Inha Oleksiivna Khmelevska
Volodymyr Olehovych Slobodianiuk

Artificial intelligence in music production: prospects for preparing future masters of musical art for creative activity in a modern professional environment

UDC 378.78.02.
DOI https://doi.org/10.24195/artstudies.2025-3.13

Inha Oleksiivna Khmelevska
Candidate of Pedagogical Sciences,
Senior Lecturer at the Department of
Musical Art and Choreography
The State Institution "South Ukrainian
National Pedagogical
University named after K. D. Ushynsky"
ORCID: 0000-0001-9129-2173

Volodymyr Olehovych Slobodianiuk
PhD Student,
The State Institution "South Ukrainian
National Pedagogical
University named after K. D. Ushynsky"
ORCID: 0009-0008-4186-2113

Стаття надійшла в редакцію: 02.06.2025
Стаття прийнята: 29.06.2025
Опубліковано: 14.10.2025

The article is devoted to the problems of integrating artificial intelligence (AI) technologies into the digital music production cycle and its significance in the context of professional training of future Masters of Musical Art. The aim of the study is to trace the evolution of AI models from early connectionist and rule-based systems to modern transformer-based architectures, to characterise the factors influencing their applicability in professional workflows for generating musical material, and to determine their advantages and limitations in terms of preparing future Masters of Musical Art for successful creative activity in the modern professional environment.
Research methods included retrospective analysis of key stages in the development of AI models, comparative analysis of recurrent neural networks, variational autoencoders, generative adversarial networks, and transformer-based systems, as well as a synthesis of literature using interdisciplinary sources in the fields of computer music, cognitive science, and music pedagogy.
The results showed that AI in music production has evolved in line with a sequential process of developing new generations of models: from early symbolic systems to deep generative architectures and multimodal text-to-music conversion technologies. With each stage of model improvement, music producers gained new tools for creativity, but significant limitations and shortcomings that remain to this day necessitate the training of professionals to critically evaluate and adapt generated material integrated into digital audio workstations (DAW). The main prospects for preparing future Masters of Musical Art for creative activity in a modern professional environment are the need to acquire technological literacy, adaptability in working with generated material, and the ability to interact with AI in the format of joint creative and critical awareness of technological and artistic limitations. The prospect for further research is to determine the essence and functions of specific skills of future Masters of Musical Art to successfully engage in creative activity in a modern professional environment, implementing a cycle of music production using AI technologies.
Keywords: *artificial intelligence, music production, digital audio workstations (DAW), deep learning models (RNN, VAE, GAN, Transformer), training of future Masters of Musical Art, technological literacy.*

Introduction. With the development of technology, the environment of professional and creative musical activity has undergone transformations, among which the emergence of the phenomenon of "music production" is quite significant. Music production is a process that, relying on technology, covers the entire cycle of creating a musical composition – from the generation of an artistic idea to its sound embodiment, i.e., it includes composition, arrangement, recording, editing, mixing, and mastering (processing and adapting the finished audio track to different playback conditions) (Collins, 2014).

In the context of professional training for future Masters of Musical Art, knowledge of music production technologies is becoming increasingly relevant, given the popularity of musical compositions generated in this area of contemporary musical creativity. Accordingly, a pressing task in the field of music education is to prepare future professionals for effective creative work in a technologically mediated professional environment (Ovcharenko et al., 2021), by ensuring their awareness of music production technologies.

A distinctive feature of modern music production is the transfer of the cycle from analogue studios to digital platforms, where almost all stages of the process can be performed "in the box" (ITB), i.e. entirely on a computer and its software environment, without the use of external (analogue) equipment for signal processing (Collins, 2014).

The ability to perform the entire music production cycle "in the box", as opposed to the "out of the box" format (where the audio signal is output from the computer to external mixers, processed, and then fed back in), is closely associated with the emergence of digital audio workstations (DAW). DAWs are software environments for music production that allow multitrack recording, MIDI sequencing, audio editing, mixing and mastering in a single interface.

The main factor that once ensured the multifunctionality of DAWs was the integration of plugins (such as VST, AU, AAX), which replaced external compressors, equalisers, reverbs, etc. in the cycle of creating and processing musical material. Today, the tools and plugins of the most popular DAW environments (such as Ableton Live, Logic Pro, FL

Studio, Pro Tools, Cubase, etc.) differ slightly from each other (for example, in terms of operating system compatibility), but they all perform the general function of a DAW: they provide the producer with a modular environment with virtual instruments and effect chains for creating and processing sound for the entire music production cycle.

Thus, a distinctive feature of DAW architecture is the ability to integrate third-party tools and instruments. Today, in addition to plugins (VST, AU, AAX), artificial intelligence (AI) technologies are increasingly integrated into DAWs (Collins, 2014; Giuliani et al., 2023; Sterne & Razlogova, 2021).

Literature review. A review of the literature showed that research on AI technology in music production focuses primarily on the development of new models or proposals for improving existing ones in the context of expanding their generative capabilities. Early work in this area was devoted to the study of connectionist creativity using AI technology. P. Todd's work (Todd, 1989) had a significant influence in this context. Todd set out to investigate whether connectionist (inspired by biological neural networks) models could mimic aspects of human musical thinking. To this end, the scientist used neural networks to generate short melodic fragments and then analysed whether the technology actually exhibited behaviour similar to musical creativity. However, Todd's research was aimed at both studying cognitive processes and inventing technologies for creating music.

Further research focused on inventing technology for generating musical material using neural network models capable of independently identifying stylistic characteristics and patterns and generating musical sequences on this basis (Mozer, 1994; Eck & Schmidhuber, 2002). With the advent of deep learning technology and, accordingly, more multilayered artificial neural network (ANN) models, models for musical composition based on latent space models and adversarial algorithms have been developed. Among these, the most suitable for use in professional music production are those capable of generating results in MIDI format, as this allows for further editing in a DAW to improve quality, eliminate errors, and enable creative adaptation (Roberts et al., 2018).

A significant shift occurred with the advent of transformer technology, which allowed ANN models to use a "self-attention" algorithm to identify and systematise a variety of musical characteristics and generate complex and lengthy musical sequences based on this "knowledge" (learned relationships) (Huang et al., 2018).

Music education research emphasises the important role of training future professionals in the field of musical art who are able to respond flexibly to the needs of the modern labour market. To this end, it is proposed to introduce interdisciplinary models of liberal arts into the educational process, such as courses in musical composition using technology (Rastrygina, 2020). The need to develop the technological competence of future professionals in the field of music education is also emphasised, which should be based on appropriate awareness of innovative technologies and the skills and abilities to apply them in music teaching (Ovcharenko et al., 2021).

Overall, the analysis of the literature showed that a relevant direction in music education is the preparation of future professionals for the creative use of technologies to ensure consistency between university education and employment prospects. Along with this, the music production industry is popular and rapidly developing in the modern professional and creative music environment. At the same time, the process of introducing AI technologies plays a key role in this industry. This process is evolutionary, and each of its significant milestones corresponds to the emergence of new classes of ANN models. Today, there is a noticeable evolution from connectionist models to those based on deep learning technology and transformer-based architecture (Vaswani et al., 2017; Tiwari, 2025). It has also been clarified that models capable of generating results in symbolic format (e.g., MIDI) retain higher adaptability for DAW workflows, while audio generators offer a higher level of complexity and conditional realism of musical material, but do not provide the ability to edit music after generation, which reduces the usability of such models in professional "in the box" music production (Civit et al., 2022; Collins, 2014).

Purpose of Article. Purpose of the article is to explore the evolution of artificial intelligence technologies in music production using digital audio workstations (DAW), examining the methodological and technological shifts that have shaped modern AI-based musical instruments, and assessing their advantages and limitations in the context of determining the prospects for preparing future Masters of Musical Art for successful creative activity in the modern professional environment.

Research Methods. To achieve the research objective, a number of methods were applied that made it possible to study the factors influencing AI technology on work processes in the field of music production. In particular, the *retrospective analysis* method made it possible to trace the historical development of AI models from rule-based systems to state-of-the-art deep learning models with transformer-based architectures. *Comparative analysis* was used to evaluate the advantages and limitations of different model families in terms of their potential to contribute to tasks in the music production cycle. In turn, the *structural-functional* analysis method made it possible to clarify how different models process music data and how this affects their practical application in DAWs.

To determine the prospects for preparing future Masters of Musical Art for creative activity in the modern professional environment, associated with the growing popularity of such a type of creativity as music production, as well as the rapid development of AI technologies and their integration into this type of activity, and the technological limitations associated with this process, *a synthesis* of the results of the analysis of scientific publications in such areas of research as music pedagogy, computer music and algorithmic composition, artificial intelligence and machine learning, neurocognitive science and music psychology, as well as other interdisciplinary reviews at the intersection of musicology, AI research and audio engineering.

Results & Discussion. The evolutionary process of AI technology development in the context of its application in music production unfolded from the stage of its functioning as *isolated prototypes*, i.e., programmes that generated music by executing a limited set of rules created by humans. Developers manually coded the compositional logic – for example, chord sequences, rhythmic patterns, etc. – which the system applied step by step. One of the foundations of early algorithmic (rule-based) music production was *probabilistic grammar* method (Quick, 2016).

Grammar in the context of musical composition is a set of rules for combining elements of musical language, in particular sounds – into intonation, phrases, chords – into harmonic structures, durations – into rhythmic patterns, etc. In language, grammar determines which sequences of words are “acceptable”. In music, grammar can determine which intervals, rhythms, or harmonic structures can follow one another. Probabilistic grammar, in turn, is an algorithm that adds a set of probabilities for the technology to execute each rule. For example, instead of determining the mandatory use of the dominant after the tonic (e.g., literally “after the C chord, only G is allowed”), the algorithm contains instructions such as “after the chord C with a probability of 0.6, G follows, with a probability of 0.3 – F, and with a probability of 0.1 – Am”; this provides “controlled randomness”, which allows AI to generate several different pieces while adhering to a single stylistic trend (Quick, 2016).

For example, in an early music project, the composition Illiac Suite for String Quartet (Hiller & Isaacson, 1979) was created entirely by AI using stochastic (probabilistic) rules to determine pitch, rhythm, and harmony. Later systems formalised this using stochastic context-free grammars or Markov chains, which are essentially simplified probabilistic grammars. These algorithms allowed for diversity and originality through the “surprise” effect, but were still very limited, as they followed a series of pre-written rules step by step without adapting the probabilities of choice to the input data (Wei et al., 2025; Zhao et al., 2025).

The next significant step in the development of AI in a broad sense, and in particular in music production, was the invention and spread of artificial neural network (ANN) technology in the 2000s and 2010s. ANNs are mathematical models programmed to function in a generalised context somewhat similar to the neural system of a biological brain – that is, to form connections between themselves, activating in response to information (signals) received from outside by “weighing” its significance. In the process of such hierarchical information analysis, “artificial neurons” (essentially mathematical functions) are combined into layers that gradually, step by step, transform the input data, determining increasingly abstract characteristics (LeCun et al., 2015). Fundamental to the work of ANNs is the process of regulating the weighting mechanism using algorithms that minimise the error between predicted and target results based on data examples created by the programmer. This process is conventionally referred to as *“machine learning”* (ML) (LeCun et al., 2015).

The integration of ML technology into the field of music production involved self-tuning in the weighting process by reducing prediction error with a focus on a set of purposefully selected music data. The first neural network models to perform these functions were recurrent neural networks (RNN). RNNs are a class of artificial neural networks designed to process sequential data. They store information from previous steps and use it to influence future results, making them suitable for modelling time-dependent data such as speech and music (Lipton et al., 2015). However, RNN technology had certain limitations, particularly due to the complexity of supporting long-term dependencies, which motivated the development of improved variants such as long short-term memory (LSTM) networks (Lipton et al., 2015).

LSTMs were developed to overcome the difficulties encountered by standard recurrent networks when storing information from long sequences. In particular, they introduced memory cells and control mechanisms that regulate the flow of information, allowing relevant context to be retained and irrelevant details to be discarded. This design made LSTMs more suitable for modelling long-term structures such as musical sequences (Hochreiter & Schmidhuber, 1997). For example, Eck & Schmidhuber (2002) applied LSTM to jazz improvisation, demonstrating that a neural network is capable of modelling musical material that continues given melodies and predicts chords while maintaining stylistic consistency. However, these systems were unable to model long-term structures and required special coding, so they could not be used in DAW (i.e., “in the box”).

The same drawback – the inability to integrate with DAW – was also characteristic of more advanced models based on deep learning technology (2010s). However, it should be noted that the capabilities of deep learning technology have caused a real revolution in the field of music production. While models based on machine learning (not deep learning) required manual configuration (characteristics such as tempo, intervals, intonation, and harmony were determined by researchers and integrated into an algorithm, which then studied musical patterns based on these characteristics), deep learning made it possible to create models that independently extracted characteristics directly from the input data. Thus, this technology contributed to the emergence of so-called multilayer neural networks, in which the lower layers study simple elements (e.g., short note patterns), and the upper layers combine them into more abstract representations (e.g., harmonic sequences). In music production, this technology made it possible to "train" models by providing them with raw MIDI files and even audio recordings as input data, allowing them to identify characteristics on their own. Thus, with the help of AI, complex musical structures (in terms of stylistic correspondence, phrasing, timbre, etc.) were generated that would have been too difficult to code manually (LeCun et al., 2015).

Among the advanced deep learning models developed in the 2010s are:

Variational autoencoders (VAE) – this type of model compresses data into a simplified but systematic "internal representation" and then reproduces it in its original form. The systematised structure on which VAE is based allows both the reproduction of input data and the creation of new variations based on stored information (Kingma & Welling, 2013);

Generative adversarial networks (GAN) – their functioning is based on the interaction of two models that work in opposition to each other – one creates new material, and the other checks how well this material corresponds to certain characteristics. Thanks to this competition, the system gradually improves until the generated results become indistinguishable from the benchmark. GANs are used to create convincing images, sounds, and music, although it should be noted that their training process can be unstable (Goodfellow, 2014).

A striking example of the application of deep learning technologies in music production is the Google Magenta project (2016), which created music generation tools such as MelodyRNN, PerformanceRNN, and MusicVAE. These models made it possible to create more original and artistically vivid musical material, but their use was mainly academic and required technical infrastructure. In particular, they could not be used as plugins or instruments integrated into DAWs and required programming or command line interfaces or Python code. Also, they were typically trained on small datasets, producing short MIDI sequences covering only a few bars (Roberts et al., 2018).

Thus, the next step in the evolution of AI in music production was the creation of tools based on deep learning models integrated into commercial services, and into DAWs. Among such tools are: LANDR (automatic mastering); Amper Music (AI-based soundtrack generation); iZotope's Neutron & Ozone (machine learning for mixing/mastering within DAW) (Sterne & Razlogova, 2021; Simpson & Groff, 2023; Hatem, 2023).

However, a fundamental breakthrough in music production and, at the same time, a radical paradigm shift led to the invention and integration of transformer architecture models into ANN (Vaswani et al., 2017). Unlike RNNs, which process data step by step, transformers rely on a self-attention mechanism that allows the model to consider all positions in the sequence simultaneously. By calculating the relevance of each token to every other token, the self-attention mechanism allows long-term dependencies to be modelled without repetition. Also, simultaneous processing of all positions (parallelisation) allows the model to be trained on very large datasets (Tiwari, 2025). Finally, positional encoding adds information about the order of the sequence, compensating for the lack of recursion (Paleti, 2024).

The transformer architecture has proven to be much more effective for music generation because it can capture global context (e.g., harmonic structure, phrase-level coherence) while simultaneously processing fine details (note timing, dynamics). This technology has been rapidly adopted in music generation systems such as Music Transformer, which uses relative positional encoding to model long musical sequences (Huang et al., 2018); MuseNet, trained on various MIDI datasets and capable of creating compositions using multiple instruments and in different styles (MuseNet, 2019); Jukebox, which combined transformers with hierarchical VQ-VAE to generate raw audio in various genres with vocal content (Dhariwal et al., 2020).

An important achievement enabled by transformer technology is the ability to generate music based on a text query (Agostinelli et al., 2023). Thanks to its ability to learn from paired data, the model has gained the ability to find correlations between text descriptions and musical patterns. For example, introducing a description such as "relaxing piano music" and corresponding musical data (MIDI or audio) during the training process allows the model to learn how words correlate with musical characteristics such as tempo, intonation, dynamics, timbre, etc. During inference (generation), the user provides a text query, and the model generates music that statistically matches the characteristics comparable to the descriptions obtained during training. An example of such a model is MusicLM with a text-to-music transformer system that generates long, coherent audio tracks based on text queries (Agostinelli et al., 2023).

However, despite the rapid pace of evolution of AI-based technologies in music production, only a few systems are currently suitable for use by professional music producers. The main problem is the prevalence of so-called “end-to-end models”, which accept user requests and return finished musical material in the form of audio signals that cannot be processed (e.g., WAV, MP3 files). Consequently, although this material may sound quite realistic, it cannot be edited – no changes can be made to the melody, instrumentation, rhythm, etc. In addition, generators of such audio often leave unwanted artefacts, such as noise, distortion, or unnatural textures, which are not easy to fix due to the closed nature of the system. In contrast, artificial intelligence models that output results in symbolic format (e.g., MIDI files) offer greater adaptability. Such results can be imported into a DAW and edited, making them more suitable for use in professional music production (Zhao et al., 2025).

Thus, to date, most AI tools for music production are designed primarily to replace human creativity rather than to support it. Even systems such as Generative Audio Workstations (GAW), which attempt to incorporate generative AI into music production workflows, are typically offered as standalone services, making it impossible to integrate them directly into existing software. One notable exception is LambDAW, a new GAW developed on top of the commercial workstation Reaper, which was created with the primary goal of providing seamless integration of programming into the workstation itself. As for other tools for composition and performance, systems such as GEDMAS and Musical Mosaicing can be mentioned, as well as other types of co-creative applications such as Reflexive Looper and Flow Machines (Giuliani et al., 2023, p. 4).

It should also be noted that the artistic characteristics of music generated by artificial intelligence technologies are still far from evoking the emotions that can be elicited by compositions created by humans and performed by live musicians. These shortcomings are particularly noticeable when AI generates musical material intended to imitate classical music and the sound of acoustic instruments. Conveying the nuances of agogics, dynamics, and timbre, which are conditioned by human perception and emotions, is still an unattainable level for generative technological creativity and, at the same time, an ambitious goal that inspires researchers to invent new algorithms for training artificial neural networks.

Thus, analysis of the literature has made it possible to identify two key trends in the contemporary professional music creation space, particularly in the field of music production: 1) the evolutionary development of the functional capabilities of the technologies involved in its cycle, in line with the stages of introduction of increasingly new and technically advanced classes of ANN models; 2) the decisive role of the adaptability of music material generated by technological means for integration into professional work processes. These trends also determine the prospects for preparing future Masters of Musical Art for creative activity in modern conditions. Among these prospects, the following should be identified:

– *technological literacy* – future professionals in the field of musical art must be familiar with AI-based music production tools. The evolution from recurrent neural networks (Eck & Schmidhuber, 2002) to multilayer generative models such as VAE (Roberts et al., 2018) and GAN (Yang et al., 2017) illustrates how rapidly the technological landscape is changing and how important it is to ensure that future professionals have a basic knowledge of these tools and their potential for use in contemporary forms of musical creativity, such as composition, sound design and music production;

– *adaptability as a professional skill* – An important factor influencing the artistic outcome of compositional creativity using AI is the format in which the musical material is generated. For example, material in symbolic format (e.g., MIDI) is much better suited for use in DAW workflows than audio files, which cannot be edited (Mitra & Zualkernan, 2025). Consequently, effectively preparing future Masters of Musical Art for creative work in the music production cycle requires teaching them to identify and utilise such adaptability while being aware of the specific limitations of current technology;

– *creative partnership with AI* – a significant drawback of most modern generative systems used in music production is their focus on replacing rather than supporting human creativity (Giuliani et al., 2023), which determines the expediency of forming in future Masters of Musical Art a value-based attitude towards live creativity, participation in compositional creativity through technology, and the desire to organise the music production cycle as a co-creative partnership using, for example, systems such as Flow Machines and MusicVAE, which provide opportunities to use AI to generate ideas, variations or accompaniments, while leaving artistic responsibility to the human creator;

– *critical analysis and assessment of artistic quality* – musical material generated by AI-based systems often has significant shortcomings, such as undesirable sound artefacts (Zhao et al., 2025) and, in general, the quality of compositions (in particular, parameters of form, dynamics, intonation, harmonic structure, etc.) does not always meet the criteria of stylistic authenticity and artistic value. This necessitates preparing students to use these systems from the perspective of critical analysis of musical material based on a established system of specific evaluation criteria.

Conclusions. The retrospective analysis shows that the integration of artificial intelligence technology into “in the box” music production is an evolutionary process. Early systems relied on fixed rules written by programmers, which allowed them to generate simple melodies and harmonies. Significant improvements came with the invention and implementation of artificial neural network (ANN) technology in music production, in which algorithms themselves determined the characteristics and adjusted the parameters of the generated musical material based on musical examples. This made it possible to imitate a certain artistic style, but with limited ability to analyse and create long musical structures.

With the advent of deep learning technology, the architecture of ANN models became much more multilayered, with each layer gaining the ability to further transform the input data. Through several levels of abstraction, the identified characteristics of individual musical elements in such models are combined into increasingly complex abstract structures, allowing patterns to be generalised, creating internal representations that reflect broader concepts such as mode, intonation, stylistic correspondence, etc.

A new era in the evolutionary development of AI was marked by the emergence of transformer architectures. Unlike recurrent networks, which process data step by step, transformers use a self-attention mechanism that considers all elements of a sequence simultaneously and determines how they are related to each other. This design allows for the identification of long-term dependencies, scaling to very large data sets, and parallel processing of information. In music production, transformers have made it possible to generate coherent compositions with a stable structure, multi-instrumental musical material, and create music based on text queries.

However, despite the rapid development of AI technology, issues such as limited user control in end-to-end models, undesirable artefacts in generated audio that cannot be processed or corrected, and the high resource requirements of modern models still hinder the widespread adoption of this technology in professional music production. In this regard, further research should focus on developing hybrid and interactive approaches that balance computing efficiency and creative adaptability and enable AI to function in collaboration with humans in professional music production.

The study identified a number of prospects in the context of preparing future Masters of Musical Art for creative activity in a modern professional environment. In particular, it was found that the changing space of professional and creative activity in the field of music production using AI technology requires the acquisition of appropriate technological literacy. The latter includes awareness of the peculiarities of the functioning of tools used to generate musical compositions, in particular AI technologies, which have significant potential to expand the possibilities of musical creativity. Another important perspective is the development of adaptability as a professional skill that ensures the ability to recognise, create and apply formats of musical material that are optimally adapted for creativity in music production, and to develop the ability to organise work with AI as a creative partnership. The latter involves using AI to generate ideas, variations or accompaniments, while being aware of one's own responsibility for the artistic result, the ability to evaluate which also constitutes a perspective on the need to develop an appropriate system of criteria for future Masters of Musical Art.

Thus, the priority direction for further research is to determine the essence and functions of the specific skills of future Masters of Musical Art to successfully engage in creative activity in a modern professional environment to implement a cycle of music production using AI technologies.

REFERENCES

1. Agostinelli, A., Denk, T. I., Borsos, Z., Engel, J., Verzetti, M., Caillon, A., ... & Frank, C. (2023). Musiclm: Generating music from text. *arXiv preprint arXiv:2301*.11325. https://doi.org/10.48550/arXiv.2301.11325
2. Civit, M., Civit-Masot, J., Cuadrado, F., & Escalona, M. J. (2022). A systematic review of artificial intelligence-based music generation: Scope, applications, and future trends. *Expert Systems with Applications, 209*, 118190. https://doi.org/10.1016/j.eswa.2022.118190
3. Collins, M. (2014). *In the Box Music Production: Advanced Tools and Techniques for Pro Tools* (1 ed.). Routledge. https://doi.org/10.4324/9780203066362
4. Dhariwal, P., Jun, H., Payne, C., Kim, J. W., Radford, A., & Sutskever, I. (2020). Jukebox: A generative model for music. *arXiv:2005*.00341. Retrieved from: https://assets.pubpub.org/2gnzbcnd/11608661311181.pdf
5. Eck, D., & Schmidhuber, J. (2002, August). Learning the long-term structure of the blues. In *International Conference on Artificial Neural Networks* (pp. 284–289). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/3-540-46084-5_47
6. Giuliani, L., De Filippo, A., & Borghesi, A. (2023). Towards Intelligent Music Production: A Sample-based Approach. In *Ceur workshop proceedings* (Vol. 3519, pp. 50–59). CEUR-WS. Retrieved from: https://hdl.handle.net/11585/952308
7. Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative Adversarial Nets. *Advances in Neural Information Processing Systems (NeurIPS), 27*. https://doi.org/10.48550/arXiv.1406.2661
8. Hatem, G. (2023). The Role of Artificial Intelligence in Music Production: A Look at Amper Music's Innovative Approach. Linkedin. Retrieved from: https://surl.lu/bffllx
9. Hiller, L. A., & Isaacson, L. M. (1979). *Experimental Music; Composition with an electronic computer*. Greenwood Publishing Group Inc. Retrieved from: https://dn790000.ca.archive.org/0/items/experimentalmusi00hill/experimentalmusi00hill.pdf
10. Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation, 9* (8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

11. Huang, C. Z. A., Vaswani, A., Uszkoreit, J., Shazeer, N., Simon, I., Hawthorne, C., ... & Eck, D. (2018). Music transformer: Generating Music with Long-Term Structure. *ICLR 2019 Conference Paper*. Retrieved from: https://openreview.net/pdf?id=rJe4ShAcF7

12. Kingma, D. P., & Welling, M. (2013). Auto-encoding variational bayes. *arXiv preprint arXiv*:1312.6114. https://doi.org/10.48550/arXiv.1312.6114

13. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521* (7553), 436–444. https://doi.org/10.1038/nature14539

14. Lipton, Z. C., Berkowitz, J., & Elkan, C. (2015). A Critical Review of Recurrent Neural Networks for Sequence Learning. *arXiv:1506.00019*. https://doi.org/10.48550/arXiv.1506.00019

15. Liu, J. (2024). Expressive MIDI-format Piano Performance Generation. *arXiv preprint arXiv*:2408.00900. https://doi.org/10.48550/arXiv.2408.00900

16. 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

17. Mozer, M. C. (1994). Neural Network Music Composition by Prediction: Exploring the Benefits of Psychoacoustic Constraints and Multi-scale Processing. *Connection Science, 6* (2–3), 247–280. https://doi.org/10.1080/09540099408915726

18. MuseNet. (2019). OpenAI. Retrieved from: https://openai.com/index/musenet/?utm_source

19. Ovcharenko, N., Merezhko, Y., Moskva, O., Neboga, O., & Kosiak, L. (2021). Technological competence of future music teachers: diagnostic and formative methods. *Amazonia Investiga, 10* (48), 238–247. https://doi.org/10.34069/AI/2021.48.12.25

20. Paleti, N. C. (2024). *Positional Encoding Explained: A Deep Dive into Transformer PE*. Medium. Retrieved from: https://medium.com/thedeephub/positional-encoding-explained-a-deep-dive-into-transformer-pe-65cfe8cfe10b

21. Quick, D. (2016). Learning production probabilities for musical grammars. *Journal of New Music Research, 45* (4), 295–313. https://doi.org/10.1080/09298215.2016.1228680

22. Rastrygina, A. M. (2020). Training of future specialists-musicians on the principles of interdisciplinary models of free arts. *Academic Notes. Series: Pedagogical Sciences*, (190), 38–42. https://doi.org/10.36550/2415-7988-2020-1-190-38-42

23. Roberts, A., Engel, J., Raffel, C., Hawthorne, C., & Eck, D. (2018a). A hierarchical latent vector model for learning long-term structure in music. In *International conference on machine learning* (pp. 4364–4373). PMLR. Retrieved from: https://proceedings.mlr.press/v80/roberts18a.html

24. Roberts, A., Engel, J., Raffel, C., Simon, I., & Hawthorne, C. (2018b). MusicVAE: Creating a palette for musical scores with machine learning. Magenta. Retrieved from: https://magenta.withgoogle.com/music-vae

25. Simpson, J., & Groff, B. (2023). The 5 Best AI Music Generators of 2023. Benjamin Groff. Retrieved from: https://www.benjamingroff.com/top-ai-music-generators/

26. Sterne, J., & Razlogova, E. (2021). Tuning sound for infrastructures: artificial intelligence, automation, and the cultural politics of audio mastering. *Cultural Studies, 35* (4–5), 750–770. https://doi.org/10.1080/09502386.2021.1895247

27. Tiwari, V. (2025). Attention Is All You Need – A Deep Dive into the Revolutionary Transformer Architecture. Towards AI. Retrieved from: https://surli.cc/ijvsla

28. Todd, P. M. (1989). A Connectionist Approach to Algorithmic Composition. *Computer Music Journal, 13*(4), 27–43. https://doi.org/10.2307/3679551

29. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention Is All You Need. *31st Conference on Neural Information Processing Systems (NIPS 2017)*. Long Beach, CA, USA. https://doi.org/10.48550/arXiv.1706.03762

30. Wei, L., Yu, Y., Qin, Y., & Zhang, S. (2025). From Tools to Creators: A Review on the Development and Application of Artificial Intelligence Music Generation. *Information, 16* (8), 656. https://doi.org/10.3390/info16080656

31. Yang, L. C., Chou, S. Y., & Yang, Y. H. (2017). MidiNet: A convolutional generative adversarial network for symbolic-domain music generation. *arXiv preprint arXiv:1703*.10847. https://doi.org/10.48550/arXiv.1703.10847

32. Zhao, Y., Yang, M., Lin, Y., Zhang, X., Shi, F., Wang, Z., Ding, J., & Ning, H. (2025). AI-Enabled Text-to-Music Generation: A Comprehensive Review of Methods, Frameworks, and Future Directions. *Electronics, 14* (6), 1197. https://doi.org/10.3390/electronics14061197

Штучний інтелект у музичному продакшні: перспективи підготовки майбутніх магістрів музичного мистецтва до творчої діяльності в сучасному професійному середовищі

Інга Олексіївна Хмелевська
кандидат педагогічних наук,
старший викладач кафедри музичного мистецтва і хореографії
ДЗ «Південноукраїнський національний педагогічний університет імені К. Д. Ушинського»
ORCID: 0000-0001-9129-2173

Володимир Олегович Слободянюк
здобувач третього (освітньо-наукового) рівня вищої освіти
ДЗ «Південноукраїнський національний педагогічний університет імені К. Д. Ушинського»
ORCID: 0009-0008-4186-2113

Стаття присвячена проблемі впровадження технологій штучного інтелекту (ШІ) до циклу цифрового музичного продакшну та її значущості в контексті професійної підготовки майбутніх магістрів музичного мистецтва. Мета дослідження – простежити еволюційний розвиток моделей штучного інтелекту від ранніх конекціоністських і заснованих на правилах систем до сучасних архітектур на основі трансформерів, охарактеризувати чинники, що впливають на їх застосовність у професійних робочих процесах генерації музичного матеріалу, визначити їхні переваги й обмеження у проєкції перспектив підготовки майбутніх магістрів музичного мистецтва до успішної творчої діяльності в сучасному професійному середовищі. Методи дослідження включають ретроспективний аналіз ключових етапів розвитку моделей штучного інтелекту, порівняльний аналіз рекуррентних нейронних мереж, варіаційних автокодувальників, генеративних змагальних мереж і систем на основі трансформерів, а також синтез літератури з використанням міждисциплінарних джерел у галузі комп'ютерної музики, когнітивних наук і музичної педагогіки.

Результати показали, що штучний інтелект у музичному продакшні еволюціонував відповідно до послідовного процесу розроблення нових поколінь моделей: від ранніх символьних систем до глибоких генеративних архітектур і мультимодальних технологій перетворення тексту на музику. З кожним етапом удосконалення моделей музичні продюсери отримували нові інструменти для творчості, але суттєві обмеження та недоліки, що зберігаються і досі, зумовлюють необхідність підготовки професіоналів до критичного оцінювання та здійснення адаптації згенерованого матеріалу, що інтегрується в цифрові аудіоробочі станції (DAW). Основними перспективами підготовки майбутніх магістрів музичного мистецтва до творчої діяльності в сучасному професійному середовищі визначено необхідність набуття технологічної грамотності, адаптивності в роботі зі згенерованим матеріалом, здатності до взаємодії зі штучним інтелектом у форматі спільної творчої та критичного усвідомлення технологічних і художньо-творчих обмежень. Перспективою подальших досліджень є визначення сутності та функцій специфічних умінь майбутніх магістрів музичного мистецтва для успішної творчої діяльності в сучасному професійному середовищі, реалізуючи цикл музичного продакшну з використанням технологій штучного інтелекту.

Ключові слова: *штучний інтелект, музичний продакшн, цифрові аудіоробочі станції (DAW), моделі глибокого навчання (RNN, VAE, GAN, Transformer), підготовка майбутніх магістрів музичного мистецтва, технологічна грамотність.*