

David Stammer

# Generative KI im Musikbusiness

Die zweite Ausgabe dieses Whitepapers wurde **unterstützt durch YouTube**. Sie beinhaltet und ergänzt neue Entwicklungen im Bereich KI und Musikbusiness im Zeitraum Dezember 2023 – Juni 2024, Übersetzungen in Englisch und Französisch, sowie ein neues Layout.

Die erste Ausgabe des Whitepapers in deutscher Sprache ist im Dezember 2023 erschienen. Es bestand einerseits aus einer Zusammenfassung der Inhalte des Future Music Camp, der Konferenz für Innovationen und Zukunftsthemen der Musikbranche an der Popakademie Baden-Württemberg. Das Future Music Camp fand am 25. & 26. Mai 2023 unter dem Schwerpunktthema „Kreative Künstliche Intelligenz im Musikbusiness“ statt. Andererseits basiert es auf Erkenntnissen von zwei studentischen Projekten im SMIX.LAB (Digitales Kompetenzzentrum der Popakademie) 2023, die KI-Werkzeuge zur Musikgeneration, Content-Erstellung und Prozessoptimierung im Marketing evaluierten.

2. Ausgabe / Juni, 2024

Autor:  
David Stammer

Ausgenommen Kapitel 3.2 (KI im Musiklabel):  
Autoren:  
Lukas Kolb & David Stammer

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

Layout:  
Brückner + Brückner

Toolboard KI & Musikproduktion:  
Nils Pastor & David Stammer

Herausgegeben von:  
Prof. Dr. Alexander Endreß (alexander.endress@popakademie.de)  
David Stammer, M.A. (david.stammer@popakademie.de)

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

**SMIX.LAB**

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

# Inhalt

1 Einleitung: KI im Musikbusiness	04
2 Definition: Generative Künstliche Intelligenz	06
3 Anwendungsfelder	08
3.1 KI in der Musikproduktion	09
3.1.1 Experimentelle Modelle / Forschung	12
3.1.2 KI-gestützte Plug-Ins	12
3.1.3 Creator-/Consumer-Tools	13
3.1.4 Spotlight: YouTube KI-Tools	16
3.2 KI im Label	18
3.2.1 Content-Erstellung	18
3.2.2 Prozessoptimierung	19
3.3 KI im Verlag	22
3.3.1 Die Bedeutung von Metadaten für KI	22
3.3.2 Lizenzierung und KI	22
3.4 KI im Live-Bereich	26
3.4.1 KI in Booking und Tourmanagement	26
3.4.2 KI bei der Live-Performance	26
3.4.3 KI bei virtuellen Live-Events	27
4 Normative Betrachtungen	30
4.1 Generative KI und Urheberrecht	31
4.2 Generative KI und Ethik	34
5 Fazit	38

# 1 Einleitung: KI im Musikbusiness

Die Musikindustrie ist im Wandel, geprägt durch die rapide Entwicklung der generativen Künstlichen Intelligenz (KI). Diese Technologie entwickelt sich mehr und mehr zu einem wichtigen Bestandteil des Musikschaffens und der Musikvermarktung. Das vorliegende Whitepaper widmet sich der Aufgabe, die vielfältigen Auswirkungen der generativen KI auf die Musikbranche zu beleuchten.

Historisch gesehen wurde die Musikindustrie schon immer von technologischen Fortschritten geformt, sei es durch die Einführung der Schallplatte, des Synthesizers oder der digitalen Musikproduktion. Die jüngste Innovation in diesem Bereich – die generative KI – stellt jedoch einen Paradigmenwechsel dar. Sie ermöglicht nicht nur neue Formen der Musikproduktion, sondern verändert auch grundlegend, wie Musik geteilt, vermarktet und erlebt wird. Sie bietet ungeahnte Möglichkeiten für Kreativität und Personalisierung und eröffnet zugleich komplexe Fragestellungen in Bezug auf Urheberrecht, Authentizität und künstlerische Integrität. Mit der fortschreitenden Entwicklung dieser Technologie stehen wir vor der Herausforderung, ihre Potenziale zu nutzen und gleichzeitig verantwortungsvoll mit ihren Implikationen umzugehen.

## Ziel und Struktur des Whitepapers

Dieses Whitepaper versucht die wesentlichen Zusammenhänge zwischen generativer KI und der Musikbranche einfach verständlich darzustellen. Von der Erörterung spezifischer Einsatzgebiete bis hin zur Diskussion rechtlicher sowie ethischer Fragestellungen bietet dieses Dokument eine umfassende Ressource für alle, die sich mit der Schnittstelle von KI und Musik(-business) auseinandersetzen möchten.

**Kapitel 2 (Definitionen)** zielt darauf ab, ein grundlegendes Verständnis der Kernkonzepte generativer KI zu vermitteln.

**Kapitel 3** beschreibt die Auswirkungen auf vier Teilbereiche der Musikbranche: **3.1. KI und Musikproduktion** beschreibt die Funktionsweise der aktuellen Modelle zur Musikgenerierung, ordnet sie in einer neuen Systematik und macht Möglichkeiten und Grenzen dieser Programme deutlich. **3.2. KI im Musiklabel** zeigt Anwendungsfelder in der Vermarktung von Musik wie Content-Erstellung, Marketing und Prozessoptimierung auf. **3.3 KI im Musikverlag** erläutert die Bedeutung von Metadaten im Zusammenhang mit KI und stellt die Frage nach zukünftigen Lizenzierungsmodellen. **3.4 KI im Live-Bereich** beleuchtet erste Ansätze zum Einsatz von generativer KI in der Live-Industrie, wie im Booking, oder der Live-Performance.

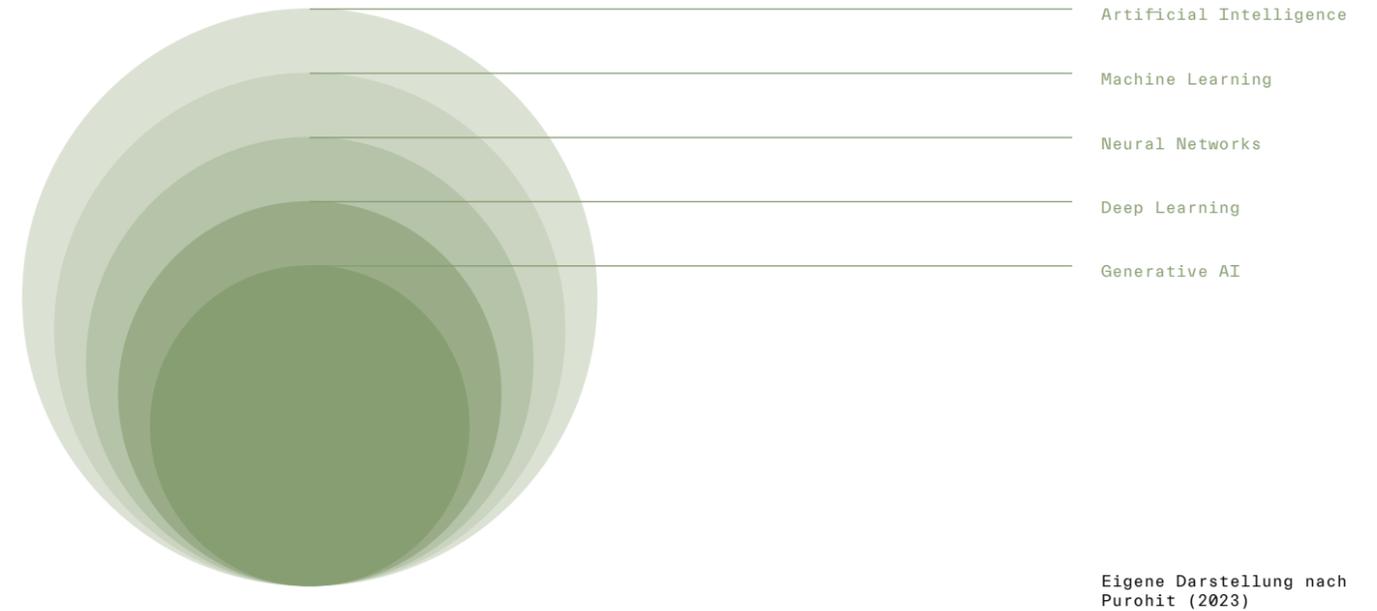
In **Kapitel 4** werden die wichtigsten rechtlichen und ethischen Fragestellungen skizziert. Neben einer Einordnung des aktuellen Diskurses werden Links zu weiterführenden Informationen und Datenbanken bereitgestellt.

# 2 Definition: Generative Künstliche Intelligenz

## Künstliche Intelligenz

Künstliche Intelligenz (KI) ist ein weitreichendes Feld der Informatik, das sich mit der Schaffung von Systemen befasst, die Aufgaben ausführen können, welche typischerweise menschliche Intelligenz erfordern. Diese Systeme sind darauf ausgelegt, komplexe Probleme zu lösen, Entscheidungen zu treffen und sogar kreativ zu sein, indem sie menschliche kognitive Prozesse wie Lernen, Schlussfolgern und Wahrnehmung nachahmen. KI-Technologien umfassen beispielsweise Maschinelles Lernen (ML), neuronale Netzwerke, natürliche Sprachverarbeitung (NLP) und Computer Vision. Diese Technologien ermöglichen es Maschinen, aus Erfahrungen zu lernen, Muster zu erkennen und Entscheidungen mit einem gewissen Grad an Autonomie zu treffen.<sup>1</sup>

<sup>1</sup> Lackes & Siepermann (2018)



## Generative Künstliche Intelligenz

Dieser Bereich der KI zeichnet sich durch die Fähigkeit aus, selbstständig neue Inhalte zu generieren, die von denen menschlicher Schöpfung kaum zu unterscheiden sind. Der Begriff „generativ“ bezieht sich hier auf die Eigenschaft der KI-Systeme, eigenständige Werke zu erschaffen, seien es Texte, Bilder, Code oder Musik.<sup>2</sup>

<sup>2</sup> Martineau (2023)

Die generative KI basiert auf fortgeschrittenen Algorithmen des maschinellen Lernens und tiefen neuronalen Netzwerken, die zusammen das Fundament dieser Technologie bilden. Diese Systeme sind darauf ausgerichtet, durch die Analyse und Verarbeitung großer Datenmengen Muster, Strukturen und komplexe Zusammenhänge zu identifizieren und zu imitieren. Ein zentrales Element dabei ist das Deep Learning, das durch mehrschichtige neuronale Netze ermöglicht wird. Jede Schicht dieser Netze ist spezialisiert auf die Extraktion und Verarbeitung bestimmter Merkmale der Eingabedaten, seien es Texte, Töne oder Bilder. Durch das Training mit einer Vielfalt an Daten lernen diese Netzwerke, spezifische Aspekte zu erkennen und zu reproduzieren, wodurch sie in der Lage sind, menschenähnliche Kreativität in Form von Sprache, Musik oder visuellen Künsten zu simulieren. Diese Fähigkeit zur Mustererkennung und -nachahmung ist es, die generative KI besonders macht und ihr ein breites Spektrum an Anwendungsmöglichkeiten eröffnet.<sup>3</sup>

<sup>3</sup> Zewe (2023)

# 3

# Anwendungsfelder

## 3.1 KI in der Musikproduktion

In den letzten Jahren hat die Entwicklung von KI-Modellen zur Musikgenerierung bedeutende Fortschritte gemacht.<sup>4</sup> Zunächst waren nur einzelne Bestandteile von Kompositionen automatisch generierbar (z.B. Melodie, Rhythmik), oder es konnte nur Musik innerhalb bestimmter Genres erzeugt werden.<sup>5</sup> Heute können einige Modelle längere Musikstücke auf Textbasis erzeugen, die von Konsument:innen nur schwer von menschlich erzeugter Musik unterscheidbar sind. Zudem existieren zahlreiche KI-Anwendungen, die für Teilschritte der professionellen Musikproduktion eingesetzt werden können.<sup>6</sup> Die Brisanz dieser Entwicklung zeigt sich auch darin, dass in einer Studie der Verwertungsgesellschaften GEMA und SACEM 73% der befragten Urheber:innen ihre wirtschaftliche Existenz gefährdet sehen.<sup>7</sup>

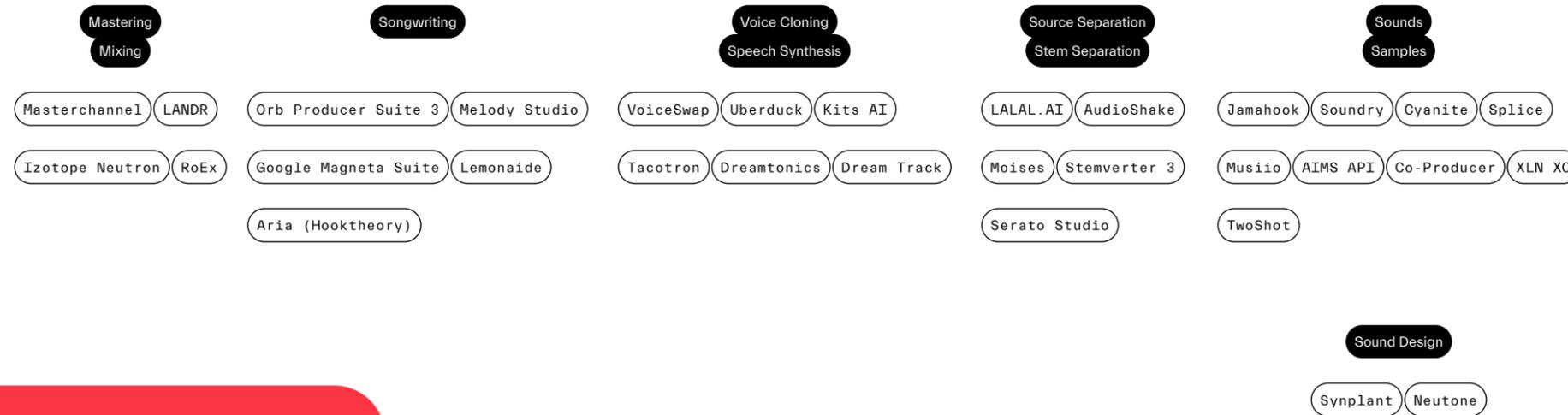
<sup>4</sup> Civit et al. (2023)

<sup>5</sup> Deruty (2022)

<sup>6</sup> Sandzer-Bell (2024)

<sup>7</sup> Goldmedia (GEMA/SACEM) (2024)

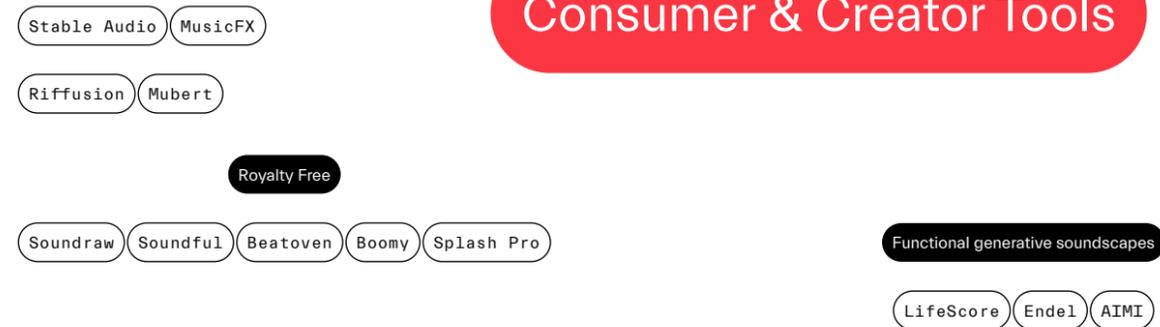
## Plug-Ins & Standalone



## Experimental & Research



## Consumer & Creator Tools



### KI-Modelle für Musikgenerierung

Es existiert noch keine einheitliche Systematik, nach der sich die Vielzahl an KI-Musik-Tools kategorisieren lassen. Einige unterscheiden nach Zielgruppe (von Konsument:innen bis professionelle Anwender:innen),<sup>8</sup> andere nach Anwendungsgebiet (z.B. Songwriting oder Mastering).<sup>9</sup>

Eine mögliche Lösung bestünde in einer Systematik, die alle bestehenden Tools nach Interaktivität und Qualität des Outputs sortiert.

Interaktivität meint dabei den Grad der Beeinflussbarkeit des musikalischen Outputs durch einen Menschen bzw. eine:n Künstler:in. Dies kann zum Beispiel durch umfangreiche Möglichkeiten für Text-Prompts zur Audiogenerierung geschehen. Hohe Interaktivität zeichnet sich auch dadurch aus, dass Tools in einem wechselseitigen co-kreativen Prozess genutzt werden, indem sie etwa neue Kompositionen auf Basis eigener Kompositionen vorschlagen, oder einen spezifischen Teilbereich der Musikproduktion automatisieren können.

Qualität beschreibt den Charakter des Outputs und kann in musikalische Qualität und Produktionsqualität unterscheiden werden: Hohe Qualität zeichnet Musik mit kohärenter Struktur, mit klaren Übergängen, Themen und Motiven aus. Sie sollte musikalisch sinnvoll sein und den Grundprinzipien von Rhythmus, Harmonie und Melodie folgen. Produktionsqualität beschreibt Ausgewogenheit im Klang sowie die Audioqualität.

<sup>8</sup> Moore & Acharya (2023)

<sup>9</sup> Edwards & McGlynn (2023)

## KI-Musik Toolboard



Das interaktive Miro-Board zeigt verschiedene KI-Musik Anwendungen, gruppiert nach Nutzungszweck und positioniert in Relation zu ihren Interaktionsmöglichkeiten sowie der Qualität des musikalischen Outputs. Zum Anklicken und Ausprobieren.

Innerhalb dieser Achsen lassen sich die drei grundlegenden Bereiche 3.1.1 Experimentelle Modelle/Forschung, 3.1.2 KI-gestützte Plugins für Muskschaffende, sowie 3.1.3 Creator-/Consumer-Tools unterscheiden, welche auf den Folgeseiten beschrieben werden.

## 3.1.1 Experimentelle Modelle / Forschung

Dieser Bereich umfasst die Modelle, die in der Forschung zu KI-generierter Musik entwickelt werden und öffentlich benutzbar sind. Der Code ist einsehbar und es existieren Forschungsberichte, die das Training und die Bewertung der Modelle beschreiben.

Aktuelle Anwendungen wie beispielsweise MusicGen<sup>10</sup> oder MusicLM<sup>11</sup>, die von Meta bzw. Google unterstützten Forschungsgruppen entwickelt werden, sind zu einem hohen Grad interaktiv: Auf Basis von Text-Input wird Musik als Audio-Waveform generiert. Teilweise können die Codes, verwendete Libraries, verwendet, angepasst und mit eigenen Daten trainiert werden. Die Audioqualität dieser Modelle wird zwar stetig besser, teilweise treten aber noch Störgeräusche oder Artefakte auf. Auch längere musikalische Zusammenhänge funktionieren häufig nicht gut.

Die Modelle sind als Grundlage für darauf basierende, weiter entwickelte Anwendungen zu sehen. Durch die in der Regel gegebene Transparenz der Trainingsdaten, detaillierte Beschreibungen der Modell-Architekturen und öffentlich zugänglichem Code fördern sie Fortschritt sowohl im akademischen Bereich als auch für kommerzielle Projekte.

## 3.1.2 KI-gestützte Plug-Ins

Für ambitionierte Hobbymusiker:innen wie professionelle Artists existiert eine Vielzahl an Möglichkeiten, Teile des Musikproduktionsprozesses mit künstlicher Intelligenz zu unterstützen. Musikproduzent:innen können Tools nutzen, die Mixing, Mastering ganz oder teilweise übernehmen, Vorschläge für Melodien, Akkord-Strukturen oder passende Loops und Samples machen. Die Outputs (wie Sounds, Midi-Noten) dieser Plug-Ins sind zu einem hohen Maß beeinflussbar und für die professionelle Weiterverarbeitung geeignet. Die Qualität des musikalischen Outputs ist, wenn die Instrumente passend eingesetzt werden, potenziell sehr hoch.

Die wichtigsten Anwendungsfelder hierbei sind:

### Songwriting

Im Bereich des Songwritings unterstützt KI die Schaffung einzigartiger Melodien und Akkordfolgen, wodurch sie Künstler:innen neue kreative Perspektiven und Inspirationen für ihre musikalischen Werke bietet. Sinnvoll eingesetzt können diese Tools einen co-kreativen Prozess zwischen Künstler:innen und KI anstoßen. Neben dem konkreten Arbeiten mit einzelnen Melodien, oder Textbausteinen, können auch Tools, die komplette Musikstücke generieren können, zur Inspiration im Songwriting genutzt werden – etwa um schnell zu prüfen, wie ein Text in verschiedenen Tonlagen auf verschiedene Instrumentale funktionieren könnte.

<sup>10</sup> Copet et al. (2023)

<sup>11</sup> Agostinelli et al. (2023)

### Mixing & Mastering

Durch den Einsatz von KI im Mixing- und Mastering-Prozess werden Pegel automatisch angeglichen, EQ-Einstellungen optimiert und die Klangqualität verbessert. Dies kann zu einer effizienteren und konsistenteren Audioproduktion führen – zumindest für Nutzer:innen, die Zeit und Kosten für menschliches Mixing/Mastering einsparen möchten und etwa auf die Bedeutung des spezifischen Klangs einer:s menschlichen Mixing- oder Mastering Expert:in weniger Wert legen.

### Voice-Cloning / Sprachsynthese

Technologien für das Klonen von Stimmen und Sprachsynthese ermöglichen die Erstellung von realistischen, anpassbaren digitalen Stimmen, die in vielen Fällen von menschlicher Sprache kaum zu unterscheiden sind. Sie können unter anderem für Deepfakes eingesetzt werden und erhielten im Zuge des viralen Songs „Heart on my Sleeve“ mit künstlichen Stimmen von Drake & The Weeknd erstmals große Aufmerksamkeit.<sup>12</sup> Eine Alternative bieten dagegen Startups an, die legal Stimmen von Artists lizenzieren, die dann beispielsweise für Demoverionen oder Toplines benutzt werden können.<sup>13</sup>

<sup>12</sup> Ingham (2023)

<sup>13</sup> Dredge (2023a)

### Stem Separation

KI vereinfacht die Trennung von einzelnen Komponenten (Stems) aus gemischten Audiospuren, wie beispielsweise die Isolierung von Gesang von Instrumenten, was eine flexiblere Bearbeitung und Neuabmischung ermöglicht. Diese Technologie wurde beispielsweise bei der Produktion des Beatles-Songs „Now and Then“ eingesetzt, für die John Lennons Stimme aus einer alten Kassetten-Aufnahme extrahiert wurde.<sup>14</sup> Auch etwa im DJ-Bereich ist eine solche Trennung in einzelne Spuren für kreatives Mixing sehr hilfreich.

<sup>14</sup> Tencer (2023)

### Sounds & Samples

Die Sound- und Sample-Suche wird durch KI-Anwendungen verbessert, indem sie Audiomerkmale und Benutzerpräferenzen von Musikstücken analysieren. Auf dieser Basis können Vorschläge für passende Sounds oder Samples für musikalische Projekte gemacht werden. In den letzten Monaten wurde dieser Bereich ergänzt um Services, die auf Prompt-Basis individuelle Sounds oder vollständige Sample Packs generieren. Dies stellt potenziell eine sinnvolle Alternative zum Suchen nach passenden Samples in Bibliotheken dar, da Samples so maßgeschneidert auf den spezifischen Kontext der Musikproduktion generiert werden können.

Zusätzlich zu den hier genannten einzelnen Anwendungsfällen, entstehen immer mehr Services, die eine Reihe von Funktionalitäten kombinieren. Diese bieten etwa Modelle für Audio-Trennung, generierte Singstimmen und KI-Mastering in einer Tool-Suite an.

## 3.1.3 Creator-/Consumer-Tools

Dieser Bereich umfasst Anwendungen, um funktionale Musik zu generieren, Ideen im musikalischen Anfangsbereich zu erhalten oder auf Prompt-Basis komplette Songs zu erzeugen. Funktionale Musik bezieht sich auf Musik, die für einen bestimmten Zweck komponiert oder ausgewählt wird, wie etwa für Werbung oder als Hintergrundmusik. Die Audioqualität dieser Modelle war zunächst ansprechend, die Interaktionsmöglichkeiten zur kreativen Weiterverarbeitung waren aber eher eingeschränkt. Durch die schnell fortschreitende Entwicklung in diesem Bereich, insbesondere durch die Möglichkeit zur Erzeugung von Lyrics und Stimmen, werden diese Modelle zunehmend als Konkurrenz zu menschlich erzeugter Musik wahrgenommen, insbesondere im Streaming-Markt.

Die folgenden vier Anwendungsfelder können unterschieden werden:

### KI-Musik mit Vocals

Innerhalb der letzten Monate kamen mit Suno und Udio zwei Modelle auf den Markt, die der Diskussion um KI-generierte Musikmodelle neue Brisanz verleihen. Erstmals können diese Modelle auf Prompt-Basis vollständige Tracks mit Gesangsstimmen in verschiedenen Sprachen und in verschiedenen Genres generieren. Dies ermöglicht es Menschen ohne musikalische Vorkenntnisse, schnell und unkompliziert (vergleichsweise) professionell klingende Songs zu erstellen. Auch wenn je nach Prompt und Genre unterschiedlich gute Ergebnisse erzielt werden können, sind die Modelle sehr interaktiv: Text / Stimme und Instrumental können unabhängig voneinander beeinflusst, eigene Lyrics eingefügt werden. Seit Juni 2024 können zudem eigene Musikstücke oder Audioaufnahmen hochgeladen und weiterverwendet werden.

Aufgrund der hohen Qualität der generierten Musikstücke im Fokus der Öffentlichkeit, wurden diese Anbieter für fehlende Transparenz in Bezug auf ihre Trainingsdaten kritisiert. Da der Output der Modelle (etwa Stimmen, oder einzelne Melodie-Linien) teilweise dem von bekannten Songs ähnelt, wurde der Vorwurf laut, die Modelle mit urheberrechtlich geschütztem Material trainiert zu haben.<sup>15 16</sup>

Eine weitere Einordnung der rechtlichen und ethischen Zusammenhänge findet in Kapitel 4 statt.

<sup>15</sup> Newton-Rex (2024a)

<sup>16</sup> Newton-Rex (2024b)

### Lizenzfreie (Hintergrund-)Musik

Dieser Bereich beschreibt Hintergrundmusik für Content-Creator:innen oder Unternehmen. Als Alternative zu Musik aus Production Music Libraries können individuelle kurze Musikstücke oder Loops erzeugt werden. Bei den meisten Modellen wird nicht künstlich Audio synthetisiert – stattdessen können entweder Genres, Stimmungen, Instrumente oder Geschwindigkeit ausgewählt werden. Diese Auswahl beeinflusst eine in den Services voreingestellte Struktur aus Loops, oder Midi-Noten, die dann neu kombiniert wird und jeweils einzigartige Musikstücke ergibt. Der Output dieser Anwendungen kann deshalb gleichförmig klingen und lässt im Nachhinein wenig interaktiven Bearbeitungsspielraum zu. Innerhalb dieses Bereichs gibt es allerdings auch Modelle, die auf Basis von Text-Prompts funktionieren und damit interaktiver sind.

### DIY Production Tools

Die Tools in diesem Bereich funktionieren ähnlich wie die zuvor beschriebenen Services. Sie besitzen allerdings weitere, darüberhinausgehende Funktionen wie die

Möglichkeit Midi-Noten zu exportieren oder diese direkt in einer Struktur aus Piano Roll und einfachem Sequencer zu bearbeiten.

### Funktionale generative Modelle

Ein weiterer Bereich beschreibt Anwendungen für funktionale Musik, etwa zur Entspannung oder Konzentration. Unternehmen wie Endel oder Lifescore kooperieren dazu unter anderem auch mit Artists, um aus deren Werken fortlaufende „Soundscapes“ zu erstellen. Als Benutzer:in solcher Services ist die direkte Interaktionsmöglichkeit eher eingeschränkt. Die generierte Musik reagiert aber etwa auf Daten wie Wetter, Tageszeit oder wenn gewünscht auch die Herzfrequenz der Benutzer:innen und passt sich dementsprechend an.

Das zu diesem Whitepaper gehörende Online-Toolboard<sup>17</sup> visualisiert 60 zum Zeitpunkt Juni 2024 verfügbare Tools zur KI-gestützten Musikgenerierung nach der beschriebenen Systematik. In der angehängten Datenbank<sup>18</sup> sind diese Tools verlinkt und nach Anwendungsgebiet sowie Funktion beschrieben.

<sup>17</sup> Toolboard KI & Musikproduktion (Stammer, 2023)

<sup>18</sup> KI-Musik-Database Pastor & Stammer (2023)

Das Toolboard und die Datenbank haben keinen Anspruch auf Vollständigkeit, sondern sollen dazu dienen den Bereich konzeptionell zu skizzieren und übersichtlich darzustellen. Für eine umfangreiche und laufend aktualisierte Datenbank ist die Arbeit des Forschungskollektivs Water and Music zu empfehlen: <https://www.waterandmusic.com/data/music-ai-market-tracker>

Update 2. Ausgabe: Einen umfassenden Marktüberblick, sowie Befragungen von Muskschaffenden zur Thematik bietet die Studie von Goldmedia im Auftrag der GEMA & SACEM (2024)

## 3.1.4 Spotlight: YouTube KI-Tools

YouTube nimmt eine aktive Rolle an der Schnittstelle zwischen Künstlicher Intelligenz (KI) und Musik ein und hat eine Reihe von Projekten und Experimenten angekündigt, die YouTube's Ansatz zur Integration von KI in die Musikindustrie verdeutlichen. Sie arbeiten eng mit Künstler:innen und der Musikindustrie zusammen, um innovative und sichere Lösungen zu entwickeln, die den kreativen Prozess bereichern und gleichzeitig den Schutz und die Integrität der Werke der Künstler:innen gewährleisten.

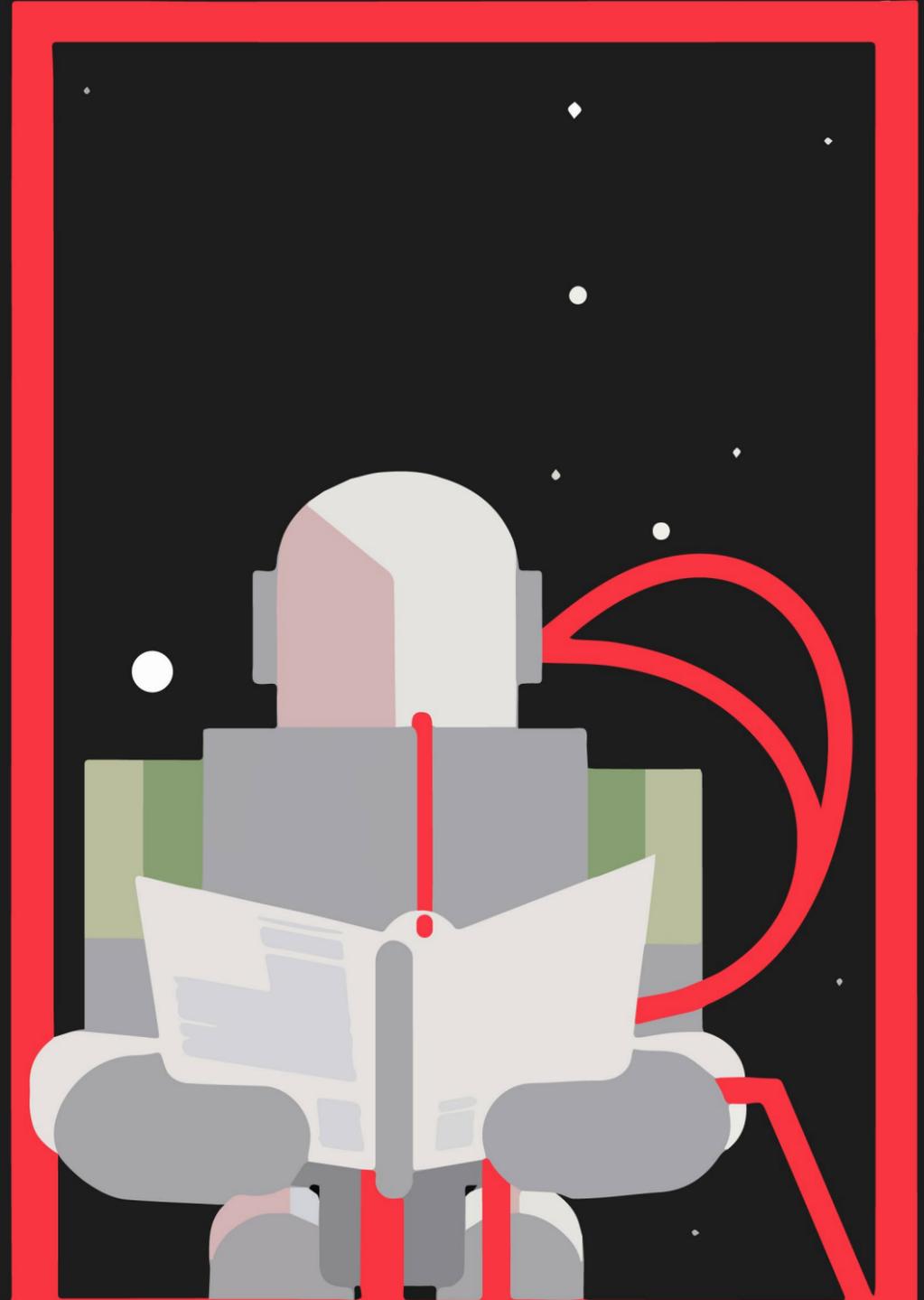
Der „Music AI Incubator“ wurde 2023 von YouTube ins Leben gerufen und umfasst mittlerweile Dutzende von Künstler:innen, Songwriter:innen und Produzent:innen weltweit. Zu den Teilnehmer:innen gehören Anitta, Björn Ulvaeus, d4vd, Don Was, Juanes, Louis Bell, Max Richter, Rodney Jerkins, Rosanne Cash, Ryan Tedder von OneRepublic, Yo Gotti, die Nachlassverwaltung von Frank Sinatra, Wyclef Jean und viele andere, die Einblicke in YouTube's generative KI-Experimente und -Forschung geben, um bei der Gestaltung und dem Testen neuer kreativer Werkzeuge zu helfen.<sup>19</sup> Gemeinsam mit Google DeepMind erkunden YouTube ein „Music AI Sandbox“: Diese Werkzeuge sollen den kreativen Prozess unterstützen, indem sie es beispielsweise ermöglichen, eine Melodie zu singen, um einen Trompetenklang zu erzeugen, oder Akkorde von einem MIDI-Keyboards in einen realistischen Vokalchor zu verwandeln. Weitere Funktionen umfassen das Erstellen von begleitenden Instrumentaltracks zu vorhandenen Vocals oder die Transformation von Audio von einem Musikstil in einen anderen. Parallel dazu hat Google DeepMind „Lyria“ angekündigt, ein fortschrittliches KI-Musikgenerationsmodell. Lyria ist in der Lage, qualitativ hochwertige Musik mit Instrumenten und Gesang zu erzeugen und bietet den Nutzer mehr Kontrolle über den Stil und die Ausführung der erzeugten Musik. Die erste Zusammenarbeit mit bekannten Künstler:innen in diesem Zusammenhang ist DreamTrack, ein Tool für Konsument:innen und Creator:innen, das eine Reihe von KI-Experimenten enthält, von denen einige Gesangsfähigkeiten einbeziehen und das in den USA mit neun renommierten Künstler:innen gestartet wurde.<sup>20</sup>

Ein weiterer wichtiger Aspekt der KI-Strategie von YouTube ist der Schutz der Arbeit der Kreativschaffenden. Die Content-ID-Technologie von YouTube hat sich als effektiv erwiesen, um Songs auf der Plattform zu identifizieren, den Rechteinhaber:innen Kontrolle über Einnahmen zu geben und ein Gleichgewicht zwischen den Interessen der Urheber:innen und der kreativen Gemeinschaft der Plattform zu wahren. YouTube plant, diese Technologie weiter auszubauen, um den Herausforderungen durch KI-generierte Inhalte zu begegnen und die Plattform sicher zu halten – etwa um Urheberrechtsverletzungen, Fehlinformationen und Spam entgegenzuwirken. Ein Beispiel dafür ist die Einführung von SynthID, einer Technologie, die ein für das menschliche Ohr unhörbares Wasserzeichen zu KI-generierten Audioinhalten hinzufügt, um deren Ursprung zu identifizieren und Missbrauch zu verhindern.<sup>21</sup>

<sup>19</sup> Dredge (2023b), Mohan (2023)

<sup>20</sup> Google DeepMind (2023)

<sup>21</sup> Google DeepMind (2024)



## 3.2 KI im Label



Künstliche Intelligenz spielt in der Vermarktung von Musik schon lange eine wesentliche Rolle, etwa in datengetriebenen Scoutingprozessen im A&Ring<sup>22</sup>, oder in den Algorithmen der Musikempfehlungssysteme der Streaming-Anbieter.<sup>23</sup>

Mit dem Aufkommen von generativer Künstlicher Intelligenz ergeben sich aber auch neue Potenziale in Bereichen wie Content-Erstellung, Promotion und Prozessoptimierung.

### 3.2.1 Content-Erstellung

Die Integration von KI in den Workflow der Content-Erstellung kann zu einer Beschleunigung und Kostenreduktion führen, da sie ermöglicht Prozesse zu automatisieren und die Komplexität von Aufgaben zu reduzieren.<sup>24</sup>

Damit verbunden stehen einerseits der Vorteil, lästige Routineaufgaben schneller zu erledigen, um mehr Zeit für wirklich kreative Arbeit mit Künstler:innen zu schaffen. Andererseits kann künstliche Intelligenz Impulse setzen, die menschliche Kreativität erweitert und verändert, statt sie zu ersetzen.<sup>25</sup>

#### Text

Die Veröffentlichung von ChatGPT im November 2022 markierte einen Wendepunkt für generative KI, indem es diese Technologie erstmals ins Rampenlicht der Öffentlichkeit rückte und weitreichende Interaktionen zwischen KI-Modellen und Nutzern ermöglichte.<sup>26</sup> Die benutzerfreundliche, konversationelle Schnittstelle erleichterte die aktive Nutzung von KI auch für Privatpersonen und führte zu einer rasanten Verbreitung mit über 100 Millionen Nutzer:innen binnen zweier Monate nach Einführung. Infolge des Erfolgs von ChatGPT wurden zahlreiche textbasierte Tools entwickelt, die auf der API großer LLMs aufbauen und Anwendungen wie KI-gestützte Copywriting-Tools und CRM-Systemintegrationen umfassen.

Labels können diese nutzen, um Inhalte zu übersetzen, Presstexte, Playlist-Pitches, oder Biografien zu schreiben oder auch Denkanstöße und Ideen für den Prozess der Content-Erstellung zu generieren. Mithilfe von „custom instructions“ können Angaben über das Unternehmen gespeichert werden, die bei jedem Chat berücksichtigt werden.<sup>27</sup> So kann gewährleistet werden, dass alle generierten Inhalte auf den Kontext und Kommunikationsstil des Labels angepasst sind. Durch die Verknüpfung mit weiteren Datenschnittstellen (z.B. Chartmetric) können diese Inhalte schnell mit relevanten Informationen zu Artist, Songs oder Zielgruppen angereichert werden.

Weiteres Marketingpotenzial für Labels besteht durch die Entwicklung eigener Chatbots, die im Sprach-/ Kommunikationsstil eigener Artists trainiert sind. Diese können in der direkten Fan-Kommunikation eingesetzt werden, oder um Texte für Social Media Posts oder Captions zu generieren.

#### Bild- und Video

KI-Tools zur Bildgenerierung basieren auf Diffusionsmodellen und können für die Zwecke des Musikmarketings sehr vielseitig eingesetzt werden. So können beispielsweise Bildgenerierungstools wie Midjourney (vgl. Midjourney, 2023), Dall.E 2

22 Rauscher (2021)

23 Knees (2021)

24 Rathore (2023)

25 Anantrasirichai & Bull (2022)

26 Berg (2023)

27 OpenAI (2023)

(vgl. OpenAI, 2023) oder Stable Diffusion (vgl. Stability AI Ltd., 2023) dazu genutzt werden, vielfältige und vor allem einzigartige Bilder zu generieren<sup>28</sup>, die beispielsweise für Artworks, Playlist-Cover, Social Media-Beiträge oder als Inhalte für Websites oder Newsletter genutzt werden können.

Durch eine komplexe Verkettung und Anpassung vieler einzelner Bildgenerierungsprozesse miteinander können Tools wie Gen-2 von Runway (vgl. Runway AI, Inc., 2023) oder Kaiber (vgl. Kaiber, 2023a) bereits kurze Videos erstellen, die für Musikvideos bzw. Visualizer, Social Media-Beiträge oder Formate wie Spotify-Canvas genutzt werden können. Da für jeden Szenenwechsel ein neues Video mit neuem Prompt (und teilweise jeweils neuen Referenzbildern, anhand derer sich das Tool in der Bildgenerierung orientiert) erstellt und mit den Videos der anderen Szenen verknüpft werden muss, stellt die Anfertigung eines kompletten Musikvideos trotz der Unterstützung durch KI heute noch einen großen manuellen Aufwand dar. Das erste Musikvideo, welches mit dem von OpenAI vorgestellten Modell Sora erstellt wurde, zeigt aber eindrucksvoll mögliche Potenziale.<sup>29</sup>

### 3.2.2 Prozessoptimierung

Neben der Möglichkeit neue Inhalte zu generieren, kann künstliche Intelligenz auch dabei helfen, bestehende Inhalte oder Prozesse einfacher oder schneller zu erstellen oder zu erledigen.

#### Transfer

Vielversprechend scheinen Werkzeuge, die bestehendes Material schnell in verschiedene Formate exportieren können. Beispielsweise kann ein YouTube-Musikvideo oder Behind-the-Scenes-Content im Querformat mit Tools wie Opus Clip in ein Video in Hochkantformat für Kurzvideoplattformen umgewandelt werden. Das Tool hebt dabei automatisch und plattformspezifisch potenziell relevante Stellen hervor.<sup>30</sup> Für internationale Marketingaktivitäten können ergänzend dazu automatisch Untertitel in verschiedenen Sprachen erzeugt werden.

Aus bestehenden Informationen wie z.B. Tourdaten können mit Tools wie Oxolo einfache Social-Media-Inhalte produziert werden, bei denen etwa Text oder Stimme automatisch generiert wird. Anbieter wie Blurb.fm erweitern automatische Content-Erstellung um zahlreiche Funktionalitäten, indem auf Basis eines Tracks und zugehöriger Metadaten eine Reihe von Assets für verschiedene Plattformen entstehen.

#### Recherche & Datenverarbeitung

Durch die zunehmende Integration von Suchmaschinen innerhalb von Chatbots wie ChatGPT wird KI-gestützte Recherchearbeit einfacher. So lassen sich im Rahmen von Promotionaktivitäten beispielsweise relevante Medien, Ausspielwege oder Kanäle identifizieren.<sup>31</sup>

Im Zuge der KI-gestützten Sentiment Analyse können Informationen aus unstrukturierten Daten wie Kommentaren unter Social-Media-Postings gewonnen werden. Hierbei wird KI eingesetzt, um in Texten ausgedrückte Emotionen zu erkennen. Die KI erfasst den Ton einer Aussage und erkennt nicht nur, ob bestimmte Wörter innerhalb einer Textgruppe eine negative oder positive Konnotation haben. Durch eine sinnvolle Verknüpfung dieser Daten mit generativen Werkzeugen können zum

28 Anantrasirichai & Bull (2022)

29 Luna (2024)

30 Loewe (2023)

31 Loewe (2023)

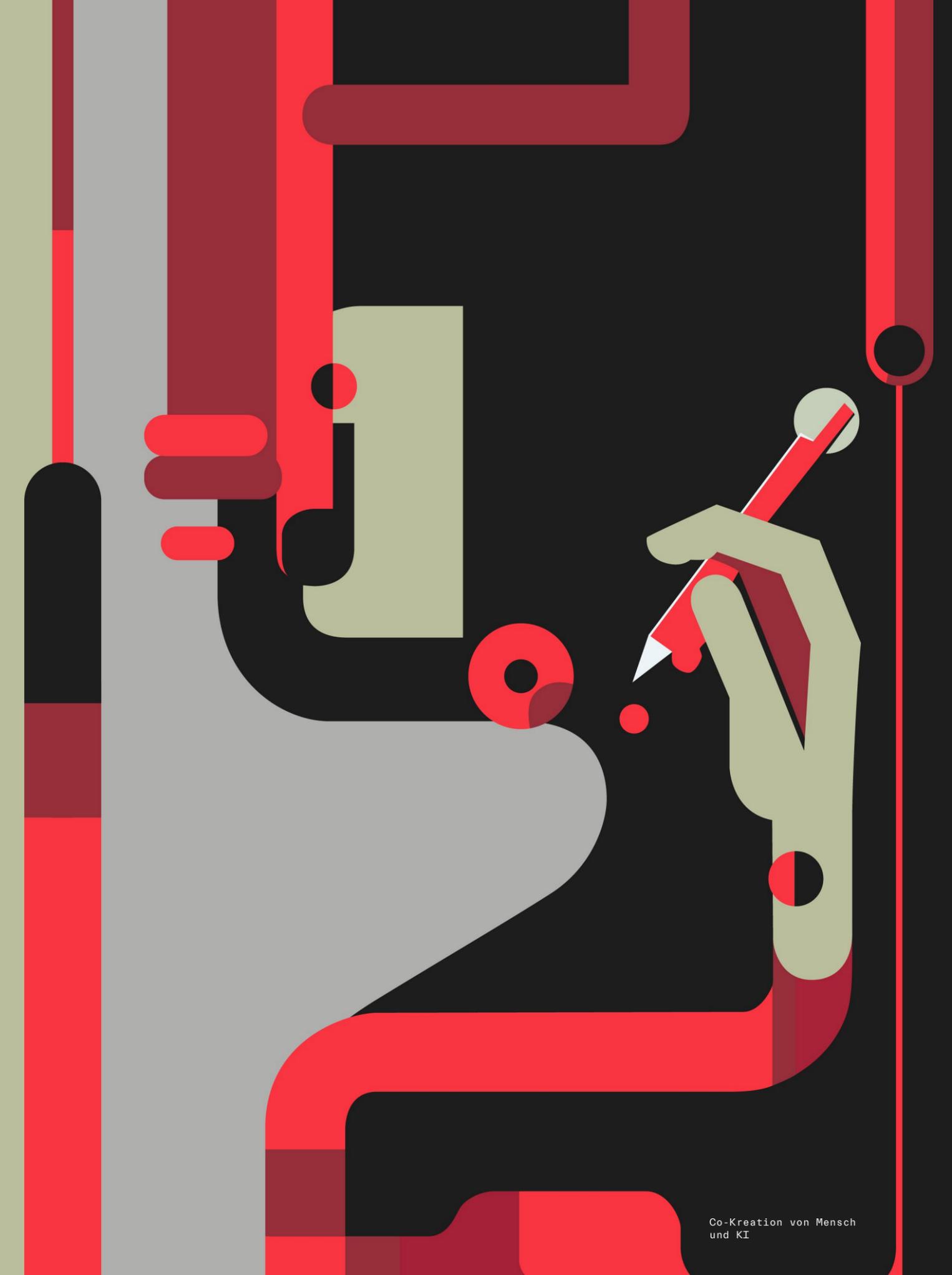
Beispiel Schlüsselinhalt für die Kommunikation oder, darauf aufbauend, konkrete Inhalte oder Ansprachen entwickelt werden.<sup>32</sup>

32 Taherdoost & Madanchian (2023)

### Steuerung von Werbeanzeigen

Auch die Steuerung von Werbeanzeigen durch KI hat sich durch fortschrittliche Datenverarbeitungsfähigkeiten wesentlich entwickelt, wobei Entscheidungen nun stärker auf Basis von Daten statt Erfahrung und Intuition getroffen werden. KI hilft dabei, Marketingbudgets zu optimieren und die Relevanz, Platzierung und Qualität von Anzeigen zu verbessern, indem sie Kampagnenleistungen und Marktentwicklungen in Echtzeit überwacht. Durch die Analyse von Nutzungsdaten und Vorhersagen zukünftigen Verhaltens kann KI das Timing von Werbeanzeigen verbessern und Retargeting effizienter gestalten. KI ermöglicht eine schnellere Anpassung an Marktveränderungen und Konsumverhaltensänderungen, indem Werbebudgets und Marketingkommunikation automatisiert verwaltet werden, was zu einem Wettbewerbsvorteil führen kann.<sup>33</sup> Trotz des Potenzials von KI in der Anzeigensteuerung gibt es derzeit keine Anwendung, die alle Aspekte umfassend integriert, obwohl Tools wie Symphony speziell für die Musikindustrie entwickelt wurden, um Werbeanzeigen automatisiert zu steuern und zu optimieren.

33 Taherdoost & Madanchian (2023)



## 3.3 KI im Verlag



Die Zusammenhänge zwischen generativer KI und der Arbeit von Musikverlagen sind vielfältig. Zunächst müssen rechtliche und ethische Aspekte (siehe Kapitel 4) beachtet werden, um Rechteinhaber:innen, Komponist:innen und Textdichter:innen vor potenziellen negativen Auswirkungen zu schützen. Insbesondere die Nutzung von generativen Text-Modellen im Songwriting und die Schutzfähigkeit dieser Inhalte spielt dabei eine Rolle.

Zwei weitere Bereiche werden im Folgenden skizziert:

### 3.3.1 Die Bedeutung von Metadaten für KI

Musik-Metadaten sind das Fundament der Arbeit von Verlagen.<sup>34</sup> Sie helfen bei der genauen Verfolgung und Verwaltung von Musikrechten, was bei generativer KI, die große Mengen an Inhalten erstellen kann, besonders wichtig ist. Korrekte Metadaten gewährleisten die korrekte Zuordnung und Vergütung der Urheber:innen, wenn ihre Werke von der KI verwendet oder gesampelt werden. Neben der Nachverfolgbarkeit bereits erfolgter Nutzungen sind Metadaten auch beim Sortieren, Kategorisieren und Auffinden von Inhalten unverzichtbar. Generative KI kann dabei zur Analyse und Generierung von Metadaten eingesetzt werden, die die Genauigkeit von Empfehlungen verbessern und es neuen und bestehenden Werken erleichtern, ihr Zielpublikum zu erreichen. KI-gestützte Analyse-Tools wie Cyanite oder Musiio übernehmen das Tagging umfangreicher Kataloge und schaffen so die Basis für Auffindbarkeit und Rückverfolgbarkeit.

<sup>34</sup> Spohn (2021)

### 3.3.2 Lizenzierung und KI

#### Lizenzfreie Musik durch KI-Tools

Die in Kapitel 3.1 beschriebenen Creator-/Consumer-Tools ermöglichen es, lizenzfreie Musik etwa für Online-Nutzungen wie Streams oder Social-Media-Content, aber auch Hintergrundmusik für Unternehmenskontexte, Events oder öffentliche Räume zu generieren. Diese Entwicklung stellt insbesondere den Bereich der Produktionsmusik vieler traditioneller Verlage auf die Probe, deren Teil des Geschäftsmodells auf diese Weise gefährdet ist. Erste Plattformen für Rechtemanagement wie Musical AI (vormals SOMMS.AI), die Kataloge von Rechteinhaber:innen verwalten und Lizenzen an KI-Unternehmen vergeben, gleichzeitig entstehende Musikgenerierungen erfassen und an die Rechteinhaber:innen zurückmelden, gewinnen in diesem Zusammenhang an Bedeutung.<sup>35</sup>

<sup>35</sup> Musical AI (2024)

#### Lizenzierung von generativer & adaptiver Musik

Ein weiteres Thema in dieser Hinsicht ist die Lizenzierung von generativer oder adaptiver Musik, die in Echtzeit entsteht und somit nicht von klassischen Urheberrechtsstrukturen erfasst wird (vgl. funktionale Musik Kapitel 3.1.).

Wie können Musikverlage mit dieser Art von Musik umgehen, die sich ständig verändert und möglicherweise außerhalb herkömmlicher Lizenzierungsmodelle liegt? Diese Frage stellt nicht nur rechtliche, sondern auch kreative und geschäftliche Herausforderungen dar, die die Branche in der Zukunft angehen muss.

#### Lizenzierung von KI-generierten Stimmen

KI-Modelle zur Generierung von Singstimmen haben in den letzten Jahren große Fortschritte gemacht und werden immer realistischer.<sup>36</sup> Artists wie Grimes stellen ihre Stimmen zur Verfügung, oder gründen Startups, die die Stimmen anderer Artists legal an Dritte lizenzieren, die damit eigene Songs schaffen können.<sup>37</sup>

<sup>36</sup> Huang et al. (2022)

<sup>37</sup> Forristal (2023) und Dredge (2023)

Auch hier stellt sich die Frage, wie Musikverlage damit umgehen, wenn kreative Schöpfungen nicht mehr klassisch als Werk (Text und Komposition) vorliegen. Auch wenn die Stimme stärker leistungsschutzrechtlich relevant ist, zeigen die Beispiele, dass in Kombination von Modellen für Text (am Beispiel einer Songwriterin trainiert) und Stimme neue Lizenzierungsformen geprüft werden sollten.



## 3.4 KI im Live-Bereich



Generative KI spielt in der Live-Branche bisher noch eine untergeordnete Rolle. Dieses Kapitel hat daher stärker explorativen Charakter, stellt erste Erkenntnisse, sowie eigene Ideen auf Basis der aktuellen Technologien vor.

### 3.4.1 KI in Booking und Tourmanagement

Booking und Tourmanagement-Prozesse sind kleinteilig und kommunikationsintensiv. Dabei müssen viele Daten, Zahlen und Informationen von verschiedenen Stellen abgefragt und für die beteiligten Artists und Crewmitglieder aufbereitet werden.

Sprachmodelle können hier einen hilfreichen Beitrag leisten, etwa durch die in ChatGPT 4 integrierten Plugins: So können Informationen aus URLs oder PDFs, wie zum Beispiel Buchungsdokumenten oder Zugtickets ausgelesen und interpretiert werden. Damit können Kosten schneller berechnet und aufgeteilt werden, Itineraries schneller für einzelne Personen erstellt und in verschiedene Dateiformate konvertiert werden.

Insbesondere durch die im November 2023 von OpenAI eingeführten „GPTs“ und „Assistants“ ergeben sich weitere Möglichkeiten: GPTs sind maßgeschneiderte Versionen von ChatGPT für bestimmte Zwecke, die in der Gestaltung und Bedienung sehr einfach sind. Durch die etwas komplexere Verknüpfung mit Programmier-schnittstellen (APIs) und eigene Daten-Pools lassen sich zudem digitale „Assistants“ erstellen, die große Teile von Booking und Tourmanagement-Prozessen unterstützen könnten.

### 3.4.2 KI bei der Live-Performance

Im Bereich der Live-Performances bei Konzerten oder Festivals sind zunächst Tools denkbar, die ähnlich wie die KI-gestützten Mixing- oder Mastering-Tools in der Musikproduktion funktionieren. Die Vision dabei ist eine Audioverbesserung in Echtzeit: KI könnte die Live-Audioqualität verbessern, indem es Hintergrundgeräusche reduziert, die Tonqualität verbessert und die Lautstärke an den Veranstaltungsort und die Größe des Publikums anpasst.<sup>38</sup>

38 Frackiewicz (2023)

KI-generierte Visuals können zur Unterstützung der Performance dienen, etwa durch vorproduzierten Content mithilfe von Werkzeugen zur Video-Generierung (vgl. Kapitel 2.2.). Zukünftig ist denkbar, dass auch in Echtzeit Licht und visuelle Effekte auf die Musik oder die Bewegung der Künstler:innen auf der Bühne reagieren. Aktuell müssen diese Komponenten manuell von Techniker:innen gesteuert werden, die die Bewegungen der Darsteller:innen auf der Bühne vorhersehen und darauf reagieren. KI-gesteuerte Systeme könnten die Bühne in Echtzeit zu analysieren, die Künstler:innen verfolgen und die Beleuchtung und das Bildmaterial autonom anpassen. Dies würde es Techniker:innen ermöglichen, sich anderen Aspekten des Produktionsprozesses zu widmen.

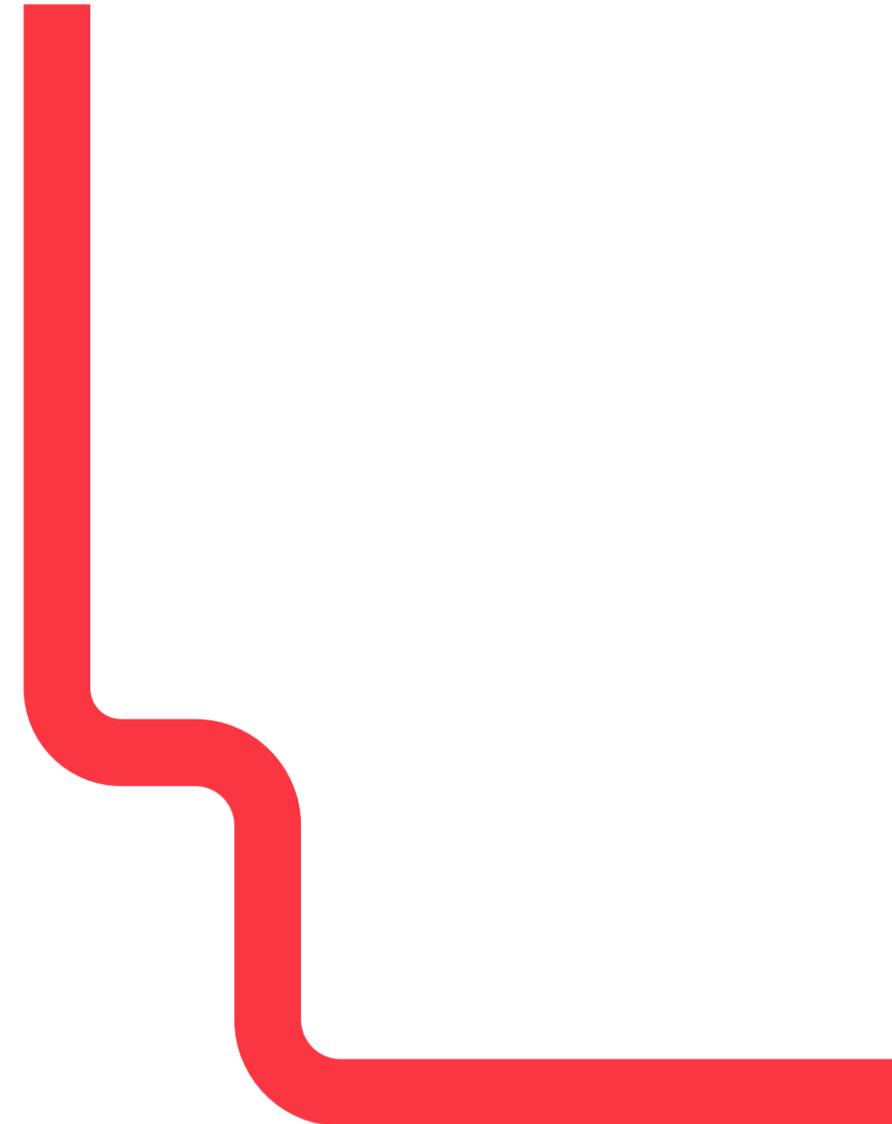
## 3.4.3 KI bei virtuellen Live-Events

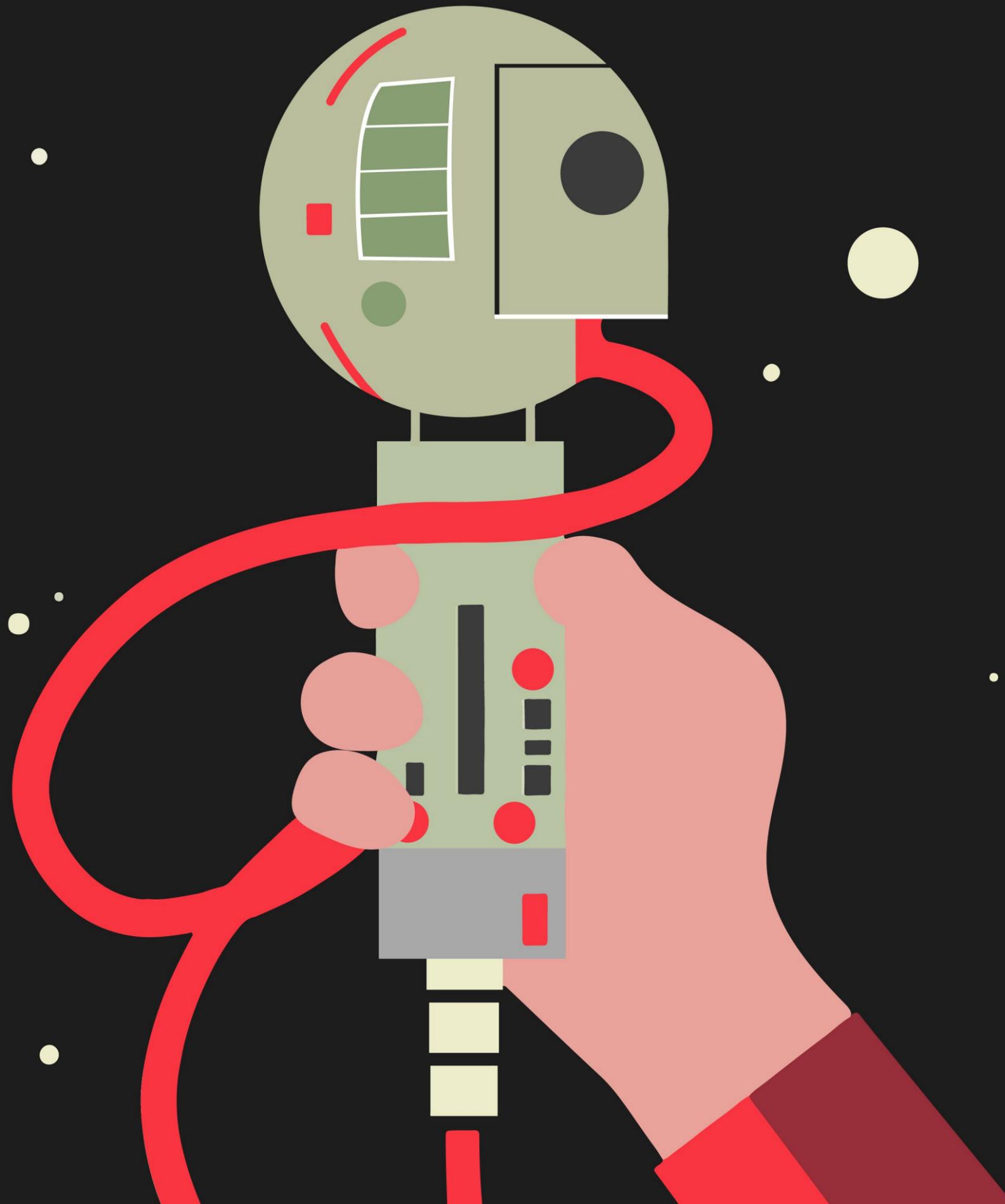
Die generative KI-basierte Bilderzeugung spielt eine entscheidende Rolle bei der Schaffung immersiver virtueller Welten. Durch die Erzeugung realistischer und vielfältiger Bilder von Landschaften, Gebäuden und Objekten tragen diese Modelle zum visuellen Reichtum und zur Glaubwürdigkeit virtueller Umgebungen bei. So kann generative KI zum Beispiel virtuelle Venues, Städte, Landschaften oder futuristische Umgebungen schaffen und den Benutzer:innen visuell fesselnde und ansprechende virtuelle Welten bieten, die sie erkunden und mit denen sie interagieren können.<sup>39</sup>

39 Chamoia et al (2023)

Generative Musik (vgl. Kapitel 2.1) mit ihrer Fähigkeit, Variationen von Musikkompositionen zu erzeugen, kann die Erfahrungen der Benutzer:innen verbessern, indem sie eine dynamische und anpassungsfähige auditive Kulisse bietet. Sie kann so programmiert werden, dass sie auf Interaktionen, sich verändernde Umgebungen oder bestimmte Ereignisse innerhalb einer virtuellen Welt reagiert. Wenn Benutzer:innen beispielsweise verschiedene Landschaften durchqueren oder verschiedenen Aktivitäten nachgehen, kann die Musik nahtlos das Tempo, die Stimmung oder das Genre wechseln und so die Atmosphäre der virtuellen Umgebung widerspiegeln und verstärken.<sup>40</sup>

40 Jin et al. (2022)





# 4

# Normative Betrachtungen

## 4.1 Generative KI und Urheberrecht

Die Diskussion um die urheberrechtlichen Implikationen KI-generierter Musik ist nicht neu, gewinnt jedoch zunehmend an Bedeutung, je fortschrittlicher die Modelle und deren musikalische Ergebnisse werden – je stärker die Grenzen zwischen menschlicher Kreativität und maschineller Generierung verschwimmen.<sup>41</sup>

41 Sturm et al. (2019)

### Rechtliche Situation

Das Urheberrechtsgesetz verlangt für urheberrechtlichen Schutz eine „persönliche geistige Schöpfung“, die eine menschlich-gestalterische Tätigkeit des Urhebers erfordert. Maschinelle Erzeugnisse, wie solche durch Künstliche Intelligenz generiert, sind gemäß der vorherrschenden Rechtsmeinung nicht schutzfähig. Technische Hilfsmittel können zwar als Werkzeug im kreativen Prozess dienen, jedoch wird bei einer Gestaltung, die maßgeblich durch technische Mittel, wie KI, geprägt ist, die Schutzfähigkeit kontrovers diskutiert. Entscheidend ist dabei, ob der menschliche Anteil an der Gestaltung noch als ausreichend erachtet wird, um dem Menschen die Urheberschaft zuzuordnen.<sup>42</sup>

42 Kitzberger (2023)

Die Frage, ob den Entwickler:innen einer Software Urheberrechte an den durch diese Software generierten Werken zustehen, hängt stark von der Einflussnahme der Programmierer:innen auf das Endprodukt ab. Werden wesentliche Gestaltungsentscheidungen von der Software autonom getroffen, schmälert dies die Möglichkeit eines Urheberrechts für die Programmierer:innen. Zugleich gewinnen Leistungsschutzrechte, besonders im Bereich der Musik, an Bedeutung, insbesondere wenn es um die urheberrechtliche Einordnung des durch KI-Programme erzeugten Outputs und des dabei verwendeten Inputs geht. In Fällen, in denen ein KI-Programm urheberrechtlich geschützte Werke reproduziert oder transformiert, bleiben die Rechte der Original-Urheber bestehen.<sup>43</sup> Aktuelle Urteile sind diesbezüglich konsistent, auch im Vergleich der Rechtsprechung zwischen USA und Europa.<sup>44</sup>

43 Fierdag (2023)

44 Hembt (2023)

### Forderungen von Verbänden / Interessensvertretungen

Aufgrund der Bedeutung des Themas, und aus Sorge, dass Präzedenzfälle geschaffen werden könnten, haben verschiedene internationale Verbände Prinzipien / Guidelines zu generativer Künstlicher Intelligenz veröffentlicht (u.a. Human Artistry Campaign, UK Music, Independent Music Publishers International Forum).

Folgende Punkte sind in den Veröffentlichungen zu KI-Guidelines und Prinzipien am häufigsten genannt:

- **Erlaubnis:** Rechteinhaber:innen müssen um Erlaubnis gefragt werden, wenn ihre Musik zu Trainingszwecken benutzt wird.
- **Transparenz / Nachverfolgbarkeit:** Es muss nachverfolgbar sein, welche Werke im Trainingsprozess genutzt werden.
- **Kennzeichnungspflicht:** KI-generierte Musik muss als solche gekennzeichnet werden.
- **Rein KI-generierte Werke:** Werke, die ohne menschlichen Schaffensprozess entstanden sind, sollen keinen urheberrechtlichen Schutz genießen.
- **Persönlichkeitsrechte der Artists & Songwriter:** Diese sollen gewahrt werden, besonders im Kontext von Deepfakes im Bereich der Stimmgenerierung.

Die wachsende Bedeutung und Komplexität dieser Thematik erfordert eine dynamische rechtliche Auseinandersetzung, um klare Rahmenbedingungen für den Einsatz von KI zu schaffen und Urheber effektiv zu schützen. Dies wird auch im KI-Gesetz der EU deutlich:

„Generative Foundation-Modelle wie ChatGPT müssten zusätzliche Transparenzanforderungen erfüllen:

- Offenlegung, dass der Inhalt durch KI generiert wurde
- Gestaltung des Modells, um zu verhindern, dass es illegale Inhalte erzeugt
- Veröffentlichung von Zusammenfassungen urheberrechtlich geschützter Daten, die für das Training verwendet wurden.“<sup>45</sup>

45 Europäisches Parlament (2023)

### Reaktionen der Musikindustrie

Wie relevant diese Forderungen sind, zeigt eine aktuelle Untersuchung zu Trainingsdatensätzen für automatische Musikgenerierung. In dieser wurde deutlich, dass in fast der Hälfte aller Datensätze online verfügbare Musik verwendet wurde, ohne die Erlaubnis der Musiker:innen einzuholen.<sup>46</sup> Diese Untersuchung betraf wissenschaftliche Modelle, die eine grundsätzliche Transparenz in Bezug auf ihre Datensätze zeigen.

46 Morreale, Sharma & Wie (2023)

Anders sieht die Situation im kommerziellen Bereich aus, weshalb die Musikindustrie zunehmend auf diese Entwicklungen reagiert. Zunächst wies Sony Music in einem Schreiben an 700 Unternehmen im KI-Sektor darauf hin, dass deren Werke (sowohl von Label- und Verlagsseite) nicht für das Training von KI-Modellen benutzt werden dürften.<sup>47</sup>

47 Dredge (2024)

Schließlich folgten Urheberrechtsklagen der Recording Industry Association of America (RIAA) an die in Kapitel 3.1. beschriebenen, derzeit bekanntesten Anbieter Udio und Suno. Die Klagen beziehen sich auf den Vorwurf, dass beide Unternehmen urheberrechtlich geschützte Musik ohne Erlaubnis verwendet haben, um ihre KI-Modelle zu trainieren. Die RIAA argumentiert, dass dies keine "faire Nutzung" darstelle und die generierten Inhalte direkt mit den ursprünglichen Werken konkurrieren, was die Musikindustrie erheblich schädige.<sup>48</sup>

48 Tencer (2024)



## 4.2 Generative KI und Ethik



Neben den urheberrechtlich kritischen Punkten sind weitere ethische Fragestellungen im Zusammenhang mit generativer Musik zu betrachten.

Eine aktuelle Untersuchung von Modellen für Musik- und Sprachgenerierung zeigt, dass nur in 10% der Fälle mögliche negative Auswirkungen von Seiten der Autor:innen genannt werden. Somit ist das Bewusstsein darüber auf Seiten vieler Entwickler:innen nur gering ausgeprägt.<sup>49</sup> Folgende Bereiche werden zusätzlich zu den bereits beschriebenen rechtlichen Punkten genannt:

<sup>49</sup> Barnett (2023)

### Biases

Viele Datensätze, mit denen Musik-Modelle trainiert werden, sind westlich geprägt, da die Forschung im Bereich „Music Information Retrieval“ vorrangig in USA und Europa stattfindet.<sup>50</sup> Dies kann Auswirkungen auf den Output von KI-Musikmodellen haben, wie z.B. die Beibehaltung des Quintenzirkels und die Unterscheidung von Dur- und Molltonarten, auch wenn das gar nicht intendiert war.<sup>51</sup> Dies spiegelt ein allgemeines Problem des maschinellen Lernens wider, bei dem Modelle oft Verzerrungen aus ihren Trainingsdaten übernehmen.

<sup>50</sup> Holzapfel, Sturm & Coeckelbergh (2018)

<sup>51</sup> Zhao et al. (2022)

### Kulturelle Aneignung

Demgegenüber besteht bei generativen Audiomodellen, die umfangreiche und vielfältige Datensätze verwenden können, die Gefahr der kulturellen Aneignung, insbesondere wenn die Daten Inhalte aus marginalisierten Kulturen enthalten. Die generierte Musik spiegelt nicht nur die Verzerrungen der Trainingsdaten wider, sondern könnte sich auch unbeabsichtigt kulturelle Elemente aneignen, ohne die jeweiligen Gemeinschaften richtig zu verstehen oder sich mit ihnen auseinanderzusetzen.<sup>52</sup> Dies unterstreicht die Notwendigkeit, die Zusammensetzung der Trainingsdaten sorgfältig zu überdenken, um zu vermeiden, dass Vorurteile und kulturelle Unsensibilität fortbestehen.

<sup>52</sup> Agostinelli et al. (2023)

### Fairness in Trainingsdaten

Auch am Beispiel von Musikempfehlungssystemen werden Probleme in Bezug auf die Input- und Output-Daten deutlich. Nutzer:innen von Streaming-Services konsumieren bis zu einem gewissen Grad das, was ihnen empfohlen wird. Diese Nutzungsdaten fungieren dann als Input für die Empfehlungsalgorithmen, die wiederum neue, ähnliche Tracks als Output empfehlen, die wiederum häufiger gehört werden.

Ein solcher Feedback-Loop zeigt, dass bestehende Ungleichheiten in Datensätzen durch Algorithmen fortgeführt oder verstärkt werden können. Dies wird am Beispiel von Gender-Repräsentation in Musikempfehlungen deutlich.<sup>53</sup>

<sup>53</sup> Ferraro, Serra & Bauer (2021)

### KI und Nachhaltigkeit

Der Ressourcenverbrauch generativer Modelle gerät immer stärker in den Fokus. Beispielsweise wurden für das Training von GPT3 in Datacentern von Microsoft 700.000 Liter Frischwasser verbraucht (durch Kühlanlagen und bei der notwendigen Stromerzeugung).<sup>54</sup>

<sup>54</sup> Li et al. (2023)

Um dem entgegenzuwirken, existieren eine Reihe von Vorschlägen:

- bestehende Modelle zu verbessern statt neue, eigene zu trainieren
- energiesparendere Rechenmethoden verwenden,
- KI-Modelle nur benutzen, wenn sie von signifikantem Wert sind

Speziell für den Bereich der Musikgenerierung schlagen Forschende vor, dass in der Bewertung der Qualität eines Musikmodells auch dessen Energiebilanz mit einfließen sollte.<sup>55</sup>

<sup>55</sup> Douwes, Esling & Briot (2021)



Für eine weitere Auseinandersetzung mit ethischen und rechtlichen Fragestellungen sowie politischen Forderungen seitens der Musikindustrie ist der „Music AI Ethics Tracker“ des Forschungsverbundes Water & Music ein guter Ausgangspunkt.<sup>56</sup> Für weiterführende Quellen und ethisch relevante Fragestellungen, die über die Musikindustrie hinausgehen, sind die Publikationen des Montreal AI Ethic Institute zu empfehlen: <https://montrealetics.ai/>

<sup>56</sup> Hu (2023)



# 5 Fazit

Das vorliegende Whitepaper hat die facettenreichen Einflüsse generativer Künstlicher Intelligenz (KI) auf die Musikbranche beleuchtet. Es hat verdeutlicht, dass generative KI weit mehr als nur ein technologisches Hilfsmittel ist; sie repräsentiert einen Paradigmenwechsel in der Musikproduktion, -vermarktung und im Erleben von Musik.

Dazu wurden die Kernaspekte generativer KI, von den technischen Grundlagen bis hin zu spezifischen Anwendungsgebieten, dargestellt. Es wurde ersichtlich, dass generative KI in der Musikproduktion eine Reihe verschiedener Funktionen und Anwendungsgebiete abdeckt: von experimentellen Modellen über professionelle Plug-Ins bis hin zu Consumer-/Creator-Tools. Im Bereich der Musiklabels und -verlage zeigt sich ein breites Spektrum an Einsatzmöglichkeiten für generative KI, von Content-Erstellung bis hin zur Prozessoptimierung. Auch im Live-Bereich deutet sich ein wachsendes Potential an, obwohl dies noch in den Kinderschuhen steckt.

Generative KI hat das Potenzial, die Musikindustrie grundlegend zu verändern. Sie bietet neue Möglichkeiten für Kreativität und Personalisierung, stellt aber gleichzeitig die Branche vor komplexe Herausforderungen: Die Diskussion etwa um Urheberrechte an KI-generierter Musik und die ethischen Implikationen von Verzerrungen in Trainingsdaten sowie kultureller Aneignung zeigen, dass die Einführung generativer KI in der Musikbranche eine sorgfältige und differenzierte Betrachtung erfordert. Es wird deutlich, dass eine verantwortungsbewusste und nachhaltige Nutzung von KI in der Musikbranche unabdingbar ist.

Abschließend lässt sich sagen, dass generative KI nicht nur eine technologische Revolution darstellt, sondern auch eine kulturelle. Sie fordert uns auf, über die Rolle von Technologie in der kreativen Arbeit nachzudenken und verlangt ein Umdenken in Bezug auf Urheberrecht, Ethik und die Beziehung zwischen Mensch und Maschine in der Kunst.

Trotz der technologischen Fortschritte muss der Fokus darauf liegen, dass der menschliche Ausdruck, die Kreativität und die Emotionen in der Musik beibehalten werden. Der Einsatz von generativer KI muss demnach immer mit dem Ziel erfolgen, die menschliche Kreativität zu unterstützen und zu erweitern, nicht jedoch sie zu ersetzen. Die Transformation durch generative KI sollte deshalb in enger Zusammenarbeit mit Muskschaffenden erfolgen. Ihre aktive Beteiligung ist unerlässlich, um eine sinnvolle Integration der KI in Prozesse der Kreativität und Vermarktung zu gewährleisten. Mit diesem Whitepaper wollen wir einen Beitrag zum Verständnis und zur kritischen Auseinandersetzung mit dem Thema leisten.



## Quellenverzeichnis

- Agostinelli, Andrea, Timo I. Denk, Zalán Borsos, Jesse Engel, Mauro Verzetti, Antoine Caillon, Qingqing Huang, u. a. „MusicLM: Generating Music From Text“. arXiv, 26. Januar 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. Januar 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'{}eal 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. Zugegriffen 29. November 2023. <https://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. Juni 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. Oktober 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 Auameur. „On the Development and Practice of AI Technology for Contemporary Popular Music Production“ 5, Nr. 1 (7. Februar 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. Oktober 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. Mai 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. Juli 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. Januar 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. Juni 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/>.
- Frackiewicz, Marcin. „How AI Is Revolutionizing the World of Concert Production and Management“. *TS2 SPACE (blog)*, 1. Mai 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. Mai 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“. Zugegriffen 30. November 2023. <https://www.twobirds.com/en/insights/2023/global/the-debate-on-ai-created-art-and-copyright-protection>.
- Holzappel, 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. Mai 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. Dezember 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. Zugegriffen 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. Oktober 2023. <http://arxiv.org/abs/2304.03271>.
- Loewe, Julian T. „Future Music Camp 2023: Kreative KI im Label“. Mannheim, 26. Juni 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. Mai 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. Juli 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“. *Andreesen 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-T0enUjnZiSwlTMfZtQuuU1VGSzuwEdiMZ1TGrTHNPj4NV936Sxm40XxWUcJnVGAesfTt0td2hIEQrinVjrb-TidZNMwM1G1sQ](https://openai.com/blog/custom-instructions-for-chatgpt?mkt_tok=NjczLVBI-Sy05NDgAAAGN4RwAUP-T0enUjnZiSwlTMfZtQuuU1VGSzuwEdiMZ1TGrTHNPj4NV936Sxm40XxWUcJnVGAesfTt0td2hIEQrinVjrb-TidZNMwM1G1sQ).
- Purohit, Anish. „AI, ML, DL, and Generative AI Face Off: A Comparative Analysis“. *Synoptek (blog)*. Zugegriffen 16. Dezember 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. Februar 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. Mai 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 (Februar 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. Juni 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>.

