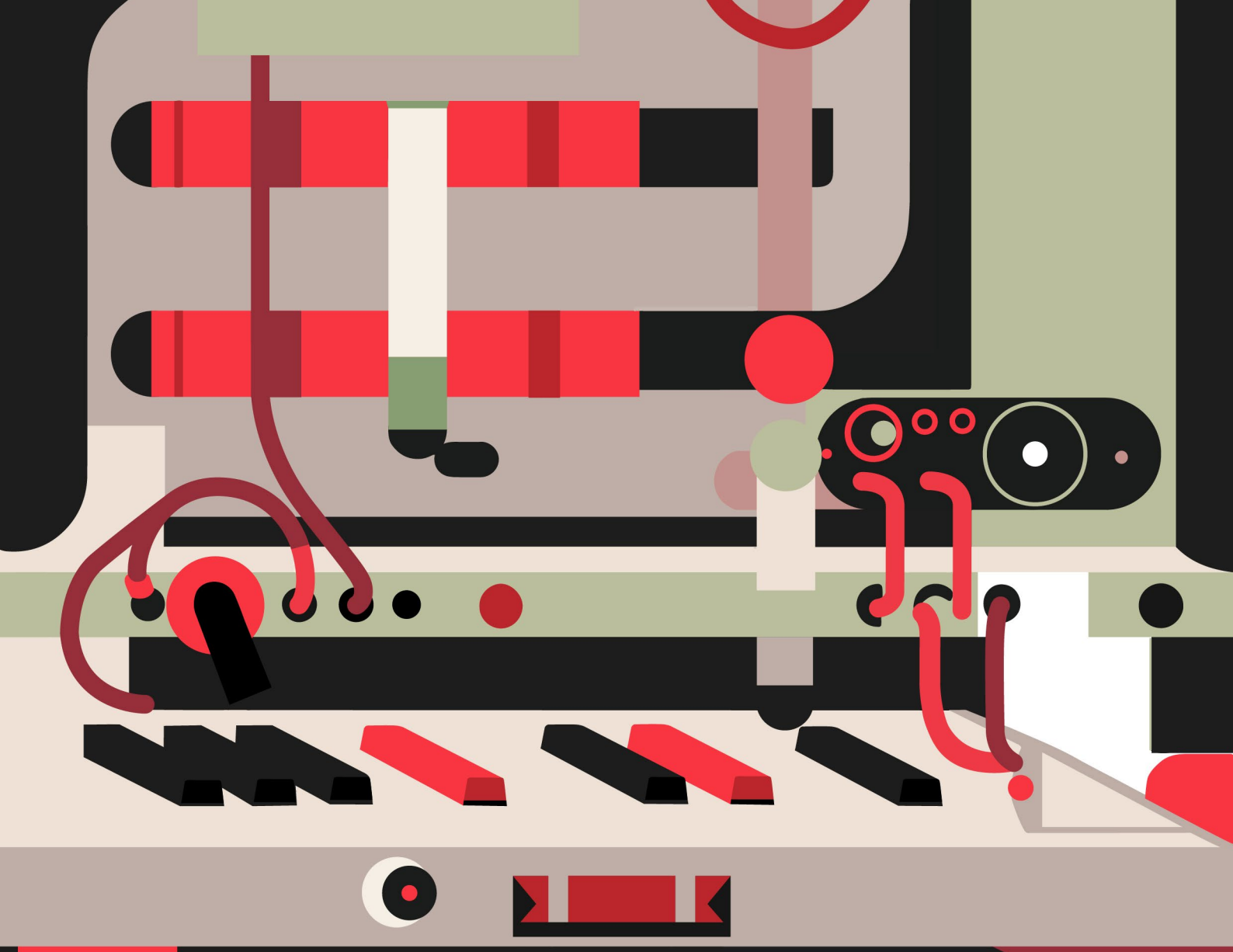


David Stammer

# Generative AI in the Music Business

POPAKADEMIEDIGITAL #6 (2nd edition) / Published by Prof. Dr. Alexander Endreß & David Stammer

The second edition of this white paper was supported by YouTube. It includes and complements new developments in the field of AI and the music business in the period from December 2023 to June 2024, as well as translations into English and French and a fresh layout.

The first edition of this white paper was first published in German in December 2023. On the one hand, it consisted of a summary of the contents of the Future Music Camp, the conference for innovations and future topics in the music industry at the Popakademie Baden-Württemberg. The Future Music Camp was held on 25 & 26 May 2023 and focused on “Creative Artificial Intelligence in the Music Business”. It is also based on the findings of two student projects at SMIX.LAB (the Popakademie Digital Competence Centre) 2023, which evaluated AI tools for music generation, content creation and process optimisation in marketing.

2nd edition/June 2024

Author:
David Stammer

Apart from Section 3.2 (AI among Music Labels):
Authors:
Lukas Kolb & David Stammer

Artwork:
Brückner + Brückner, Nils Pastor & David Stammer (Midjourney)

Layout:
Brückner + Brückner

AI toolboard & music production:
Nils Pastor & David Stammer

Published by:
Prof. Alexander Endress (alexander.endress@popakademie.de)
David Stammer, M.A. (david.stammer@popakademie.de)

Popakademie Baden-Württemberg Foundation
Hafenstraße 33
68159 Mannheim
www.popakademie.de

SMIX.LAB

POPAKADEMIE BADEN-WÜRTTEMBERG | UNIVERSITY OF POPULAR MUSIC AND MUSIC BUSINESS

# Content

# 1 Introduction: AI in the Music Business

The rapid evolution of generative artificial intelligence is transforming the music industry. As a technology, it is increasingly becoming a mainstay in the creation and marketing of music. This white paper is intended to shed light on the various ways in which generative AI is impacting the music industry.

Viewed historically, the music industry has always been shaped by technological progress, be it the introduction of vinyl records, synthesisers or digital music production. But the most recent innovation in this field – generative AI – marks a paradigm shift. Not only does it enable new forms of music production, it also brings fundamental changes to the ways in which music is shared, marketed and experienced. It comes with unimagined opportunities for creativity and personalisation and at the same time points to complex issues in the areas of copyright, authenticity and artistic integrity. As this technology continues to develop, we will be faced with the challenge of harnessing its potential while simultaneously approaching its implications in a responsible manner.

## Aim and structure of the white paper

This white paper seeks to present the material connections between generative AI and the music industry in a comprehensible form. From discussing specific fields of application to elaborating on legal and ethical issues, this document provides a comprehensive resource to anyone interested in exploring the points of intersection between AI and music (business).

Chapter 2 (Definitions) is intended to provide a basic understanding of core concepts within generative AI.

Chapter 3 outlines the impacts on four specific areas of the music industry: 3.1. AI and Music Production describes how current models for music generation work, groups them into a new system and highlights the opportunities and limitations of these programs. 3.2. AI among Music Labels indicates areas in which the technology can be used to market music, including content creation, marketing itself and process optimisation. 3.3 AI among Music Publishers explains the significance of metadata in connection with AI and raises the question of future licensing models. 3.4 AI in the Live Segment highlights initial approaches to the use of generative AI in the live industry, such as for bookings or in live performances.

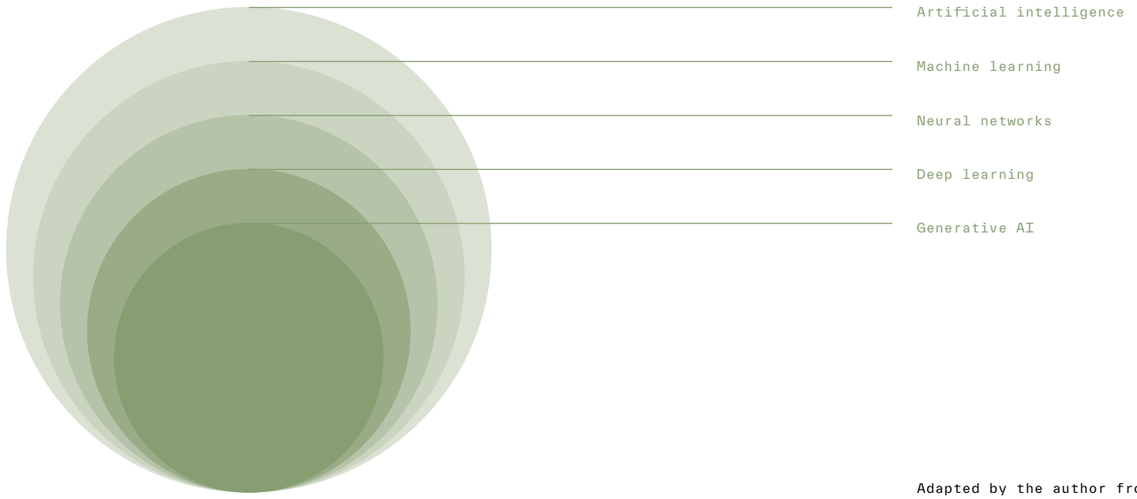
Chapter 4 outlines the most important legal and ethical issues. Aside from establishing categories within current debates, it also provides links to additional information and databases.

# 2 Definition: Generative Artificial Intelligence

## Artificial intelligence

Artificial intelligence (AI) is a broad field within computer science that deals with the creation of systems that can perform tasks that typically require human intelligence. These systems are designed to solve complex problems, make decisions and even develop creative potential by replicating human cognitive processes such as learning, deducing and perceiving. Among the AI technologies are, for example, machine learning (ML), neural networks, natural language processing (NLP) and computer vision. These technologies enable machines to learn from experience, detect patterns and reach decisions with a certain degree of autonomy.[1]

1 Lackes & Siepermann (2018)




Adapted by the author from a diagram by Purohit (2023)

## Generative artificial intelligence

A hallmark of this area of AI is the autonomous ability to generate fresh content that is barely distinct from creative human output. In this regard, the term ‘generative’ refers to the autonomous capacity of AI systems to create works, including texts, images, code or music.[2]

2 Martineau (2023)

Generative AI technology builds on a combined foundation of advanced machine learning algorithms and deep neural networks. These systems are designed to analyse and process large volumes of data and, on this basis, to identify and replicate patterns, structures and complex relationships. At the heart of this process is deep learning, which is enabled by multilayer neural networks. Each layer in these networks is specialised in the extraction and processing of certain attributes in the input data – including texts, sounds and images. A huge variety of data is used to train these networks, which then learn to detect and replicate specific aspects, enabling them to emulate human creativity in the form of speech, music or visual arts. It is this capability to detect and replicate patterns that sets generative AI apart and taps into a wide range of potential applications.[3]

3 Zewe (2023)

# 3
# Fields of Application

## 3.1 AI in Music Production

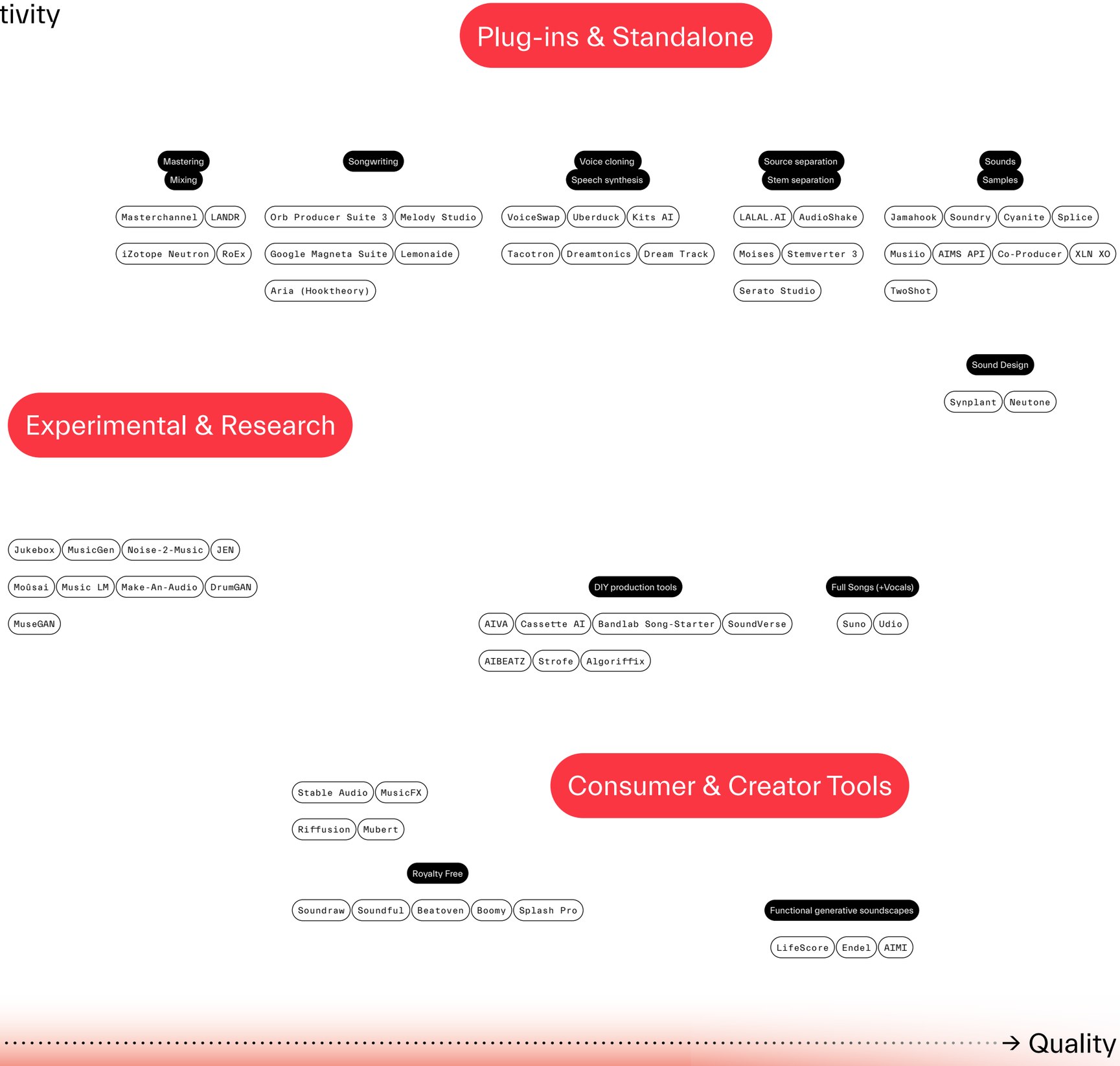
Recent years have seen significant progress in the development of AI models for music generation.[4] Initially, it was only possible to generate individual components of compositions automatically (e.g. melody, rhythm), or only music within certain genres.[5] Since then, however, some models have acquired the capability to create longer pieces of lyric-based music that consumers find difficult to distinguish from those composed by humans. What is more, there are numerous AI applications that can be used for individual stages within professional music production.[6] The deep controversy underlying this trend is reflected in the fact that 73% of copyright holders surveyed in a study by the collecting societies GEMA and SACEM believe that their livelihoods are at risk.[7]

4 Civit et al. (2023)

5 Deruty (2022)

6 Sandzer-Bell (2024)

7 Goldmedia (GEMA/SACEM) (2024)




## AI models for music generation

A uniform system to categorise the wide variety of AI music tools has not yet been established to date. Some differ according to the target audience (from consumers to professional users),[8] and others depending on the field in which they are used (e.g. for songwriting or mastering).[9]

A system that categorises all current tools based on their interactivity and output quality might provide a potential solution.

Here, interactivity is defined as the degree to which musical output can be influenced by a person, that is, an artist. This might include, for example, extensive options for text prompts in audio generation. Another hallmark of interactivity among tools is their suitability for use in a reciprocal co-creative process, for example by suggesting new compositions based on a person's own compositions or by automating a particular area of music production.

Quality describes the quality of output and can be subdivided into musical quality and production quality. High-quality music will possess a coherent structure with clear transitions, themes and motifs. It should be meaningful from a musical perspective and adhere to fundamental tenets of rhythm, harmony and melody. Production quality describes balance in the sound and audio quality.

8 Moore & Acharya (2023)

9 Edwards & McGlynn (2023)

## AI Music Toolboard

The interactive Miro board shows various AI music applications, grouped according to their intended use and positioned in relation to their interaction options and the quality of musical output. Click on them and experiment.

The three basic areas of 3.1.1 Experimental Models/Research, 3.1.2 AI-powered Plug-ins for Music Professionals and 3.1.3 Creator/Consumer Tools – which are described on the following pages – can be distinguished along these axes.

# 3.1.1 Experimental Models/ Research

This area includes models that are developed in research on AI-generated music and are publicly usable. The code is accessible, and there are research reports describing how the models are trained and evaluated.

Current applications such as MusicGen[10] or MusicLM[11], which are developed by research groups supported Meta or Google, are highly interactive: text input is used to generate music as an audio waveform. Some of the codes and libraries can be used, customised and trained with the user's own data. Although the audio quality of these models is improving steadily, some noise disturbance or artefacts may occur nonetheless. More extended passages are often substandard as well.

The models should be perceived as a foundation for building more advanced applications. But with the usual transparency of training data, detailed descriptions of the model architectures and the publicly accessible code, they do promote progress in academia and commercial projects alike.

10 Copet et al. (2023)
11 Agostinelli et al. (2023)

# 3.1.2 AI-powered Plug-ins

Ambitious hobby musicians and professional artists can all draw on artificial intelligence to assist with many areas of the music production process. Music producers can use tools that take complete or partial control of mixing and mastering, and make suggestions for melodies, chord structures or suitable loops and samples. These plug-ins generate output (such as sounds, MIDI notes) that is highly variable and suitable for professional processing. Provided the instruments are used correctly, the quality of musical output is potentially excellent.

The most important fields of application are:

Songwriting

AI supports the creation of unique melodies and chord progressions in the area of songwriting, presenting artists with fresh creative perspectives and inspiration for their musical works. Used meaningfully, these tools can establish a co-creative process between the artist and AI. Aside from the actual work on individual melodies or text components, the tools are also able to generate complete musical pieces as inspiration in the songwriting process – for instance to review how a text might work in a variety of keys and with different instrumental arrangements.

Mixing & mastering

The use of AI in the mixing and mastering process automatically adjusts levels, optimises EQ settings and improves sound quality. This can create a more efficient and consistent audio production – at least for users who are keen to save time and costs otherwise expended on human mixing/mastering and who view the specific sound produced by a human expert in mixing or mastering as less important.

Voice cloning/synthesis

Technologies used to clone and synthesise voices can be used to generate realistic, adaptable digital voices that are, in many cases, barely distinguishable from human speech. They can be used for deep fakes, among other things, and first attracted a lot of attention in the wake of the viral song Heart on my Sleeve with voices replicating those of Drake & The Weeknd.[12] By contrast, there are start-ups that legally licence artists' voices, which can then be used for demo versions, toplines or in other areas.[13]

12 Ingham (2023)
13 Dredge (2023a)

Stem separation

AI simplifies the separation of individual components (stems) from mixed audio tracks. This may include isolating vocals from instruments, which adds flexibility to the editing and remixing processes. Examples in which this technology has been used include the Beatles song Now and Then, for which the voice of John Lennon was extracted from an old cassette recording.[14] This kind of separation into individual tracks is also very helpful for DJs to create creative mixes, for example.

14 Tencer (2023)

Sounds & samples

AI applications analyse audio attributes and the preferences that users express for particular pieces of music to improve the process of searching for sounds and samples. They use this information to suggest suitable sounds and samples for musical projects. In recent months, services have been added to this segment that draw on prompts to generate individual sounds or entire sample packs. This may potentially offer a meaningful alternative to the search for suitable samples in libraries, as the samples generated can be tailored to the specific context of the music production.

Aside from the use cases listed here, an increase in the number of services that combine a selection of functionalities has also been noted. They may, for example, offer models for audio separation, voice generation and AI mastering in a single tool suite.

# 3.1.3 Creator/Consumer Tools

This segment comprises applications for the generation of functional music, initial musical ideas or complete songs based on prompt input. Functional music means pieces that are composed or selected for a specific purpose, including for commercials or as background music. While the audio quality produced by these models was initially appealing, the interactive options for creative development were largely limited. But rapid progress in this area – due in particular to the possibility of generating lyrics and vocals – has increasingly given these models the status of perceived competitors to human-composed music, especially in the streaming market.

The following four fields of use can be distinguished:

## AI music with vocals

Two models – Suno and Udio – have been introduced to the market in recent months, fuelling controversy in the discussion about AI-generated music models. These models are the first of their kind to enable the generation of complete tracks, including vocals, in various languages and genres, based entirely on prompts. They mean that people will no longer require prior knowledge of music to create (comparatively) professional sounding songs quickly and easily. And although the quality of the models varies depending on the prompt and the genre, the models are highly interactive. Users can manipulate the texts/vocals and instrumental arrangements as separate components and even add their own lyrics. Adding to this since June 2024 has been the ability for users to upload and continue working on their own pieces of music or audio recordings.

These providers have been criticised for a lack of transparency with regard to their training data, as the compositions generated in this way are viewed by the public as high quality. Given that the output produced by these models (including voices or individual melody lines) sometimes bears similarities to well-known songs, the providers have been accused of using protected material for training purposes.[15] [16]

Chapter 4 contains a more detailed exploration of legal and ethical relationships.

15 Newton-Rex (2024a)

16 Newton-Rex (2024b)

## Royalty-free (background) music

This area covers background music for content creators or companies. Short, personalised pieces or loops can be generated as an alternative to compositions from production music libraries. The majority of models do not synthesise sound artificially and instead allow users to select genres, moods, instruments or tempo. This selection influences a structure of loops or MIDI notes preconfigured in the services, which are combined once again to create unique pieces of music. Therefore, these applications can produce output that sounds monotonous, and the scope for downstream interaction is low. But there are models in this segment that operate on the basis of text prompts and are hence more interactive.

## DIY production tools

The tools in this area work in a similar way to the services described earlier. However, they also provide additional functions such as the option to export MIDI notes or edit them directly in a structure comprising a piano roll and a simple sequencer.

## Functional generative models

Another segment contains applications for functional music, e.g. to improve relaxation or concentration. Companies such as Endel and Lifescore also collaborate with artists to create continuous ‘soundscapes’ from their works. These services tend to offer users only limited interactivity. But the music generated does respond and adapt to metrics such as the weather, time of day or, if selected, the user’s heart rate.



The online toolboard included in this white paper[17] provides a visual representation of 60 tools for AI-powered music generation that were available in June 2024 and adhere to the systematic structure outlined here. Links to these tools are included in the attached database[18], along with descriptions according to the areas in which they are used and their function.

17 AI toolboard & music production (Stammer, 2023)

18 AI music database Pastor & Stammer (2023)

The toolboard and the database do not claim to be exhaustive, but are intended to outline the segment from a conceptual perspective and provide a clear presentation. For an extensive and constantly updated database, we recommend the work of the Water and Music research collective: https://www.waterandmusic.com/data/music-ai-market-tracker
Update 2nd edition: The Goldmedia study commissioned by GEMA & SACEM (2024) provides a detailed market overview, as well as surveys on this topic conducted with music creators

# 3.1.4 Spotlight: YouTube AI Tools

YouTube is taking an active role at the intersection of artificial intelligence and music and has announced a series of projects and experiments that demonstrate YouTube's approach to integrating AI into the music industry. They are working closely with artists and the music industry to develop innovative and secure solutions that enrich the creative process while ensuring the protection and integrity of artists' works.

The "Music AI Incubator" was launched by YouTube in 2023 and now includes dozens of global artists, songwriters and producers. Participants include Anitta, Björn Ulvaeus, d4vd, Don Was, Juanes, Louis Bell, Max Richter, Rodney Jerkins, Rosanne Cash, Ryan Tedder of OneRepublic, Yo Gotti, the estate of Frank Sinatra, Wyclef Jean and many others who provide insights into YouTube's generative AI experiments and research to help shape and test new creative tools.[19] Together Google DeepMind and YouTube are exploring a "Music AI Sandbox": These tools are intended to support the creative process by, for example, enabling you to sing a melody to create a trumpet sound, transform chords from a MIDI keyboard into a realistic vocal choir. Other functions include creating accompanying instrumental tracks to existing vocals or transforming audio from one musical style to another. In parallel, Google DeepMind has announced "Lyria", an advanced AI music generation model. Lyria is able to create high-quality music with instruments and vocals and offers users more control over the style and execution of the music produced. The first collaboration with major artists in this context is DreamTrack, which is a consumer and creator tool, featuring a set of AI experiments, some incorporating vocal capabilities, which has been launched in the US with nine renowned artists.[20]

19 Dredge(2023b), Mohan (2023)

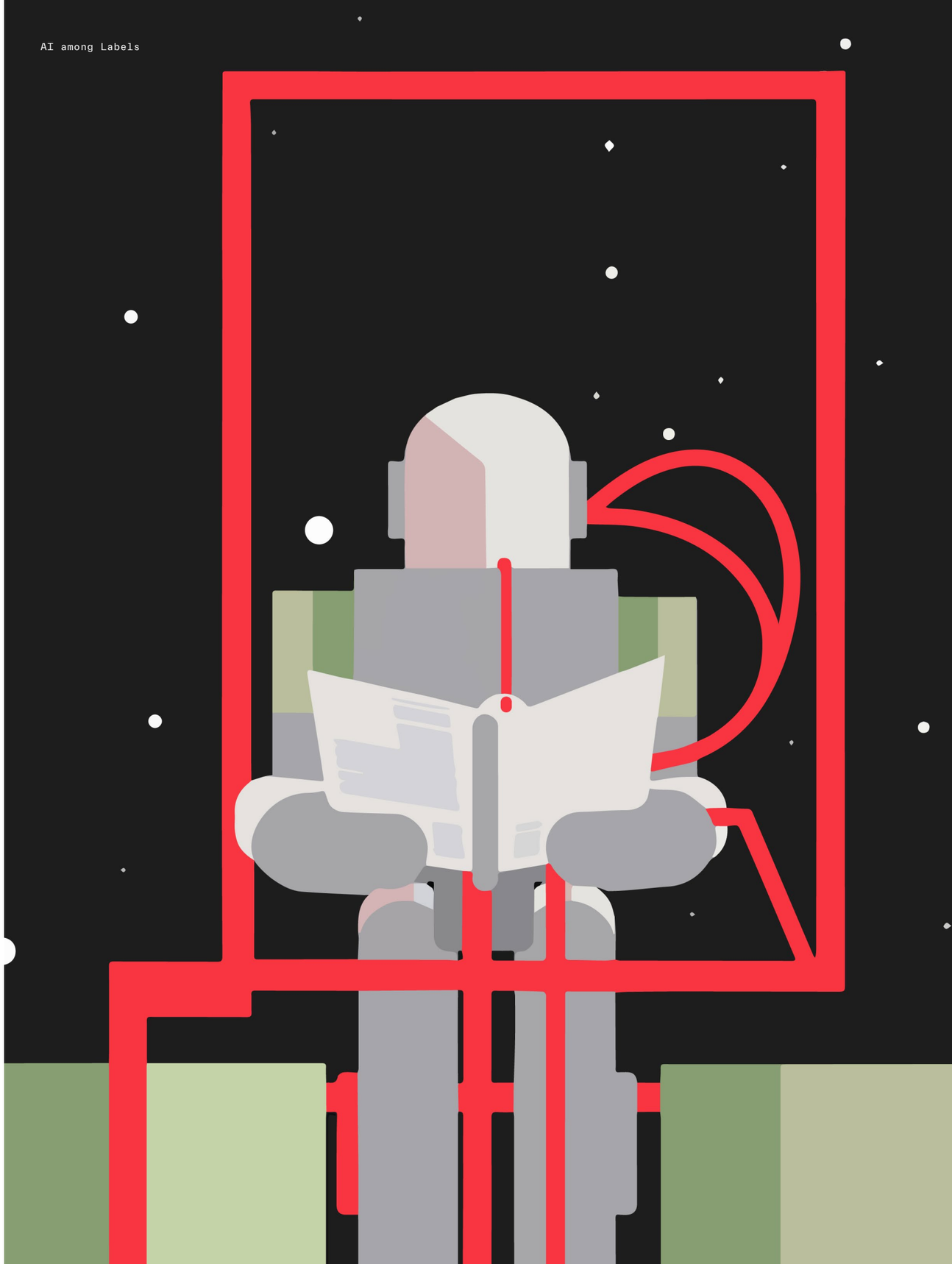
20 Google DeepMind (2023)

Another important aspect of YouTube's AI strategy is the protection of the creative work of the artists and creators. The Content ID YouTube's technology has proven to be effective in identifying songs on the platform, giving rights holders control over revenue, and maintaining a balance between the interests of creators and the creative community on the platform. YouTube plans to continue to expand this technology to address the challenges of AI-generated content and keep the platform safe - for example, to counteract copyright infringement, misinformation, and spam. One example of this is the introduction of SynthID, a technology that adds a watermark to AI-generated audio content that is inaudible to the human ear to identify its origin and prevent abuse.[21]

21 Google DeepMind (2024)







# 3.2 AI among Labels

Artificial intelligence has played a major role in the marketing of music for a long time. Use cases include data-driven scouting processes in A&R[22], or the introduction of algorithms to music recommendation systems among streaming providers.[23]

But the emergence of generative artificial intelligence has opened the door to fresh potential such as content creation, promotion and process optimisation.

22 Rauscher (2021)

23 Knees (2021)

## 3.2.1 Content Creation

Integrating AI into the content creation workflow can speed up the process and cut costs, as it enables process automation and reduces the complexity of tasks.[24]

Firstly, this has the benefit of completing tedious routine tasks more quickly, which frees up time for genuinely creative work with artists. Secondly, artificial intelligence can expand and change – not replace – the scope of human creativity.[25]

24 Rathore (2023)

25 Anantrasirichai & Bull (2022)

### Text

The release of ChatGPT in November 2022 marked a turning point for generative AI, bringing this technology into the public spotlight for the first time and enabling far-reaching interactions between AI models and users.[26] What is more, the user-friendly, conversational interface has made it easier for private individuals to draw on AI as an active resource and has led to rapid adoption with over 100 million users inside of two months after its launch. A slew of text-based tools were developed in the wake of ChatGPT's success. They are built on the APIs for large LLMs and include applications such as AI-powered copy writing tools and CRM system integration.

Labels can use these to translate content, write press releases, playlist pitches or biographies, or even generate inspiration and ideas for the content creation process. 'Custom instructions' are used to save information about the company, which is taken into account in each chat.[27] This guarantees that all generated content reflects the label's particular circumstances and style of communication. Relevant information on artists, songs or target audiences can quickly be added to this content using links to other data interfaces (e.g. Chartmetric).

Labels can exploit additional marketing potential by developing proprietary chatbots, which are trained in the language/communicative style of their contracted artists. They can be used to communicate directly with the fan base or to generate copy for social media posts or captions.

26 Berg (2023)

27 OpenAI (2023)

### Image and video

AI tools for image generation build on diffusion models and are highly versatile in many areas of marketing music. For example, image generation tools such as Midjourney (cf. Midjourney, 2023), Dall.E 2 (cf. OpenAI, 2023) or Stable Diffusion (cf. Stability AI Ltd., 2023) can be used to generate varied and, above all, unique images[28] that are suitable for use in artwork, playlist covers, social media posts or as content for websites or newsletters, to name just a few.

Even now, tools such as Gen-2 from Runway (cf. Runway AI, Inc., 2023) or Kaiber (cf. Kaiber, 2023a) can draw on complex link and adaptation processes to create short videos that are suitable for use in music videos or visualisers, social media posts or formats such as Spotify canvases. Each scene change requires a new video with a new prompt (and in some cases new reference images, which the tool draws on for image generation), which must then be linked to the videos for the other scenes. This means that producing a complete music video still involves very significant manual effort, despite the support provided by AI. But the first music video created with the Sora model presented by OpenAI impressively demonstrates the inherent potential.[29]

28 Anantrasirichai & Bull (2022)

29 Luna (2024)



## 3.2.2 Process Optimisation

Besides the possibility to generate fresh content, artificial intelligence can also facilitate or speed up processes for the creation or completion of existing content.

### Transfer

Tools that rapidly convert existing material into a variety of formats appear auspicious. For example, tools such as Opus Clip can turn a YouTube music video or behind-the-scenes content in landscape format into a portrait format clip for short video platforms. The tool automatically highlights potentially relevant points in the recording to meet the needs of specific platforms.[30]
Subtitles can also be auto-generated in various languages for international marketing activities.

Tools such as Oxolo can be used to extract simple social media content from current information such as tour dates, and also enables the auto-generation of texts or narration. Providers such as Blurb.fm add numerous functionalities to automatic content creation. To do so, they draw on a track and associated metadata to produce a range of assets for different platforms.

30 Loewe (2023)

### Research & data processing

AI-powered research has been made significantly easier by the increasing integration of search engines in chatbots such as ChatGPT. For instance, relevant media, dissemination pathways or channels can be identified as part of promotional activities.[31]

When completing AI-powered sentiment analysis, information can also be extracted from unstructured data such as comments under social media postings. Here, AI is used to detect emotions expressed in text. AI captures the underlying tone of a statement and recognises not only whether certain words within a text group have a negative or positive connotation. This can be used, for instance, to develop key content for communication and specific content or forms of address by establishing meaningful links to generative tools.[32]

31 Loewe (2023)

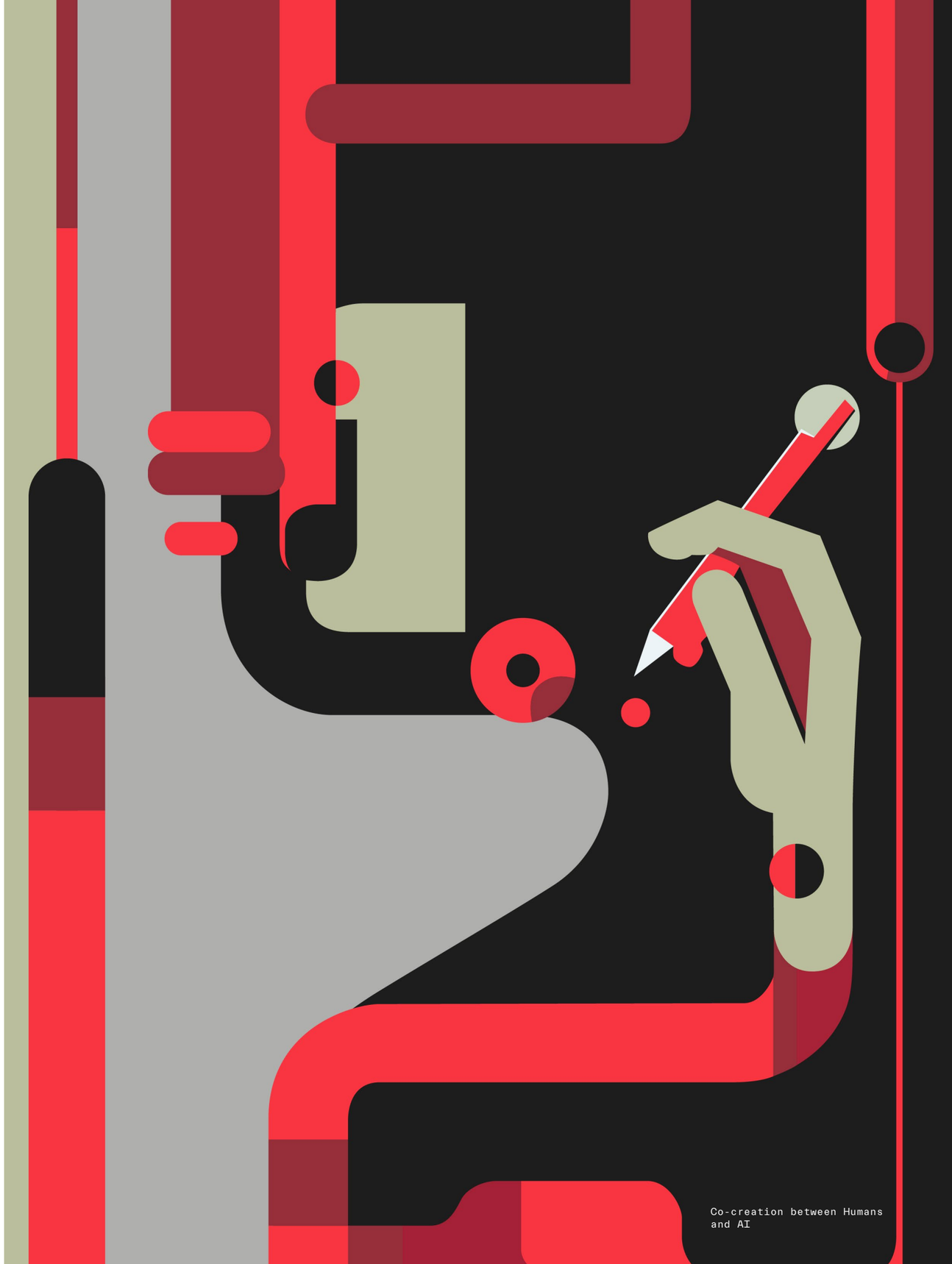
32 Taherdoost & Madanchian (2023)

### Advertisement control

The control of adverts by AI has also evolved to a significant extent to advanced data processing capabilities. Decisions tend now to be made on the basis of data rather than experience and intuition. AI helps in the optimisation of marketing budgets and

monitors campaign performance and market developments in real time to leverage the relevance, placement and quality of ads. Artificial intelligence can improve the timing of advertisements and enhance the efficiency of retargeting by analysing usage data and predicting future behaviour. AI automates the management of advertising budgets and marketing communications to enable faster adaptation to market shifts and changes in consumer behaviour, which can produce a competitive edge.[33] Despite the potential of AI for advertisement control, there is currently no application that completely integrates all aspects, although tools such as Symphony were specifically developed for the automatic control and optimisation of ads in the music industry.

33 Taherdoost & Madanchian (2023)



# 3.3 AI among Publishers

There are numerous points of intersection between AI and the work of music publishers. Firstly legal and ethical aspects (cf. Chapter 4) must be considered to protect rights holders, composers and lyricists from potential negative repercussions. The use of generative text models in songwriting and the eligibility of this content for protection are particularly noteworthy aspects in this context.

Two further areas will be outlined in the following:

## 3.3.1 The Significance of Metadata for AI

Publishers base their work on music metadata.[34] This information helps in the precise tracking and administration of music rights – a particularly important aspect for generative AI due to its capacity to generate huge volumes of content. Correct metadata ensures correct allocation and remuneration of authors when their works are used or sampled by AI. Besides the tracking of historical use, metadata is also indispensable for sorting, categorising and locating content. In this context, generative AI can be used to analyse and deliver metadata that improves the accuracy of recommendations and makes it easier for current and new works to reach their target audience. AI-powered analysis tools such as Cyanite or Musiio take charge of tagging the extensive catalogues and in doing so lay the foundation for recoverability and traceability.

34 Spohn (2021)

## 3.3.2 Licensing and AI

Royalty-free music from AI tools

The creator/consumer tools described in section 3.1 enable the generation of royalty-free music for online use such as in streams or social media content, as well as background music for corporate contexts, events or public spaces. For the production music departments in many traditional publishers, this trend presents a particular challenge as it jeopardises their very business model. Initial platforms for rights management – including Musical AI (formerly SOMMS.AI) – that manage the catalogues of rights holders and grant licences to AI companies, while simultaneously recording use and reporting back to the rights holders, are becoming increasingly important in this area.[35]

35 Musical AI (2024)

Licensing of generative & adaptive music

The licensing of generative or adaptive music that is created in real time and is therefore not covered by traditional copyright structures (cf. Functional Music, section 3.1.) is another relevant issue in this regard.



How can music publishers handle this type of music with its continuous changes that may be exempt from conventional licensing models? This issue poses not only legal, but also creative and business challenges that the industry will need to address going forward.

Licensing of AI-generated vocals

AI models used to generate vocals have come on in leaps and bounds in recent years and are becoming increasingly realistic.[36] Artists like Grimes are making their voices available or establishing start-ups that lawfully grant licenses for the voices of other artists to third parties who can use them in the creation of their own songs.[37]

Here, also, it is worth asking how music publishers are meant to respond when creative output is no longer delivered as a conventional work (lyrics and composition). Even if the vocals possess greater relevance in regard to neighbouring rights, the examples demonstrate the need to examine new forms of licensing, combined with models for lyrics (trained using a specific songwriter) and vocals.

36 Huang et al. (2022)

37 Forristal (2023) and Dredge (2023)

# 3.4 AI in the Live Segment

Generative AI has played a fairly minor role in the live segment thus far. It follows, therefore, that this chapter is somewhat exploratory and draws on current technologies to present initial insights and the author's personal ideas.

## 3.4.1 AI in Booking and Tour Management

Booking and tour management processes consist of innumerable subtasks and require a lot of communication. Within them, significant amounts of data, figures and information must be obtained from sources and then prepared for the artists and crew members involved.

Language models can provide valuable support in this context, for instance with the plug-ins integrated in ChatGPT 4. They enable the extraction and interpretation of information from URLs or PDFs, including booking documents or train tickets. Costs can then be calculated and shared more quickly. What is more, it would speed up the process of creating itineraries for specific persons and converting them into various file formats.

The GPTs and Assistants introduced by OpenAI in November 2023 will yield fresh opportunities in particular: GPTs are customised versions of ChatGPT that are built for specific purposes and are very easy to design and use. Despite the slightly more complex integration into programming interfaces (APIs) and proprietary data pools, digital 'assistants' could also be created to assist with large parts of the booking and tour management processes.

## 3.4.2 AI in Live Performances

Tools that work in a similar way to others for AI-powered mixing or mastering in music production are tentatively conceivable for the segment of live performances at concerts or festivals. The vision is to improve sound quality in real time: AI could improve the live audio experience by reducing background noise, improving sound quality and adjusting the volume to the venue and audience size.[38]

38 Frąckiewicz (2023)

AI-powered visuals can enhance the performance, for example with video generation tools to create pre-produced content (cf. section 2.2.). Going forward, it would be conceivable for the lighting and visual effects to respond to the music or the movement of artists on stage in real time. At present, these components are controlled manually by technicians who anticipate and react to the movements of the performers on stage. AI-powered systems could analyse the stage in real time, track the artists and adjust the lighting and visuals autonomously. Technicians would then be free to concentrate on other aspects of the production process.



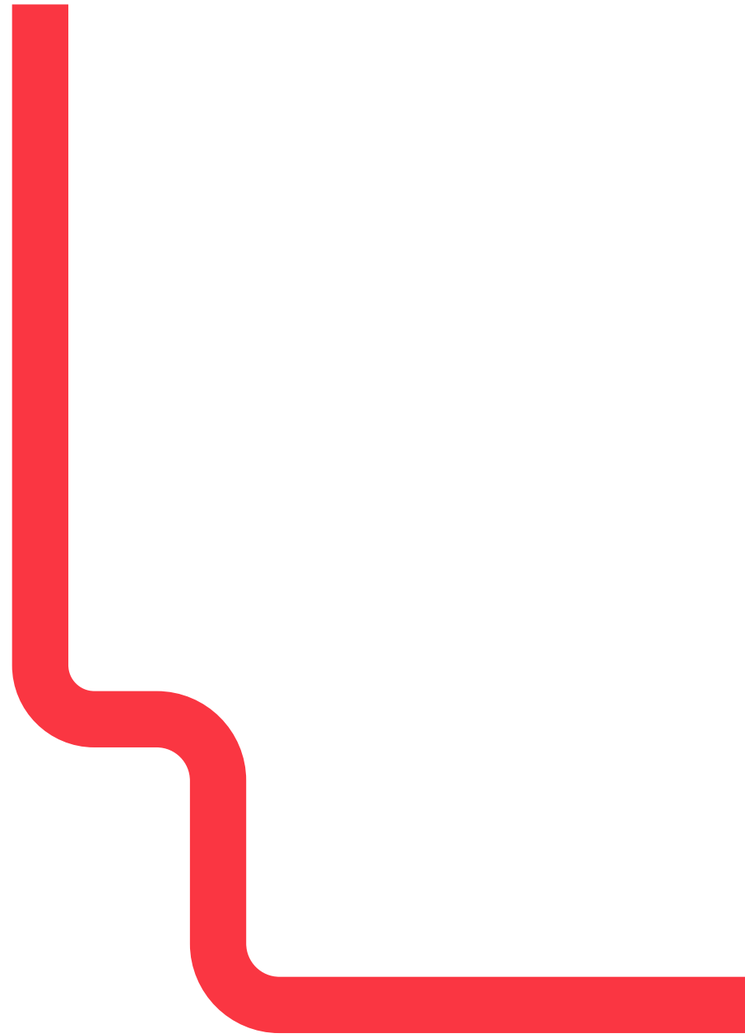
## 3.4.3 AI in Virtual Live Events

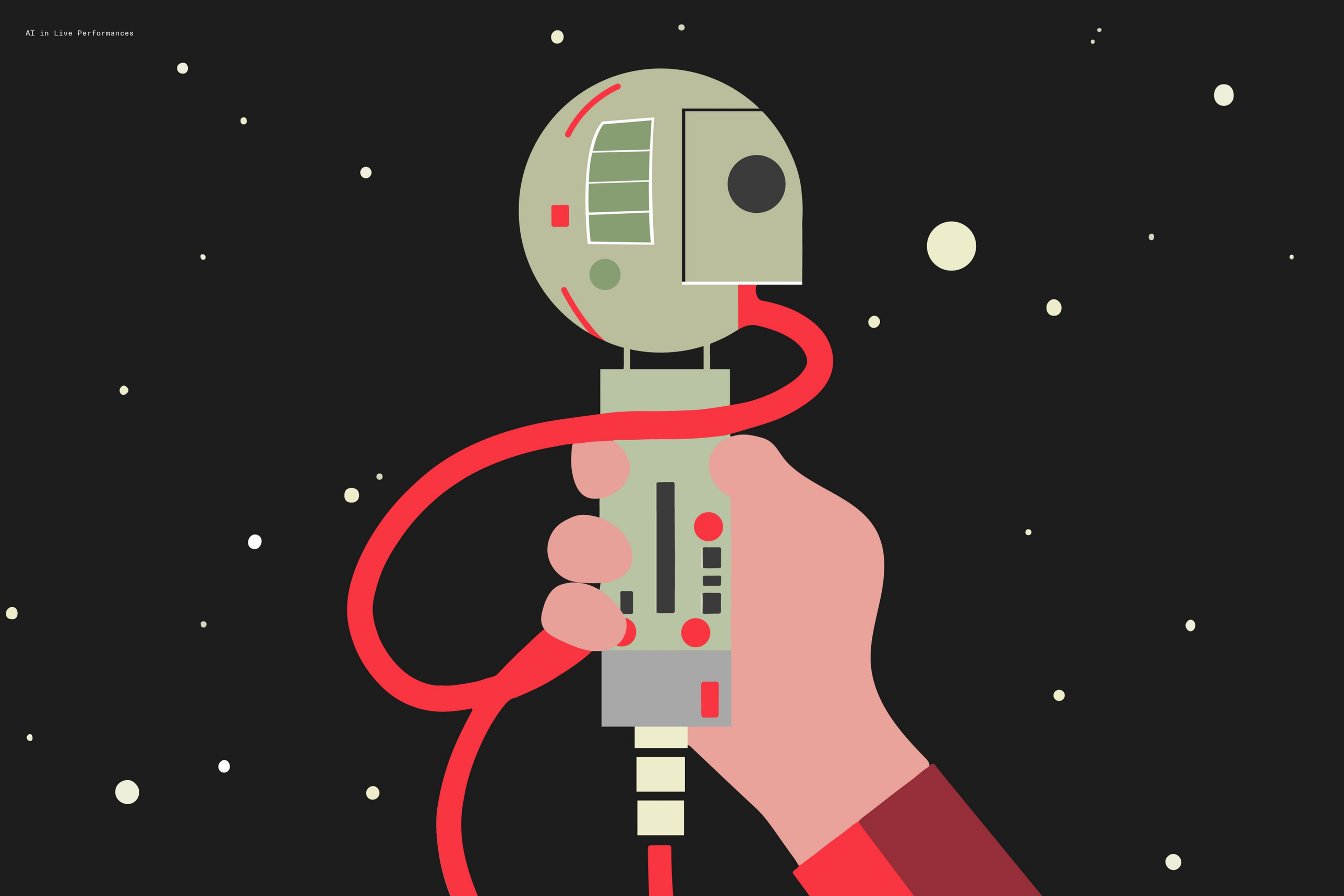
Generative AI-based image generation plays a vital role in the creation of immersive virtual worlds. These models contribute to the visual splendour and credibility of virtual environments by creating realistic and multifaceted images of landscapes, buildings and objects. For example, generative AI can create virtual venues, cities, landscapes or futuristic environments and provide users with visually enthralling, engaging and interactive virtual worlds to explore.[39]

39 Chamola et al. (2023)

With its capacity to create variations of musical compositions, generative music (cf. section 2.1) can enrich the user experience by providing a dynamic and adaptable auditory backdrop. It can be programmed to respond to interactions, changing environments or certain events within a virtual world. For instance, the music can seamlessly change tempo, mood or genre to reflect and reinforce the atmosphere of the virtual environment as users traverse different landscapes or engage in a variety of activities.[40]

40 Jin et al. (2022)





# 4

# Normative Considerations

## 4.1 Generative AI and Copyright

While the discussion of how AI-powered music will affect copyright is not new, it is becoming increasingly pressing as the models and their musical output become increasingly advanced, blurring the boundaries between human creativity and machine generation.[41]

41 Sturm et al. (2019)

Legal situation

To qualify for protection, the Copyright Act (UrhG) requires a “personal intellectual creation”, which demands that the author – as a human – engages in creative activity. According to prevailing legal opinion, machine outputs – generated for instance by artificial intelligence – are not eligible for protection. Although technology can be used as a tool in the creative process, the eligibility for protection of creative output that is shaped materially by technical means – as is the case with AI – is the subject of controversial debate. The assessment will essentially hinge on whether the human contribution to the creative output is viewed as sufficient to award copyright to a person.[42]

42 Kitzberger (2023)

The question of whether the software developers are entitled to copyrights to the works generated by this software depends heavily on the extent to which the programmers influence the end product. If the software is tasked with reaching autonomous decisions on materials aspects of the creative output, this would reduce the qualification of programmers to obtain copyright. At the same time, ancillary copyrights, particularly in the field of music, are becoming increasingly important, especially to classify the copyright status of output generated by AI programs and the input used in the process. The rights of the original authors remain in force whenever an AI program reproduces or transforms copyright-protected works.[43] There have been consistent rulings in this regard, including when comparing between US and European case law.[44]

43 Fierdag (2023)

44 Hembt (2023)

Demands from associations/stakeholder groups

In view of the importance of the topic and out of concern that precedents might be established, various international associations have published principles/guidelines on generative artificial intelligence (including the Human Artistry Campaign, UK Music, Independent Music Publishers International Forum).

The following points feature most frequently in publications on AI guidelines and principles:

- Permission: Permission must be obtained from rights holders if their music is used for training purposes.
- Transparency/traceability: The issue of which works are used in the training process must be traceable.
- Compulsory labelling: AI-generated music must be labelled as such.
- Purely AI-generated works: Works that were created without a human creative process should not be afforded copyright protection.
- Personal rights of the artists & songwriters: They must be upheld, especially in the context of deep fakes in the area of voice generation.

The growing importance and complexity of this topic will required dynamic legal debate across many layers to create clear framework conditions for the use of AI and to ensure effective protection for authors. This is made clear by the EU AI Act:

“Generative AI, like ChatGPT, [...] will have to comply with transparency requirements:

- Disclosing that the content was generated by AI
- Designing the model to prevent it from generating illegal content
- Publishing summaries of copyrighted data used for training”[45]

45 European Parliament (2023)

## Response from the music industry

A recent study on training datasets used in automatic music generation highlights the pressing relevance of these requirements. It revealed that almost half of all datasets contained music available online that was used without obtaining the permission from the artists.[46] The study investigated scientific models that reveal basic transparency with regard to their datasets.

46 Morreale, Sharma & Wie (2023)

The situation in the commercial segment is different, which is increasingly prompting a response to this development from the music industry. Sony Music was the first to respond, sending a letter to 700 companies in the AI sector stating that their works (from both the label and publishing side) must not be used for training AI models.[47]

47 Dredge (2024)

This was followed by copyright lawsuits filed by the Recording Industry Association of America (RIAA) against Udio and Suno, currently the most prominent providers as described in section 3.1. The lawsuits relate to the allegation that both companies failed to obtain permission before using copyrighted music to train their AI models. RIAA argues that this does not constitute ‘fair use’ and that the generated content is in direct competition with the original works, thereby inflicting significant harm on the music industry.[48]

48 Tencer (2024)

# 4.2 Generative AI and Ethics

Aside from the critical issues under copyright law, there are other ethical questions to be considered in connection with generative music.

A recent study of models for music and language generation reveals only 10% mention potentially negative impacts. It follows that many developers seem barely aware of this aspect.[49] The following areas are mentioned in addition to the points listed earlier:

49 Barnett (2023)

## Biases

Research in the field of music information retrieval primarily takes place in the USA and Europe, so many of the datasets used to train music models tend to be western.[50] This might potentially impact the output of AI music models, such as retaining the circle of fifths and the distinction between major and minor keys, even if doing so is entirely unintentional.[51] This is a fundamental problem in machine learning, where models often inherit biases from their training data.

50 Holzapfel, Sturm & Coeckelbergh (2018)

51 Zhao et al. (2022)

## Cultural appropriation

By contrast, generative audio models – which can draw on extensive and diverse data sets – run the risk of cultural appropriation, especially if the data contains content from marginalised cultures. Not only does the music generated reflect the inherent biases in the training data, it might furthermore inadvertently appropriate cultural elements without properly understanding or engaging with the communities in question.[52] This highlights the need for a careful consideration of how training data is composed to avoid perpetuating biases and cultural insensitivity.

52 Agostinelli et al. (2023)

## Fairness in training data

The example of music recommendation systems aptly demonstrates problems in relation to input and output data as well. To a certain degree at least, users of streaming services will consume what is suggested to them. This usage data is then fed back into the recommendation algorithms as input, which then serve up new but similar tracks as output. These tracks benefit from heavier rotation.

A feedback loop like this shows how algorithms can perpetuate or amplify prevailing inequalities in datasets. This becomes clear in the example of gender representation in music recommendations.[53]

53 Ferraro, Serra & Bauer (2021)

## AI and sustainability

Resource consumption among generative models is attracting increasing attention. For example, 700,000 litres of fresh water were used to train GPT3 in Microsoft data centres (by cooling systems and for the necessary power generation).[54]

54 Li et al. (2023)

A number of proposals have been submitted to counter this trend:

- to improve current models instead of training new, proprietary ones
- to use more energy-efficient computing models
- to use AI models only when they yield significant value



In the field of music generation specifically, researchers propose that a music model's energy balance should also be considered when assessing its quality.[55]

55 Douwes, Esling & Briot (2021)

The Music AI Ethics Tracker from the Water & Music research network is a good place to start for a more detailed examination of ethical and legal issues as well as political demands from the music industry.[56] We recommend publications by the Montreal AI Ethic Institute for additional sources and questions of ethical relevance: https://montrealethics.ai/

56 Hu (2023)

# 5 Summary

This white paper has shed light on the multifaceted influences of generative artificial intelligence (AI) on the music industry. It has become clear that generative AI is far more than just a technological tool, rather a paradigm shift in music production, marketing and how it is experienced.

The paper also presented the core tenets of generative AI, from its technical fundamentals to specific use cases. It has made evident that generative AI in music production covers a range of different functions and areas of application, from experimental modelling tools to professional plug-ins and consumer/creator tools. There is a broad range of potential applications for generative AI among music labels and publishers, from content creation to process optimisation. Although still at a nascent stage, the potential within the live segment is growing as well.

Generative AI may have capacity to turn music industry on its head. It offers fresh opportunities for creativity and personalisation, but presents the industry with complex challenges as well: the debate about copyright in AI-generated music and the ethical implications of distorted training data and cultural appropriation demonstrate that the roll-out of generative AI in the music industry requires careful and nuanced consideration. Evidently, it will be indispensable to ensure conscientious and sustainable use of AI in the music industry.

In conclusion, it is reasonable to say that generative AI represents a technological as well as a cultural revolution. It challenges us to ponder the role of technology in creative work and demands a fresh mindset in regard to copyright, ethics and the relationship between humans and machines in art.

For all the advances in technology, the focus must be placed on retaining human expression, creativity and emotion in music. Hence, generative AI must always be used to support and enhance – but never to replace – human creativity. The transformation brought about by generative AI should therefore take place in close collaboration with music creators. Their active participation is indispensable to ensuring the meaningful integration of AI into creative and marketing processes. This white paper is intended as our contribution to understanding and appraising this topic from a critical perspective.

Sources


Agostinelli, Andrea, Timo I. Denk, Zalán Borsos, Jesse Engel, Mauro Verzetti, Antoine Caillon, Qingqing Huang, u. a. „MusicLM: Generating Music From Text“. arXiv, 26 January 2023. http://arxiv.org/abs/2301.11325.

Anantrasirichai, Nantheera, und David Bull. „Artificial Intelligence in the Creative Industries: A Review“. Artificial Intelligence Review 55, Nr. 1 (1 January 2022): 589–656. https://doi.org/10.1007/s10462-021-10039-7.

Andres Ferraro, Xavier Serra, & Christine Bauer (2021). Break the loop: gender imbalance in music recommenders. Proceedings of the 6th ACM SIGIR Conference on Human Information Interaction and Retrieval (CHIIR '21). Canberra, ACT, Australia, 14–19 March, pp 249–254. DOI: 10.1145/3406522.3446033

Barnett, Julia. „The Ethical Implications of Generative Audio Models: A Systematic Literature Review“. In Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society, 146–61. Montr\'{e}al QC Canada: ACM, 2023. https://doi.org/10.1145/3600211.3604686.

Berg, Janine, David Bescond, Pawel Gmyrek, und International Labour Organization. Research Department,. Generative AI and Jobs: A Global Analysis of Potential Effects on Job Quantity and Quality. Geneva: ILO, 2023. https://doi.org/10.54394/FHEM8239.

Chamola, Vinay, Gaurang Bansal, Tridib Kumar Das, Jiacheng Wang, Rihcard Yu, Sherali Zeadally, Amir Hussain, u. a. „Beyond Reality: The Pivotal Role of Generative AI in the Metaverse“. arXiv. Accessed on 29 November 2023. https://ar5iv.labs.arxiv.org/html/2308.06272.

Copet, Jade, Felix Kreuk, Itai Gat, Tal Remez, David Kant, Gabriel Synnaeve, Yossi Adi, und Alexandre Défossez. „Simple and Controllable Music Generation“. arXiv, 8 June 2023. https://doi.org/10.48550/arXiv.2306.05284.

Daniel Tencer. „The Beatles' Last Song 'Now and Then' to Be Released November 2“. Music Business Worldwide, 26 October 2023. https://www.musicbusinessworldwide.com/beatles-song-now-and-then-to-be-released-november-2/.

Deruty, Emmanuel, Maarten Grachten, Stefan Lattner, Javier Nistal, und Cyran Aouameur. „On the Development and Practice of AI Technology for Contemporary Popular Music Production“ 5, Nr. 1 (7 February 2022): 35–49. https://doi.org/10.5334/tismir.100.

Douwes, Constance, Philippe Esling, und Jean-Pierre Briot. „Energy Consumption of Deep Generative Audio Models“. arXiv, 13 October 2021. https://doi.org/10.48550/arXiv.2107.02621.

Dredge, Stuart. „Beat Check: Music Rightsholders Reach Their ‚or What?' Moment with AI“. Music Ally, 16 May 2024. http://musically.com/2024/05/16/beat-check-music-rightsholders-reach-their-or-what-moment-with-ai-2/.

Dredge, Stuart (a). „DJ Fresh Enlists Artists for New Voice-Swapping AI Service“. Music Ally, 6 July 2023. https://musically.com/2023/07/06/dj-fresh-enlists-artists-for-new-voice-swapping-ai-service/.

Dredge, Stuart (b). „YouTube Launches 'Music AI Incubator' with UMG on Board“. Music Ally, 21 August 2023. https://musically.com/2023/08/21/youtube-launches-music-ai-incubator-with-umg-on-board/.

Edwards, Dave, und Declan McGlynn. „Creative AI for Artists: Track 80+ Tools - Water & Music“, 20 January 2023. https://www.waterandmusic.com/data/creative-ai-for-artists/.

Fierdag, Prof. Dr. Hanno. „Future Music Camp 2023: Künstliche Intelligenz & Urheberrecht“. Future Music Camp, Mannheim, 26 June 2023. https://www.youtube.com/watch?v=jJDC9dRQtDY.

Forristal, Lauren. „CreateSafe, the Company behind Grimes' Voice Cloning Tool, Launches New AI Tools“. TechCrunch (blog), 16 November 2023. https://techcrunch.com/2023/11/16/createsafe-the-company-behind-grimes-voice-cloning-tool-launches-new-ai-tools/.

Frąckiewicz, Marcin. „How AI Is Revolutionizing the World of Concert Production and Management“. TS2 SPACE (blog), 1 May 2023. https://ts2.space/en/how-ai-is-revolutionizing-the-world-of-concert-production-and-management/.

Goldmedia (GEMA / SACEM). „Studie: KI und Musik“. gema.de, 2024. https://www.gema.de/de/aktuelles/ki-studie.

Google DeepMind. „SynthID“, 14 May 2024. https://deepmind.google/technologies/synthid/.

Google DeepMind. „Transforming the Future of Music Creation“. Google DeepMind, 16 November 2023. https://deepmind.google/discover/blog/transforming-the-future-of-music-creation/.

Hembt, Simon. „The Debate on AI-Created Art and Copyright Protection“. Accessed on 30 November 2023. https://www.twobirds.com/en/insights/2023/global/the-debate-on-ai-created-art-and-copyright-protection.

Holzapfel, Andre, Bob L. Sturm, und Mark Coeckelbergh. „Ethical Dimensions of Music Information Retrieval Technology“. Transactions of the International Society for Music Information Retrieval 1, Nr. 1 (21 September 2018): 44–55. https://doi.org/10.5334/tismir.13.

Huang, Rongjie, Chenye Cui, FeiYang cHEN, Yi Ren, Jinglin Liu, Zhou Zhao, Baoxing Huai, und Zhefeng Wang. „SingGAN: Generative Adversarial Network For High-Fidelity Singing Voice Generation“. In Proceedings of the 30th ACM International Conference on Multimedia, 2525–35. MM '22. New York, NY, USA: Association for Computing Machinery, 2022. https://doi.org/10.1145/3503161.3547854.

Ingham, Tim. „Universal Music Group Responds to 'Fake Drake' AI Track: Streaming Platforms Have 'a Fundamental Responsibility to Prevent the Use of Their Services in Ways That Harm Artists'“. Music Business Worldwide, 17 April 2023. https://www.musicbusinessworldwide.com/universal-music-group-responds-to-fake-drake-ai-track-streaming-platforms-have-a-fundamental-responsibility/.

Isakova, Tatsiana. „Ad Optimization Spend with AI: Big Overview“. InData Labs, 25 May 2023. https://indatalabs.com/blog/ad-spend-optimization-with-ai.

Jin, Cong, Fengjuan Wu, Jing Wang, Yang Liu, Zixuan Guan, und Zhe Han. „MetaMGC: a music generation framework for concerts in metaverse“. EURASIP Journal on Audio, Speech, and Music Processing 2022, Nr. 1 (13 December 2022): 31. https://doi.org/10.1186/s13636-022-00261-8.





Kitzberger, Ralf. „KI und Urheberrecht: Das ist die aktuelle Lage in Deutschland - Backstage PRO“, 2023. https://www.backstagepro.de/thema/ki-und-urheberrecht-das-ist-die-aktuelle-lage-in-deutschland-2023-04-28-RPKHDmW0s6.

Knees, Peter. „Datengestützte Empfehlungssysteme - Kuratiertes Musikangebot“. In Musikwirtschaft im Zeitalter der Digitalisierung: Handbuch für Wissenschaft und Praxis, herausgegeben von Alexander Endreß und Hubert Wandjo, 1. Auflage., 403–4018. Nomos Handbuch. Baden-Baden: Nomos, 2021.

Lackes, Richard, und Markus Siepermann. „Definition Künstliche Intelligenz (KI)“. Text. https://wirtschaftslexikon.gabler.de/definition/kuenstliche-intelligenz-ki-40285. Springer Fachmedien Wiesbaden GmbH. Accessed on 10 November 2023. https://wirtschaftslexikon.gabler.de/definition/kuenstliche-intelligenz-ki-40285.

Li, Pengfei, Jianyi Yang, Mohammad A. Islam, und Shaolei Ren. „Making AI Less ‚Thirsty': Uncovering and Addressing the Secret Water Footprint of AI Models“. arXiv, 29 October 2023. http://arxiv.org/abs/2304.03271.

Loewe, Julian T. „Future Music Camp 2023: Kreative KI im Label“. Mannheim, 26 June 2023. https://www.youtube.com/watch?v=gQlzeG2J-MI.

Luna, Elizabeth de. „Watch the First Major Music Video Generated by OpenAI's Sora“. Mashable, 3 May 2024. https://mashable.com/article/ai-music-videos.

Martineau, Kim. „What Is Generative AI?“ IBM Research, 2023. https://research.ibm.com/blog/what-is-generative-AI.

Miguel Civit, Javier Civit-Masot, Francisco Cuadrado, und María José Escalona. „A systematic review of artificial intelligence-based music generation: Scope, applications, and future trends“. Expert Systems With Applications 209 (1 July 2022): 118190-118190. https://doi.org/10.1016/j.eswa.2022.118190.

Mohan, Neil. „Our Principles for Partnering with the Music Industry on AI Technology“. YouTube Blog. Partnering with the music industry on AI, 21 August 2023. https://blog.youtube/inside-youtube/partnering-with-the-music-industry-on-ai/.

Moore, Justine, und Anish Acharya. „The Future of Music: How Generative AI Is Transforming the Music Industry“. Andreessen Horowitz, 9 November 2023. https://a16z.com/the-future-of-music-how-generative-ai-is-transforming-the-music-industry/.

Morreale, Fabio, Megha Sharma, und I.-Chieh Wei. „Data Collection in Music Generation Training Sets: A Critical Analysis“, 2023. https://researchspace.auckland.ac.nz/handle/2292/65322.

Musical AI. „Musical AI - Rights Management Platform for AI“, 2024. https://www.wearemusical.ai/.

Newton-Rex, Ed. „Suno Is a Music AI Company Aiming to Generate $120 Billion per Year. But Is It Trained on Copyrighted Recordings?“ Music Business Worldwide, 2 April 2024. https://www.musicbusinessworldwide.com/suno-is-a-music-ai-company-aiming-to-generate-120-billion-per-year-newton-rex/.

———. „Yes… Udio's Output Resembles Copyrighted Music, Too.“ Music Business Worldwide, 18 April 2024. https://www.musicbusinessworldwide.com/yes-udios-output-resembles-copyrighted-music-too/.

OpenAI. „Custom Instructions for ChatGPT“, 2023. https://openai.com/blog/custom-instructions-for-chatgpt?mkt_tok=NjczLVBI-Sy05NDgAAAGN4RwAUP-T0enbUjnZiSwlTMfZtQuuU1VGSzuwEdiMZlTGrTHNPj4NV936Sxm4OXxWUcjnvGAesfTt00td2hIEQrinVjrb-TidZNWmxM1G1sQ.

Purohit, Anish. „AI, ML, DL, and Generative AI Face Off: A Comparative Analysis“. Synoptek (blog). Accessed on 16 December 2023. https://synoptek.com/insights/it-blogs/data-insights/ai-ml-dl-and-generative-ai-faceoff-a-comparative-analysis/.

Rathore, Dr Bharati. „Future of AI & Generation Alpha: ChatGPT beyond Boundaries“. Eduzone: International Peer Reviewed/Refereed Multidisciplinary Journal 12, Nr. 1 (10 February 2023): 63–68.

Rauscher, Ryan. „Data-driven Scouting“. In Musikwirtschaft im Zeitalter der Digitalisierung: Handbuch für Wissenschaft und Praxis, herausgegeben von Alexander Endreß und Hubert Wandjo, 1. Auflage., 403–4018. Nomos Handbuch. Baden-Baden: Nomos, 2021.

Sandzer-Bell, Ezra. „Audio to Audio AI: Melody-to-Song, Style Transfer & More“. AudioCipher, 29 May 2024. https://www.audiocipher.com/post/audio-to-audio.

Spohn, Kolja. „Wie Metadaten zur Wertschöpfung beitragen“. In Musikwirtschaft im Zeitalter der Digitalisierung: Handbuch für Wissenschaft und Praxis, herausgegeben von Alexander Endreß und Hubert Wandjo, 2021.

Sturm, Bob L., Maria Iglesias, Oded Ben-Tal, Marius Miron, und Emilia Gómez. „Artificial intelligence and music : open questions of copyright law and engineering praxis“. The Artist and Journal of Home Culture 8, Nr. 3 (6 September 2019): 115. https://doi.org/10.3390/arts8030115.

Taherdoost, Hamed, und Mitra Madanchian. „Artificial Intelligence and Sentiment Analysis: A Review in Competitive Research“. Computers 12, Nr. 2 (February 2023): 37. https://doi.org/10.3390/computers12020037.

Tencer, Daniel. „Major Record Companies Sue Suno, Udio for 'Mass Infringement' of Copyright“. Music Business Worldwide, 24 June 2024. https://www.musicbusinessworldwide.com/major-record-companies-sue-ai-music-generators-suno-udio-for-mass-infringement-of-copyright/.

Zewe, Adam. „Explained: Generative AI“. MIT News | Massachusetts Institute of Technology, 9 November 2023. https://news.mit.edu/2023/explained-generative-ai-1109.

Zhao, Yizhou, Liang Qiu, Wensi Ai, Feng Shi, und Song-Chun Zhu. „Vertical-Horizontal Structured Attention for Generating Music with Chords“. arXiv, 17 November 2020. http://arxiv.org/abs/2011.09078.