

## **Musik-Automat**

Mit der KI im musikalischen Frage-Antwort-Spiel

Antragsteller: Prof. Jörn Arnecke, Hochschule für Musik FRANZ LISZT Weimar

### **Warum bewerben Sie sich um ein Fellowship (persönliche Motivation)?**

Die letzten Monate haben eine explosionsartige Entwicklung der KI erbracht. Die Möglichkeiten, aber auch die erheblichen Risiken faszinieren und beängstigen zugleich. Forschung und Lehre werden sich ändern (müssen) – in Ausmaßen, die wir noch gar nicht absehen. Zugleich eröffnen diese Entwicklungen inhaltliche Optionen, welche die Musiktheorie sprunghaft voranbringen können.

Die Arbeit mit KI ist in besonderem Maße von Interdisziplinarität geprägt. Wir wollen diese in unserem Fellowship-Projekt umsetzen: Der Antragsteller – ein Musiktheoretiker – möchte hierzu mit einem Komponisten/Programmierer und einem Informatiker (mit langjähriger Erfahrung in maschinellem Lernen und in Deep-Learning-Modellen) zusammenarbeiten. Anregungen und Sichtweisen aus anderen Disziplinen sind gerade bei der KI, die in hohem Maße gesamtgesellschaftliche Relevanz aufweist, unerlässlich. Der Austausch mit anderen Projektverantwortlichen sowie mit dem eTeach-Netzwerk und dem Netzwerk Musikhochschulen 4.0 verspricht daher, das Projekt zu beleben – wie auch umgekehrt unser Vorhaben eine neue Facette für KI-Diskussionen liefern kann.

### **Was veranlasst Sie zu der geplanten Lehrinnovation? Welches Problem soll bearbeitet werden? Inwieweit handelt es sich dabei um ein zentrales Problem in der Lehre im jeweiligen Studienfach?**

Musiktheorie im Hauptfach (Master of Music) ist eine Disziplin, in der Musik vergangener Zeiten analysiert und künstlerisch nachgeformt, also neu komponiert, wird. Zunehmend werden solche Arbeitsergebnisse auch aufgeführt, etwa wenn Fragmente bekannter Komponisten vervollständigt werden; in Weimar wurden jüngst in einem studentischen Projekt Skizzen Franz Liszts in unterschiedlicher Weise fortgeführt, und die verschiedenen Lösungen wurden einander gegenübergestellt. Ein anderes Präsentationsformat, das sich etabliert und dem Publikum neue Mitwirkungsmöglichkeiten verschafft, ist ein Rätselkonzert: Unter historische Kompositionen werden neue Kreationen gemischt – das Publikum ist aufgerufen zu erkennen, welche Stücke tatsächlich aus der vergangenen Epoche stammen und welche heute dem alten Stil nachempfunden wurden.

Diese Arbeitsweise setzt eine genaue Kenntnis der zu imitierenden Stile voraus. Komponist\*innen haben seit jeher viel improvisiert, um auf Ideen zu kommen. Hier setzt das Konzept des Musik-Automaten an: Musiker\*innen (und Musiktheoretiker\*innen) sollen dazu animiert werden, im musikalischen Frage-Antwort-Spiel Einfälle zu kreieren – in gewisser Weise vergleichbar mit dem ChatGPT-Format, aber speziell entwickelt für musikalisches Material. Dieses Vorgehen passt besonders gut zu historischen Stilen und zu einfachen Tänzen, weil hier das syntaktische Prinzip der Imitation und Weiterführung grundlegend gilt. Das Standard-Modell für einen Ländler, wie ihn etwa Franz Schubert schrieb, sieht 16 Takte vor, die in zweimal acht Takte aufgeteilt sind, die sich

wiederum in zweimal vier Takte gliedern.<sup>1</sup> Diese Folgen reizen zum Frage-Antwort-Spiel, sowohl in der eröffnenden als auch in der reagierenden Rolle.

Der Musik-Automat greift dabei spielerisch zwei Kernbereiche der Lehre auf und macht sie kreativ nutzbar: das Erfinden von musikalischen Ansätzen einerseits und das Entwickeln vorhandener Ideen andererseits.

### **Welche Ziele verfolgen Sie mit der geplanten Lehrinnovation? Was ist daran neuartig?**

Die Interaktion zwischen Mensch und Maschine stellt einen wesentlichen Baustein dar für das Verständnis der Prozesse einer KI und ihrer möglichen Auswirkungen. Den meisten Modellen für die Musikerzeugung fehlt bislang noch die Möglichkeit, den Menschen in den Prozess einzubeziehen; eine Ausnahme bildet das *Human-in-the-Loop-System*.<sup>2</sup> Das Ziel der bislang vorgelegten Modelle ist überwiegend die Generierung von musikalischem Material an sich und nicht eine Analyse des Prozesses und der vorgeschlagenen melodischen Fortsetzungen.

Unser System ist stattdessen auf einen kollaborativen Prozess ausgerichtet: Die menschliche Auswahl und das Feedback werden verwendet, um das Modell zu einer besseren Auswahl von Melodien zu führen. Das Dialogische wird also in den künstlerischen Bereich transferiert. Die geplante Lehrinnovation besteht unter diesen Voraussetzungen aus einer benutzerfreundlichen App zur Komposition von Melodien über Akkordfolgen, die für Musikinteressierte aller Kenntnisstufen zugänglich sein soll.

Unser Hauptziel ist die Erforschung neuer Methoden zur Musikdatenverarbeitung mit neuronalen Netzen und die Schaffung einer leistungsfähigen, interessanten und lehrreichen App, die sowohl den kreativen Prozess fördert als auch pädagogische Zwecke erfüllt: Ein solches Programm eignet sich hervorragend für die Lehre, weil es die praktische Anwendung theoretischer Musikkonzepte auf ansprechende und interaktive Weise ermöglicht. Ein KI-Mensch-Diskurs dient zudem als Anregung für kritisches Denken und Diskussionen im Unterricht, u. a. über stilistische Nuancen, Kompositionstechniken und den menschlichen künstlerischen Ausdruck.

Die App wird in einem Webbrowser implementiert und ermöglicht den Nutzer\*innen, interaktiv Noten einzugeben und „live“ zu komponieren, wodurch sie einen Dialog mit der Maschine führen können – fast wie ein Spiel, bei dem Mensch und Maschine die musikalischen Sätze des anderen beginnen, fortsetzen und beenden können. Obwohl das Programm sowohl Akkordfolgen als auch Melodien generieren kann, liegt der Schwerpunkt des Projekts auf der menschlichen Interaktion im kreativen Prozess. Bislang kann KI in der Musik z.B. Stile nachahmen (DeepBach<sup>3</sup>) oder Sätze nach historischen Regeln schreiben (KATI – die Kantionalsatz-Maschine, vom Antragsteller gemeinsam mit dem Komponisten und Programmierer Alex James Vaughan und dem Musiktheoretiker Philipp Schmidt entwickelt<sup>4</sup>). Die Möglichkeit, Frage und Antwort in enger zeitlicher Folge aufeinander reagieren zu lassen, bringt die menschliche Fähigkeit der Spontaneität und Intuition hervor und fördert sie.

---

<sup>1</sup> <https://doi.org/10.31751/659>

<sup>2</sup> Als Beispiel: <https://arxiv.org/abs/2010.03190>

<sup>3</sup> <https://arxiv.org/abs/1612.01010>

<sup>4</sup> <https://digitale-musiktheorie.de>, außerdem in Vorbereitung: Jörn Arnecke / Philipp Schmidt, »Laboratory Kantional-satz Machine: Presentation, Testing, Discussion«, GMTH Proceedings: <https://www.gmth.de/proceedings.aspx>

Im Idealfall kann die Anwendung den Prozess des Frage- und Antwort-Spiels auch im Nachhinein aufzeigen und diesen Weg der Entstehung für den Unterricht nutzbar machen. KI soll also hier nicht den Menschen ersetzen, sondern seine Fähigkeiten trainieren. Ein solcher Automat existiert bisher nicht; die Kombination aus einem *Human-in-the-Loop-Design*, einer pädagogischen Ausrichtung des Systems, der Schnittstelle und des Feedbacks sowie der eingeschränkten Generierung wäre neuartig.

Mit maschinellern Lernen können bereits recht gut Akkordfolgen hergestellt werden<sup>5</sup>, melodische Erfindung ist schon deutlich schwieriger, die Kombination verschiedener komplexer Phänomene (Melodie – Harmonie – Syntax – Form) bislang unmöglich. Unser Ziel ist es nicht, der KI Komponieren beizubringen, sondern den menschlichen Anteil am kreativen Vorgang zu erhalten und zu fördern: Syntax und Form folgen vorgegebenen Mustern; für die konkrete Ausgestaltung besorgt die KI die Basis (Harmonie), der Mensch aber gibt die individuelle Ausprägung hinzu (Melodie). Diese kann er konzentrierter entwickeln, weil ihm Vorarbeiten hierzu von der KI abgenommen werden. Außerdem hält der Mensch den Prozess in der Hand: Ob er bzw. sie einen ganzen Vierton schreibt oder aber nur einen Takt komponiert und drei von der KI ergänzen lässt, das kann spontan entschieden werden. Weiterhin kann der Mensch die KI anhalten, um eine musikalische Idee selbst fortzusetzen: Das musikalische Chatsystem soll ähnlich flexibel sein wie ein sprachliches.

### ***In welche Studiengänge und -abschnitte soll die geplante Lehrinnovation implementiert werden? Handelt es sich dabei um den Pflicht-, Wahlpflicht oder Wahlbereich?***

Die Entwicklung der App wird in einen verpflichtenden Kurs für Musiktheoriestudierende (Hauptfach) integriert und hierbei erprobt und optimiert. Im Sinne eines forschungsnahen Lernens erleben die Studierenden damit, nach welchen Kriterien die KI trainiert und verbessert wird.

Ziel ist es, ein lehrreiches und anregendes pädagogisches Werkzeug zu bauen, das auch für Musiktheorie im Nebenfach eingesetzt werden kann: Dieses Fach (Harmonielehre) durchlaufen alle künstlerischen Studierenden mehrere Semester lang an der Musikhochschule, eine breite Anwendbarkeit wäre also gewährleistet.

### ***Wie lassen sich nach Erprobung der Lehrinnovation Erfolg und eventuelle Risiken beurteilen?***

Die musiktheoretischen Lehrveranstaltungen werden turnusgemäß evaluiert. Weiterhin besteht aufgrund der engen persönlichen Zusammenarbeit mit Studierenden die Möglichkeit, in themenzentrierten Einzelgesprächen Rückmeldungen der Hauptfach-Studierenden zu sammeln.

Der Projektablauf ist weiterhin so geplant, dass verschiedene technische Modelle ineinandergreifen und aufeinander aufbauen, sodass ein Abschnitt des Projektes funktionieren muss, bevor der nächste folgt: Nach dem Training der KI mit einer Vielzahl von Musikdaten besteht der nächste Schritt in der Feinabstimmung ihrer Fähigkeiten durch Benutzerinteraktion und Feedback. In diesem Arbeitsschritt spielen die Studierenden – als künstlerisch Forschende – eine wichtige Rolle: Insbesondere werden sie aufgefordert, die von der KI erzeugten Kompositionen zu beurteilen. Sie

---

<sup>5</sup> Eine eingeschränkte Generierung (z. B. auf der Grundlage einer Akkordfolge) ist zwar vergleichsweise selten, existiert aber: <https://ieeexplore.ieee.org/abstract/document/9376975>

sollten dann musikalische „Regeln“ oder Algorithmen entwickeln, um die Ergebnisse der Maschine automatisch zu bewerten.

Beispielsweise könnte ein Algorithmus entwickelt werden, um zu überprüfen, ob ein Takt Dauern enthält, die sich auf die richtige Anzahl von Schlägen summieren. Wenn ein 4/4-Takt Dauern von mehr als vier Schlägen enthält, würde der Algorithmus die KI darüber informieren, dass dieses Ergebnis nicht korrekt ist. Ähnliche Überprüfungen können für parallele Quinten oder Oktaven, falsche Notenschreibweise, untypische Rhythmen oder Synkopen für ein bestimmtes Genre usw. durchgeführt werden. Durch die kontinuierliche Bewertung soll die Leistung der KI verbessert werden.

Die Einbeziehung eines zusätzlichen *reinforcement learning* hat sich zwar in den jüngsten großen Sprachmodellen, wie den neuesten Generationen von GPT, als grundsätzlich erfolgreicher Ansatz erwiesen, ist aber technisch nach wie vor schwierig auszuführen. In der Tat stellt der Aspekt des *reinforcement learning* in unserem vorgeschlagenen Projekt das größte Risiko dar: Die Integration von Transformator-Architekturen mit Techniken des *reinforcement learning* ist relativ neu, und es gibt nur wenige standardisierte, öffentlich verfügbare Tools, um diese komplexe Aufgabe zu bewältigen.

Wenn wir diese Methodik jedoch erfolgreich umsetzen, könnte das Ergebnis alle bisherigen Errungenschaften auf diesem Gebiet übertreffen. Außerdem haben wir als Musikhochschule Zugang zu den notwendigen musikalischen Ressourcen, dem Fachwissen und den Studierenden, um eine Maschine auf diese Weise zu trainieren.

### **Wie soll die geplante Lehrinnovation verstetigt werden?**

Bei einer guten Qualität der Anwendung wird der Einsatz im Haupt- und Nebenfachunterricht ein kompetenzförderndes spielerisches und dauerhaft nutzbares Element sein, das von den Studierenden auch zum Üben und zum Ausbilden ihrer musikalischen Reaktionsfähigkeit eingesetzt werden wird.

### **Auf welche Lehr-Lern-Situationen – auch in anderen Disziplinen oder Hochschulen – kann die geplante Lehrinnovation übertragen werden?**

Die geplante Lehrinnovation kann auf andere Disziplinen oder Hochschulen ausstrahlen: Zusätzlich zum Fach Musiktheorie stellen wir uns an der Hochschule für Musik FRANZ LISZT Weimar Anwendungen in Komposition, Alte Musik und in den Instrumentalstudiengängen vor. Die menschenzentrierte Herangehensweise und der Fokus auf die Interaktion zwischen Mensch und Maschine könnten auch auf andere kreative Disziplinen angewendet werden, in denen KI und menschliche Expertise kombiniert werden.

An der TU Ilmenau würde sich die Konzeption/Implementierung des Modells gut für Master- und Bachelorprojekte sowie für studentische Gruppenprojekte im Institut für Medientechnik, für den Master of Science Medientechnik (wegen des Projektschwerpunkts auf der Arbeit mit Musikdaten) und auch für den Bachelor of Science und Master of Science Informatik (wegen der Verwendung von Deep-Learning-Architekturen) eignen.

### **Was versprechen Sie sich vom Austausch mit anderen Fellows des Programms für sich persönlich und für Ihr Projekt?**

Der Antragsteller kennt den Austausch, der durch das Fellowship möglich ist, gut aus eigenem Erleben. Bei der ersten Förderung, an der er teilnehmen durfte, waren jedoch durch Corona die persönlichen Begegnungen nur eingeschränkt möglich; die eTeach-Tagungen hingegen, an denen er auch nach dem Fellowship teilnahm, boten ein interessantes Forum mit wichtigen inhaltlichen Impulsen. Auch diese Erfahrungen motivieren ihn, mit einem zweiten Projekt in einen vertieften Diskurs einzusteigen.

Da davon auszugehen ist, dass aufgrund der aktuellen Entwicklung mehrere Fellowship-Projekte Berührungspunkte mit KI haben werden, wird hier auch ein fachbezogener Dialog möglich sein. Anwendungen zum künstlerischen Ausdruck bedeuten immer noch eine besondere Herausforderung für die KI, insofern hoffen wir auf Unterstützung – und glauben, spezielle Aspekte in den Diskurs einbringen zu können.

### **Über welche Expertise zur Durchführung des Vorhabens verfügen Sie? Wie sind Sie insbesondere mit der von Ihnen geplanten Lehrinnovation innerhalb Ihrer Hochschule organisatorisch eingebunden und vernetzt?**

Der Antragsteller hat – gemeinsam mit dem Komponisten und Programmierer Alex James Vaughan – durch das Projekt der „Kantionalatz-Maschine“ (KATI) unter Beweis gestellt, dass er in der Lage ist, technische Visionen für musikalische Aufgaben umzusetzen: KATI schreibt nach gegebenen Melodien Choräle, die sich stilistisch auf die Zeit um 1600 beziehen, und verwendet dazu ein Regelwerk der damaligen Zeit. Dieses Projekt wurde sowohl der eTeach-Community (Jahrestagung 2022) als auch der allgemein interessierten Öffentlichkeit (Lange Nacht der Wissenschaft 2022) vorgestellt; beim Kongress der Gesellschaft für Musiktheorie (GMTH) in Salzburg 2022 wurde die „Kantionalatz-Maschine“ den Fachvertreter\*innen präsentiert.

Durch seine Funktion als Leiter des Zentrums für Musiktheorie, die sich aus seiner Professur für Musiktheorie und Gehörbildung ergibt, kann der Antragsteller inhaltliche Impulse für seine hochschulweit wirkende Abteilung setzen. Hierdurch ist eine breite Information der Fachkolleg\*innen über Neuentwicklungen und ihre Potenziale gewährleistet, auch in die Bereiche Alte Musik, Instrumentalpädagogik hinein und – demnächst – zur neuen Professur für Musikpädagogik in digitalen Kontexten.

Der als Komponist und Programmierer vorgesehene Alex James Vaughan verfügt über mehr als zwölf Jahre Erfahrung in der professionellen Front-End-Website-/Webapp-Entwicklung und hat musikalische Tools, Algorithmen und KI entwickelt, die sowohl national als auch international Anerkennung gefunden haben, wie z. B. sein Programm „Boids Composer“, das vom Mozartfest Würzburg in Auftrag gegeben wurde, und sein Programm „Labyrinth“ (eine orchestrale musikgenerierende KI), dessen Alpha-Version erstmals vor 200 Wissenschaftlern des Pharmaunternehmens Roche öffentlich vorgestellt wurde. Darüber hinaus verfügt Alex James Vaughan über Abschlüsse in Komposition (bis hin zum Konzertexamen) und über umfangreiche Lehrerfahrung in zeitgenössischen Kompositionstechniken. Als Komponist gewann er zahlreiche Preise.

Der als technischer Berater eingeplante Andrew McLeod (Fraunhofer-Institut Ilmenau) hat in Informatik promoviert und verfügt über neun Jahre Erfahrung mit maschinellern Lernen und Deep-Learning-Modellen für musikalische – einschließlich metrischer – Strukturen, Stimmentrennung,

harmonischer Progressionen und Transkription. In letzter Zeit hat sich sein Schwerpunkt auf generative Modelle verlagert, und er hat Unterstützung für dieses Projekt von seinem Institut sowie durch andere laufende Projekte, die dieselben Tools und Techniken für die Generierung verwenden. Dazu gehören zwei aktuelle und zwei geplante Studierendenprojekte zu generativen Modellen für Musik, sowohl für symbolische (d.h. Partitur-Notation) als auch für Audio.

Über die Einbindung von Andrew McLeod soll ohne formelle Kooperation eine Zusammenarbeit ausgebaut werden zwischen der Hochschule für Musik FRANZ LISZT Weimar und der TU Ilmenau, denn das Fraunhofer-Institut arbeitet eng mit den Studierenden der TU Ilmenau zusammen und berät sie bei ihren Diplomarbeiten. Dieser Aspekt ist für das Projekt auch deshalb interessant, weil sich darüber technisch interessierte studentische Assistent\*innen finden lassen, die an einer Musikhochschule sehr selten sind.

Andrew McLeod kann als Berater auch deshalb zusätzliche Expertise bereitstellen, weil er in der Vergangenheit intensiv mit Deep-Learning-Modellen für Musik gearbeitet hat und über ein fundiertes Wissen verfügt über mögliche Datendarstellungen und Architekturen, die funktionieren. Außerdem hat er Erfahrung mit und Zugang zu einer breiten Palette von Datensätzen, die sehr hilfreich wären, um sie als Trainingsdaten für das aktuelle Projekt zu verwenden.

## Musik-Automat: Arbeitsplan

Oktober 2023	<p>Beginn der Datenerfassung für Akkord- und Melodiedaten sowie für Musikstrukturdaten.</p> <p>Beginn der Verarbeitung, Reinigung und Formatierung der Daten.</p>
November 2023	<p>Beginn erster Experimente mit generativen neuronalen Netzwerkarchitekturen (z.B. Transformers, GANs, VAEs).</p> <p>Diskussionen mit Studierenden der Musiktheorie über verschiedene Möglichkeiten der Verarbeitung musikalischer Daten. (<i>Was sind die grundlegenden Bausteine in der Musik? Wie kann komplexes musikalisches Material für eine Maschine vereinfacht werden? Auf welche Weise können musikalische Parameter voneinander getrennt werden, z.B. Trennung von Tonhöhe und Rhythmus?</i>)</p>
Dezember 2023	<p>Beginn von Experimenten mit <i>embedding spaces</i> – Vermittlung von musikalischen Beziehungen an Maschinen. Diskussionen mit Studierenden der Musiktheorie über das maschinelle Verständnis von Beziehungen zwischen musikalischen Einheiten. (<i>Stimmen wir mit der Interpretation der musikalischen Beziehungen durch die Maschine überein? Wie verstehen wir und die Maschine diese Beziehungen unterschiedlich?</i>)</p>
Januar / Februar 2024	Fortgesetzte Experimente, Training und Feinabstimmung generativer Modelle.
Februar/ März 2024	Beginn der Entwicklung einer webbasierten interaktiven Schnittstelle in HTML, CSS und JavaScript, die es dem Benutzer ermöglicht, musikalisches Material „on-the-fly“ im Browser zu komponieren.
April / Juli 2024	<p>Beginn von Experimenten im Bereich des <i>reinforcement learning</i> (dies wird möglicherweise einer der anspruchsvollsten Aspekte des Projekts sein), die es der Maschine ermöglichen, auf der Grundlage der Rückmeldungen der Nutzer*innen / Studierenden zu lernen und sich zu verbessern.</p> <p>Die Studierenden werden ermutigt, die Ergebnisse des Programms kritisch zu bewerten und musikalische „Regeln“ oder Algorithmen zu entwerfen, die im Verstärkungsprozess zur weiteren Feinabstimmung des Modells verwendet werden können.</p>
August 2024	Fertigstellung einer Alphaversion des gesamten Systems – ein „ <i>Human-in-the-Loop</i> “-Programm, das in der Lage ist, Akkordfolgen und Melodien in verschiedenen Tonarten, Taktarten, Genres und Komplexitätsgraden zu erzeugen, und über den Browser zugänglich ist.
September 2024	Weiterentwicklung, Feinabstimmung, Beseitigung von Fehlern und Benutzerfreundlichkeit der Schnittstelle. Allgemeine Vorbereitung für die öffentliche Veröffentlichung im Internet.
Oktober 2024	Fertigstellung der Beta-Version. Experimente und Tests der Software mit Studierenden und Fakultätsmitgliedern im Hinblick auf die Benutzerfreundlichkeit sowie das künstlerische, akademische und pädagogische Potenzial.
November 2024	Öffentliche Freigabe der Betaversion. Kontinuierliche Aktualisierungen und Verbesserungen des Systems.
Dezember 2024	Fertigstellung des vollwertigen Online-Programms.