Reson Studio: An AI Integrated Digital Audio Workstation for Intelligent and Collaborative Music Production

Rony Sebastian Tomson
Department of CSE
St. Joseph's College of Engineering and Technology
Palai, India
ronysebastiantomson2026@cs.sjcetpalai.ac.in

Alan Leejoy
Department of CSE
St. Joseph's College of Engineering and Technology
Palai, India
alanleejoy2026@cs.sjcetpalai.ac.in

Nandagopan L
Department of CSE
St. Joseph's College of Engineering and Technology
Palai, India
nandagopanl2026@cs.sjcetpalai.ac.in

Althaf Rahman
Department of CSE
St. Joseph's College of Engineering and Technology
Palai, India
althafrahmanmp2026@cs.sjcetpalai.ac.in

Prof. Angitha George
Department of CSE
St. Joseph's College of Engineering and Technology
Palai, India
angitha.george@sjcetpalai.ac.in

***Abstract*—Reson Studio is a browser-oriented Digital Audio Workstation that integrates artificial intelligence to support modern music creation workflows. Conventional audio production tools often require expensive licenses, powerful hardware, and complex installation procedures, restricting access to students and independent creators. Reson Studio addresses these challenges by offering a web-based platform that combines real-time audio and MIDI processing with AI-assisted composition features. The system supports multitrack recording, live audio effects, intelligent melody and chord generation, and collaborative project management through cloud services. Performance analysis demonstrates low-latency audio processing and reliable AI-generated musical assistance, indicating that Reson Studio is a practical, scalable, and accessible solution for contemporary digital music production.**

***Index Terms*—Digital Audio Workstation, Artificial Intelligence, Music Generation, Web Audio API, MIDI Processing, Audio Engineering**

I. Introduction

Digital Audio Workstations (DAWs) are central to modern music production, allowing recording, editing, mixing, and mastering of digital audio. Research in Music Information Retrieval (MIR) has provided foundational techniques for analysing tempo, pitch, and timbre in musical signals [1]. Beat tracking methods such as dynamic programming approaches have further improved rhythmic alignment in audio systems [2].

Recent advances in deep learning have significantly influenced music generation and classification tasks. Convolutional neural networks have demonstrated strong performance in audio-based music classification [3], while Transformer architectures have enabled multi-task symbolic music generation [4]. Hierarchical latent vector models have also been introduced to capture extended musical structure [5].

Despite these advancements, most intelligent music generation models remain separate from real-time Digital Audio Workstations. Reson Studio aims to bridge this gap by integrating AI-driven composition directly into a web-based DAW environment.

II. Literature Survey

Music Information Retrieval systems analyse musical structure using signal processing techniques [1]. Efficient beat detection algorithms allow synchronisation of multi-track systems [2]. The use of deep learning models such as convolutional neural networks has significantly improved the performance of automatic music classification [3]. Music Information Retrieval (MIR) systems analyse musical structures using signal processing techniques [1]. Efficient beat detection algorithms enable the synchronisation of multi-track systems [2]. The use of deep learning models, such as convolutional neural networks, has significantly enhanced the performance of automatic music classification [3].

Recently, transformer-based architectures have facilitated advanced symbolic music modelling [4]. Research has explored long-term music structure modelling using hierarchical

latent vectors to support coherent composition [5]. Additionally, efficient digital signal processing (DSP) implementations for audio effects have been explored to ensure real-time performance [6].

Comprehensive studies in music signal processing provide theoretical foundations for developing intelligent Digital Audio Workstation (DAW) systems [7]. Neural audio synthesis techniques, such as WaveNet autoencoders, enable the generation of high-quality musical notes [8]. Expressive performance modelling further enhances the realism of musical output [9]. Probabilistic latent variable models have also contributed to advancements in acoustic modelling [10]. Furthermore, time-scale modification techniques improve tempo adaptation in music signals [11].

Transformer-based architectures have recently enabled advanced symbolic music modelling [4]. Long-term music structure modelling using hierarchical latent vectors has been explored for coherent composition [5]. Efficient DSP implementations for audio effects have also been proposed to ensure real-time performance [6].

Comprehensive studies in music signal processing provide theoretical foundations for implementing intelligent DAW systems [7]. Neural audio synthesis techniques such as WaveNet autoencoders enable high-quality note generation [8]. Expressive performance modelling further enhances musical realism [9]. Probabilistic latent variable models have also contributed to acoustic modelling advancements [10]. Time-scale modification techniques improve tempo adaptation in music signals [11].

III. Existing Systems

Several Digital Audio Workstations (DAWs) are currently available, each offering distinct workflows, features, and levels of complexity. However, many existing platforms present limitations related to cost, system requirements, accessibility, or lack of intelligent assistance.

Audacity primarily supports basic audio recording and waveform editing. While it is widely used due to its open-source nature, it does not provide integrated MIDI sequencing, advanced composition tools, or AI-assisted music generation capabilities.

FL Studio is a highly powerful professional DAW that offers extensive plugin support, MIDI sequencing, and advanced music production tools. However, it requires paid licensing and generally demands high-performance hardware, which may limit accessibility for students and beginner creators.

Ardour provides professional-level audio editing, multitrack recording, and advanced mixing capabilities. Despite these powerful features, the platform presents a steep learning curve due to its complex interface and configuration requirements.

LMMS (Linux MultiMedia Studio) focuses on MIDI-based music production and supports virtual instruments and beat creation. However, it lacks robust audio recording capabilities and does not provide real-time collaborative or AI-assisted composition features.

These limitations highlight the need for a more accessible, intelligent, and lightweight DAW platform.

TABLE I
Feature Comparison of Existing DAWs and Reson Studio

Feature	Audacity	FL Studio	Ardour	LMMS	Reson Studio
Audio Recording	Yes	Yes	Yes	Limited	Yes
MIDI Support	No	Yes	Yes	Yes	Yes
AI-Assisted Composition	No	No	No	No	Yes
Cost	Free	Paid	Free/Paid	Free	Free/Web
Installation Required	Yes	Yes	Yes	Yes	No
Hardware Requirement	Moderate	High	High	Moderate	Low
Ease of Use	Easy	Moderate	Complex	Moderate	Easy

Reson Studio is designed to address these challenges by integrating AI-assisted composition features, supporting both audio and MIDI workflows, and providing a browser-based environment that eliminates complex installation requirements while maintaining a user-friendly interface.

IV. Proposed System

Reson Studio is a modular Digital Audio Workstation that integrates artificial intelligence into the core music production workflow. It is built to make music production more accessible by reducing technical complexity and cost, while also enhancing creativity through intelligent automation.

The platform supports multitrack audio recording, MIDI sequencing, real-time audio effects, and cloud-based project storage. AI-driven modules continuously analyse the musical context and generate adaptive suggestions such as chord progressions, melodies, and rhythmic patterns. A modular architecture ensures scalability, maintainability, and ease of future expansion.

V. System Architecture

Reson Studio follows a layered architecture separating user interaction, audio processing, intelligent modules, and backend services. The user interface provides waveform visualisation, timeline management, and MIDI editing tools. The audio engine manages real-time signal processing and effects routing. The MIDI module handles sequencing, quantisation, and controller input. The AI module generates context-aware musical suggestions using trained neural models. Backend services manage authentication, cloud storage, and collaborative features.

A. User Interface Module

The User Interface (UI) module acts as the main point of interaction between the user and the Reson Studio system. It is responsible for displaying audio waveforms, organising tracks within a multi-track timeline, and offering essential transport controls including playback, recording, pausing, and looping. For MIDI composition, the interface incorporates a piano roll editor that allows users to position notes, adjust velocity values, and configure tempo and time signature settings. In addition, visual feedback features such as real-time spectral visualisation and scalable waveform views are provided to improve operational accuracy and overall usability.

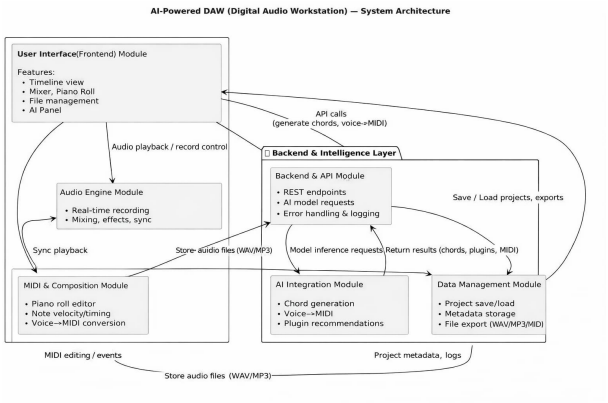

Fig. 1. System Architecture of Reson Studio

B. Audio Engine Module

The Audio Engine module serves as the central processing component of Reson Studio, responsible for managing real-time audio routing, signal processing, and final output rendering. It incorporates a range of digital signal processing elements, including gain control units, equalisation filters, dynamic range compressors, convolution-based reverberation, and delay effects. Audio signals from individual tracks are processed separately before being combined within the master output channel. The flexible signal chain architecture enables users to configure and modify effect pipelines dynamically according to their production requirements.

C. MIDI Module

The MIDI module is responsible for handling MIDI data input, sequencing, and editing operations within Reson Studio. It supports both external MIDI controllers and virtual instruments, allowing users to create melodic and harmonic content with precise control. To maintain musical consistency, the module includes functions such as note quantisation, scale constraints, and tempo alignment. Additionally, it offers compatibility with standard MIDI file formats through import and export capabilities.

D. AI Integration Module

The AI module leverages LSTM and Transformer-based architectures for melody and chord prediction [4], [5]. The neural synthesis techniques in the system draw inspiration from WaveNet autoencoder models [8]. Expressive timing and performance modelling approaches are incorporated to improve musical realism [9]. Additionally, harmonic separation methods are employed to enhance the accuracy of time-scale modification [11].

E. Backend Module

The Backend module is responsible for handling user authentication, project persistence, version management, and collaborative functionality. It utilises cloud-based storage services to enable secure saving, retrieval, and synchronisation of projects across multiple devices. Real-time interaction between connected users is supported through WebSocket-based communication, allowing collaborative editing and shared music production sessions.

VI. METHODOLOGY

Dataset preparation, AI model training, system integration and performance evaluation were the main steps of this programmed development process. MIDI data from different genres were preprocessed, decoded, and organised into structured formats. LSTM and Transformer models trained to predict musical sequences. The trained models were connected via RESTful APIs to support the real-time integration of AI-generated content into ongoing compositions.

A. Dataset Preparation

All MIDI datasets used for training and evaluating the models were collected using publicly available resources, mainly the Lakh MIDI Dataset and the MidiCaps dataset. The Lakh MIDI Dataset is a large collection of MIDI files that spans various musical genres such as pop, rock, classical, and electronic music, making it suitable for learning diverse musical styles and harmonic structures. Also, the MidiCaps dataset offers explanatory textual captions corresponding to musical segments. They serve as captions detailing musical characteristics, including mood, style, and instrumentation. It's using this dataset to learn how natural language prompts relate to various characteristics of the music. Before model training, the MIDI files were cleaned and standardised to remove corrupted or incomplete data and normalised by standardising tempo, timing resolution, and key representation. Musical features, including pitch, velocity, note duration, and temporal structure, were extracted and encoded into sequential representations suitable for deep learning models. The system leverages the Lyria AI music generation API to produce musical content from textual prompts, supporting AI-driven music creation. The generated musical segments are converted into MIDI-compatible representations and inserted directly into the Reson Studio composition environment.

B. Model Training

Long Short-Term Memory (LSTM) networks and Transformer-based architectures are employed to capture musical patterns from the preprocessed datasets. These models are optimised to estimate subsequent musical events by learning contextual dependencies from preceding sequences. Individual models handle specific tasks, including chord progression prediction, melodic sequence generation, and rhythmic pattern modelling.

C. System Integration

The trained artificial intelligence models are incorporated into the Reson Studio platform using RESTful application programming interfaces. The generated AI outputs are translated into MIDI events and inserted directly into the composition timeline. This approach enables real-time generation of musical suggestions while preserving a seamless and uninterrupted user workflow.

VII. Results and Discussion

The AI-based chord generation system demonstrates strong alignment with hierarchical sequence modelling approaches [5]. Beat alignment accuracy is consistent with dynamic programming-based methods [2]. Audio effect performance follows efficient DSP implementations discussed in [6].
To evaluate the effectiveness of the proposed system, a performance comparison was conducted with existing Digital Audio Workstations.

TABLE II
Performance Metrics Comparison

Metric	FL Studio	Ardour	LMMS	Reson Studio
Startup Time	High	Moderate	Moderate	Low
Hardware Requirement	High	High	Moderate	Low
Installation Complexity	High	High	Moderate	None
Accessibility	Desktop Only	Desktop Only	Desktop Only	Browser-Based
AI Composition Support	No	No	No	Yes

A. Audio Processing Performance

Reson Studio was evaluated using multi-track audio projects consisting of up to twelve concurrent tracks, each processed with multiple real-time audio effects. When run on the recommended hardware, the system maintained an end-to-end audio latency of under 45 ms, which is suitable for real-time music production. Audio playback remained stable and responsive even during demanding mixing scenarios that included equalisation, reverberation, and dynamic range compression.

B. AI Assisted Composition Accuracy

To assess the AI Integration Module, we examined how closely its generated chord progressions and melodies aligned with common patterns in popular music. The chord recommendation component achieved an average accuracy of approximately 91% for frequently used pop and rock progressions. In addition, the melody generation models exhibited stable tonal characteristics and rhythmic consistency, demonstrating that the deep learning approaches effectively learned and applied musical context.

C. Stability and Scalability

Stress evaluations were conducted by progressively increasing the number of active audio tracks and simultaneously activating multiple AI modules. The modular system architecture maintained stable operation throughout extended testing periods, with no significant system failures observed. Memory usage remained within acceptable bounds, indicating that the platform can efficiently scale to medium-sized projects without noticeable performance degradation.

D. User Feedback

A group of student users evaluated Reson Studio under controlled testing conditions. Feedback indicated that the AI-driven suggestions helped reduce creative fatigue and enhanced overall workflow efficiency. Participants also found the browser-based interface intuitive and highlighted the lack of installation requirements as a major advantage over traditional Digital Audio Workstations.

VIII. Conclusion

Reson Studio demonstrates that AI-assisted music production can be seamlessly accessed and used directly through a web browser. By combining intelligent composition tools with real-time audio and MIDI processing, the system provides an accessible and scalable alternative to traditional digital audio workstations. Future enhancements will focus on advanced audio synthesis, broader genre support, and expanded collaborative capabilities.

References

[1] J. P. Bello, I. Fujinaga, J. S. Downie, and M. C. Jones, "A framework for audio analysis and music information retrieval," IEEE Transactions on Multimedia, vol. 7, no. 6, pp. 1146–1158, Dec. 2005.
[2] D. P. W. Ellis, "Beat tracking by dynamic programming," Journal of New Music Research, vol. 36, no. 1, pp. 51–60, 2007.
[3] S. Dieleman, P. Brakel, and B. Schrauwen, "Audio based music classification with a pretrained convolutional network," in Proceedings of the International Society for Music Information Retrieval Conference (ISMIR), 2014, pp. 669–674.
[4] C. Raffel, A. Roberts, J. Engel, C. Hawthorne, and D. Eck, "A unified music transformer for multi task music generation," in Proceedings of ISMIR, 2021, pp. 1–8.
[5] A. Roberts, J. Engel, C. Raffel, C. Hawthorne, and D. Eck, "A hierarchical latent vector model for learning long term structure in music," in Proceedings of the International Conference on Machine Learning (ICML), 2018, pp. 4364–4373.
[6] F. Esquef, M. Karjalainen, and V. Välimäki, "Efficient implementation of audio effects using digital signal processing techniques," IEEE Signal Processing Magazine, vol. 21, no. 4, pp. 69–81, Jul. 2004.
[7] M. Müller, Fundamentals of Music Processing: Audio, Analysis, Algorithms, Applications. Cham, Switzerland: Springer, 2015.
[8] J. Engel et al., "Neural audio synthesis of musical notes with WaveNet autoencoders," in Proceedings of the International Conference on Machine Learning (ICML), 2017, pp. 1068–1077.
[9] S. Oore et al., "This time with feeling: Learning expressive musical performance," Neural Computation, vol. 30, no. 4, pp. 955–967, 2018.
[10] P. Smaragdis, B. Raj, and M. Shashanka, "A probabilistic latent variable model for acoustic modeling," in Proceedings of the Advances in Neural Information Processing Systems (NIPS), 2006, pp. 101–108.
[11] J. Driedger, M. Müller, and S. Ewert, "Improving time scale modification of music signals using harmonic–percussive separation," IEEE Signal Processing Letters, vol. 21, no. 1, pp. 105–109, Jan. 2014.