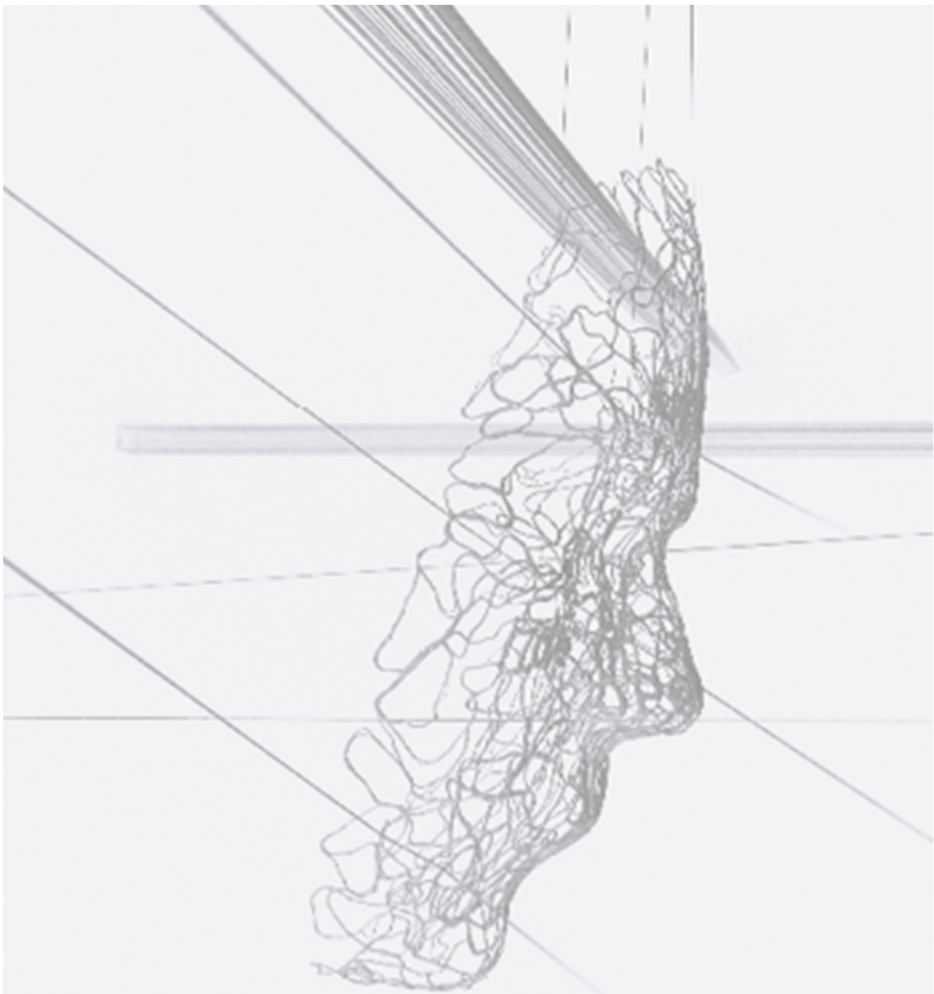
I N S Δ M

JOURNAL OF CONTEMPORARY MUSIC, ART AND TECHNOLOGY



Artificial Intelligence:
Duality in Applications of Generative AI and Assistive AI in Music

David Atanacković

INSAM Journal of Contemporary Music, Art and Technology
No. 12, July 2024, pp. 12–31.
https://doi.org/10.51191/issn.2637-1898.2024.7.12.12

INSAM

DOI https://doi.org/10.51191/issn.2637-1898.2024.7.12.12

David Atanacković*
Faculty of Music, University of Arts
Belgrade, Serbia

ARTIFICIAL INTELLIGENCE: DUALITY IN APPLICATIONS OF GENERATIVE AI AND ASSISTIVE AI IN MUSIC

Abstract: This paper explores the multifaceted role of Artificial Intelligence in the field of music, more specifically, examining the positives and negatives of generative and assistive capacities. Artificial Intelligence (AI) in music involves the application of computational techniques to various aspects of music creation, production and consumption. In the domain of assistive AI, the concentration is on how machine learning could potentially help musicians in the area of composition and performance to enhance their musical creativity. The paper will discuss an interesting collaborative effort between pure human creativity and computational assistance covering an explanation for a vast number of tools using generative as well as assistive artificial intelligence models. In addition, the paper will address the concerns facing the music industry while this technology keeps on improving, the potential drawbacks and ethical considerations. It opens the question of authenticity and emotional depth, and when or if this new technology could be able to replicate it. Further explanation in the paper will consider music examples with a focus on music styles assisted and generated by the use of artificial intelligence, from pop to classical music. With a thorough analysis of the aforementioned subject, the paper aims to provide a detailed perspective on the constant evolution of AI tools used in music with highlights on the need for a balanced approach. In providing

* Author's contact information: david.atanackovic@gmail.com.

a detailed perspective on the evolving landscape of AI tools in music, this study adopts a methodological approach that involves comprehensive analysis of both the benefits and challenges associated with these innovative gadgets. The paper contributes to the ongoing discussion on the intersection of technology and artistic expression. By examining the potential benefits and challenges with these innovative models, the paper signifies ongoing discourse on the impact of technology on artistic expression.

Keywords: Artificial Intelligence (AI), Composition, Technology, Music, Collaboration.

Introduction

Artificial Intelligence (AI), a term coined by emeritus Stanford Professor John McCarthy in 1955, was defined by him as "the science and engineering of making intelligent machines" (Manning 2020). In this ever growing and relatively new field, the distinction has been made between two constantly evolving branches of the domain, called Generative AI and Assistive AI. In its essence, the Generative part of Artificial Intelligence technology covers the possibility of generating various types of content, which in its expansion could potentially come close to the possibilities, creation, and talent that mimics a human being. One of the most well-known examples is image generation. The service providers that are popular in this area of expertise used machine learning in fusion with the neural networks concept to train their servers on a vast amount of image datasets. On the other hand, the counterpart of Generative AI, called Assistive AI, powers the enhancement of human creativity, efficiency and knowledge improvement. On the subject of finance, it helps to optimize investment strategies, analyze the current state of market trends and detect suspicious activities. Even though not fully accepted or utilized, these tools can be found in the world of music as well and have recently became very popular for creating full songs, compositions of various genres and helping music enthusiasts and professionals. The place for improvement is still astonishing, merging traditional ways, innovation and assistance in the creative process. However, the potential of AI to assist and generate content raises the question of ethical considerations not only in music, but also in various other fields.

Methodology

This study involves an exploration of the current state of Generative and Assistive AI, concentrating on the music profession. It considers the potential,

limitations and ethical implications of the above-mentioned technology. With thorough research of existing literature, it explores the connection and differences between Generative AI and Assistive AI, and its potential impact on human creativity and workflow. This collection of data relies heavily on investigating existing gadgets and artificial intelligence tools by focusing on their advantages and drawbacks, the potential for improvement and advantages as well as their effectiveness. The tools, which are discussed in the paper, were chosen based on certain criteria, mainly focusing on software with the latest and most up-to-date technology advancements, the amount of features and possibilities within these generative and assistive AI platforms, their ease of use and their level of popularity. The detailed analysis shows existing patterns, themes and trends in the AI world involving music creativity. The study will conclude with concepts of possible expansions of AI implications and an examination of how to navigate the technology ethically and responsibly.

Chapter I
The Soundscape of Generative AI

One of the primary subjects of generative AI in music involves composition at the forefront. At the heart of music generative tools lie a diverse array of algorithms and techniques designed to mimic and augment the creative process. All the tools employ machine-learning models, such as recurrent neural networks (RNN), variational autoencoders (VAE) and generative adversarial networks (GAN). They include a diverse array of tools and platforms, each offering unique approaches to algorithmic creativity and sonic exploration. All these tools could be divided into the sections of music components they specialize in, from melody generation, harmonic progression, rhythmical patterns and sound synthesis. In the last few years, there has been an expansion of these software and tools and numerous companies have been competing with each other in this domain, thus making it impossible to cover all of them.

Generative AI tools

MusicVAE and MuseNet

In the first chapter, we will focus on some of the most advanced up-to-date generative music tools. One such is called **MusicVAE**. Researchers at Google Magenta developed it. MusicVAE employs variational autoencoders (VAE) to learn the underlying structure of musical data and generate music compositions

based on learned patterns and styles. MusicVAE has advanced in such a way that enables users to explore and create a vast number of musical possibilities, from familiar melodies to avant-garde compositions. Similarly, OpenAI's MuseNet leverages deep neural networks to compose original music across a wide range of genres and styles. By training on a diverse dataset of musical scores, MuseNet learns to capture the nuances of different musical traditions and genres, generating compositions that exhibit both creativity and coherence. This online tool captures jazz improvisations and classical compositions by feeding its algorithm with a proper composer and performer's work. However, its focus is generating simple melodies and short motifs.

Soundraw and AIVA

Another notable device is Soundraw, a platform that allows users to generate royalty-free music tailored to their specific needs and preferences. By leveraging machine learning to analyze user inputs and musical parameters, Soundraw generates custom compositions, offering a seamless solution for content creators, multimedia producers or even filmmakers seeking soundtracks for their projects. In the same category, among the pioneers in this domain is AIVA, an AI music generation assistant that allows users to generate new songs in more than 250 different styles in a matter of seconds. It is suitable for seasoned professionals as well as complete beginners in the field. What is more, it allows consumers to create their own style models by uploading audio or MIDI files to influence the algorithm. The composition can be then downloaded in multiple file formats. However, its focus is generating mainly instrumental music.

Suno AI and MelodyRNN

Similarly, Suno AI not only generates music from the text input, but it can also generate suitable lyrics for the song, as shown by a few examples on their official website. Another suitable melodic generator might be Magenta's MelodyRNN that excels at generating melodic sequences. Similarly to the previously mentioned software, it leverages deep learning algorithms to analyze melodic contours present in existing musical compositions, enabling it to generate new melodies that adhere to established stylistic conventions while also exploring novel variations and permutations.

Harmonic and rhythmic AI generators

When it comes to generating harmonic progressions, a component that plays a crucial role in shaping the tonal and emotional landscape of a composition,

generative tools such as ChordChord and ToneGym Chord Progression Generator are adept at generating harmonic sequences that complement and support melodic lines. These tools employ machine-learning algorithms to analyze the harmonic structures, chord progressions and voice leading principles prevalent in different musical genres and styles, enabling them to generate new progressions that are both harmonically rich and aesthetically pleasing. From the aspect of the rhythmic music component, GroovePizza specializes in generating rhythmic patterns and drum loops that potentially form the foundation of a musical composition. Analyzing the rhythmic motifs, syncopated grooves and polyrhythmic interactions present in various musical traditions and cultures enables the tool to generate new patterns that are both rhythmically diverse and dynamically engaging.

Generative sound synthesis and transformation tools

In addition, generative AI is also capable of synthesizing sounds and timbres that contribute to the overall sonic palette of a composition. Platforms such as Google's Nsynth and OpenAI's GPT-3 based SynthGPT analyze the spectral characteristics, harmonic overtones and transient dynamics and shaping present in different acoustic and synthetic instruments and sound sources, enabling them to synthesize new sounds with the fusion of realism and expressiveness. One of the more advanced ideas derived from ChatGPT and generative art Discord servers (Midjourney), which SynthGPT utilizes, is text-to-prompt. This allows users to describe the sound they want to use in their composition and production that results in a fully created synth patch. However, it is not yet updated to the point at which SynthGPT allows the creatives to change some of its parameters.

Another interesting addition to sound synthesis-based AI is Magenta's Tone Transfer. Users can upload any audio file and material that is analyzed and later synthesized to some acoustic instruments such as Saxophone, Violin, Flute, Cello etc. Tone Transfer is also available as a VST (Virtual Studio Technologies) plugin, which music professionals can use in their DAW (Digital Audio Workstation) of choice.

One of the more advanced technologies with the same text-to-prompt idea is used in the Synplant plugin, a VST instrument highly used in modern electronic music. In its new version (2.0), not only is it possible to describe the sound and get the full patch, but also producers and composers are allowed to shape the sound using more advanced parameters such as filtering, shaping the length of the sound using ADSR envelope (Attack, Decay, Sustain and Release) and more.

Even though SynthGPT and Tone Transfer (along with Synplant) could be thought of as generative AI tools, due to the fact of their text prompt or sound

analytic possibilities, they can also potentially be used as assistive tools for composers and producers.

Personalized text-powered music generation tools

MusicLM

Language processing advancements are offering a unique approach to music creation through conversational interaction with the machine. One example is MusicLM. By interpreting user inputs and preferences, MusicLM generates personalized tracks tailored to individual tastes. It is a model generating music form text descriptions such as 'a calming violin melody backed by a distorted guitar riff'. MusicLM casts the process of conditional music generation as a hierarchical sequence-to-sequence modeling task, and it generates music at 24kHz (kilo Hertz) with mono compatibility that remains consistent over several minutes. MusicLM can be conditioned on both text and melody in that it can transform whistled and hummed melodies according to the style described in a text caption (Google, n.d.).

MusicGen

Certain extensions of the MusicLM model can be traced in its more advanced version called MusicGen. In the words of its authors, it is:

> a powerful single Language Model (LM) redefining the boundaries of conditional music generation, with the ability to create high-quality music by taking cues from text descriptions or melodies. Extensive studies have confirmed the superior performance of MusicGen compared to existing approaches (MusicGen n.d.).

One of the biggest advancements in transition from the MusicLM model to the one used in MusicGen is the possibility of creating music compositions and productions with not only mono, but with stereo compatibility as well.

Stable Audio

Another good example of text prompt-based tools is Stable Audio, a music generation product built by Stability AI (Stable Audio 2024). Users can utilize their website to create original tracks that can potentially be used later for their personal projects. Audio files from the website are allowed to be downloaded as

stereo files with a sampling frequency rate of up to 44.1 kHz, which is a current industry standard, besides 48 kHz, more preferred in the video, film and video game industry.

Lyria

Beyond all the aforementioned products, one of the most currently advanced technologies in the scope of generative AI music is Google's Lyria. In partnership with YouTube in the focus of Shorts, Lyria represents two AI experiments designed to open a new playground for creativity. The first one is called Dream Track, which is an experiment in YouTube Shorts still in its trial period, designed to help deepen connections between artists, creators and fans through music creation. The second is a set of music AI tools that Google is designing with professional artists, songwriters and producers to help bolster their creative processes:

> Within the experiment, a limited set of creators will be able to use Dream Track for producing a unique soundtrack with the AI-generated voice and musical style of artists including Alec Benjamin, Charlie Puth, Charli XCX, Demi Lovato, John Legend, Sia, T-Pain, Troye Sivan and Papoose. Each participating artist has partnered with us and will have a hand in helping us test and learn to shape the future of AI in music (Google DeepMind 2023).

Considering the target of this new form of created material is the audience and their stronger relationship with the artists, it would be for fans around the globe to create songs tailored to their needs using these available algorithms.

> Dream Track users can simply enter a topic and choose an artist from the carousel to generate a thirty-second soundtrack for their Short. Using our Lyria model, Dream Track simultaneously generates the lyrics, backing track and AI-generated voice in the style of the participating artist selected (Google DeepMind 2023).

Assistive AI tools by Lyria will be discussed later on.

Udio

One of the latest text-powered tools in the scope of generative AI in music is Udio. On 10th April, 2024, the developers publicly released a Beta version of their model. Even though it is still in its early stages, Udio has proven to be one of the most advanced text-to-music tools at the moment. Critics praised the developer's ability to create higher quality and realistic sounding vocals. It allows its users to

create 600 songs a month free of charge with a vast majority of music genres to choose from. It allows users to describe what kind of song or instrumental they want to produce with the additional feature to choose between an instrumental or vocal version. The lyrics may be generated by the service as well or can be written by the users themselves. It became very popular and widespread, especially on social media, where people started creating pages and channels with music they generated on this platform and gaining millions of views. With the possibility of an additional payment to this service, users have all the copyright on finished song materials. However, this raised a question of the possibility that developers at Udio trained their material on copyrighted data.

Chapter II
Implementation of Assistive AI in music

As stated previously, the idea behind Assistive AI in music is to create a symbolic relationship between human and machine to foster collaboration and accelerate the creative process. In this way, the machine would behave as a personal assistant to established music professionals, music producers, audio engineers, sound designers, performers and even music enthusiasts. The pioneers behind some of these tools, as mentioned earlier, are the creatives from IRCAM, dating to the late 1990s to the beginning of the 2000s. They have developed assistive music software such as OpenMusic and Orchidea, which intrigued and inspired primarily established and well-known contemporary classical composers to create their original work, like Brian Ferneyhough, Marco Stroppa, Kaija Saariaho, Tristan Murail and more. The software are available either as plugins (standalone or for DAW) or are designed to be controlled via another interactive software called MaxMSP.

Assistive AI tools

AI tools by IRCAM

Orchidea functions as a composition assistant or computer-aided orchestration tool, giving orchestral or arrangement solutions based on user input. Music composers may 'upload' an orchestral solution they currently have and by analyzing that specific situation, Orchidea could potentially propose numerous different orchestration and arrangement choices based on the composer's input. In the words of the IRCAM developers on Orchidea:

> Assisted orchestration can be thought as the process of searching for the best combinations of orchestral sounds to match a target sound under specified metric and constraints. Although a solution to this problem has been a long-standing request from many composers, it remains relatively unexplored because of its high complexity, requiring knowledge and understanding of both mathematical formalization and musical writing. Orchidea is an evolution of the Orch tools and performs static and dynamic assisted orchestration. Orchidea performs mono-objective optimization on various features and requires a little number of parameters for orchestration (Cella 2019).

OpenMusic takes a similar approach, offering a visual programming environment for computer assisted composition and analysis, empowering users to explore complex musical structures.

Emergent Drums

Another company called Audialab recently launched a drum programming and drum sample synthesis tool called Emergent Drums. Since rhythm lies at the heart of musical expression, Emergent Drums harnesses the power of AI to imbue compositions with dynamic, expressive drum patterns. By leveraging advanced neural networks trained on vast repositories of rhythmic data, Emergent Drums generates intricate grooves, fills and even polyrhythms with human-like fluidity and nuance. Whether constructing pulsating beats for electronic dance tracks or intricate percussion arrangements for cinematic soundscapes, this tool offers a versatile toolkit for rhythm exploration and experimentation. What is more, Emergent Drums offers a way to upload a composer's or producer's own samples, analyze them and create numerous variations of the same sample which could later be used. The only downside is that it does not work well with samples of longer duration.

Magenta bundle

Google's Magenta created a diverse array of tools for assistive AI, that could even be downloaded as plugins and incorporated into a musician's DAW of choice, or they could be installed as standalone software. In some new versions, these tools are collected in a bundle called Magenta Studio, a plugin built on Magenta's open-source models. The preferred way of usage is incorporation with Ableton Live. Assistive tools available in the bundle are Continue, Interpolate, Groove and Drumify. In its essence, it is a MIDI plugin which lets users apply existing models to Ableton Live's MIDI clips in Session View. One innovative feature, similar to the creation of sample variation in Emergent Drums, is la-

beled Temperature. All plugins in this bundle have this slider, which is basically a parameter used for controlling randomness: higher values produce more variation or creative chaos, while lower values are more conservative in their predictions. One drawback of the software, specifically its melody generation, is that it is limited to monophonic (one voice or one note at a time) input. As stated previously, all the tools employ machine-learning models, which is the same with the Magenta Studio bundle.

First, the plugin called Continue uses the predictive power of recurrent neural networks or RNN to generate notes that are likely to follow drum beats or melodies. This means its algorithm should be fed with a pre-existing musical motif, whether it is a groove or a melodic structure. After the processing, an existing MIDI clip can be extended by up to 32 measures. This could be especially helpful for adding variation to a drum beat or creating new material for a melodic track. It picks up on factors such as durations, key signatures and timing. By increasing the temperature parameter, it can produce more random outputs.

Second, the plugin called Generate is similar to Continue, but it generates a four-bar phrase with no necessary input. Composers can choose the number of variations, which can be especially helpful for breaking creative blocks. It can also be used as a source of inspiration. Unlike previous RNN machine-learning models, Generate uses Variational Autoencoder or VAE that has been trained on millions of melodies and rhythms. Generate chooses a random combination of these summarized qualities and decodes it back to MIDI to form a completely new MIDI clip which can later be edited in Ableton.

The third plugin, called Interpolate, has been described by the developers. It "takes two drum beats or two melodies as inputs. It then generates up to 16 clips, which combine the qualities of these two clips. It is useful for merging musical ideas or creating a smooth morphing between them" (Magenta n.d.). Similarly to Generate, Interpolate also utilizes Variational Autoencoder.

The fourth plugin is specialized in drum clips and patterns. It is called Groove and it "adjusts the timing and velocity of an input drum clip to produce the 'feel' of a drummer's performance. This is similar to what a 'humanize' plugin does, but achieved in a totally different way" (Magenta n.d.). The way Groove was programmed with such precision is through fifteen hours of prerecorded material with real drummers. Many digital audio workstations have the ability to create innovative timing and velocity adjustments, especially to drum parts since they can tend to become overly robotic in the arrangement. This parameter is usually called **swing,** which tries to move the time of the accentuated or non-accentuated beats away from the grid in a way they do not exactly align with the tempo. This parameter is also adjustable in **quantization** section in many digital audio workstations.

The fifth plugin is called Drumify and it:

> creates grooves based on the rhythm of any input. It can be used to generate a drum accompaniment to a bassline or a melody or to create a drum track from a tapped rhythm. It works best with performed inputs, but it can also handle quantized clips…We extract a rhythm from each performance by removing the pitches and velocities, while keeping the precise timing details. When you provide an input sequence, be it a melody, bassline, chord progression or a drum pattern – we extract a rhythm in the same way and have the model turn it into a groove (Magenta n.d.).

Lyria

Another advanced assistive AI option for music creators is Google's DeepMind project called Lyria, already mentioned in the context of generative AI (Dream Track). Beyond experimenting with YouTube shorts to get a deeper connection between artists and their fans, Google is currently developing a set of music AI assistive tools with the help of music industry professionals like songwriters, artists and producers. These audio demo examples can be heard on Google's DeepMind website (Google DeepMind, 2023). The idea is to analyze the singing or recorded voice of an artist or a composer as well as having the ability to import MIDI information and transform it to a suitable medium like ensembles, orchestras, choirs, solo instruments or to add an instrumental accompaniment to a vocal track. This means that, for example, it is possible to transform a beatbox into a drum loop, singing into an orchestral score or MIDI piano information into a vocal choir, which is shown on a few examples.

The developers at DeepMind stated:

> With our music AI tools, users can create new music or instrumental sections from scratch, transform audio from one music style or instrument to another and create instrumental and vocal accompaniments. This work draws on our history of research and experimentation with AI and music and we'll continue testing our music AI tools with incubator participants throughout their development (Google DeepMind 2023).

Assistive AI audio production tools

It is no secret that assistive AI software can be found in audio production. Many audio professionals like audio, mixing or mastering engineers and sound designers are utilizing some of the software to their advantage. One of the pio-

neers in this section is company called Izotope, well-known for their post-production audio effects. In their mixing and mastering suite called Ozone, they created a feature called Master Assistant. The Master Assistant helps users to quickly achieve optimal mastering settings by analyzing audio and providing customized suggestions for equalization, compression and numerous other parameters. One similar array of audio tools, called Gullfoss by Soundtheory, also specializes in the mastering process. The idea is to use AI-driven software to optimize the clarity, balance and dynamics of an audio recording. By analyzing audio signals in real-time and applying targeted equalization, compression and enhancement, Gullfoss ensures optimal sonic quality across a wide range of playback systems. Various tools in the domain of podcast editing have also been developed as AI-driven by analyzing audio and creating a vocal or speech separation from the ambience. This is especially helpful if audio was previously recorded in an environment with a lot of echo and reverberation. One such tool is called GOYO, an AI voice separator tailored for reducing reverb, ambience and noise from vocal and speech recordings.

Generative audio workstations (GAW)

In addition, a few companies emerged to create futuristic digital audio workstations that are leveraging artificial intelligence in music creation in their software, also known as Generative Audio Workstations or GAW. Some of the most advanced Generative Audio Workstations are RipX and Wavtool. Some of the greatest assets of these workstations, that these companies emphasize, are options called stem separation and chat bots, which are not available in a traditional DAW setting.

The idea of stem separation is the possibility to create multiple audio files from one main stereo file. Based on machine learning, it can analyze the mixed or mastered stereo version of an audio file, its frequency spectrum and isolate various elements of the mix into a set of separate audio files. This is very useful for music creatives and is often utilized by music composers and producers to create a remix of popular songs.

Even though the pioneers in this region are the developers at Izotope with their tool called music rebalance, which made vocal isolation, instrument separation, level adjustment and enhancement possible, the above-mentioned generative workstations are a bit more advanced in that aspect with more control and capabilities.

For instance, the RipX developers came up with their unique 'Rip' format that is in connection with audio waveforms. For this precise reason, the software is available to make audio stem separation along with numerous other features

such as changing the tempo, mute/solo the stems, add innovative effects, equalization or even transposing stems to a new scale. This is an extension of basic functions that Izotope already utilized; however, the main difference is the fact that there are significantly fewer artefacts like clicks and pops appearing in audio files after the stem separation process than the one used in Izotope's RX. In addition, various effects could assist music composers and producers to make original music from those existing stems.

Unlike RipX that had its focus in terms of stem separation to stand out from the competition, Wavtool is the first GAW to employ a personal assistant or a chat bot called Conductor, powered by ChatGPT-4. Similar to how the text in GPT works, it is used to describe and ask the assistant to create chord progressions, basslines, melodies and even mix or master the tracks suitable to the style composed and produced. It is possible to create drum loops from the written text, receive the suggestions on harmonies and post processing, stem separation, converting audio files to MIDI files, time stretching manipulation and more.

The current state of Generative and Assistive AI in music

Even though it is evident that artificial intelligence is a technology where important companies such as Google keep on investing, seeing it as potentially highly lucrative, at least in music, it is still fairly new and not fully accepted or even understood. Although it can help to generate content and assist people in various fields, in music, whether generative or assistive, it still has plenty of drawbacks and it has not been perfected yet. The results given by the above-mentioned GAW software and additional AI-based tools still tend to sound overly robotic, plain and generic and the audio quality given by AI, containing less or more unwanted artefacts, does not still completely match with the current audio or music industry standards.

One of the most well-known audio engineers, producer, author and educator, Bobby Owsinski, explains the drawbacks in audio quality using AI models for mastering in his book *The Musician's Ai Handbook*:

> Another thing to keep in mind is that many Ai mastering platforms have a maximum resolution of 44.1kHz/16bit. While that works fine for a CD or for submission to a streaming distributor, it may not satisfy the needs of a record label or high-resolution distributor like Apple Music or Tidal, who require at least 24bit files with a sampling rate at 96kHz or higher (Owsinski 2023, 96–104).

In addition, there are significant limitations of ideas produced by current AI music technology, and after a few generated variations of a melody, chord

progressions, drum loops and similar, it starts to sound very repetitive. Furthermore, the engines behind AI are still not powerful enough to evoke emotions of a real human playing and a professional performer's feel. It is especially difficult in the domain of classical music or any form that uses acoustic instruments exclusively. Many companies produced virtual instruments for DAW software that replicate the sound for better or worse, but it is never identical to a real performer, recorded or listened in a good acoustical space. The imperfection in a performer's live playing or studio setting is a quality which separates them from a machine. Relying solely on AI to create something original and intriguing without any intervention from a music professional is still highly unlikely.

As generative and assistive artificial intelligence continue their advancements in technology, art and various other fields, it is essential to consider the ethical implications and societal impacts of these technologies.

Chapter III
Ethical considerations and future directions

Questions of authorship, ownership, copyright and accountability loom large in the age of AI-driven creativity, raising concerns about the role of humans in the creative process and the potential for algorithmic bias and discrimination. Furthermore, as generative AI tools become increasingly sophisticated and autonomous, they may challenge traditional notions of artistic authenticity and originality, blurring the distinction between human and machine creativity.

While AI algorithms can undoubtedly inspire, augment and amplify human creative endeavors, they also raise fundamental questions about the nature of creativity itself and the relationship between technology and culture. Looking ahead, the future of generative AI in music composition and beyond is ripe with possibilities and challenges. As researchers and practitioners continue to push the boundaries of AI-driven creativity, it is crucial to foster interdisciplinary dialogue and collaboration, ensuring that these technologies are developed and deployed in ways that are ethical, equitable and empowering.

AI Schubert

In the domain of classical music, there was a project developed by Huawei Company in 2019 to demonstrate the technology in their smartphones. They set up a challenge to 'complete' one of the most famous classical music works – Franz Schubert's Symphony no. 8. With the existing two movements and fragments of final two movements, they decided to use AI models to complete the composition. Online portal Classic.fm wrote:

> Engineers fed music in the form of data, into the phone's dual Neural Processing Unit – so the AI had information about the timbre, pitch and meter that Schubert liked to use in his melodies. The AI then created melodies from that information and composer Lucas Cantor chose his favorites. He then orchestrated those melodies and turned them into the final two movements to complete Schubert's Unfinished Symphony (Davis 2019).

Again, this is an example of how AI was used as a tool for inspiration, but a real human composer did the orchestration and, partially, composition. This was, however, an example of music that is not protected by copyright law, due to the fact that it expires 70 years after the author's passing and immediately becomes public domain.

AI implementation in popular music

The main turning point that aroused more concerns on authorship and copyright happened in 2023, when a song called 'Heart on my sleeve' appeared using AI-trained voices of some of the biggest artists today, more specifically Drake and The Weeknd. An anonymous music producer going by the name of Ghoswriter produced it. The beat and all the instrumental parts were done using virtual instruments and samples inside a DAW; however, the vocal part was trained using AI tools. The song has been pulled from multiple streaming platforms since it violated intellectual property rights. Universal Music Group, the record label where both of these artists are signed, said in a media statement: "The training of generative AI using our artist's music represented both a breach of our agreements and a violation of copyright law" (Axios 2023).

The case of Midjourney and Stability AI

Another similar violation happened in the field of images and graphic art. Namely, illustrators and cartoonists Sarah Andersen, Kelly McKernan and Karla Ortiz are suing creators of AI art generators called Midjourney and Stability AI. The artists claim that these organizations have violated the rights of "millions of artists" (Vincent 2023) by training their AI engine on a couple of billion images without the authors' consent. In the meantime, in the USA (District of Columbia), online magazine 'Vulture' wrote: "A federal court ruled on the August 18th that AI-generated artwork cannot be copyrighted on the grounds that copyright law only extends to human beings" (Davis 2023). This happened after the US Copyright Office refused to accept Stephan Thaler's image generated by AI as an original work. Judge Beryl A. Howell stated that "human authorship is an essential part of a valid copyright claim" (Guy 2023).

In this way, human artists could be protected and favored by laws against the unethical usage of AI software. However, this raises the question of how or if the authorities will be able to distinguish and know whether a piece of art was created by a human or a machine and how someone or something would detect if a machine was trained on works still protected by copyright law.

SynthID

Developers at Google's DeepMind came up with an idea of a watermark engraved in audio files to recognize whether a work was fully generated using AI technologies. On the company's website is a further and more detailed explanation:

> Our team is also pioneering responsible deployment of our technologies with best-in-class tools for watermarking and identifying synthetically generated content. Any content published by our Lyria model will be watermarked with SynthID, the same technology toolkit we're using for identifying images generated by Imagen…SynthID embeds a watermark into AI-generated audio content that's inaudible to the human ear and doesn't compromise the listening experience…The watermark is designed to maintain detectability even when the audio content undergoes many common modifications…SynthID can also detect the presence of a watermark throughout a track to help determine if parts of a song were generated by Lyria (Google DeepMind 2023).

Employment concerns

Universal law on AI and intellectual properties is still considered and not yet established in every country in the world, thus leaving many future concerns that should be addressed and discussed. Therefore, it is still difficult to determine the level of impact, whether positive or negative, that AI may potentially have on music and in other art domains. In the book by Mark Coeckelbergh called *AI Ethics,* the author writes about "A famous report by Benedikt Frey and Michael Osborne (2013) predicts that 47 percent of all jobs in the United States could be automated. Other reports have less dramatic figures, but most predict that job loss will be significant" (Coeckelbergh 2020, 137).

Even though the future is still uncertain, it is no secret that AI will affect the art workforce as well. For instance, many professional music composers and producers are earning a living by working for music libraries that are promoting and specializing in synchronization music for film trailers, TV series and advertisement. In this process, musicians are earning royalties and advance fees provided by music supervisors and various clients. The implementation, rise

and fast improvement of AI models raise the question of whether these kind of jobs could potentially become obsolete, the main reason being the cost effective nature of generative AI platforms with no need of greater advance fees. In addition, clients may get finished songs and compositions within seconds or minutes, which makes the whole process significantly faster. The process and the quickness of generative AI is something that most music composers are not able to compete with.

However, generative AI is only able to produce and simulate pre-existing music material, only by analyzing different patterns and musical structures following certain rules. Generative AI is unable to be creative to the point of which it can produce something that is unheard of and completely original. In that sense, human creation and originality have a great advantage over this technology by breaking these rules and constantly moving the boundaries. This may lead to an increasing need for original and groundbreaking works, becoming more valuable due to the impact of AI on the music market; yet, this point of view might be utopian.

Conclusion

The continuation in the evolution of AI technologies and its potential effect on the music industry may become transformative, moving away from the traditional ways of creating, promoting and distributing all musical content. Firstly, generative AI systems are enabling autonomous creation of music composition and availability of the technology to not only music professionals, but also expanding the process to amateurs and enthusiasts. On the other hand, the area of assistive tools promises to help and enhance the creative process for human musicians, providing suggestions and assistance with music composition, sound synthesis and audio production. The state of the current technology has come a long way, introducing the possibility of creating melodies, harmonies, and chord progressions. Text-based audio synthesis could help transform the way music is made. Although the primary focus of AI is to enhance and empower human creativity, it is of utmost importance to discuss the ethical considerations of its usage. The first concern is the potential loss of human creativity and using artificial intelligence as a potential replacement for human artists. This could potentially raise the question of the state of employment in all creative industries Finally, addressing these questions should be done in collaboration and discussion with various professionals in the field of art, technology, art business professionals and society, ensuring that AI in art is promoted and used responsibly and ethically.

List of References

Axios. 2023. "AI copyright in spotlight after platforms pull 'fake Drake' song." https://www.axios.com/2023/04/19/ai-fake-drake-weeknd-song-streaming-services-removed.

Cella, Carmine-Emanuele. 2019. "Orchidea." IRCAM forum. https://forum.ircam.fr/projects/detail/orchidea/.

Clancy, Martin. 2022. *Artificial Intelligence and Music Ecosystem*. CRC Press.

Coeckelbergh, Mark. 2020. *AI Ethics.* The MIT Press.

Davis, Wes. 2023. "AI-generated art cannot be copyrighted, rules a US federal judge." ClassicFM, 6 February, 2019. https://www.theverge.com/2023/8/19/23838458/ai-generated-art-no-copyright-district-court.

Davis. Elizabeth. 2019. "Schubert's 'Unfinished' Symphony completed by artificial intelligence". https://www.classicfm.com/composers/schubert/unfinished-symphony-completed-by-ai/.

Deruty E., Grachten M., Lattner, S. Nistal, J. and Aouameur, C. 2022. "On the Development and Practice of AI Technology for Contemporary Popular Music Production." Transactions of the International Society for Music Information Retreival.

Dubnov, Shlomo and Greer, Ross. 2024. *Deep and Shallow: Machine Learning in Music and Audio*. CRC Press.

Google DeepMind. 2023. "Transforming the future of music creation." November 16, 2023 https://deepmind.google/discover/blog/transforming-the-future-of-music-creation/.

Google. *n.d.* "MusicLM: Generating Music From Text." https://google-research.github.io/seanet/musiclm/examples/.

Guy, Zoe. 2023. "AI-Generated Art Is Not Copyrightable, Judge Rules." Vulture, August 21, 2023. https://www.vulture.com/2023/08/ai-art-copyright-ineligible.html.

Hit 'N' Mix. 2024. "RipX DAW – The first of its kind AI DAW." https://hitnmix.com/ripx-daw/.

Lange, Cameron. 2020. "A Study of Artificial Intelligence for Creative Uses in Music". BA Thesis. Faculty of the School of Engineering and Applied Science, University of Virginia.

Magenta. *n.d.* "Magenta Studio – Ableton Live Plugin". https://magenta.tensorflow.org/studio/.

Manning, Christopher. 2020. "Articial Intelligence Definitions". Stanford University.

Miller, Arthur I. 2019. *The Artist in the Machine: The World of AI-Powered Creativity*. MIT Press.

MusicGen. *n.d.* "MusicGen: Simple and Controllable Music Generation". https://ai.honu.io/papers/musicgen/.

Owsinsky, Bobby. 2023. *The Musician's AI Handbook: Enhance And Promote Your Music With Artificial Intelligence.* Burbank: Bobby Owsinski Media Group.

Stable Audio. 2024. "Stable audio – Generative AI for music and sound fx". https://www.stableaudio.com/.

Vincent, James. 2023. "AI art tools Stable Diffusion and Midjourney targeted with copyright lawsuit." The Verge, January 16, 2023. https://www.theverge.com/2023/1/16/23557098/generative-ai-art-copyright-legal-lawsuit-stable-diffusion-midjourney-deviantart.

Wavtool. *n.d.* "WavTool". https://wavtool.com/.

ARTIFICIAL INTELLIGENCE: DUALITY IN APPLICATIONS OF GENERATIVE AI AND ASSISTIVE AI IN MUSIC
(summary)

In the realm where technology fuses with creativity, showing the dynamic intersection of technology and artistic expression, a new wave of innovation has emerged, bringing forth a plethora of AI-powered tools designed to revolutionize the music-making process. These tools, ranging from virtual assistants to advanced composition engines, promise to augment the capabilities of musicians, offering unprecedented avenues for exploration and expression, but also raising numerous questions on ethical considerations, copyright and the duality between human and machine creation.

Artificial Intelligence has permeated various domains, especially in the last few years, finding its path and usage in fields such as information technologies, game development, graphic art and photography as well as the audio industry, including music. The applications of this powerful technology in the art of music exhibit a duality between Generative AI and Assistive AI. The two ways of its development are exemplified through platforms and software made for this purpose. Each of these platforms is contributing uniquely to the landscape of AI development in music. It is set to revolutionize the industry in a similar manner, for instance, to the beginnings of recording studios or synthesizer development.

Generative AI in music refers to systems that autonomously compose or generate musical content. For example, it utilizes deep learning algorithms to create original compositions, mimicking the style of various artists and composers or even generating entirely original music. These systems analyze vast datasets of existing music to understand patterns, harmonies and structures, enabling them to compose music.

On the other hand, the main purpose of Assistive AI in music focuses on enhancing human creativity by providing tools for composition, production and performance. This includes AI-driven software for music production, virtual instruments, mixing and mastering tools. The idea behind developing some of these tools is not new and they have existed for quite some time. For example, IRCAM (*Institut de Recherche et Coordination Acoustique Musique*) developed a vast amount of software which hold the idea of AI algorithms. It is no secret that a few famous contemporary composers used some of this software as assistive tools for composing their original works.

In addition, developers created tools with a primary focus in language modeling applied to music. These innovations display the potential of AI to understand and respond to musical input.

Moreover, there are AI applications geared towards audio processing and synthesis. This technology is of great value to audio professionals, especially audio engineers seeking to improve the fidelity of their recordings. AI also contributes to music recommendation systems, personalized playlists and adaptive music streaming platforms, enhancing user experience and engagement. Companies such as Spotify or Pandora leverage AI algorithms to analyze listening habits and preferences, curating tailored playlists for individual users.

Overall, the two-faced crossover between Generative AI and Assistive AI in music reflects a nuanced approach to leveraging this technology for creative expression and innovation, reshaping how industry professionals compose or produce music and how listeners consume the final product.

Article received: February 2, 2024
Article accepted: May 17, 2023
Review article