

Composing AI

Formen, Ideen, Visionen

Miriam Akkermann

Abstract: *Die Vorstellung, dass Musik mithilfe von Computern erzeugt werden kann, fasziniert Komponierende seit Mitte des 20. Jahrhunderts. So hat die Einbindung entsprechender Techniken entgegen dem nun neu entfachten Interesse an KI in der Musik bereits eine längere Tradition. Gerade der Einsatz von Machine Learning-Techniken gibt den bekannten Visionen jedoch derzeit neuen Antrieb, verbindet eine rein KI-generierte Musik doch verschiedene Ziele, Wünsche und Erwartungen sowie eine Neugierde von Komponierenden und Publikum. Demgegenüber stehen eine mehr oder weniger komplexe und responsive technische Anlage der Programme sowie sehr unterschiedliche KI-Verfahren, deren realer Spielraum nicht immer den zugeschriebenen Freiräumen entspricht. In dem Artikel werden an drei Beispielen verschiedene Formen der Einbindung von KI-Verfahren in musikalische Arbeiten vorgestellt und angelehnt an eine Kategorisierung unterschiedlicher Systemkonzeptionen wird kurz diskutiert, welche ›Freiheiten‹ in KI-basierten musikalischen Arbeiten implementiert und welche zugeschrieben sind.*

Einleitung

Die Einbindung von Prozessen aus dem Bereich ›Künstliche Intelligenz‹ in musikalische Arbeiten scheint – zumindest mit Blick auf die Kombination der Schlagworte ›Musik‹ und ›KI‹ – ein eher junges Phänomen zu sein. Dies täuscht darüber hinweg, dass sich Komponistinnen und Komponisten bereits seit Mitte des 20. Jahrhunderts mit (Rechen-)Verfahren aus dem Forschungsbereich der ›Künstlichen Intelligenz‹ befassen und diese in verschiedenster Art und Weise in ihren künstlerischen Arbeitsprozessen verwenden. Bieten regelbasierte generative Prozesse umfassende Möglichkeiten, um eine schier

endlose Vielfalt an Ergebnissen zu produzieren, so üben gerade Ansätze eine große Faszination aus, die analytische und generative Elemente verbinden, versprechen diese doch nicht nur eine große Anzahl an verschiedenen, sondern viele, den Anforderungen angemessene Ergebnisse und ein Wissen über bis dahin vielleicht verborgene Regeln, die aus der nun größeren Anzahl an untersuchten Exempeln herausgearbeitet werden können. Die Einbindung solcher kombinierter Prozesse bringt jedoch auch verschiedenste neue Herausforderungen mit sich, die von sehr praktischen Fragen, beispielsweise nach den Ein- und Ausgabeformaten, bis hin zu komplexen Debatten über Handlungs- und Entscheidungszuschreibungen reichen und die implizit immer wieder um die dem Bereich KI ebenso wie dem der Musik inhärenten Frage nach Autor:innenschaft bzw. Autonomie und Einflussnahme einzelner Personen kreisen. Denn in der Beschreibung oder Präsentation der Musikstücke verschwimmen – gewollt oder ungewollt – oft die Handlungszuordnungen, insbesondere wenn Computerprogramme als Teil der Komposition eingebunden werden. Doch was kann ein KI-basiertes Programm in der Musik leisten? Wie sind diese Programme aufgebaut und in den kompositorischen Prozess sowie die Aufführungsumgebungen eingebettet? Und welche Handlungszuschreibungen, Erwartungen und Wünsche gibt es hinsichtlich der künstlerischen Autonomie und Kreativität?

Basierend auf einer schlaglichtartigen historischen Annäherung werden im Folgenden exemplarisch Formate, Ideen und Visionen vorgestellt, die zwar kein umfängliches Bild möglicher KI-Einbindungen in musikalische Arbeiten zeichnen, mithilfe derer es aber möglich ist, den Stellenwert von KI-Ansätzen in musikalischen Kompositionen differenziert zu beleuchten. Dies bietet wiederum einen Ausgangs- und Anknüpfungspunkt für eine weiterführende Debatte über KI in künstlerischen Prozessen und lädt zu einer Reflexion darüber ein, welche Unterscheidungen bei der Einbindung von KI-Konzepten möglich sind und welche Zuschreibungen sinnvoll sein können.

Frühe KI-Konzepte in kompositorischen Prozessen. Drei Beispiele

Kompositionen, die im weitesten Sinne Methoden aus dem KI-Bereich einbinden, werden aus musikwissenschaftlicher Perspektive im 20. Jahrhundert zumeist als Algorithmische Kompositionen bezeichnet und der Elektroakustischen Musik bzw. (heute vornehmlich) der Computermusik zugeordnet. Wurden bis in die 1990er Jahre vor allem statistische und stochastische Pro-

zesse sowie nicht-lernende generative Regelsysteme in die Arbeitsprozesse implementiert, beispielsweise zur sogenannten Partitursynthese (Erstellung von Notation) sowie zur unmittelbaren Klangerzeugung, kommen heute auch komplexe generative Verfahren in Musikanwendungen zum Einsatz, die auf Deep-Learning-Methoden basieren. Dies erfolgt oft mit dem Ziel, eine möglichst autonome, also vom Menschen unabhängig ablaufende Musik- bzw. Klangerzeugung zu ermöglichen. Zudem versprechen die neuen ebenso wie die frühen Techniken die Produktion einer größeren Vielfalt an Varianten als die, die ein Mensch im gleichen Zeitraum produzieren könnte. Die Zielstellungen, unter denen die verschiedenen Verfahren eingebunden werden, sind dabei ebenso unterschiedlich wie der Umfang, in dem die Komponistinnen und Komponisten in den Synthese- und Selektionsprozess eingreifen. Die Selektion, also die Einordnung und entsprechende Auswahl (zusammen)passender bzw. gewünschter Elemente aus den entstehenden (klingenden) Möglichkeiten, ist hierbei eine der zentralen Herausforderungen, die sich durch die gesamte Bandbreite musikalischer Kreationen mit generativen Verfahren zieht.

Eine der ersten musikalischen Arbeiten, bei der eine Einbettung früher KI-Techniken erfolgte, ist die *Illiac Suite* (1955–1956) von Lejaren A. Hiller und Leonard M. Isaacson (vgl. Hiller und Isaacson 1958). Dieses Stück, das auch als die erste Computermusikkomposition angesehen wird, basiert auf einer computergestützten Partitursynthese (vgl. Supper 1995a: 967). Mithilfe des Computerprogramms *ILLIAC I* wurde die Notation für ein Streichquartett erstellt, wobei jedem der vier Sätze der Suite ein anderes Experiment zugrunde gelegt ist, mit dem Hiller und Isaacson erproben wollten, »wie mit einem Computer eine Komposition generiert werden kann« (ebd.: 972). Der erste Satz folgt 16 verschiedenen Regeln des Kontrapunkts; der zweite erweitert diese auf eine vierstimmige Polyphonie und die Möglichkeit, verschiedene Stile zu erhalten; im dritten Satz finden serielle Verfahren Verwendung, um Rhythmus- und Dynamikelemente zu erzeugen, wohingegen die Parameter Tempo und Metrum von den Komponisten bestimmt sind; im vierten Satz bestehen die »Kompositionsregeln [...] [aus] einer Folge abhängiger Zufallsgrößen« (Supper 1995b: 74); die musikalischen Regelwerke werden hierbei durch sogenannte Markov-Ketten ersetzt, ein stochastisches Rechenverfahren, das dazu verwendet wird, Aussagen über das Eintreten eines Ereignisses in der Zukunft in Abhängigkeit von der Gegenwart zu treffen (vgl. Akkermann 2015: 31). Jede Note wird damit entsprechend ihrer Wahrscheinlichkeit in Abhängigkeit von der vorangegangenen Note berechnet. Die Erstellung der Noten erfolgte durch das Computer-

programm jeweils in einzelnen Experimenten, wobei mithilfe des Programms nicht nur eine einzelne Komposition, sondern eine ganze Klasse an Kompositionen berechnet werden konnte, die jeweils einem entsprechenden Stil folgten (vgl. Supper 1995a: 974). Die finale Selektion der generierten Stimmen und ihre Zusammenstellung in den jeweiligen Sätzen der Suite oblag jedoch den Komponisten und auch die Interpretation der Noten ist an Musiker:innen gebunden.

Während in der Illiac Suite weiterhin Menschen die computergenerierten Aufzeichnungen zum Klingen bringen, umfasst George E. Lewis' *Voyager* eine Kombination aus (digitalen) Klanganalyse- und Klangausgabeelementen, die es erlauben, (hörbaren) Klang zu erfassen, auf Regeln hin zu analysieren sowie dessen Erzeugung zu steuern. Das interaktive System *Voyager*, das 1987 erstmals auf der Bühne präsentiert wurde, ist auf eine Ko-Improvisation mit einem menschlichen Improvisierenden ausgelegt. Hierbei wird die erklingende Improvisation des Menschen mithilfe eines Pitch-to-MIDI-Trackings in Echtzeit in Tonhöheninformation umgewandelt und auf musikalische Regeln hin analysiert, um dann eine Art ›Antwort‹ darauf zu erzeugen. Dies geschieht basierend auf einem sogenannten *set-phrasе-behaviour*, einem kombinierten Regelset, das sich aus zwei Elementen zusammensetzt: den aus dem zu hörenden Klang herausgearbeiteten Regeln sowie einem eigenen, unabhängigen Regelwerk, das nicht auf den analysierten Klang Bezug nimmt und innerhalb dessen von sogenannten zellulären Automaten Klangstrukturen produziert werden (vgl. Lewis 2000). Die Problematik, dass digitale Berechnungen von Klängen als Klangsynthese in Echtzeit in den 1980er Jahren noch eine größere technische Herausforderung darstellten, da sie an umfangreiche Rechereinheiten mit limitierten Berechnungskapazitäten gebunden waren (vgl. Akkermann 2020: 125ff.), wurde umgangen, indem die Ausgabe der berechneten Tonhöheninformation per MIDI erfolgte, was die Klangausgabe mittels eines MIDI-gesteuerten (akustischen) Instruments (Disklavier) erlaubte – eine Lösung, die bis heute noch genutzt wird. *Voyager* ist damit (technisch) in der Lage, verschiedene Regelsets aus Klängen – Improvisationen – herauszuarbeiten und nachzuahmen sowie genuin eigene Klangsequenzen zu produzieren, die auf inhärenten Regelwerken beruhen. Lewis greift hierbei als Komponist selbst nicht aktiv in die Klangproduktion bzw. die Selektion der produzierten Sequenzen ein, sondern das Programm ist so konzipiert, dass es während einer Aufführung als eigenständige improvisierende Einheit in Interaktion mit einem Menschen musiziert. Eine Limitierung der Klang-erzeugung ist nur durch die Auswahl des per MIDI gesteuerten Instruments

gegeben – dieses ist zumeist, wie auch bei dem Konzert 2020 in Berlin, ein Disklavier.

Ausgehend von der Idee, ein Computerprogramm zu erstellen, das einen ganz persönlichen musikalischen Stil (erkennen und er-)lernen kann, um dann – quasi mit ihm zusammen – neue Stücke zu erarbeiten, entwickelte der Komponist David Cope ab den 1980er Jahren das Programm *Experiments in Musical Intelligence* (EMI), auch *Emmy* genannt (vgl. *Computer History Museum* o.J.). Dieses generative System, das eigenständig beim Kompositionsprozess, nicht aber bei der Ausführung einer Improvisation mitarbeiten sollte, ist darauf ausgelegt, mithilfe von per MIDI getrackten Inhalten Regelsets abzuleiten und dann – ebenfalls per MIDI – notierte Musiksequenzen zu produzieren. Cope (1999: 79) beschreibt den Ablauf wie folgt:

Experiments in Musical Intelligence composes by first analyzing the music in its database and then using the rules it discovers there to create new instances of music in that style.

EMI folgt damit im weitesten Sinne einem Prinzip, das auch heute bei KI-basierten Verfahren zum Einsatz kommt: Cope wählt bestimmte Kompositionen aus (database), aus denen von EMI jeweils die gemeinsamen Regeln herausgearbeitet werden (EMI ›lernt‹ die Regeln), die dann wiederum als Grundlage für die Berechnung von Tonhöhen und Tondauern dienen (die Regeln werden angewandt, um neue Musik im selben Stil zu produzieren). Die hierbei inhärente ›musikalische Intelligenz‹ sieht Cope in der Simulation musikalischen Denkens. Es wird, wenn man so möchte, ein Regelset erstellt, das die (charakteristischen) Merkmale eines Kompositionsstils, die aus einem gewählten Kompositionsdatensatz resultieren, nachstellen kann. Anders gesagt: Charakteristische Handlungskomponenten der Komponierenden werden auf komplexe Regelsets innerhalb von EMI übertragen (vgl. Cope 1992: 69). Dadurch entstehen musikalische Sequenzen, die einem vorgegebenen musikalischen Vorbild folgen, also bestimmte musikalische Charakteristika reproduzieren, ohne dabei jedoch exakte Kopien der Vorlagen zu sein. Die Wahl der im Datensatz enthaltenen Musikstücke kann hierbei stark das Resultat beeinflussen: Je ähnlicher die im Korpus enthaltenen Stücke sind und je prägnanter die für einen Stil als charakteristisch erachteten Elemente in Erscheinung treten, umso besser entspricht das Ergebnis den impliziten Vorgaben. Neben Copes eigenen Kompositionen dienen insbesondere Arbeiten bekannter Komponisten aus früheren Jahrhunderten als Vorlage für Musikproduktionen von EMI. Hierbei entstanden Musikstücke im Stil von unter anderem Béla Bartók, Johannes Brahms,

Johann Sebastian Bach und Frédéric Chopin, aber auch Wolfgang Amadeus Mozart und Sergei W. Rachmaninow. Während einige Stücke von Menschen aufgeführt wurden, ließ Cope andere per Disklavier erklingen, insbesondere dann, wenn die generierte Notation für Menschen unspielbar schien. Copes Rolle ist hierbei (immer) weniger die eines Komponisten, sondern eher die eines Programmentwicklers und Organisators, der den Inhalt der Datenbank wählt, die Resultate einem angemessenen Klangkörper – Musiker:innen oder Disklavier – zuführt und die entsprechenden Klangerzeugnisse präsentiert, erklärt und ggf. vermarktet. Das initiale Interesse von Cope scheint sich hierbei im Laufe der Zeit zu verlagern, weg von einem Programm, das ihn selbst unterstützt, hin zu einem, das musikalische Eigenheiten erkennen, erlernen und in ähnlicher Art und Weise (in Notation) produzieren kann.

Formen der Einbindung

Die angeführten Beispiele zeigen exemplarisch und in Abstufung drei grundlegende Einsatzrichtungen auf, mit denen Rechenprozesse in die Erarbeitung von Musikstücken eingebunden werden können:

- zur Analyse von vorhandenen Kompositionen (Notation) oder Audio bzw. live gespielten Klängen: diese werden in Echtzeit oder in vorgelagerten Prozessen auf Tonhöhen, Tonabfolgen und formale Strukturen hin analysiert, um daraus Regeln oder Muster abzuleiten;
- zur Partitursynthese: im kompositorischen Arbeitsprozess wird die Notation nach jeweils festgelegten Prozessen erzeugt; und
- zur Klangsteuerung bzw. -erzeugung: basierend auf Regeln, Mustern oder Notation werden Tonhöhenabfolgen (zumeist im Aufführungskontext) zum Klingen gebracht, beispielsweise mithilfe von digital gesteuerten Instrumenten oder direkt in Form von Klangsynthese.

Diese Formen der Einbindung können mit unterschiedlichen Rechenprozessen und in verschiedenen Kombinationen eingesetzt werden, wodurch Kompositionen entstehen, die sich – kompositorisch oder klanglich – auf algorithmische Prinzipien stützen, auch wenn dies auf den ersten Blick nicht immer zu erkennen ist. Die Komplexität der hierbei implementierten Rechenprozesse ist von der Art der Einbindung unabhängig und kann von einfachen Regelwerken bis hin zu ineinandergreifenden Systemen reichen.

Philippe Pasquier, der wissenschaftlich und künstlerisch zu Künstlicher Intelligenz forscht und auch selbst KI-basierte Systeme für musikalische Arbeiten entwickelt, unterscheidet mit Blick auf die Einbindung von generativen Prozessen in musikalische Arbeiten fünf Systemkonzeptionen nach deren praktischer Anlage:

- Systeme, die mit einem initialen Input autonom ablaufen, und solche, die keinen initialen Input benötigen;
- Systeme, die in Echtzeit ablaufen, und solche, die dies nicht tun;
- Systeme, die nach der Art des ›Wissens‹ – den Informationen, die den Rechenprozessen zugrunde liegen – gegliedert sind;
- Systeme, die nach der Art ihres Inputs bzw. Outputs – in Form von Zeichen/Notation, Audio oder hybrid – kategorisiert werden; sowie
- nichtkorpusbasierte Systeme, die ihre Ausgabe generieren, ohne musikalische Informationen als Eingabe zu erhalten, und korpusbasierte Systeme, die Wissen aus einem bereitgestellten Korpus von Musikstücken oder Ausschnitten aus musikalischen Arbeiten extrahieren (vgl. Pasquier et al. 2017: 6).

Diese grundlegenden Systemkonzeptionen beziehen sich auf unterschiedliche Umsetzungsebenen und umfassen die Kombination aus bestimmten Programmen und den im gesamten Set-up angelegten Interaktionsmöglichkeiten, wie auch in den vorgestellten Beispielen gut zu sehen ist. Während bei allen drei Beispielen die zugrunde liegenden Programme eines initialen Inputs in Form von Voreinstellungen oder vorgegebenen Inhalten bedürfen, ist nur Voyager in Echtzeit im Einsatz. Voyager greift wiederum zum einen – wie EMI – auf ein »Wissen« aus Klang bzw. Musikstücken zurück, und zum anderen sind in das Programm – vergleichbar mit ILLIAC I – eigenständige Regelwerke zur Generierung von Klangsequenzen eingebettet.

Etwas weniger eindeutig wird die Unterscheidung nach Input und Output: Zwar ist ILLIAC I insofern klar zu verorten, als das Programm einen symbolischen Output generiert, so bieten EMI und Voyager jedoch technisch MIDI-Schnittstellen zur Ein- und Ausgabe, wenn auch mit unterschiedlicher Intention. EMI bietet die Möglichkeit, die in MIDI formatierte Notation über ein MIDI-gesteuertes Musikinstrument wiedergeben zu lassen, das intendierte Ergebnis sind jedoch notierte Musikstücke, die mithilfe von MIDI erstellt werden. Bei Voyager hingegen wird MIDI als Brücke zwischen der Klingeingabe (Pitch-to-MIDI-tracking) und der Klangusgabe (z.B. MIDI-Piano) verwen-

det. Intendierter Input wie auch Output sind hier Audio bzw. hörbarer Klang. Dies ist in Pasquiers Typologie, in der zwischen In- und Output diskreter symbolischer Natur (zum Beispiel Musiknotation, MIDI-Dateien oder MIDI), Audio (in Form einer Tonspur oder Klang in Echtzeit) und hybriden Formen unterschieden wird, nicht präzise abbildbar. EMI und Voyager sind nach ihrer technischen Anlage beide Systeme mit symbolischem In- und Output – auch da ein direktes Pitch-Tracking ohne den Umweg MIDI in den 1980er Jahren, also zur Entwicklungszeit von Voyager, nicht problemlos in Echtzeit umgesetzt werden konnte. Dieses (technische) Problem stellt sich heutzutage nicht mehr, denn iterative wie interaktive Systeme können mit unterschiedlichen In- und Outputformaten umgesetzt werden. In Michael Youngs Piano Prosthesis (2008) wird beispielsweise der Audioinput direkt analysiert und mit einem komponierten generativen Musiksystem und Machine-Learning-Prozessen verbunden, um daraus sowohl Regeln abzuleiten als auch neue Klänge zu produzieren (vgl. Blackwell et al. 2012: 170). Ein in diesem Sinne hybrides System, das beide Arten als In- und Output verarbeiten kann, ist beispielsweise das unter anderem von Shlomo Dubnov an der University of California San Diego (UCSD) und am Institut de Recherche et Coordination Acoustique/Musique (IRCAM) entwickelte Audio Oracle, ein Meta-Creation-System, das dafür eingesetzt werden kann, verschiedene Variationen einer Audioaufnahme zu erzeugen (vgl. Dubnov et al. 2011).

Pasquier unterscheidet zudem zwischen zwei generellen Grundanlagen hinsichtlich der Beschaffenheit des Inputs: nichtkorpusbasierte und korpusbasierte Systeme. Nichtkorpusbasierte Systeme, wie sie durch die Einbindung der Programme ILLIAC I und Voyager entstehen, sind darauf angewiesen, dass vor ihrer Nutzung (dynamische oder andere) Werte verschiedenen Parametern zugeordnet werden. Bekannte nichtkorpusbasierte Systeme sind beispielsweise die algorithmischen Kompositionssysteme Projekt 1 und Projekt 2 des Komponisten Gottfried Michael Koenig. Für diese Systeme ist charakteristisch, dass die Personen, die die initialen Werte eingeben, mit dieser Eingabe immer auch Teile ihres jeweiligen musikalischen Wissens und/oder ihrer ästhetischen Beurteilungen beisteuern und somit den nachfolgenden Prozess – implizit und explizit – beeinflussen.

Korpusbasierte Systeme wie Copes EMI sind dagegen nicht zwingend an das Wissen der programmnutzenden Person gebunden, sondern benötigen eine (angemessene) Anzahl an Musikinformationen, entweder in Form von symbolischer Notation oder in Form von Audiodaten, um neue, den vorgegebenen Musikstücken ähnliche Ergebnisse zu produzieren. Während in nichtkorpus-

basierten Systemen die Ergebnisproduktion also von den formalisierten Angaben (z.B. Wertauswahl) einzelner Personen abhängig ist, sind in korpusbasierte Systeme vorprogrammierte Algorithmen implementiert, die auf der Grundlage bereits entstandener oder live entstehender Musik arbeiten, welche wiederum für diesen Zweck ausgewählt bzw. zusammengestellt wird.

Diese grundlegende Unterscheidung zwischen der Anlage nichtkorpusbasierter und korpusbasierter Systeme spiegelt sich, wie der Informatiker Alexander Waibel (2021) darlegt, implizit auch in der Unterscheidung zwischen dem Verständnis von ›Artificial Intelligence‹ (AI) der »Dartmouth Conference« 1956 und den heute häufig synonym verwendeten Begriffen ›Machine Learning‹ und ›Neural Networks‹ (›Deep Learning‹) wider. So sind für Waibel bei den algorithmischen Systemen, die in den 1950er Jahren diskutiert wurden, bereits Aspekte wie Automatisierung, Selbstoptimierung, Abstraktion, Zufälligkeit und Kreativität zu finden. Während in den 1980er Jahren sowohl statistische wie auch neuronale Machine-Learning-Ansätze entstanden, die aufgrund der noch beschränkten Rechenleistung vergleichbare Ergebnisse erzielten, erlauben die ›Neural Networks‹ ab den 2010er Jahren aufgrund der nun zur Verfügung stehenden Rechnerleistung die Verarbeitung großer Datenmengen (vgl. Waibel 2021). Dies macht sie gerade für korpusbasierte Systeme attraktiv.

Entsprechende Entwicklungen sind auch im Bereich der Musik zu sehen. So werden zwar weiterhin neue nichtkorpusbasierte algorithmische (Kompositions-)Systeme entwickelt, jedoch sind in den im 21. Jahrhundert entstehenden korpusbasierten Systemen zumeist neuronale Netzwerke und keine statistischen Netzwerke mehr zu finden. Beliebt sind hier insbesondere ›Recurrent Neural Networks‹ (RNNs), die nicht nur einmal mit dem vorgegebenen Korpus arbeiten, sondern nach dem Durchlaufen aller vorgegebenen Inhalte auf sich selbst rekurren und so mehrmalige Durchläufe vornehmen. Erzeugt wird dabei eine Art von Gedächtnis, da bereits gelernte Regeln nicht wieder vergessen (überschrieben) werden, sondern durch die wiederholten Durchläufe immer wieder ›in Erinnerung gerufen‹ und erneut angewandt werden können (vgl. dazu Abolafia 2016).

Ideen und Visionen

Die Art und Weise, wie Prozesse aus den Bereichen KI, besonders Verfahren von ›Machine Learning‹, in musikalische Systeme eingebettet werden, ist ebenso vielfältig wie die Ideen, Annahmen und Erwartungen, die mit einer

solchen Einbindung einhergehen. Für Pasquier unterscheidet sich beispielsweise »computational creativity« als programmiertechnischer Ansatz dadurch von anderen rationalen, problemlösungsorientierten KI-Prozessen, dass es sich hierbei häufig um Probleme handelt, für die zuvor keine optimale Lösung definiert ist (vgl. Pasquier et al. 2017: 2). Es geht also nicht darum, ein Problem möglichst schnell oder effizient zu lösen, sondern auch darum, dass die Ergebnisse eine spezifische Qualität in Bezug auf die Anforderungen aufweisen – selbst wenn dies ggf. mehr Rechenaufwand erfordert. Pasquier wie auch sein Forschungskollege Kivanç Tatar beziehen sich dabei auf den Soziologen Herbert A. Simon, der KI im Kontext von Entscheidungen als »the science of having machines solve problems that do require intelligence when solved by humans« (Simon 1960; zitiert nach Tatar 2019: 56) definiert.

Dabei, so scheint es, überschneiden sich nun zwei unterschiedliche Zielvorstellungen im Hinblick auf Anwendungen in der Musik: zum einen der Wunsch, Prozesse, die in ihrer Ausführung zu komplex oder zu langwierig sind, um sie per Hand von einem Menschen umzusetzen, auf Computerprozesse auszulagern und dadurch eine Fülle an Möglichkeiten zu generieren, aus der der Mensch dann auswählen kann; und zum anderen das Interesse daran, Prozesse zu entwickeln, die so angelegt sind, dass sie – mehr oder weniger selbstständig – die Balance zwischen einer geforderten Regeltreue und einer erwarteten kreativ-künstlerischen Freiheit einhalten können und dabei (autonom) Ergebnisse produzieren, die so auch von einem Menschen hätten erzeugt werden können. Beide Vorstellungen beinhalten die Idee, dass mithilfe der Systeme Ergebnisse erzeugt werden, die in den vom Menschen a priori angelegten Bewertungsrahmen passen. Bei der zweiten Zielvorstellung schwingt zudem die Vision mit, dass ein Ergebnis entsteht, bei dem der Mensch keine weiteren Selektionen oder Anpassungen mehr vornehmen muss, das aber dennoch menschengemachten Ergebnissen qualitativ ebenbürtig ist.

Der Vision von einer weitgehend autonom ablaufenden, generativen Musikproduktion wird aktuell in vielen Projekten nachgespürt. Unter den aktuellsten Projekten stechen zwei hervor: Beethoven X (2021) und I'll Marry You, Punk Come (2019/20), ein Projekt, das zum Gewinnersong des AI Song Contest 2020 gekürt wurde.

Beethoven X. The AI Project verfolgte das Ziel, die 10. Sinfonie Ludwig van Beethovens, die dieser aufgrund seines Todes nicht selbst fertigstellen konnte, mithilfe von dafür entwickelten korpusbasierten Systemen zu vollenden. Die verfügbaren Sinfonien Beethovens sollten als Basis dienen, um daraus cha-

rakteristische Kompositionstechniken herauszuarbeiten und darauf aufbauend die bisher unvollendete Sinfonie fertig zu komponieren. Dieses Projekt, das verschiedenste Disziplinen – unter anderem Komposition, Musikwissenschaft und Informatik – zusammenbrachte und in Kooperation großer Institutionen und Unternehmen, darunter die Rutgers University, Google, Telekom und BMG, umgesetzt wurde, stellte mit großem medialen Aufwand heraus, welche Möglichkeiten für ein solches Unterfangen heutzutage zur Verfügung stehen. Es fällt jedoch auf, dass das Projekt zwar medial in Form von Trailern und einer Webseite präsentiert und viel Werbung für das Ergebnis – ein Konzert und ein Tonträger – gemacht wurde, die breitere Öffentlichkeit jedoch keinen Einblick in die Projektabläufe bekommt und auch über den Aufbau des verwendeten Systems sowie die darin implementierten (Selektions-)Prozesse und Lernmethoden sehr wenig bekannt ist.

Ein ähnliches Ziel, nämlich die automatische Erstellung eines Musikstücks und dessen (medienwirksame) öffentliche Präsentation, wurde auch in dem von Magenta 2019/20 ausgeschriebenen AI Song Contest 2020 verfolgt, bei dem nur Musikstücke eingereicht werden durften, die rein generativ und ohne menschliches Zutun entstanden. Ausgeschrieben von Magenta, eine seit 2016 bestehende Forschungsgruppe bei Google, die Tools und Plug-ins im Bereich ›AI Music Creation‹ entwickelt und diese kostenfrei zur Verfügung stellt, werden im Rahmen des Contests aber nicht nur die finalen Musikstücke präsentiert und prämiert, sondern auch die verwendeten Prozesse und Iterationen wurden in die Bewertung miteinbezogen; Eingriffe durch Selektionsprozesse und ein finales Editing der Klangergebnisse waren durch das Anforderungsprofil des Contests ausgeschlossen. Eine Jury und die über die Magenta-Webseite abstimmende allgemeine Hörer:innenschaft kürten I'll Marry You, Punk Come vom Team Dadabots x Portrait XO, bestehend aus CJ Carr, Zack Zukowski und dem Team Portrait XO, zum Gewinnersong. Das ursprünglich in den USA angesiedelte Projekt Dadabots, das 2012 initiiert wurde und bereits mit einem 24/7 streamenden generativen KI-Metal-Kanal auf YouTube auf sich aufmerksam machen konnte, arbeitete sowohl direkt mit Audiodateien als auch mit Konzepten aus text-to-speech-Anwendungen, darunter RNNs mit ›unsupervised learning‹. So entstehen in dem Song immer wieder Passagen, die stilistisch klar verortbar sind, wenngleich im Text, der ebenfalls generiert ist, wie auch in der Musik einige Wendungen durchaus in ihrer (stilistischen) Abfolge überraschen. I'll Marry You, Punk Come steht damit sinnbildlich für das, was oft bei der Einbettung von KI in Musik imaginiert wird, nämlich für ein Musikstück, das von einer artifiziellen Entität,

die im Sinne Pasquiers kreativ und in gewisser Weise auch autonom, also unabhängig von menschlichem Zutun oder Eingreifen, operiert, produziert wurde.

Reflexion

In den voranstehend angeführten Beispielen wird deutlich, dass in den meisten Fällen, in denen KI-Prozesse in Musikproduktionen implementiert werden, diese inhärenten, von Menschen gesteuerten Selektionsmechanismen unterliegen. Nicht nur müssen die hierbei verwendeten Prozesse auf verschiedene Arten und Weisen auf möglichst viel oder möglichst präzises Wissen von Menschen zugreifen, sondern es stellt sich dabei auch das Problem, dass eine KI nicht auf derart abstrahierte Bewertungskriterien zurückgreifen kann, wie Menschen es tun. Die Vision, durch die Einbettung von KI-Prozessen in kompositorische oder interpretative Abläufe Ergebnisse zu erhalten, die menschliches Handeln und menschliche Entscheidungen ersetzen, ist damit noch nicht erreicht. Selbst die Möglichkeit, große Datenmengen mithilfe Künstlicher Intelligenz verarbeiten zu können, ist laut Alexander Waibel (2021) zwar ein zentrales Alleinstellungsmerkmal, aber auch eine zentrale Unzulänglichkeit, denn im Gegensatz zum Menschen verfüge eine KI nicht über die Fähigkeit, von wenigen Beispielen zu lernen und aus den derart gewonnenen Einsichten abstrakte Schlüsse mit Blick auf eine Gesamtheit zu ziehen. Was können KI-basierte Ansätze in der Musik also leisten? Sind klare, systematisch erfassbare Ähnlichkeiten in Musikstücken vorhanden, so können korpusbasierte Systeme aus diesen Stücken durchaus generative Regeln ableiten, die zu ähnlichen klanglichen Ergebnissen wie im Auswahlkorpus führen. Ebenfalls können generative Systeme auf der Grundlage statistischer Berechnungen einer bestimmten Klanglichkeit ähnliche Höreindrücke produzieren. Jedoch korrespondieren die Zuschreibungen dessen, was ein ›Computer‹ kann und welche ›Freiheiten‹ eine KI hat, nicht zwingend mit den implementierten Regeln oder dem eingebetteten Wissen; vielmehr werden sie den Systemen von außen angetragen, was durch eine entsprechende Präsentation seitens der Erstellenden noch forciert werden kann. Sind die Mechanismen versteckt oder derart komplex, dass sie nicht (mehr) einfach nachvollzogen werden können, so scheinen Handlungszuschreibungen auch einfache Erklärmodelle zu sein, in denen sich Konzeption und Umsetzung mit Erwartungen und Wünschen hinsichtlich der Autonomie und der künstlerischen Kreativität dieser Systeme

mischen. In anderen Worten: angesichts zu großer Komplexität wird (gerne) zu – einerseits einfachen und andererseits erwünschten – Erklärungen gegriffen, bei denen der Maschine oder den Systemen Autonomie und künstlerische Kreativität zugeschrieben wird. Inwieweit es dabei wirklich möglich ist, alle implementierten Prozesse zu verstehen, ist ebenso zu diskutieren, wie die Auswirkungen, die es haben kann, wenn in größerem Maße Musik produziert wird, die bereits vorhandenen (und für die Lernkorpora ausgewählten) Musikstücken immer ähnlicher wird – eine Frage, die sowohl eine neue Debatte über Kanonbildung und ästhetische Werte eröffnet als auch erneut den Stellenwert der Selektion hervorhebt, die mit einer Einordnung der Inhalte (als passend oder unpassend) immer einhergeht. Dies zeigt noch einmal deutlich, dass die Einordnung und auch die finale (ästhetische) Bewertung eines Musikstücks weiterhin den (zuhörenden oder adressierten) Menschen obliegt, zumindest sofern die Musik, die entsteht, ein menschliches Publikum als Bewertungsinstanz intendiert.

Literatur

- Abolafia, Dan. 2016. A Recurrent Neural Network Music Generation Tutorial. <https://magenta.tensorflow.org/2016/06/10/recurrent-neural-network-generation-tutorial>. Zugegriffen: 22. November 2021.
- Akkermann, Miriam. 2015. Instrument oder Komposition? David Wessels Contacts Turbulents. In *Reflexion – Improvisation – Multimedialität. Kompositionsstrategien in der Musik des 20. und 21. Jahrhunderts*, Hg. Christian Storch, 95–108. Göttingen: Universitätsverlag Göttingen.
- Akkermann, Miriam. 2017. Zwischen Improvisation und Algorithmus. David Wessel, Karlheinz Essl und Georg Hajdu. Schliengen: Edition Argus.
- Akkermann, Miriam. 2020. (Musik)instrument (im) Computer. In *Brückenschläge zwischen Musikwissenschaft und Informatik. Theoretische und praktische Aspekte der Kooperation*, Hg. Stephanie Acquavella-Rauch, Andreas Münzmay und Joachim Veit, 125–140. Detmold: Musikwissenschaftliches Seminar der Universität Paderborn und der Hochschule für Musik Detmold.
- Blackwell, Tim, Oliver Bown und Michael Young. 2012. Live Algorithms: Towards Autonomous Computer Improvisers. In *Computers and Creativity*, Hg. John McCormack und Marc d'Inverno, 147–204. Heidelberg: Springer.

- Computer History Museum. o.J. Cope, David Oral History. Artefact Details. <https://www.computerhistory.org/collections/catalog/102738612>. Zugegriffen: 13. März 2021.
- Cope, David. 1992. Computer Modeling of Musical Intelligence in EMI. *Computer Music Journal* 16: 69–83.
- Cope, David. 1999. Facing the Music: Perspectives on Machine-Composed Music. *Leonardo Music Journal* 9: 79–87.
- Dubnov, Shlomo, Gerard Assayag und Cont, Arshia. 2011. Audio Oracle Analysis of Musical Information Rate. *Proceedings of the Fifth IEEE International Conference on Semantic Computing*, 567–571. <https://doi.org/10.1109/ICSC.2011.106>.
- Goodyer, Jason. 2021. How an AI Finished Beethoven's Last Symphony and What That Means for the Future of Music. *BBC Science Focus Magazine*. <https://www.sciencefocus.com/news/ai-beethovens-symphony/>. Zugegriffen: 29. April 2022.
- Hiller, Lejaren A. und Isaacson, Leonard M. 1958. Musical Composition with a High-Speed Digital Computer, *JAES* 6 (3): 154–160. <https://www.aes.org/e-lib/browse.cfm?elib=231>.
- Hoffmann, Peter. 2000. The New GENDYN Program. *Computer Music Journal* 24: 31–38.
- Hoffmann, Peter. 2002. GENDY3 von Iannis Xenakis: Eine Höranalyse. In *Konzert – Klangkunst – Computer. Wandel der musikalischen Wirklichkeit*, Hg. Institut für Neue Musik Darmstadt, 255–272. Mainz: Schott.
- Koenig, Gottfried Michael. 1979. Projekt Eins – Modell und Wirklichkeit. *Musik und Bildung* 11/12: 752–756.
- Koenig, Gottfried Michael. 1983. Integrazione estetica di partiture composte mediante elaboratore – Segmente 99–105. *Bollettino 3 – Bollettino del Laboratorio* 27/28: 29–34.
- Koenig, Gottfried Michael. 1993a. Zu Projekt 1 – Version 1. In *Ästhetische Praxis. Texte zur Musik 1968–1991. Quellentexte zur Musik des 20. Jahrhunderts*, Hg. Musikwissenschaftliches Institut der Universität des Saarlandes, 12. Saarbrücken: Pfau.
- Koenig, Gottfried Michael. 1993b. Tabellen, Graphiken, Klänge. Ein Computer-Programm für kompositorische Strategien: Projekt 3. In *Ästhetische Praxis. Texte zur Musik 1968–1991. Quellentexte zur Musik des 20. Jahrhunderts*, Hg. Musikwissenschaftliches Institut der Universität des Saarlandes, 315–319. Saarbrücken: Pfau.

- Koenig, Gottfried Michael. 1999. Project 1 Revisited: On the Analysis and Interpretation of PR1 Tables. In *Navigating New Musical Horizons*, Hg. Otto Laske, 53–70. Santa Barbara: Praeger.
- Lewis, George. 2000. Too Many Notes: Computers, Complexity and Culture in »Voyager«. *Leonardo Music Journal* 10: 33–39.
- Mirianda, Eduardo. 2007. Cellular Automata Music: From Sound Synthesis to Musical Forms. In *Evolutionary Computer Music*, Hg. Eduardo R. Miranda und John Al Biles, 170–193. London: Springer.
- Pasquier, Philippe, Arne Eigenfeldt, Oliver Bown und Shlomo Dubnov. 2017. An Introduction to Musical Metacreation. *Computers in Entertainment* 14: 1–14. <https://doi.org/10.1145/2930672>.
- Schiff, Joel L. 2007. *Cellular Automata: A Discrete View of the World*. Hoboken: John Wiley & Sons.
- Simon, Herbert A. 1960. *The New Science of Management Decision: The Ford Distinguished Lectures*. New York: Harper & Brothers.
- Supper, Martin. 1995a. Computermusik. In *MGG – Die Musik in Geschichte und Gegenwart*, Hg. Ludwig Finscher, Sp. 967–987. Kassel und Stuttgart: Bärenreiter.
- Supper, Martin. 1995b. Elektroakustische Musik, Elektroakustische Musik ab 1950, Live-Elektronik. In *MGG Online*, Hg. Lorenz Lütteken. <https://www.mgg-online.com/mgg/stable/14626>. Zugegriffen: 15. Februar 2020.
- Supper, Martin. 1997. *Elektroakustische Musik & Computermusik*. Hofheim: Wolke.
- Tatar, Kivanç. 2019. Musical Agents: A Typology and State of the Art Towards Musical Metacreation. *Journal of New Music Research* 48: 56–105.
- Waibel, Alexander. 2021. AI in the Service of Humanity. <https://www.youtube.com/watch?v=HwtrWwd8WSY>. Zugegriffen: 29. April 2022.

