

INNA ANTIPINA

ORCID iD: <https://orcid.org/0000-0003-3845-9629>

*PhD, Lecturer at the Department of Theory and History of Culture,
Head of the Scientific and Information Department
of the Hero of Ukraine Myroslav Skoryk Center for Musical Ukrainian Studies
at the Ukrainian National Tchaikovsky Academy of Music
(Kyiv, Ukraine)
innaantipina92@gmail.com*

ANALYSIS OF LIMITATIONS AND CHALLENGES IN THE USE OF ARTIFICIAL INTELLIGENCE IN THE MUSIC INDUSTRY

This study examines the current state of artificial intelligence (AI) technologies in the field of musical creativity and analyzes the key limitations that define the boundaries of their application in the artistic context. The research explores the functioning of generative neural networks — Transformers, Generative Adversarial Networks (GANs), Recurrent Neural Networks (LSTMs), and autoencoders — in synthesizing musical structures of varying complexity. It is found that even the most advanced algorithms show a significant dependence on training data corpora, leading to repetitive motifs and a loss of stylistic uniqueness. The study emphasizes the limited capacity of AI models to reproduce emotional expression, thematic development, and the depth of artistic intention. Particular attention is given to the interaction between composers or performers and generative systems in real-time. The prospects of developing interactive human-machine interfaces, capable of responding instantly to creative input and forming a shared creative environment, are analyzed. The potential of multimodal models, which can simultaneously analyze sound, text, and visual signals, is highlighted, opening new avenues for interdisciplinary musical creativity. Socio-cultural implications of AI integration into the music industry are explored, including transformations in the notions of authorship, authenticity, and creativity. The risks of cultural homogenization due to the dominance of Western training datasets are examined, emphasizing the importance of developing locally oriented Ukrainian musical data corpora. It is argued that the advancement of nationally focused generative systems can help preserve cultural diversity and expand opportunities for artists' creative self-realization. The study concludes that artificial intelligence should not be seen as an alternative to human creativity, but rather as a tool for its transformation and a means of establishing new forms of interaction between technological and artistic thinking.

Keywords: *artificial intelligence, generative models, music industry, variability, interactivity, creativity, cultural adaptation, emotional expression, authenticity, neural networks.*

Problem statement... Artificial intelligence (AI) is revolutionizing numerous fields, including creative domains such as visual art, literature, and music. In the music industry, AI is employed to automate a wide range of tasks, including composition, music recommendation, transcription, and production. Through advanced information technologies such as deep neural networks, generative adversarial networks (GANs),

and Transformers, AI enables the creation of music that imitates human compositions, thus opening new possibilities for musicians, composers, and audiences. However, despite these achievements, AI-based music generation faces significant limitations that hinder its full creative potential. This article focuses on analyzing three key aspects: variability, interactivity, and creativity.

Variability in music generation implies the AI's ability to produce diverse compositions that are structurally coherent and encompass a broad range of genres and styles. Nevertheless, modern AI models often encounter challenges in this domain. Recent scholarly work, such as "Artificial Intelligence in Music: Recent Trends and Challenges" (Mycka, Mańdziuk, 2025), indicates that AI tends to generate music that is repetitive or excessively dependent on its training data. This limitation can be attributed to the difficulties in modeling long-term dependencies, such as the creation of coherent sections (e.g., verses and choruses), as well as the insufficient diversity of datasets such as MAESTRO (classical piano music) or Lakh MIDI Dataset (pop and rock music). Moreover, legal restrictions related to copyright further complicate access to diverse musical materials suitable for machine learning, thereby narrowing AI's potential to generate truly unique compositions.

Interactivity in the context of AI-based music generation implies the system's ability to respond to user actions in real time, for example, during live performances or collaborative music creation. As noted in "Music and Artificial Intelligence" (2013), there exist tools such as Continuator, Piano Genie, and Magenta Studio, which support interactive composition, enabling AI to react to live performances or improvisations. However, these systems face significant technical challenges, including data-processing delays (latency) and the difficulty of generating musically appropriate responses that align with the user's expectations and intentions. For instance, in the study "Applications and Advances of Artificial Intelligence in Music Generation: A Review", the authors emphasize that diffusion-based models such as Riffusion and Moûsai remain limited in real-time operation due to the lengthy processes of training and generation, which complicate their application in interactive contexts such as gaming environments or live performances (Chen, Huang, Gou, 2024).

Creativity in AI-generated music has become the subject of both philosophical and practical debate. The central question concerns whether AI is capable of producing genuinely original compositions, or whether it merely recombines existing patterns — interpolation rather than extrapolation. The study "A Survey on Artificial Intelligence for Music Generation: Agents, Domains and Perspectives" demonstrates that contemporary AI models do not perceive music in the same way as the human brain and lack the intrinsic creative component, which restricts their ability to produce truly innovative works (Hernandez-Olivan, Hernandez-Olivan, Beltrán, 2022). Furthermore, the inherent subjectivity of musical evaluation complicates the conceptualization and measurement of creativity, as human perception is profoundly shaped by cultural context, aesthetic norms, and individual preferences. This subjectivity also raises complex practical and ethical issues, particularly concerning the attribution of authorship in AI-generated music and its legal recognition within existing frameworks of intellectual property rights.

Analysis of recent studies and publications shows that a growing number of scholars worldwide are exploring the specificities of AI applications in the music industry. In particular, J. Mycka and J. Mańdziuk (2025) propose classifying current research on artificial intelligence in music into three major domains: music classification, music generation, and recommendation systems. In their systematization, the authors review leading conference proceedings and journal articles (covering 2017–2023), analyze different music representations (waveform, MIDI, piano-roll, and notation), and demonstrate how the choice of representation influences the types of musical properties that can be modeled. Their work serves as a valuable technical map of the field, especially for researchers investigating the variability of generated content and its dependence on data representation and model architecture. At the same time, the authors identify several open problems: the dominance of deep neural networks imposes new requirements for quality evaluation, while reliable objective metrics for assessing long and expressive musical forms comparable to human creativity remain underdeveloped. Thus, while the review highlights notable technological progress, it also underscores the need for interdisciplinary approaches

that combine technical, performative, and ethical perspectives.

In their article by C. Hernandez-Olivan, J. Hernandez-Olivan and J. R. Beltran (2022), provide a comprehensive overview of contemporary approaches to music generation using artificial intelligence, identifying the key components involved in this process: datasets, models, interfaces, users, and the resulting musical output. This framework facilitates a systematic understanding of how these elements interact to influence the final outcome—specifically in terms of variability, quality, and expressive potential. The authors further distinguish between two principal domains of AI-based music generation: symbolic representations (e.g., MIDI or musical notation) and audio-based representations (e.g., waveforms). They emphasize that the choice of representation domain critically determines both the range of variations the system can produce and the degree of control it affords over musical parameters such as dynamics, timbre, and structural organization.

Hernandez-Oliván et al. (2022) further outline a range of current challenges, foremost among them the creation of models capable of maintaining long-term musical structure, combining styles and genres, and adapting to the cultural and aesthetic expectations of listeners. A major obstacle remains the evaluation of creativity in AI-generated music — objective metrics often fail to capture emotional or contextual dimensions of perception, while subjective assessments can be inconsistent or difficult to standardize. The authors recommend future research focused on improving user interfaces, integrating core musical principles (harmony, rhythm, structure) into model design, and developing more universal and genre-diverse datasets that reflect a broader range of musical styles and cultural contexts. However, as this publication dates back to 2022, it does not account for the latest breakthroughs of 2023–2025, particularly in diffusion models, foundation models, and multimodal generation.

In the article by Y. Chen, L. Huang and T. Gou (2024), a comprehensive review of AI-driven music generation technologies is presented, with a particular focus on three model categories: symbolic generation, audio generation, and hybrid systems that combine both approaches. This classification enables a detailed comparison of the capabilities and limitations inherent to each model type: symbolic models offer greater

controllability over musical structure and notational precision, whereas audio models achieve more accurate representation of timbre, dynamics, and expressive nuances. Hybrid approaches, while capable of integrating the advantages of both methods, typically introduce higher computational complexity and increased requirements for data processing and resource allocation. This technological perspective is crucial for analyzing the variability of AI-generated music, as it determines the types of musical variations possible — from minimal stylistic changes to profound innovations in musical structure and sound design. Chen et al. (2024) also emphasize the practical applications of generative models in scenarios requiring real-time interaction or integration with other disciplines, such as multimodal interfaces and systems that combine visual and auditory modalities. However, the authors note that the evaluation of such systems remains a challenge due to the absence of unified quality metrics, particularly when it comes to assessing emotional expressiveness and creativity. As they demonstrate, emotional impact depends not only on the model's architecture or training methodology but also on the dataset, the format of musical representation, and the context of use. This opens promising directions for further research on interactive approaches that give users greater control over the output and on studies that combine objective and subjective evaluation criteria for a more comprehensive validation of musical creativity.

The aim of this article is to provide a critical analysis of the limitations and emerging challenges associated with the application of artificial intelligence (AI) in the music industry, with a particular emphasis on the aspects of variability, interactivity, and creativity in the process of music generation.

To achieve this aim, the following **objectives** were defined:

- to explore the ability of AI algorithms to generate diverse musical forms, styles, and structures, as well as to determine the boundaries of their originality in comparison with human creativity.
- to analyze the potential for interaction between composers or performers and generative models in real time, as well as the impact of interactive technologies on the collaborative creative process between humans and machines.

- to evaluate AI's potential as an instrument of artistic innovation, to identify the criteria of "creativity" in the musical context, and to trace how artificial intelligence can expand the boundaries of traditional musical thinking.

- to investigate the obstacles associated with the development of AI systems capable of dynamically interacting with users or other systems during the music creation process.

Presentation of the main research material... Artificial intelligence (AI) opens new horizons for the music industry, enabling the automation of music creation, processing, and dissemination. Despite significant technological achievements, current AI systems still face numerous technical constraints that affect their capacity to generate musical material meeting aesthetic, emotional, and cultural expectations. It can now be stated with confidence that AI is revolutionizing the music industry by enabling the algorithmic creation of music. However, such aspects as variability (the ability to generate diverse music), interactivity (the ability to respond to user actions in real time), and creativity (the ability to produce original compositions) remain complex and only partially resolved.

One of the primary technical limitations of AI lies in its inability to fully reproduce the emotional depth intrinsic to human musical creativity. AI algorithms such as Transformers — as used in OpenAI Jukebox or Suno AI — rely on the statistical analysis of large datasets, allowing the generation of melodies, harmonies, and rhythms that conform to established stylistic or genre conventions. However, these models lack subjective experience and emotional context — the very elements that human composers infuse into their works through personal expression. For instance, while systems like AIVA are capable of producing music in the style of classical composers such as Bach or Beethoven, they cannot replicate the intuitive understanding of emotional nuance that arises from human experience. Consequently, music generated by AI, although technically accurate, often appears emotionally superficial, limiting its resonance with listeners who seek genuine affective depth and artistic sincerity.

Another significant limitation is AI's tendency toward repetitive patterning.

Algorithms such as Transformers and GANs are trained on large datasets containing recurring musical structures characteristic of specific genres. This often results in the generation of music that sounds predictable or overly similar to existing compositions. For instance, L. Wang, Z. Zhao, H. Liu, J. Pang, Y. Qin and Q. Wu (2022) note that modern AI systems such as Jukebox frequently produce works that replicate genre clichés, such as standard chord progressions in pop music. This limitation arises from the fact that AI models rely on statistical probability rather than creative innovation, which makes it difficult for them to produce truly original compositions that transcend the boundaries of their training data.

The processing of large volumes of data represents another major technical challenge. Contemporary AI systems based on latent diffusion models (for example, Stable Audio 2.0 by Stability AI) require enormous computational resources to analyze and generate audio data. Large datasets comprising thousands of hours of music are essential for training such models, yet they often contain heterogeneous formats (MIDI files, audio recordings, samples), which complicates processing and model optimization. For example, Transformer-based models such as Music Transformer perform efficiently with symbolic note sequences but struggle with raw audio data due to its high computational complexity. Moreover, insufficient data quality or diversity may lead to bias in generated outputs, such as the predominance of Western musical styles over local traditions.

Cultural adaptation of AI remains a significant challenge. Most modern AI models are trained on data that predominantly reflect Western musical culture (e.g., classical, pop, or electronic music). Consequently, they struggle to generate compositions consistent with local or underrepresented cultural traditions, such as Ukrainian folk or ethnic music. O. Sadovenko (2024) highlights the risk of cultural erosion caused by stylistic homogenization driven by AI. For example, models like Suno AI can generate jazz melodies reminiscent of Miles Davis, but they have limited capacity to reproduce the distinctive features of regional musical traditions due to the lack of corresponding data in their training sets.

Different AI algorithms employed in the music industry each possess unique

strengths and limitations that affect their overall performance and creative potential:

- Transformers, such as Music Transformer or musicful.ai, are highly effective in generating structured musical sequences (e.g., the AABA form in pop music) owing to their attention mechanism, which enables the modeling of long-term dependencies in musical patterns. However, they demand substantial computational resources and may produce predictable structures due to over-reliance on training data. Their limited capacity for real-time improvisation also makes them less flexible compared to professional human composers.

- GANs, used in systems such as AIVA, can produce highly realistic instrumental parts that are difficult to distinguish from human performances. Nevertheless, they often suffer from overfitting, where the generator reproduces a narrow set of patterns, thereby reducing originality. In addition, GANs are technically complex to train and require careful balancing between the generator and discriminator components.

- Autoencoders, employed in systems such as WavTool or LANDR, are effective tools for editing audio data (e.g., modifying tempo or instrumentation) and generating variations of existing tracks. However, their ability to generate entirely new compositions is limited because their primary objective is data reconstruction rather than creative generation.

- Recurrent neural networks (RNNs) — including LSTM and GRU architectures used in projects like Google Magenta — are well-suited for real-time improvisation thanks to their sequential processing capability. However, they are less efficient than Transformers at modeling long-term dependencies and have lower computational efficiency, making them less suitable for generating complex, large-scale compositions. For example, an LSTM can generate short jazz-style fragments but cannot maintain the polyphonic complexity found in Bach's works.

The technical limitations of AI are clearly evident in practical applications. For example, the composition “Hello World!” by Melomics, created by the Iamus computer cluster, demonstrates AI's capacity to generate new harmonic structures and emulate aesthetic elements of music, yet it is often perceived as technically impressive

but emotionally detached. Similarly, “Daddy’s Car” by Flow Machines imitates the style of The Beatles, reproducing the characteristic 1960s AABA pop structure; however, its predictability reflects the algorithm’s limitation in producing genuinely original ideas. These examples illustrate that AI can accurately replicate the structural characteristics of a genre, generating compositions that align with listener expectations. In contrast, the Holly Herndon Proto project, which employs the Spawn model, pushes creative boundaries by generating nonlinear, improvisational structures in real time, blending human vocals with machine processing. Such experiments expand the frontiers of traditional compositional forms but are often perceived as less predictable and more challenging to integrate into mainstream musical culture.

Rhythmic aspects are also influenced by artificial intelligence. For instance, Choir Transformer, employing a transformer-based architecture, generates complex polyphonic rhythms reminiscent of Bach’s style through the mechanism of relative positional attention. This enables the creation of music with high harmonic and rhythmic precision; however, the complexity of such structures may render them difficult for general audiences to perceive. The Skygge project (“Hello World!”) demonstrates AI’s capacity for stylistic hybridization, blending elements of pop, rock, and electronic music. This expands genre boundaries, yet may simultaneously lead to the loss of distinct stylistic identity.

One of the key cultural challenges lies in the risk of musical style homogenization. As O. Sadovenko (2024) notes, AI systems trained predominantly on Western musical data may contribute to stylistic unification, thereby threatening the cultural erosion of local traditions, such as Ukrainian folk music. For example, Suno AI effectively generates jazz melodies in the style of Miles Davis, yet demonstrates limited ability to reproduce the distinctive characteristics of ethnic genres due to training data constraints. This poses a challenge for the preservation of cultural diversity in music.

The issue of authenticity is central to discussions surrounding AI-generated music. Traditionally, authenticity has been associated with human experience and emotion; however, as I. Antipina (2024) observes, AI can produce new interpretations

of musical genres, prompting a re-evaluation of this concept. For instance, the composition "Hello World!" by Melomics, though technically impressive, is often perceived as less authentic due to the absence of human intentionality. Listeners may appreciate the technical mastery of such works while simultaneously sensing a lack of emotional depth, which diminishes their cultural significance.

The audience's perception of AI-generated music also presents a challenge. Studies indicate that listener responses are ambivalent: some admire its technical sophistication, while others perceive a loss of human emotionality. For instance, music produced by AI may evoke a weaker emotional response if listeners are aware of its machine origin, raising questions regarding the psychological impact of such compositions.

Thus, the aesthetic and cultural challenges of AI in the music industry involve balancing between the imitation of traditional forms and the creation of innovative structures, the risk of stylistic homogenization, difficulties in preserving authenticity, and the audience's ambiguous perception. Overcoming these challenges requires the development of AI capable of accounting for cultural contexts and emotional dimensions, as well as fostering an ongoing dialogue between technology and human creativity.

AI is increasingly utilized in the music industry for tasks such as automatic music transcription, composition recommendation, music creation, and production assistance. Technologies such as deep neural networks, generative adversarial networks (GANs), and transformers enable the creation of music that imitates human compositions. However, as noted in "A Survey on artificial intelligence for Music Generation: Agents, Domains, and Perspectives", these systems possess significant limitations that impede their full-scale implementation (Hernandez-Olivan et al., 2022).

Variability in music generation presupposes AI's ability to create diverse compositions that are neither repetitive nor overly dependent on training data. Research shows that contemporary models often struggle with modeling long-term dependencies in music, resulting in compositions with insufficient structural coherence. For instance, models may produce monotonous music due to the limitations of datasets such as

MAESTRO (classical piano music) or the Lakh MIDI Dataset (pop and rock), which fail to encompass all genres. Furthermore, copyright restrictions limit access to diverse musical data, complicating model training and hindering broader artistic applicability.

Interactivity in AI systems implies the ability to respond to user actions in real time — for instance, to improvise or adapt a composition based on input data. However, contemporary systems often possess complex or insufficiently personalized interfaces, which complicates their use. Research indicates a need to develop interfaces that account for the user's level of musical expertise, as well as the challenges of integrating AI into collaborative creative processes.

AI creativity remains a subject of both philosophical and practical debate. Many researchers argue that current AI models merely interpolate data rather than create genuinely novel content. For example, "A Survey on Artificial Intelligence for Music Generation: Agents, Domains and Perspectives" notes that AI does not perceive music as the human brain does and lacks a genuine creative component (Hernandez-Olivan et al., 2022). Evaluating AI-generated music also presents difficulties due to the subjectivity of perception and the absence of universal metrics. Subjective assessments, such as listening tests, depend on cultural context and the listener's background, whereas objective metrics — such as Pitch Class or Polyphonic Ratio — exhibit weak correlation with human perception.

Although the primary focus lies on variability, interactivity, and creativity, ethical issues such as copyright ownership of AI-generated music are equally significant. Studies highlight unresolved legal questions concerning authorship: whether it belongs to database creators, model developers, or end users. This issue may be considered an additional factor influencing technological development.

Conclusions.

The integration of artificial intelligence (AI) into the music industry introduces unprecedented creative possibilities while simultaneously posing substantial aesthetic and cultural challenges. These challenges encompass transformations in musical structure, harmony, rhythm, stylistic diversity, authenticity, and perceptual experience. Drawing on the analyzed material, this section explores the ways in which AI

transforms the aesthetic dimensions of music, heightens the risk of stylistic homogenization, complicates the preservation of cultural authenticity, and alters listener perception. Particular attention is given to illustrative examples such as “Hello World!” by Melomics and “Daddy’s Car” by Flow Machines.

1. Contemporary AI algorithms demonstrate notable progress in generating diverse musical forms, styles, and structures; however, the limits of their originality remain constrained. Despite advanced architectures such as transformers, GANs, and recurrent neural networks, their output frequently relies on statistical regularities in training datasets, leading to repetitive motifs and patterning. To enhance variability, it is essential to expand representative data corpora — particularly those reflecting national musical traditions — and improve encoding methods capable of modeling long-term structural dependencies within compositions.

2. Interactivity remains one of the key challenges in the development of generative systems. Current technologies provide only limited user–model interaction in real time, thereby restricting the potential for genuine human–machine co-creativity. Future research should focus on designing adaptive interfaces capable of responding dynamically to performers’ musical actions and on creating environments in which AI functions not merely as a passive tool but as an active creative partner in composition or improvisation.

3. Artificial intelligence today can serve as a catalyst for artistic innovation, expanding the boundaries of traditional musical thinking. Despite its lack of true intentionality, generative models can produce new combinations of sounds, structures, and timbres that introduce unconventional compositional approaches. In this context, the criteria of creativity require redefinition: instead of emulating human experience, emphasis should shift toward integrating AI as a co-author capable of generating new artistic possibilities at the intersection of algorithmic and human imagination.

4. Research on obstacles associated with developing AI systems that dynamically interact with users has revealed the need to improve technical and cognitive mechanisms of interaction. Existing limitations — such as processing latency, insufficient adaptability, and restricted responsiveness — reduce the

effectiveness of AI in creative processes. A promising direction involves creating multimodal systems capable of analyzing context, performer intent, and emotional parameters of music in real time. Such solutions could transform human–algorithm interaction into a genuine dialogue, enhancing the aesthetic depth and dynamism of musical creation.

Technical limitations of AI in the music industry include its inability to fully reproduce emotional depth, the repetitiveness of generated patterns, challenges in processing large data volumes, and weak adaptation to cultural contexts. Algorithms such as transformers, GANs, autoencoders, and LSTM networks possess unique strengths but also face significant drawbacks — including high computational complexity, overfitting, and limited improvisational ability — that hinder the full integration of AI into musical creativity. Overcoming these barriers requires algorithmic improvement, expansion of training datasets, and the development of models that integrate cultural specificity and emotional dimensions into musical production.

Prospects for further research... Future research in the field of artificial intelligence (AI) for music generation is directed toward addressing existing limitations and enhancing the sophistication of creative synthesis. A primary objective lies in developing more efficient music encoding methods that can reduce model training time while enabling more precise modeling of long-term dependencies within musical structures. Another promising direction involves designing generative systems capable of producing coherent motifs and integrated thematic phrases that emulate the structural and expressive logic of human composition.

Particular attention should also be devoted to genre specificity—namely, the adaptation of models to a wide range of musical styles and traditions, including national and folk genres, which would contribute to greater variability and cultural richness in the results. Equally significant role is to be played by the advancement of human–computer interaction (HCI) interfaces, which should evolve toward being more intuitive, adaptive, and conducive to creative collaboration between users and neural systems.

Finally, an essential continuing objective concerns the refinement of evaluation methodologies for AI-generated music, particularly the development of universal metrics capable of more accurately capturing human perception, aesthetic judgment, and the emotional resonance of musical works.

References

1. Antipina, I., 2024. Musical heritage of Ukraine during digital technology epoch: trends and innovations [Muzychna spadshchyna ukrainy v epokhu tsyfrovyykh tekhnolohii: tendentsii ta innovatsii]. *Chasopys Natsionalnoi muzychnoi akademii Ukrainy imeni P. I. Chaikovskoho*, 4(65), pp.7–24. [https://doi.org/10.31318/2414-052X.4\(65\).2024.324587](https://doi.org/10.31318/2414-052X.4(65).2024.324587)
2. Chen, Y., Huang, L. and Gou, T., 2024. Applications and advances of artificial intelligence in music generation: a review. *arXiv*. <https://doi.org/10.48550/arXiv.2409.03715>
3. Hernandez-Olivan, C., Hernandez-Olivan, J. and Beltran, J., 2022. A survey on artificial intelligence for music generation: agents, domains and perspectives. *arXiv*. <https://doi.org/10.48550/arXiv.2210.13944>
4. Music and artificial intelligence, 2013. *Wikipedia*, [online]. Available at: <https://en.wikipedia.org/wiki/Music_and_artificial_intelligence> [accessed: 03 May 2025].
5. Mycka, J. and Mańdziuk, J., 2025. Artificial intelligence in music: recent trends and challenges. *Neural Computing & Applications*, 37, pp.801–839. <https://doi.org/10.1007/s00521-024-10555-x>
6. Sadovenko, O., 2024. The musical generation of AI: the cultural implications of the intersection of transformative technologies with human creativity [Muzychna heneratsiia ShI: kulturni naslidky peretynu transformatsiinykh tekhnolohii z liudskoio kreatyvniosti]. *Visnyk Mariupolskoho derzhavnoho universytetu. Seriya: filosofii, kulturolohiia, sotsiologiia*, 28, pp.135–142. <https://doi.org/10.34079/2226-2830-2024-14-28-135-142>
7. Wang, L., Zhao, Z., Liu, H., Pang, J., Qin, Y. and Wu, Q., 2022. A review of intelligent music generation systems. *arXiv*. <https://doi.org/10.48550/arXiv.2211.09124>

ІННА АНТИПІНА

ORCID iD: <https://orcid.org/0000-0003-3845-9629>

докторка філософії, викладачка кафедри теорії та історії культури,
начальниця науково-інформаційного відділу

Центру музичної україністики імені Героя України М. Скорика
Національної музичної академії України імені П. І. Чайковського

(Київ, Україна)

innaantipina92@gmail.com

АНАЛІЗ ОБМЕЖЕНЬ ТА ВИКЛИКІВ ВИКОРИСТАННЯ

ШТУЧНОГО ІНТЕЛЕКТУ В МУЗИЧНІЙ ІНДУСТРІЇ

Розглянуто сучасний стан розвитку технологій штучного інтелекту (ШІ) у сфері музичної творчості та проаналізовано ключові обмеження, що визначають межі їхнього застосування в художньому контексті. Досліджено особливості роботи генеративних нейронних мереж — трансформерів, генеративно-змагальних мереж (GANs), рекурентних нейронних мереж (LSTM) і автокодувальників — у процесі синтезу музичних структур різного рівня складності. Визначено, що навіть найновіші алгоритми демонструють значну залежність від навчальних корпусів даних, унаслідок чого створювані твори мають тенденцію до повторюваності мотивів і втрати стилістичної унікальності. Підкреслено

проблему обмеженої здатності моделей відтворювати емоційну експресію, динаміку розвитку теми та глибину художнього задуму. Окрему увагу приділено питанню взаємодії композитора чи виконавця з генеративними системами у режимі реального часу. Досліджено перспективи розвитку інтерактивних інтерфейсів «людина–машина», які дозволяють миттєво реагувати на творчі дії користувача та формувати спільний креативний простір. Визначено потенціал мультимодальних моделей, здатних одночасно аналізувати звук, текст і візуальні сигнали, що відкриває нові шляхи для міждисциплінарної музичної творчості. Досліджено соціокультурні наслідки інтеграції ШІ у музичну індустрію, зокрема зміну уявлень про авторство, автентичність і креативність. Досліджено ризики культурної уніфікації, спричиненої домінуванням західних навчальних баз, та наголошено на необхідності створення локально орієнтованих українських корпусів музичних даних. Обґрунтовано, що розвиток національно спрямованих генеративних систем сприятиме збереженню культурної різноманітності та розширенню можливостей творчої самореалізації митців. Зроблено висновок, що штучний інтелект є не тільки альтернативою людській творчості, скільки інструментом її трансформації та нової взаємодії між технологічним і художнім мисленням.

Ключові слова: штучний інтелект, генеративні моделі, музична індустрія, варіативність, інтерактивність, креативність, культурна адаптація, емоційна експресія, автентичність, нейронні мережі.

Стаття надійшла до редакції 04.05.2025 р.
Отримано після доопрацювання 11.05.2025 р.
Прийнято до друку 15.06.2025