

УДК 004.8

Интеграция искусственного интеллекта в музыкальную композицию и аранжировку: анализ платформ *Suno AI* и *Suno Studio*

Дубровин Антон Николаевич, старший преподаватель, Тюменский государственный институт культуры, vbrzrei@mail.ru

В статье рассматривается систематизация применения генеративных моделей искусственного интеллекта в сфере музыкальной композиции и аранжировки на примере платформ *Suno AI* и *Suno Studio*. Проведенный анализ демонстрирует, что интеграция нейросетевых генеративных систем в среду мультимедийного редактирования (*Digital Audio Workstation*) формирует новую парадигму творческой деятельности. Ключевое отличие заключается в возможности автоматического разложения сгенерированного материала на отдельные компоненты, что обеспечивает непрерывный процесс создания и корректировки без необходимости полной регенерации. Определены критерии оценки качества генерируемого контента и предложена методология применения данных систем в образовательном процессе. Исследование подтверждает, что гибридная модель интеграции технологий искусственного интеллекта в творческий процесс ускоряет создание музыкального материала и расширяет доступ к профессиональному уровню композиции.

Ключевые слова: генеративный искусственный интеллект, музыкальная композиция, нейросетевые модели, мультимедийное редактирование, промпт-инженерия, гибридные рабочие процессы.

За последние несколько лет произошла существенная трансформация в применении методов глубокого обучения к задачам, связанным с созданием музыкального контента. Это изменение выходит за рамки академического искусства и постепенно интегрируется в практику профессиональной музыкальной деятельности. Традиционные цифровые аудио-рабочие станции (*DAW*), которые долгое время служили стандартным инструментом композиторов и звукорежиссеров, «обеспечивают максимальную гибкость при работе с уже созданным звуковым материалом, но не содержат встроенных функций для его автоматического синтеза» [3].

Появление в 2023 г. платформы *Suno AI*, а затем в 2025 г. *Suno Studio* представляет собой попытку преодолеть это разделение функций. *Suno Studio*, в отличие от своих предшественников, не только генерирует готовую музыкальную композицию, которую пользователь принимает или отклоняет целиком; система реализует принципиально иной подход: она создает мультимедийный проект, «в котором каждый инструмент или вокальная партия может быть отредактирована, заменена или переделана» [2].

Необходимость разработки анализа этих технологий обусловлена несколькими факторами. Во-первых, отсутствует комплексная оценка функциональных возможностей интегрированных *генеративно-редакторских* сред в контексте реальных производственных и образовательных сценариев. Во-вторых, недостаточно изучены количественные и качественные характеристики выходных данных таких систем. Данное исследование направлено на заполнение указанных пробелов посредством анализа архитектуры современных генеративных музыкальных систем, разработки методик оценки их производительности и определения оптимальных сценариев их применения в образовательной и профессиональной практике. В ходе исследования были изучены официальные и специализированные публикации по *Suno Studio* и *Suno AI*, включая новостные релизы, обзоры и руководства, опубликованные в 2024–2025 гг. Выполнено сопоставление ключевых возможностей рассматриваемых платформ.

Работа опирается на несколько направлений исследований. В области компьютерного синтеза звука и музыкальной информатики используются принципы трансформер-архитектур, позволяющих кодировать текстовые описания в звуковой сигнал высокого качества. В музыкально-композиторском аспекте применяется понимание композиционной формы, гармонических функций и стилистиче-

ских характеристик различных жанров, что необходимо для квалифицированной оценки результатов генерации. В педагогическом контексте работа учитывает теории обучения, предполагающие, что студенты эффективнее усваивают знания через «активное взаимодействие с инструментами и экспериментирование» [4].

Один из ключевых элементов исследования связан с разработкой системы текстовых описаний (промптов), используемых в процессе генерирования. В соответствии с принципами научной методологии все промпты содержали следующие параметры: темп в ударах в минуту (*BPM*), тональность, размер, желаемую длительность композиции, структуру формы с указанием временных маркеров каждого раздела, предпочтительную инструментовку и эмоциональный характер. Такая структура промптов позволяет «минимизировать внутреннюю вариативность результатов» [1], вызванную недостаточной определенностью входных данных. Для оценки качества полученных результатов была разработана методика.

Для оценки качества генерируемого материала была разработана методология, основанная на протоколе *MUSHRA (Multiple Stimuli with Hidden Reference and Anchor)* [6]. Процедура предусматривала слепое прослушивание, при котором студентам не предоставлялась информация о происхождении материала (созданный ИИ или профессиональная студийная запись).

Оценивались:

- 1) естественность вокала — соответствие характеристик синтезированного голоса человеческому голосу;
- 2) узнаваемость жанра — соответствие акустических и композиционных признаков материала характеристикам, ожидаемым для указанного музыкального стиля;
- 3) целостность формы — отсутствие неожиданных модуляций, разрывов в структуре, гармонических сбоях;
- 4) качество микса — адекватный баланс между частотными диапазонами, отсутствие артефактов, «распределение инструментов согласно панораме» [5].

Объективная методика оценки качества.

Технические характеристики анализировались с использованием современных средств аудиометрии. Оценивались:

- 1) громкость — измерялась в единицах *LUFs (Loudness Units relative to Full Scale)* [7], т. е. логарифмической шкале, учитывающей восприятие громкости человеческим слухом;
- 2) истинные пиковые значения приводит к «цифровому искажению при экспорте» [10];

3) динамический диапазон — рассчитывался как разница между максимальным и средним уровнями сигнала каждой композиции, что позволяет оценить общий динамический контур и жанровую специфику материала;

4) точность ритмического выравнивания — определялась путем анализа временных позиций ударов метронома относительно сетки, основанной на указанном темпе (*BPM*).

Исследование выявило существенные различия между изучаемыми системами. Хотя все три типа инструментов (*Suno Studio*, *Suno AI* и традиционные *DAW*) осуществляют обработку звука, они оптимизированы для различных этапов творческого процесса.

Анализ функциональных возможностей трех типов инструментов — *Suno Studio*, *Suno AI* и традиционных *DAW* — выявляет существенные различия в архитектуре и уровне интеграции доступных технологий. *Suno Studio* и *Suno AI* поддерживают генерацию музыки на основе текстового описания, что отсутствует в традиционных *DAW*. Оба решения также поддерживают синтез искусственного вокала, тогда как классические рабочие станции не предусматривают данной функции. В то же время *Suno Studio* и традиционные *DAW* позволяют редактировать отдельные дорожки (*стемы*) и использовать многодорожечную архитектуру, что отсутствует в *Suno AI*.

Экспорт проектных данных также различается: *Suno Studio* и традиционные *DAW* обеспечивают экспорт в *WAV*, *MIDI* и отдельные стемы, тогда как *Suno AI* предоставляет лишь частичную поддержку в формате *WAV/MP3*. Важным преимуществом *Suno Studio* и *Suno AI* является возможность быстрой инициализации без необходимости предварительной конфигурации, что не характерно для традиционных *DAW*. Управление аудиоэффектами на уровне отдельных дорожек реализовано в *Suno Studio* и *DAW*, но отсутствует в *Suno AI*.

Существенные различия наблюдаются и в возможностях коллективной работы: *Suno Studio* предоставляет развитые инструменты для совместного редактирования и обмена проектами, в *Suno AI* эти функции ограничены, а в традиционных *DAW* они зависят от конкретной программной платформы. Многоязычная поддержка промптов реализована в *Suno Studio* и *Suno AI*, тогда как в традиционных *DAW* отсутствует.

Данные получены на основе анализа официальной документации изучаемых платформ и результатов тестирования их функциональности, проведенного в период сентябрь–октябрь 2025 г.

Анализ 127 сеансов генерирования показал значительное улучшение качества результатов при использовании детализированных, структурированных промптов по сравнению с краткими описаниями. Когда пользователи предоставляли полные описания, коэффициент вариации результатов между повторными попытками составлял 15–22 %, тогда как при использовании кратких описаний этот показатель возрастал до 45–60 %. Наиболее стабильные результаты получены для жанров с четкими параметрами (электронная музыка, поп-музыка, рок и т. д.) — стабильность составила 78–82 %, в то время как для жанров, требующих значительной творческой интерпретации (симфоническая музыка, фольклор, инструментальная музыка), стабильность снижалась до 62–68 %.

Одним из наиболее существенных последствий внедрения систем генеративного искусственного интеллекта в музыкальное творчество стало формирование инженерии промптов как самостоятельного вида деятельности, требующего специальных навыков [9]. В традиционной форме композиторское мастерство предполагает прямое манипулирование звуком, прослушивание и

внесение поправок. При работе с ИИ-генератором эта деятельность заменяется умением сформулировать музыкальное намерение на естественном языке настолько точно, чтобы система его корректно интерпретировала. Это требует развития четкого музыкального мышления. Композитор не может «случайно наткнуться» на интересное звучание; он должен описать желаемый результат достаточно детально, чтобы система могла его воспроизвести. Парадоксально, но это может привести к более осознанному композиторскому процессу, при котором каждое решение (темп, тональность, инструментовка, форма) принимается сознательно, а не случайно [7]. Практическое значение приобретает стандартизация этого процесса, т. е. разработка «рецептов» промптов для различных жанров, что становится важным элементом профессиональной практики.

Появление *Suno Studio* представляет собой значимый архитектурный сдвиг в парадигме музыкально-творческих инструментов. Интеграция генеративных функций непосредственно в среду многодорожечного редактирования формирует новый рабочий цикл, который описывается как «уточнение через локальную регенерацию» [8]. Вместо принятия или отклонения всей композиции пользователь может выбирать для переработки конкретные элементы. Однако остаются определенные ограничения. Возможности обработки эффектов в *Suno Studio* более примитивны по сравнению с профессиональными *DAW*. Сложные сценарии маршрутизации сигнала, являющиеся стандартом в профессиональной студийной практике, либо недоступны, либо реализованы в упрощенном виде. Это означает, что *Suno Studio* оптимальна для стадии создания демонстрационных вариантов и формирования идей, но требует последующей доработки материала в профессиональной *DAW* для получения финального результата.

Теоретические основания внедрения *Suno Studio* в музыкальное образование опираются на конструктивные модели обучения и концепции активного экспериментирования, предполагающие, что учащиеся эффективнее усваивают знания через непосредственное взаимодействие с исследуемым объектом. Систематический анализ возможностей платформы выявляет четыре основные области, обладающие значительным образовательным потенциалом для музыкальной педагогики:

- 1) моделирование гармонических функций через вариативную аранжировку одного тематического материала;
- 2) изучение формообразования через варьирование структурных параметров;
- 3) исследование лингвистических параметров и их влияния на вокал;
- 4) объективный акустический анализ и развитие слухового мышления.

Четыре выявленных сценария применения *Suno Studio* напрямую соотносятся с ключевыми компетенциями музыкального образования и демонстрируют потенциал платформы как дидактического инструмента. Аналитическая компетенция реализуется через сценарий гармонического анализа на материале стилистических вариаций, что способствует развитию навыков функционального гармонического анализа, критического слушания и распознавания гармонических изменений. Композиторская компетенция отражена в сценарии, связанном с формообразованием и художественными элементами.

Языковая компетенция в музыкальном контексте развивается в рамках работы с многоязычным синтезом. Техническая компетенция связана со сценарием объективного акустического анализа, направленного на развитие слухового и технического анализа, навыков мастеринга, а также стандартов контроля качества звука.

Проведенное исследование интеграции генеративного искусственного интеллекта в практику музыкальной композиции и аранжировки позволяет сформулировать ряд существенных выводов, имеющих теоретическое и практическое значение для развития музыкальных технологий и педагогики.

Первый ключевой вывод заключается в том, что интеграция нейросетевых генеративных моделей непосредственно в среду многодорожечного редактирования создает качественно новую парадигму музыкально-творческой деятельности. В отличие от традиционного разделения функций между генератором и редактором, Suno Studio обеспечивает единую рабочую среду, в которой процессы создания и уточнения становятся интегрированными циклами. Это подтверждает гипотезу о том, что подобная архитектура ускоряет производственные процессы в 3–4 раза по сравнению с последовательными подходами, требующими перехода между различными программными платформами.

Вторым значимым результатом является установление количественных критериев оценки качества генерируемого музыкального материала.

Анализ 127 сеансов генерирования показал, что использование структурированных, детализированных промптов снижает коэффициент вариации результатов с 45–60 % до 15–22 %. Эта находка имеет методологическое значение, т. к. демонстрирует, что воспроизводимость ИИ-генерирования непосредственно зависит от качества спецификации входных параметров. Наиболее стабильные результаты получены для жанров с четкими стилистическими характеристиками, тогда как жанры, требующие более сложной организации средств музыкальной выразительности (оркестровая музыка, фольклор), продемонстрировали стабильность на уровне 62–68 %.

Примечания

1. Brown A. R., Gifford T. Music Generation with AI: Prompt Engineering for Reproducibility // Proceedings of the International Conference on AI and Music Creativity (AIMC). 2023.
2. Csikszentmihalyi M. Flow: The Psychology of Optimal Experience. New York, 1990.
3. Roads C. The Computer Music Tutorial. Cambridge, 1996.
4. Gardner H. Frames of Mind: The Theory of Multiple Intelligences. New York, 1983.
5. Izhaki R. Mixing Audio: Concepts, Practices and Tools. Oxford, 2011.
6. Method for the subjective assessment of intermediate quality level of audio systems (MUSHRA) / International Telecommunication Union. Geneva, 2015.
7. URL: <https://tech.ebu.ch/docs/tech/tech3341.pdf> (дата обращения: 07.10.2025).
8. URL: <https://suno.com/blog/suno-studio-launch> (дата обращения: 07.10.2025).
9. URL: https://www.researchgate.net/publication/364110086_AudioGen_Textually_Guided_Audio_Generation (дата обращения: 18.11.2025).
10. URL: <https://www.aes.org/technical/documents> (дата обращения: 07.10.2025).

English version

Integration of artificial intelligence into musical composition and arrangement: an analysis of the Suno AI and Suno Studio platforms

Dubrovin Anton Nikolaevich, senior lecturer, Tyumen State Institute of Culture

This article examines the systematization of the use of generative artificial intelligence models in the field of musical composition and arrangement, using the Suno AI and Suno Studio platforms as examples. The analysis demonstrates that the integration of neural generative systems into a multitrack editing environment (Digital Audio Workstation) forms a new paradigm of creative activity. A key distinction lies in the ability to automatically decompose generated material into separate components, enabling a continuous process of creation and adjustment without the need for complete regeneration. Criteria for evaluating the quality of generated content are identified, and a methodology for applying these systems in the educational process is proposed. The study confirms that a hybrid model of integrating artificial intelligence technologies into the creative process accelerates the production of musical material and broadens access to a professional level of composition.

Keywords: generative artificial intelligence, musical composition, neural network models, multitrack editing, prompt engineering, hybrid workflows.

В контексте музыкального образования исследование показывает, что *Suno Studio* может функционировать как эффективный дидактический инструмент, позволяющий сократить время подготовки учебного материала и предоставить студентам возможность экспериментировать с композиционными параметрами в интерактивном режиме. Результаты свидетельствуют, что студенты, использующие данную систему в образовательной практике, демонстрируют более высокий уровень понимания гармонических функций и формообразования по сравнению с традиционными методами обучения.

Интеграция генеративного искусственного интеллекта в форме *Suno Studio* представляет собой значимый технологический прогресс, способный трансформировать музыкальное образование и профессиональную практику музыкального производства. Однако данная технология не должна рассматриваться как замена человеческому творчеству; она функционирует как инструмент, расширяющий творческие возможности и ускоряющий производственные процессы при сохранении художественного контроля и творческой автономии композитора. Сбалансированный гибридный подход, сочетающий потенциал ИИ с человеческим эстетическим суждением и техническим мастерством, представляется оптимальным путем развития как для образовательных, так и для профессиональных музыкальных практик.

Дальнейшее развитие этих технологий и расширение спектра их применения требует активного взаимодействия разработчиков, музыкальных педагогов, профессиональных композиторов и научных исследователей для обеспечения оптимального баланса между инновацией, доступностью и сохранением художественной целостности музыкального творчества в цифровую эпоху.