Systems*, 35, 1376-1388. <https://doi.org/10.52202/068431-0101>
- Yunoki, I., Berreby, G., D'Andrea, N., et al. (2023). Exploring AI music generation: A review of deep learning algorithms and datasets for undergraduate researchers. In *International Conference on*

Human-Computer Interaction (pp. 102-116). Cham: Springer Nature Switzerland.
https://doi.org/10.1007/978-3-031-49215-0_13

Zheng, Q. (2026). Machine memory intelligence: Inspired by human memory mechanisms. *CAAI Transactions on Intelligent Systems*, 21(1), 1.

Zulić, H. (2019). How AI can change/improve/influence music composition, performance and education: Three case studies. *INSAM Journal of Contemporary Music, Art and Technology*, 2, 100-114.
<https://doi.org/10.51191/issn.2637-1898.2019.2.2.100>