

AI-Anwendungen in der Musikproduktion und Audioverarbeitung

Linus Graeter
Matrikelnummer: 45957
Ton-Seminar
Wintersemester 2024-25

Inhaltsverzeichnis:

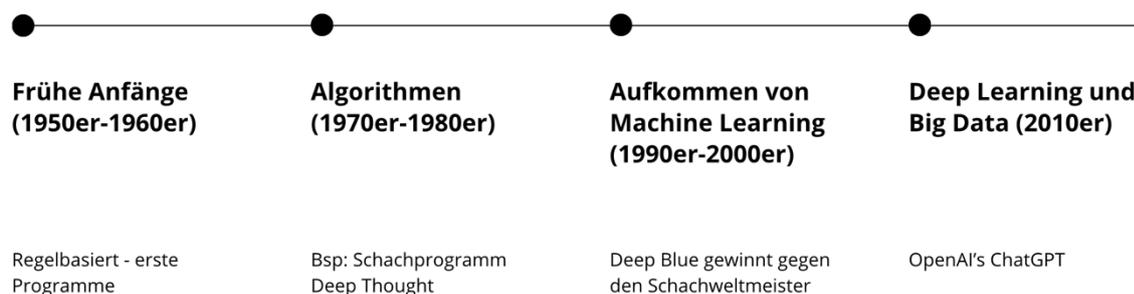
1. Grundlagen der AI
 - 1.1 Grundlegende Funktionsweisen der AI
 - 1.2 Historische Entwicklung
 - 1.3 Gesetze
 - 1.4 AI im Kontext von Audio
2. Ai Tools - Audio-Editing und -Verarbeitung
 - 2.1 Stem Separation
 - 2.2 Voice Changer
 - 2.3 Audio Repair
 - 2.4 Dynamic-Resonance-Suppressor
 - 2.5 Mix und Mastering
3. Audioerzeugende AI-Tools
 - 3.1 AI-Synthese
 - 3.2 Song Generation
 - 3.3 Text to Speech
 - 3.4 Midi Generation
4. Workflow Tools
 - 4.1 AI-Finder
 - 4.2 Sample-Finder
 - 4.3 Speech to Text
5. Anwendungserfolge
6. AI-Erkennung
7. Chancen und Herausforderungen

1. Grundlegende Funktionsweisen der AI

1.1 Grundlegende Funktionsweise der AI

Künstliche Intelligenz (AI) basiert auf der Fähigkeit von Maschinen, menschenähnliche Denk- und Lernprozesse zu simulieren. Zentrale Konzepte sind dabei maschinelles Lernen und neuronale Netze, welche es Algorithmen ermöglichen, aus großen Datenmengen Muster zu erkennen und Vorhersagen zu treffen. Durch Iterationen von Trainingsprozessen mit vorgegebenen Trainingsdaten verbessern AI-Modelle kontinuierlich ihre Leistung, indem sie neue Informationen verarbeiten und ihre Entscheidungsstrategien anpassen. Die Entscheidungsstrategien sind in Form von Gewichtungen in einem Neuronalen Netz implementiert und variabel einstellbar. AI findet Anwendung in zahlreichen Bereichen, von Spracherkennung, Bildverarbeitung, autonomem Fahren, medizinischer Diagnostik aber auch Audiotbearbeitung sowie Audiosynthese.

1.2 Historische Entwicklung



In den **1950er- und 1960er-**Jahren legten Wissenschaftler wie Alan Turing und John McCarthy die theoretischen Grundlagen der Künstlichen Intelligenz, wobei McCarthy den Begriff „Artificial Intelligence“ prägte. Frühe AI-Programme, wie z.B. das Schachprogramm von Allen Newell und Herbert A. Simon oder ELIZA, ein einfacher Chatbot, basierten auf regelbasierten Systemen und symbolischer Logik, hatten jedoch noch große Einschränkungen in ihrer Leistungsfähigkeit.

In den **1970er und 1980er** Jahren erlebte die Künstliche Intelligenz bedeutende Fortschritte, insbesondere durch die Entwicklung von Expertensystemen, die es Computern ermöglichten, menschliches Fachwissen in spezifischen Bereichen zu replizieren.

Diese Ära war jedoch auch von sogenannten "AI-Wintern" geprägt, Phasen der Ernüchterung aufgrund überzogener Erwartungen und begrenzter praktischer Anwendungen.

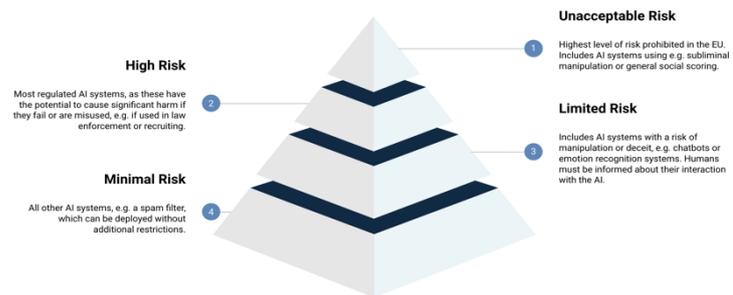
In den **1990er** Jahren erlebte die Künstliche Intelligenz einen bedeutenden Aufschwung, insbesondere durch die Entwicklung von Support Vector Machines (SVMs) und rekurrenten neuronalen Netzen (RNNs), die die Analyse großer Datenmengen und das Erlernen von Mustern ermöglichten.

Diese Fortschritte legten den Grundstein für die AI-Anwendungen der **2000er** Jahre, in denen die Verfügbarkeit riesiger Datenquellen und die zunehmende Rechenleistung die Entwicklung von Deep-Learning-Algorithmen vorantrieben, die heute in vielen Bereichen wie Bild- und Spracherkennung sowie maschineller Übersetzung eingesetzt werden.

Seit den **2010er** Jahren hat die Künstliche Intelligenz durch Fortschritte im maschinellen Lernen, insbesondere im Deep Learning, erhebliche Entwicklungen erfahren, die zu bedeutenden Durchbrüchen in Bereichen wie Bild- und Spracherkennung führten. Diese Entwicklungen haben die AI von einer spezialisierten Technologie zu einem integralen Bestandteil vieler Lebensbereiche gemacht.

1.3 Gesetze

Die Europäische Union hat im August 2024 das weltweit erste Gesetz zur Regulierung von Künstlicher Intelligenz verabschiedet, das sogenannte AI-Gesetz oder „AI Act“. Dieses Gesetz zielt darauf ab, den Einsatz von AI-Systemen zu harmonisieren und sicherzustellen, dass sie die Grundrechte der Bürger schützen. Es verbietet bestimmte Anwendungen, die als unannehmbares Risiko gelten, wie etwa AI-Systeme zur sozialen Bewertung von Personen, und fordert von Unternehmen, ihre AI-Systeme nach Risikostufen zu klassifizieren und entsprechende Maßnahmen zu ergreifen.



1.4 AI im Kontext von Audio

Künstliche Intelligenz wird in der Audiotbearbeitung und -erzeugung genutzt, um Klänge zu analysieren, zu verarbeiten und neu zu erzeugen. Dafür setzen AI-Systeme auf mathematische Modelle und Algorithmen, insbesondere neuronale Netze, die Muster in Audiodaten erkennen. Häufig werden Audiosignale in Spektrogramme umgewandelt, die zeigen, wie sich Frequenzen über die Zeit verändern. Neuronale Netze lernen dann mithilfe von Optimierungsverfahren wie dem Backpropagation-Algorithmus, bestimmte Merkmale im Klang zu identifizieren. Zur Erzeugung neuer Audiodaten kommen generative Modelle wie GANs oder RNNs zum Einsatz, die aus Trainingsdaten lernen und neue Klänge synthetisieren. Diese Technologien ermöglichen Anwendungen wie Sprachsynthese, Musikkomposition oder Geräuscherkennung mit hoher Genauigkeit.

2. Ai Tools - Audio-Editing und -Verarbeitung

2.1 Stem Separation

Stem Separation ist ein Verfahren der Audiotbearbeitung, bei dem ein Audiosignal mithilfe von AI-Algorithmen und Signalverarbeitungstechniken in seine einzelnen Bestandteile wie Gesang, Schlagzeug, Bass und weitere Instrumente zerlegt wird. Dies geschieht durch spektrale Analyse und Deep-Learning-Modelle, die Muster und Frequenzbereiche erkennen, um die verschiedenen Klangquellen voneinander zu trennen.

In den unten aufgelisteten Anwendungen lassen sich also per Knopfdruck ganze Songs in ihre Bestandteile aufteilen. Meistens werden die Optionen: Gesang, Instrumental, Drums, Bass, Gitarre, Piano, Streicher, Bläser oder Melodisch angegeben. Oftmals sind die Ergebnisse aber nicht sauber voneinander getrennt, gerade bei Elementen, die denselben Frequenzraum belegen.

Anwendungen:

- LALAL.AI
- Voice AI
- Gaudio Studio
- Traktor Pro

2.2 Voice Changer

Ein Voice Changer ist ein Audiotool, das die Eigenschaften einer Stimme in Echtzeit oder nachträglich verändert, indem es Parameter wie Tonhöhe, Formanten, Timbre oder Modulation anpasst. Dies geschieht durch Signalverarbeitungstechniken und AI-Modelle, die Frequenzbereiche manipulieren, um die Stimme beispielsweise höher, tiefer oder verzerrt klingen zu lassen.

Anwendungen:

- LALAL.AI
- Voice AI
- Eleven Labs

2.3 Audio Repair

Audio-Repair-Anwendungen wie iZotope RX nutzen spezialisierte Algorithmen und maschinelles Lernen, um Störungen wie Rauschen, Knistern oder Clipping aus Audiosignalen zu entfernen. Dabei analysieren sie das Frequenzspektrum und die zeitlichen Muster des Signals und versucht unerwünschte Artefakte gezielt zu isolieren und zu reduzieren, ohne die ursprüngliche Klangqualität zu beeinträchtigen.

Anwendungen:

- Izotope RX
- Clarity Vx

Ein **De-Reverb** reduziert oder entfernt unerwünschten Raumhall aus einer Audioaufnahme, indem es die zeitlichen und spektralen Eigenschaften des Signals analysiert. Mithilfe von Algorithmen für Signalverarbeitung und maschinelles Lernen trennt es den direkten Klanganteil von den Reflexionen und unterdrückt diese gezielt, ohne die natürliche Klangqualität zu stark zu beeinträchtigen.

Anwendungen:

- Adobe Audition
- Adobe Enhance Speech v2

Ein **Smart-Gate** ist ein Gate, welches mithilfe von AI, unterschiedliche Elemente eines Audiosignals wie z.B. Snare und Kick erkennt. Dadurch kann er auf bestimmte Tonquellen wie eine Snare eingestellt werden, und öffnet nur wenn der Pegel im Snare spezifischen Frequenzraum über einen gewissen Schwellenwert steigt.

Anwendungen:

- SmartGate – Sonible

2.4 Dynamic-Resonance-Suppressor

Ein Dynamic-Resonance-Suppressor wie Soothe analysiert in Echtzeit das Frequenzspektrum eines Audiosignals und reduziert automatisch störende Resonanzen, ohne das gesamte Signal zu beeinträchtigen. Durch spektrale Analyse und dynamische Filter erkennt das Plugin überbetonte Frequenzbereiche und passt die Dämpfung dynamisch an, um einen ausgewogenen und transparenten Klang zu erzielen.

Anwendungen:

- AI-Clarity
- Curves Equator
- Soothe 2
- Smooth Operator

2.5 Mix und Mastering

AI-gestütztes **Mixing** nutzt maschinelles Lernen und Algorithmen, um den Mischprozess zu automatisieren und zu optimieren. Dabei analysiert die AI verschiedene Aspekte eines Audiosignals, wie Lautstärke, Frequenzverteilung und Dynamik, und passt automatisch die Pegel, EQ-Einstellungen und Effekte an, um eine ausgewogene und professionelle Mischung zu erzeugen.

Anwendungen:

- Nectar
- Neutron
- CryoMix
- Waves Nx Virtual Mix Room

AI-gestütztes **Mastering** verwendet maschinelles Lernen, um den finalen Audio-Mastering-Prozess zu automatisieren und zu optimieren. Die AI analysiert einen einstellbaren Abschnitt der Audiodatei und passt automatisch Parameter wie Lautstärke, Dynamikbereich, EQ und Stereobreite an, um das bestmögliche Klangbild für verschiedene Wiedergabegeräte und Formate zu erzielen, wobei sie sich an professionellen Mastering-Standards orientiert.

Anwendungen:

- Ozone,
- Landr Mastering

3. Audioerzeugende AI-Tools

3.1 AI-Synthese

Audio-Synthese mit AI nutzt maschinelles Lernen, um neue Klänge zu erzeugen, die auf bestehenden Audiodaten basieren. Modelle wie Generative Adversarial Networks (GANs) oder rekurrente neuronale Netze (RNNs) lernen die Struktur und Eigenschaften von Klängen, indem sie große Mengen an Audiodaten analysieren, und erzeugen dann neue, realistische oder kreative Soundscapes, die oft auch komplexe musikalische oder akustische Merkmale aufweisen. Anwendungen wie Syntplant 2 schaffen es so aus einer Audiodatei einen modulierbares Synth Preset zu erzeugen.

Anwendungen:

- Syntplant 2

3.2 Song Generation

Song-Generation mit AI basiert auf Algorithmen, die Musikstile, Harmonien, Melodien, Struktur und Rhythmen analysieren, um neue Songs zu erzeugen. Deep-Learning-Modelle, wie rekurrente neuronale Netze (RNNs) oder Transformer-basierte Modelle, lernen aus umfangreichen Datensätzen von bestehenden Songs, um eigenständig neue musikalische Strukturen zu erstellen, die menschlichen Kompositionen ähneln oder kreative Variationen bieten. Über Prompts lassen sich Attribute wie: Gesang, Genres, Stil, bestimmte Instrumente, Geschwindigkeit oder Stimmung manuell vorgeben.

Anwendungen:

- Suno
- Loudly
- AIVA

3.3 Text to Speech

Text-to-Speech (TTS) mit KI wandelt geschriebenen Text in gesprochene Sprache um, indem es natürliche Sprachmuster und -töne simuliert. AI-Modelle, wie Deep-Learning-basierte neuronale Netze, lernen aus großen Sprachdatenbanken, um eine realistische und ausdrucksvolle Aussprache zu erzeugen. Anwendungsbeispiele sind virtuelle Assistenten wie Siri und Alexa, Barrierefreiheitstechnologien für Sehbehinderte oder Navigationssysteme, die sprachliche Anweisungen geben.

Anwendungen:

- ElevenLabs
- Google
- Canva

3.4 Midi Generation

Midi-Generation mit KI verwendet maschinelles Lernen, um musikalische Strukturen wie Melodien, Akkorde und Rhythmen in MIDI-Daten zu übersetzen, die von digitalen Musikinstrumenten oder Software-DAWs verwendet werden können. KI-Modelle lernen aus großen Datensätzen von Musikstücken, um eigene, harmonische und rhythmische Kompositionen zu erzeugen. Anwendungsbeispiele umfassen die automatische Erstellung von Begleitmusik, Unterstützung beim Songwriting oder die Integration in Musikproduktionssoftware, die kreative Impulse für Musiker liefert.

Anwendungen:

- AIVA
- Lemonaide VST
- HookPad's Aria

4. Workflow Tools

4.1 AI Finder

Mittlerweile gibt es auch AI Tools die einem helfen, mit anderen AI's besser zurechtzukommen oder diese erst zu finden. Gerade da mittlerweile konstant neue AI-Tools erscheinen und es schwer ist dabei den Überblick zu behalten.

Anwendungen:

- Futuretools
- Buzzmatic

4.2 Sample Finder

AI schlägt Samples vor, die zu deinem Projekt passen könnten, was oft viel Zeit ersparen kann und sogar als Inspirationsquelle genutzt werden kann.

Die AI analysiert dabei deine Samples und organisiert/sortiert diese nach BPM, Instrument, Style und vielen anderen Attributen.

Anwendungen:

- Jamahook

4.3 Speech to Text

Speech-to-Text (STT) mit AI wandelt gesprochene Sprache in geschriebenen Text um, indem es akustische Merkmale wie Tonhöhe, Lautstärke und Frequenzen analysiert und mit bestehenden Sprachmodellen abgleicht. Mithilfe von maschinellem Lernen und neuronalen Netzen, die auf großen Sprachdaten trainiert werden, kann die KI auch unterschiedliche Akzente, Geräusche und Kontextinformationen erkennen, um genaue Transkriptionen zu erstellen.

Anwendungen:

- OpenAI
- Google Assistant

5. Anwendungserfolge

“Now And Then” (Beatles)

Ein bekanntes Beispiel ist der Beatles-Song "Now and Then", bei dem AI verwendet wurde, um die Stimme von John Lennon aus alten, verrauschten Aufnahmen zu isolieren und klanglich zu verbessern. So konnte der Song trotz der schlechten Aufnahmesituation in den 1970er Jahren in hoher Qualität veröffentlicht werden.

„Verknallt in einen Talahon“

Das Werk "Verknallt in einen Talahon", ein von AI-generierter Song, hat es als erster seiner Art in die deutschen Charts geschafft. Die AI "Suno" übernahm dabei die Komposition und Produktion des Songs, was uns zeigt, dass künstliche Intelligenz nicht nur alte Musik restaurieren, sondern auch neue Hits erschaffen kann. Diese Entwicklungen verdeutlichen, wie leistungsfähig AI in der Musikproduktion geworden ist.

6. AI-Erkennung

Kann der Mensch noch AI-generierte Musik von “handgemachter Musik“ unterscheiden? Zum einen verfügen viele Menschen, wie Musiker und Tontechniker, über ein geschultes Gehör, um kleine Feinheiten in Melodie, Dynamik und Ausdruck zu erkennen. Menschliche Musik kann oft emotionale Tiefe und spontane kreative Entscheidungen enthalten, die AI-Modelle bislang nur schlecht nachahmen können. Zudem neigen AI-generierte Stücke dazu, Muster aus bestehenden Songs zu übernehmen, was in vielen Fällen zu einer Vorhersehbarkeit führt.

Auf der anderen Seite sind moderne AI-Modelle, wie etwa neuronale Netze, in der Lage, hochkomplexe musikalische Strukturen nachzubilden und quasi neue Kompositionen zu erschaffen, die sich kaum von menschlichen Songs unterscheiden lassen. Tests und Experimente wie mein Quiz, welches ich in meiner Ton-Seminar Präsentation durchgeführt habe, zeigen aber, dass viele Menschen Schwierigkeiten haben, den Unterschied zu hören. In meinem Fall haben die Teilnehmer in zwei von drei der Vergleichen per Abstimmung den AI Song als echten Song identifiziert. Ob Menschen AI-Musik von menschlicher Musik unterscheiden können, hängt zwar von der Qualität der AI-Produktion und dem geschulten Gehör des Zuhörers ab, kann aber mittlerweile selbst die erfahrensten Ton-Menschen täuschen.

7. Chancen und Herausforderungen

Die Integration von Künstlicher Intelligenz in die Musik- und Audioproduktion bringt sowohl bedeutende Chancen als auch Herausforderungen mit sich. AI hat das Potenzial, kreative Prozesse nachzubilden, aber es gibt natürlich Bedenken hinsichtlich der Auswirkungen auf die künstlerische Authentizität und die Arbeitswelt.

Chancen:

Automatisierung und Effizienzsteigerung:

AI kann viele zeitraubende Aufgaben in der Musikproduktion automatisieren, wie etwa das Mischen, Mastern oder das Erstellen von Musikkompositionen. Dies ermöglicht es Musikproduzenten, schneller und effizienter zu arbeiten. Beispielsweise ermöglicht AI-gestütztes Mastering wie in iZotope Ozone eine präzise Anpassung der Audioqualität, was früher viele einzelne manuelle Eingriffe erforderte.

Kreative Unterstützung:

AI-basierte Tools können als kreative Partner dienen, indem sie neue Ideen und musikalische Strukturen generieren. Ob Midi Noten, Songtexte oder gleich ein ganzer Song durch diese neuen Tools lässt sich viel Zeit einsparen und neue Inspiration geben.

Barrierefreiheit:

Für Anfänger oder Musiker mit wenig technischer Erfahrung bieten AI-Tools wie automatisierte Mixing- und Mastering-Plattformen eine leicht zugängliche Möglichkeit,

qualitativ hochwertige Musik zu erstellen. Auch braucht es oft nicht mehr das akustisch perfekte Studio oder das teuerste Mikrofon, wenn die Audiodateien mithilfe von AI einfach aufgewertet werden können. Dies hat das Potenzial, die Musikproduktion und der damit verbundene Spaß für ein breiteres Publikum zugänglich zu machen.

Herausforderungen:

Künstlerische Authentizität und Originalität:

Ein zentrales Thema bei der Verwendung von AI in der Musikproduktion ist die Frage nach der Originalität. AI-generierte Musik kann so realistisch und stilistisch überzeugend klingen, dass es immer schwieriger wird, den Unterschied zu "menschengemachter" Musik, zu erkennen. Dies führt zu dem Problem, dass AI die kreative Vielfalt der menschlichen Kunst beeinträchtigen oder gar ersetzen könnte.

Arbeitsmarkt und Berufsrollen:

Mit dem zunehmenden Einsatz von AI in der Musikproduktion könnte die Nachfrage nach traditionellen Fähigkeiten, wie beispielsweise dem manuellen Mixing und Mastern, sinken. Das könnte Auswirkungen auf die Beschäftigung in der Musikindustrie haben, da weniger Menschen für Aufgaben benötigt werden, die von AI übernommen werden könnten. Andererseits entstehen durch neue AI-Anwendungen auch neue Berufe, die mit der Entwicklung und dem Management solcher Technologien verbunden sind. Die Arbeitsweise im Audio-Bereich wird sich dadurch in den kommenden Jahren voraussichtlich stark verändern.

Datenabhängigkeit:

AI-Systeme sind stark auf große Datenmengen angewiesen, um effektiv zu lernen und zu arbeiten. Dabei besteht die Gefahr, dass die verwendeten Datensätze besonders bestimmte Musik-Arten oder westliche kulturelle Eigenschaften widerspiegeln, die die Musikgeneration beeinflussen können. Wenn AI nur mit bestimmten Arten von Musik oder bestimmten Stilen trainiert wird, kann dies dazu führen, dass kreative Möglichkeiten eingeschränkt und weniger vielfältige Musik produziert wird.

Rechtliche und ethische Fragen:

Die Nutzung von AI in der Musik wirft Fragen hinsichtlich Urheberrechts und Eigentum auf. Wer besitzt die Rechte an Musik, die von einer KI komponiert wird? Gehört sie den Herstellern der AI-Anwendung, der Person, die die AI mit Prompts steuert oder doch den vielen Künstler*innen mit deren Werken die AI trainiert worden ist. Diese Fragen sind noch nicht vollständig geklärt und könnten zu kontroversen und rechtlichen Auseinandersetzungen führen.

Quellenangabe

1. Grundlagen der AI

1.1 Grundlegende Funktionsweisen der AI

Künstliche Intelligenz (KI) und maschinelles Lernen

<https://www.iks.fraunhofer.de/de/themen/kuenstliche-intelligenz.html>

1.2 Historische Entwicklung

Wikipedia: „Geschichte der Künstlichen Intelligenz“

https://de.wikipedia.org/wiki/Geschichte_der_kuenstlichen_Intelligenz

1.3 Gesetze

<https://digital-strategy.ec.europa.eu/de/policies/regulatory-framework-ai>

1.4 AI im Kontext von Audio

AI sound recognition: Fundamentals of audio data

<https://www.tensorflow.com/post/audio-analysis-with-machine-learning>

2. Ai Tools - Audio-Editing und -Verarbeitung

2.1 Stem Separation

STEM SEPARATION, FUNKTIONSWEISE, TOOLS & SOFTWARE

<https://www.amazona.de/feature-stem-separation-funktionsweise-tools-software/>

2.2 Voice Changer

Das Arbeitsprinzip von Voice Changer

<https://multimedia.easeus.com/de/voice-changer-tipps/wie-funktioniert-ein-stimmwechsler.html#:~:text=Das%20Arbeitsprinzip%20von%20Voice%20Changer,-Stimmveränderer%20sind%20faszinierende&text=Sie%20funktionieren%20durch%20Manipulation%20der,Stimme%20auf%20einzigartige%20Weise%20verändern.>

2.3 Audio Repair

Ai Voice Repair Techniques

<https://www.restack.io/p/ai-voice-repair-answer-cat-ai>

2.4 Dynamic-Resonance-Suppressor

WHAT IS A DYNAMIC RESONANCE SUPPRESSOR? HOW TO FIX HARSH FREQUENCIES IN YOUR MIX

<https://babyaud.io/blog/dynamic-resonance-suppressor#:~:text=A%20dynamic%20resonance%20suppressor%20is,balance%20of%20the%20original%20material.>

2.5 Mix und Mastering

Mixing AI: Abmischen mit dem Einsatz von künstlicher Intelligenz (KI)
<https://tonstudio-wissen.de/abmischen-kuenstlicher-intelligenz/>

3. Audioerzeugende AI-Tools

3.1 AI-Synthese

<https://cdm.link/the-road-to-synplant-2/>

3.2 Song Generation

CREATING MUSIC WITH AI MUSIC GENERATORS

<https://soundful.com/de/how-to-guide-creating-music-with-ai-music-generators/#:~:text=AI%20music%20generators%20are%20software,songs%20depending%20on%20their%20capabilities.>

3.3 Text to Speech

Text-To-Speech (TTS)

<https://mindsquare.de/knowhow/text-to-speech/>

3.4 Midi Generation

The Complete Guide to AI MIDI Generators in 2025

<https://blog.staccato.ai/leading-ai-midi-tools-2023>

4. Workflow Tools

4.1 AI-Finder

<https://www.futuretools.io/>

4.2 Sample-Finder

<https://www.waves.com/plugins/cosmos-sample-finder>

4.3 Speech to Text (StT)

What is speech to text?

<https://www.ibm.com/think/topics/speech-to-text>

Abbildungsverzeichnis

<https://www.trail-ml.com/blog/eu-ai-act-how-risk-is-classified>