

# Fauxtomation, Cherry-Picking, Blackboxing

Über Künstliche Intelligenz in der Musik

Genoël von Lilienstern

**S**teckt hinter sogenannter Künstlicher Intelligenz in Wirklichkeit Billiglohnarbeit? Sind Menschen, die von der Revolution des Machine Learning in der Musik schwärmen Agenten der CIA? Wieviel CO2 wird ausgestoßen, wenn ein Computer lernen muss wie Beethoven zu komponieren? Genoël von Lilienstern ist für die Positionen auf der Suche nach der »Künstlerischen Intelligenz«.

In den letzten fünf Jahre hat es im Bereich KI-basierter Musik-Software große Fortschritte gegeben. Es zeichnen sich experimentelle Interaktionsräume ab, in denen Computer-Modelle in der Lage sind, einen erlernten Musikstil als direkt hörbaren Klang zu reproduzieren. Der auf Medien und Ästhetik bezugnehmende, kritische Diskurs innerhalb der Neuen Musik wäre prädestiniert dafür, sich mit der Künstlichen Intelligenz als technologischer Realität, aber auch als Ideologie auseinanderzusetzen.

Fast jeder kennt die seit 2015 existierenden Deep-Dream-Grafiken. In ihnen interpretieren KI-Computer-Modelle visuelle Formen in Bildern entsprechend eines Vorwissens neu. Dieses Wissen speist sich aus einer größeren Menge visueller Trainingsdaten: In den Zügen eines menschlichen Gesichtes erkennt ein Modell, dessen Datenbank vor allem Tierbilder enthält, beispielsweise vorrangig Hundewelpen und Katzenbabys und errechnet daraufhin eine neue, bizarre Version des

Originalbildes. Auf Webseiten wie Deep Dream Generator kann man mittlerweile Stil-Transfers und Deep-Dream-Grafiken sehr einfach selbst erstellen. Nach einer anfänglichen Faszination des Ultimativ-Neuen ist diese Form der Deep-Dream-Grafiken nun ein weithin bekannter Internetmanierismus geworden.

Bildererkennung und Textgenerierung gehören zu den Anwendungen, für die KI-Software-Environments wie Tensorflow, Torch oder Keras ursprünglich ausgelegt sind. Ein unterhaltsames Beispiel für das automatisierte Schreiben von Texten ist der Nonsens-Text »Harry Potter and the portrait of what looked like a large pile of ash« aus den Botnik Studios (2017). Sämtliche Zutaten des Harry-Potter-Universums tauchen in ihm auf, jedoch in wilder Verdrehung jeglicher Sinnzusammenhänge.

Der Kurzfilm *Zone Out* von Regisseur Oscar Sharp und KI-Wissenschaftler Ross Goodwin basiert nicht nur auf einem, von einem Textgenerator geschriebenen Drehbuch, sondern greift als szenisches Ausgangsmaterial auf Public-Domain-Filme zurück. Die Gesichter der Originalschauspieler werden mittels Faceswapping den Dialogen des Drehbuchs angepasst.

Besonderes mediales Echo fand in letzter Zeit auch die Berechnung fiktiver Gesichter



oder die einen Höchstpreis erzielende Auktion des ersten KI-Gemäldes.

## KI und Audio

Wie aber sieht es mit der Generierung von musikalischen Strukturen und Klang aus? Das Prinzip des Errechnens von Midi-Noten auf Grundlage datenbasierter Trainings ist schon seit längerem bekannt, beispielsweise durch die Arbeiten des US-amerikanischen Komponisten David Cope. Sehr viel neuer hingegen ist das Training von Modellen auf Sample-Ebene, also im Bereich direkter Analyse und Generierung von Klang. Das Aufspüren von Mustern in zeitlich voneinander getrennten Samples ist komplizierter und kleinteiliger als die Analyse von Wahrscheinlichkeiten in der Anordnung von räumlich benachbarten Grafikpixeln. Die bestehenden Modelle sind bisher kaum für diesen Zweck optimiert. Dementsprechend frisch und spannend ist dieses Forschungsfeld.

»Ariana Grande vs. Sample-RNN« heißt ein im Februar 2019 auf Youtube gepostetes Video. Dort kann man nachvollziehen, wie der Ablauf bei KI-basierter SampleRNN-Audio-Generierung ist: Zu hören ist ein Lernmodell im Training mit Gesangsarrangements der Popsängerin Ariana Grande. Nach mehreren Durchläufen ist eine neue Version dessen zu hören, was das Modell bisher verstanden hat. Es gibt graduelle Annäherungen an das Ziel, das Ausgangsmaterials klanglich zu imitieren. Zu 100 Prozent erreicht wird dies allerdings nicht. Im Video sind die recht bizarren Zwischenstationen dieses Lernprozesses zu hören, die mit ihren lallenden und asthmatischen Passagen einen gewissen Unterhaltungswert haben.

Auf der Webseite des KI-Musikprojekts DadaBots von CJ Carr und Zack Zukowski finden sich zahlreiche Versuche der Neuberechnung von

Musik bekannter Musikstile. So ist *Deep! The Beatles* ein Wirrwarr eines vage beatlesk klingenden Geräuschklangs, der das Ergebnis des Versuchs einer Resynthese ist.

Einiges an Aufmerksamkeit erregte vor einigen Monaten der von den DadaBots generierte, niemals endende Livestream *Relentless Doppelgaenger*, in welchem ohne Unterbrechung neue Deathmetal-Musik der Band Archspire generiert zu werden scheint.

Auffällig ist, dass vor allem die Resynthese situationistisch-intensiver Musikgenres wie Punk, Hardcore oder Metalmusik überzeugend klingt. Momentformen scheinen sich für diese Art der Synthese besonders gut zu eignen, da Intensität, schnelle Wechsel und Unvorhersehbarkeit stiltypisch für sie sind. Die Reproduktion von größeren musikalischen Zusammenhängen erfordert hingegen exponentiell mehr Rechenleistung. Die Generierung von formal schlüssig klingenden Mahler-Symphonien dürfte also wahrscheinlich noch länger auf sich warten lassen – Death-Metal, Serialismus, Impro-Jazz oder Lachenmannsche Klänge stellen hingegen kein Problem dar. Eine provokante These, die man hieran knüpfen könnte, wäre, dass sich die Qualität von Musik am Rechenaufwand ihrer KI-Resynthese ablesen lässt.

In einer Live-Situation funktioniert das Prinzip der momentanen Intensität und Imitation in Tomomi Adachis *Tomomibot*. In dieser Klang-Performance, die er mit Andreas Dzialocha und Marcello Lussana entwickelte, reagiert ein KI-Modell dialogisch auf Dada-Klanggesten, die Tomomi Adachi live improvisiert.

Es handelt sich dabei aber nicht um eine Live-Generierung von Klang auf



Sample-Ebene, sondern um eine Live-Auswahl von vorausgenommenen Klangschnipseln, die dem improvisierten Klang am ehesten entsprechen.



In einem eigenen Versuch mit intelligentem Body-Target-Mapping experimentiere ich mit von Deep-Dream-Modellen

bekanntem Stiltransfers. Auf die verschiedenen Variationen der Sarabande aus Händels *Klaviersuite in d-Moll* werden, von Mozart bis Nancarrow, Klänge von Stücken aus anderen Musikepochen gemappt.

### KI-Komposition im Vergleich zum algorithmischen Komponieren

Um besser zu verstehen, was bei der SampleRNN-

Synthese passiert, ist es hilfreich, KI-basierte Klangexperimente von der algorithmischen Komposition abzugrenzen. Zu Beginn einer algorithmischen Komposition legt die musiktheoretisch informierte Programmierer\*in oder Komponist\*in eine Reihe von Materialien und Regeln fest, anhand derer musikalische Abstraktionen generiert und transformiert werden sollen. Das kann beispielsweise ein einfaches Klangmotiv sein, eine Struktur von Akzenten oder eine Steuerungsanweisung für ein automatisiertes Sampling. Diese Grundelemente können sich wuchernd fortspinnen, von einer übergeordneten Struktur gelenkt werden oder, wie etwa beim Rewrite-Prinzip, im Verlauf durch neue Informationen ersetzt werden. Am Ende steht eine Struktur mit formalem Verlauf, die anschließend häufig in einem Notationsprogramm oder einem Audioeditor von Hand fertig bearbeitet wird.

Anders ist der Ausgangspunkt bei der Arbeit mit KI, die als Neuronales Netzwerk auf einen Datensatz zurückgreift. Hier stehen zu Beginn keine musikrelevanten Regeln oder Strukturen – das Modell weiß nichts. Stattdessen liest die KI, in gewichteten Schichten von Perzeptoren (Layers), diese Daten aus. In den iterativen Durchlesevorgängen (Epochen) werden die Gewichtungen immer weiter in Richtung der betrachteten Daten verschoben, bis es sie annähernd originalgetreu imitieren kann; das heißt das Netzwerk bildet Strukturen aus, die dem Wissen, das es betrachtet, ähnlich werden.

Das musiktheoretische Wissen einer menschlichen Intelligenz ist dabei immer noch wichtig, denn der Computer hat von sich aus, vor Beginn des Ausleseprozesses, keinerlei musikalische Ideen oder künstlerische Fantasie. Auf die Sichtung der Ergebnisse folgt in der Regel ein vielfaches Nachjustieren der verschiedenen Parameter und Aktivierungsfunktionen, nach dem Prinzip Versuch-Irrtum. Die Methoden werden so lange

händisch geändert, bis das Ergebnis gewissen Grunderwartungen der Komponist\*in entspricht. Der Workflow ist also »messy« und gute Ergebnisse sind eher die Ausnahme als die Regel. Keinesfalls darf man glauben, dass KI eine so genaue Vorstellung von einem spezifischen künstlerischen Grat hätte, also dem feinen Zusammenspiel von Klang, Timing, Stilistik etc., um direkt beim ersten Mal befriedigende Ergebnisse entstehen zu lassen; es sei denn, das Material, mit dem man hantiert, ist so banal und unterkomplex, dass ein Scheitern kaum möglich ist.

Gemeinsam ist algorithmischen und KI-Kompositions-Tools also immer noch der Werkzeugcharakter. Auch die Hingabe der Komponist\*in, an in ihrer Komplexität unvorhersehbare Abläufe, ist Merkmal beider Arbeitsweisen. Die Überraschung (oder die Langweile), mit der man auf die Ergebnisse reagiert, motiviert das Anbringen kleiner Systemänderungen, um es der eigenen, aktualisierten Vorstellung, dessen, was klingen soll, schrittweise anzunähern. Auch das Auswählen von besonders interessanten Ergebnissen (Cherry-Picking), die man dann zum Werk zusammenführt und attraktiv herrichtet, ist beiden Arbeitsweisen gemeinsam.

### Quick and cheap backing tracks

Im Bereich kommerzieller Musikproduktionen werden immer wieder KI-basierte Innovationen diskutiert, die versprechen, praktischer, schneller und billiger eine Melodie zu komponieren, Begleitharmonien hinzuzurechnen oder Gesangslinien mit Beats zu unterlegen. Wahrscheinlich ist es demnächst auch wirklich so, dass man in sein Handy singt und mithilfe einer App sofort so klingen kann wie ein Song von Billie Eilish oder Kanye West.

Mitunter mischt sich in die Werbebotschaften für diese Apps der hyperkapitalistisch-messianische Duktus des Silicon-Valley-Sprechens. So wird auf einem Techblog geschwärmt:

© eroGANous Neurographie Serie, Mario Klingemann, 2018



»The future could be about not just music composed by an A.I., but music being composed just for you – based on data about your musical preferences, your physical habits, even the beat of your heart.«<sup>1</sup>

»I actually think 10 years from now, you won't be listening to music« postulierte der Venture-Kapitalist Vinod Khosla während eines öffentlichen Chats der Super Session der Creative Destruction Labs. Stattdessen, so glaubt er, werden wir kundengerechte Songs hören, die für jedes Individuum automatisch designed werden, maßangefertigt für das Gehirn, die Hörvorlieben und spezifischen Bedürfnisse. Solche Szenarien klingen unheimlich, weniger wegen ihrer Schöne-Neue-Welt-Aura, sondern wegen der Vorstellung Programmierer\*innen solcher Software und die dazugehörigen Start-Up-Investor\*innen könnten tatsächlich glauben, anhand von Gehirnstromdaten, Marktforschungsergebnisse und Benutzer\*innenvorlieben das ganze Spektrum individual-menschlicher Bedürfnisse abdecken zu können.

Ein ernsthaftes Problem könnte auch sein, dass diejenigen, die sich im Besitz der besten Möglichkeiten für eine musikalische KI-Revolution befinden, besonders wenig Ahnung von (um den Begriff einfach mal zu benutzen) guter Musik haben.

Ein eindrucksvolles Beispiel für dieses Defizit ist *Google AI gets Artsy, composes a Song*.

Klanglich deutlich opulenter werden in Ted Talks Softwares wie AIVA vorgestellt, die eine einfache melodische Idee in hollywoodtaugliche Orchestraltracks umwandeln.

Von unsauberen Arbeitsabläufen oder Trial-and-Error-Verfahren wird dort aber tunlichst nicht gesprochen, sondern vielmehr mit sanftem Schauer der perfekte Genius der KI hervorgehoben, der all das entworfen und errechnet hat.



Ein Beispiel, welches deutlich macht, dass Vorstellungen vom komponierenden KI-Genius mindestens verfrüht sind, ist der als solcher beworbene »erste AI Popsong« *Daddy's car* von 2016.

An diesem Song sind nur die Melodie und Harmonik von einem Computer kompiliert und nicht, wie es in den meisten Meldungen hieß, gänzlich alles. Das Arrangement und der Text stammen aus der Feder des Komponisten Benoit Carré. Der Song wurde anschließend konventionell in einem Studio eingesungen und produziert. In Projekten wie diesen fungiert das Schlagwort KI als Clickbait. Wer wirklich wissen möchte, wie der Stand der Neuberechnungen von musikalischen Legenden ist, höre sich die zuvor genannten Beispiele der Dadabots an.

Eine weitere Variante »zukunftsweisender« Musik-KIs sind Modelle, in denen zwar ein Computer Entscheidungen trifft, dabei aber aus einer limitierten Anzahl von im Vorhinein kompatibler Versatzstücke zurückgreift, so dass nie etwas schief gehen kann – vergleichbar mit einem Kind, das nur auf schwarzen Klaviertasten spielt. Künstliche Intelligenz mit schwachem IQ sozusagen.

#### Artifiziell artifizielle Kompositionen

Sehr häufig ist Künstliche Intelligenz nicht halb so autonom wie sie zum Verkauf angepriesen wird. Das Label KI ist eine sprachliche Attribution, die eine promotionförderliche Aura von Hype evoziert. Es ist das Eintrittsticket für Dilettant\*innen, Goldgräber\*innen, Start-Up-Blender\*innen und Künstler\*innen in den Zirkus der Aufmerksamkeitsökonomie. Aber auch in der Neuen Musik hat die Faszination für die Intelligenz von Computern



Spuren hinterlassen und das schon vor vielen Jahrzehnten. Die *Illiad Suite* (1957) von Lejaren Hiller war über lange Zeit hinweg ein wichtiges und häufig aufgeführtes Werk im Kontext von computergestützter Komposition. Würde man sich diese Abfolge von Streichquartett-Sätzen ohne das Wissen ihrer technologischen Genese anhören,

**Fauxtomation beschreibt den Dissens zwischen dem Mythos einer technisierten Zukunft, in der menschliche Arbeit redundant geworden ist, und der Realität, in der mehr Arbeit als jemals zuvor von Menschen geleistet wird.**

könnte man leicht zum Urteil kommen es mit den ungelenk-modernistischen Übungen eines Kompositionsamateurs zu tun zu haben. Doch das Stück zehrt von seinem Ruhm eine der allerersten computergenerierten Kompositionen zu sein. Auch in der Kunstmusik gibt es Experimente, bei denen man zweifeln kann, ob es sich wirklich um Ergebnisse handelt, die von Anfang bis Ende vom Computer berechnet sind, oder vielmehr um eine gezielte, menschliche Auswahl einiger weniger, geglückter Ergebnisse (Cherry-Picking), die noch dazu von Hand poliert, geschönt und vervollständigt sind.

#### Large Chunks

David Cope, einem der Pioniere computergestützter, musikalischer Kreativität, wurde lange Zeit mit Skepsis entgegengetreten. Klassikliebhaber\*innen sind computergenerierte Stilkopien per se suspekt, Musiktheoretiker\*innen monierten satztechnische Uneleganz und formale Schwächen. Trotz allem kann niemand in Abrede stellen, dass seine Stilkopien im ersten Moment tatsächlich wie Bach, Mozart etc. klingen. Präziser ist die Kritik an der Arbeit Copes, wenn man den Fokus auf die konkrete, computergestützte Generierung der Musik legt: Große musikalische Versatzstücke bestehender Werke werden nach dem Prinzip des

Musikalischen Würfelspiels Kirnbergers (1721–1783) zusammengefügt. In Kirnbergers Würfelspiel werden, nicht ohne süffisanten Unterton, musikalische Standardfloskeln, gemäß eines harmonischen Gerüst, aneinander gereiht. Ein eher banales, aber einigermaßen schlüssig klingendes Ergebnis kommt dabei immer heraus.

Eine gewisse Leidenschaft von Cope-Skeptikern lässt sich in den Kommentarspalten unter seinen Werken auf YouTube beobachten: Einzelne Nutzer\*innen führen dort in längeren Listen auf, welche Takte, Taktgruppen und ganze Satzteile aus bestehenden Werken von Bach, Mozart oder Rachmaninoff entnommen sind. Zudem befinden sich diese meist nicht am Anfang der Stücke, sondern in den weiter hinten liegenden Teilen der Kompositionen. Letztendlich ist es also die Leistung des Computers, diese Passagen per Zufall auszuwählen und ihre Anschlussfähigkeit zu überprüfen. Zwar hat in dem Fall ein Computer die Entscheidungen getroffen, ebenso gut aber ließe sich diese Methode an einem Nachmittag mit Schere, Klebstoff und einem Würfel realisieren. Dass David Cope seine Datenbanken 2003 vernichtete und seine Methode so der genaueren Nachvollziehbarkeit entzogen hat, dürfte die Zweifel seiner Skeptiker\*innen nicht eben zerstreut haben.

Es muss aber nicht darum gehen, KI-Kunst der Scharlatanerie zu überführen. Nirgendwo steht schließlich, dass Künstler\*innen nicht lügen dürfen. Man kann das schillernd-kontingente Feld potentiell artifiziell-artifizieller Komposition auch als reflexives Dispositiv verstehen. In ihm manifestieren sich Fragen darüber, was wir als künstlerisch genuin und wertvoll empfinden. Unter der

Bedingung welchen Vorwissens sind wir bereit, einem Musikstück einen Wert zuzusprechen und uns ihm hinzugeben?

Das eingangs genannte Beispiel eines endlosen Streams neugenerierter Death-Metal-Musik spielt übrigens auch mit der Illusion einer technologischen Realität, die so aber derzeit, selbst mit größeren Rechnern, noch nicht in Real-Time realisierbar

## Alle Bereiche, in denen der Gebrauchswert von Musik vor dem ästhetischen Wert der Kunstmusik oder dem identifikatorischen Aspekt von Popmusik steht, werden mit billig produzierten KI-Klängen geflutet.

ist. Tatsächlich handelt es sich dabei um einen mehrstündigen Loop, dessen Wiederholung innerhalb des impulsiv, spontanen Klangmaterials nicht auffällt. Die Vorstellung einer unendlich sich neu generierenden Musik ist faszinierend und wahrscheinlich in wenigen Jahren möglich. Derzeit ist es aber noch eine *Black-Mirror*-Utopie, deren Existenz CJ Carr und Zack Zukowski hier suggerieren.

### Fauxtation und Hysterie

Im korrelationalen Verhältnis zur Faszination für alles Computerberechnete und KI-Erdachte stehen die Ängste, die das Feld von Automatisierung, Robotisierung und technologischem Fortschritt generell begleiten. Im Februar 2018 machte die Meldung Schlagzeilen, dass der von Elon Musks Firma OpenAI entwickelte Textgenerator GPT-2 zu gefährlich sei, um vollständig veröffentlicht zu werden. François Chollet, der Entwickler der KI-Software Keras kritisierte dies als Anheizen sensationalistischer Panikmache bei gleichzeitiger Verschleierung der eigenen Methoden (Nichtveröffentlichung).

Das Lancieren von Ängsten und die Überzeichnung der technologischen Potentiale machen sich große Unternehmen darüber hinaus gezielt mit einem Mechanismus zunutze, den die kanadische Aktivistin Astra Taylor »Fauxtation« nennt. Mit

Drohungen wie: »Sollte der Mindestlohn wirklich auf 15 Dollar angehoben werden, müssen wir alle Serviceleistungen automatisieren« werden politische Entscheidungen gelenkt und forciert. Fauxtation beschreibt den Dissens zwischen dem Mythos einer technisierten Zukunft, in der menschliche Arbeit redundant geworden ist, und der Realität, in der mehr Arbeit als jemals zuvor

von Menschen geleistet wird. Ein Beispiel dafür sind die Selbstzahler-Terminals bei McDonalds, die scheinbar die Arbeit einer Servicekraft ersetzen, in Wirklichkeit aber die Kund\*innen zur unbezahlten Mitarbeiter\*in machen. Taylor verwendet dafür auch den Begriff »artifizielle-artifizielle« Intelligenz, was den Umstand beschreibt, wenn sich hinter der KI in Wirklichkeit menschliche Handarbeit versteckt.<sup>2</sup>

### Blackbox CurAltor

CurAltor nennt sich die Musikkritik-Software des britischen Komponisten Nick Collins, mit der für die Donaueschinger Musiktage 2019 Stücke für ein Klavierkonzert ausgewählt wurden. Künstlicher Intelligenz kuratorische Entscheidungen anzuvertrauen klingt ungewöhnlich und provokant. In den Diskussionen zu diesem Vorhaben (beispielsweise auf dem Facebook-Thread der Ankündigung) fiel oft die Frage: Nach welchen Kriterien wird die Musik ausgewählt? Hat die KI ästhetische Präferenzen? Mit welcher Musik wurde das Modell trainiert? Darauf wurde geantwortet: Das solle vorerst nicht offengelegt werden, da im realen Leben Kurator\*innen ja auch autonom ihre Entscheidungen gemäß ihrer subjektiven Kriterien

träfen. So sehr dieser Altherrenkommentar an sich schon kritikwürdig erscheint, führt er zudem zu einem wichtigen Punkt: Eine KI hat ja keine subjektiven Kriterien. Alle Einstellungen und Parameter, sowie die Trainingsdaten sind zuvor von einem Menschen festgelegt worden – was nicht heißt, dass das Resultat dieser KI-Kuratierung von vorne herein klar ist, aber es zeigt, wie die KI als Black Box Entscheidungen verschleiern könnte, die in Wirklichkeit subjektiv-menschlich sind.

Dem Computer-Kurator werden die gleichen Eigenschaften wie einem Menschen zugesprochen und damit Machthierarchien gefestigt, anstatt die Chance zu ergreifen mit den neuen Mitteln auch eine neue Praxis zu begründen. Progressiver wäre es in dem Fall zu sagen: Im Gegensatz zu einer menschlichen Privatperson braucht mein Computer-Kurator keine Autonomie und Privatsphäre, wir praktizieren Transparenz und dieses und jenes sind die Kriterien, nach denen entschieden wird.



© eroGANous Neurographie Serie, Mario Klingemann, 2018

## Digitaler Aktivismus

Für diese Problematik der scheinbar zu komplizierten und deshalb demokratischen Prozessen entzogenen Technologien, nicht nur im Bereich der KI, sondern in digitalen Belangen generell, setzt sich die politische Bewegung Diem25 ein. Ihr »Grünbuch Technologische Souveränität« umreißt die Prinzipien der Demokratisierung von Technologie und Innovation.

»Make neural nets uncool again« ist der Slogan der Webseite fast.ai. Die selbsterklärte Mission der Betreiber ist es den »coolen« Exklusivstatus von KI und Deep Learning zu unterlaufen und Code und Ressourcen für Jedermann anzubieten. Auch die Frage nach der Arbeit mit KI bei begrenzten

Computer-Ressourcen gehört zur Agenda von fast.ai. Denn Monopolstellungen in diesem Bereich werden auch dadurch begünstigt, indem bei der Berechnung großer Datenmengen sowohl Privatpersonen als auch



Universitäten schnell an ihre Grenzen stoßen, wohingegen Unternehmen wie Google, Amazon oder Facebook über gigantische Rechnerpools verfügen, um die sehr aufwändigen Modelle mit Millionen von Parametern laufen zu lassen.

In diesem Zusammenhang stellt sich auch die Frage nach der ökologischen Bilanz von KI-Berechnungen. Eine Studie des MIT kommt zu dem Ergebnis, dass die Berechnungen für Transformer, ein Spracherkennungsmodell, den CO<sub>2</sub>-Fußabdruck äquivalent zum Spritverbrauch von fünf US-PKW, bezogen auf ihre gesamte Lebensdauer haben. Im Vergleich zum menschlichen Gehirn ist KI zwar effektiv, aber ineffizient. Softwares für selbstfahrende Fahrzeuge müssen, im Modell, bis zu 50.000 mal gegen einen Baum fahren, bis sie verstehen, dass das keine gute Idee ist.

## Kritik der Werkzeuge in der Neuen Musik

Es gibt also in der Diskussion im Zusammenhang mit KI eine Vielzahl von Gründen eine kritische Haltung einzunehmen, von ideologischen, politischen bis hin zu ökologischen. Das passt gut zur Tradition der Neuen Musik, für die das Hinterfragen des Verhältnisses von Werkzeugen und deren kulturhistorischen, gebrauchspraktischen Inskribierungen zum Selbstverständnis gehört. Es ist ein Fragen im Sinne von Hegels Verständnis des Begriffs, der verstanden wird als die Wirklichkeit, die sich in einer Sache manifestiert. Ein Cello versteht man unter diesen Vorzeichen nicht einfach als neutralen Klangerzeuger, sondern sieht die zugehörige verfestigte Arbeit und Geschichte, mit allen Facetten von der puren, hölzernen Materialität, über die Instrumentenbauer\*in, zur unterbezahlten Orchestermusiker\*in, zum pathetischen Vibrato von »Le cygne«.

Und dieses Bewusstsein hilft auch beim Umgang mit elektronischen und digitalen Technologien. Deswegen legte man beispielsweise zu der Zeit, als ich mit dem Studium elektronischer Musik begann, besonderen Wert auf das Erlernen von Audioprogrammiersprachen wie C-Sound – und nicht so sehr auf kommerziellere Programme wie Cubase, in denen bestimmte Arbeitsweisen und Klangbilder von vorneherein nahegelegt wurden. Der heraufdämmernden Epoche neuer Technologien, denen Zwecke und Ideologien auf noch viel komplexere Weise einschreibbar sind, sollte man mit diesem kritischen Bewusstsein für die Werkzeuge und ihren Gebrauch entgegentreten.

## KI-Musik in der näheren Zukunft

Die Auseinandersetzung mit Künstlicher Intelligenz ist nicht nur lohnenswert und spannend, sondern notwendig, um zu verstehen, mit welcher Art von Technologie man es zu tun hat. In näherer Zukunft kann man im Feld der KI-Musik mit zahlreichen Phänomenen rechnen. Alle Bereiche, in denen der Gebrauchswert von Musik vor dem äs-



© eroGANous Neurographie Serie, Mario Klingemann, 2018

thetischen Wert der Kunstmusik oder dem identifikatorischen Aspekt von Popmusik steht, werden mit billig produzierten KI-Klängen geflutet: Die Klangbänder im Hintergrund von Imagefilmen, Scores für Hollywood-Schnulzen, Supermarktbeschallungen, hyperkommerzielle Housemusik für Techno-Festivals, Musik in Lounges und Aufzügen. Marktforschungsstudien können dann direkt in den perfekten Klienten\*innenklang übersetzt werden.

Etwas weniger öde kann es zeitweise im Bereich experimenteller Musik aussehen. Gerade den auf surrealistischen Collagetechniken basierenden, postmodernen Musikströmungen steht eine

weitere Blütezeit bevor. Die Analyse und Rekombination vorgefundener Klänge im Stil von John Oswald wird auf einem vielfach komplexeren Niveau möglich sein. Songs einer Musikgruppe in ein anderes, konträres Musikidiom zu überführen, früher mühsam als Shred im Heimstudio eingespielt, wird als Stiltransfer zum Standard-Gag.

Eine Heerschaar getreuer Co-KI-Apps werden helfen, die nötigen, aber anstrengenden Teil-Operationen von Neue Musik-Komponist\*innen praktisch abzuarbeiten, beispielsweise beim langwierigen Orchestrieren oder dem zähen Ausarbeiten harmonischer Verläufe. Man wird darauf zurückgreifen, aber möglicherweise lieber darüber

schweigen, denn KI wird vielleicht recht bald zum Synonym für »billig produziert« geworden sein.

Als Selfmade-Act mit intelligenten Arpeggiatoren kann jeder mit ein paar Handgriffen wie Carry D, Snoop Dogg oder Helmut Lachenmann klingen, aber das wird nach anfänglichem Unterhaltungswert niemanden sonderlich interessieren, denn natürlich ist Musik mehr als ihr reines Klangbild. Sie ist ein Interaktionsraum aus Klang, Reflexion, körperlicher Performance, emotionaler Ambivalenz, Statement, Lebenserfahrung: Dinge, die man in ihrem Zusammenwirken unter dem Begriff »Weltwissen« subsumieren kann. In der Abwesenheit dieses Weltwissens liegt die Limitierung der hochspezialisierten KI-Systeme: Sie sind für ganz partikuläre Zwecke trainiert, haben aber kein Wissen davon, was *die Welt* ist. Alles, was bereits existiert und sich in ein Datenformat überführen lässt, wird die KI nachempfinden und replizieren können. Doch sie wird, wenn wir über ästhetisch wirklich bedeutsame Kunst sprechen, in diesem Nachschaffen lange Zeit immer nur an zweiter Stelle stehen. Sicher wird die originelle Rekombination konträrer Trainingsdaten für überraschende, unterhaltsame, bizarre oder auch schöne Momente sorgen. Aber um kreativ etwas ganz neues zu schaffen, müsste die KI selbst über ein realistisches Modell der Welt verfügen, das dann auch noch, um dem Menschen überlegen zu sein, schneller als unsere Welt laufen müsste. Vielleicht wird die Menschheit irgendwann ein solches extrem energieaufwändiges, extrem komplexes KI-Modell zum Laufen bringen – vielleicht aber wird sie sich auch, an dem Tag, an dem es möglich wäre, lieber mit ganz anderen Fragen beschäftigen. Beispielsweise solchen des nackten Überlebens.

1. <https://medium.com/s/story/music-created-by-artificial-intelligence-is-better-than-you-think- ce73631e2ec5>

2. <https://ainowinstitute.org/symposia/videos/fauxtomation.html>

**Genoël von Lilienstern** ist Komponist von Experimentalmusik und Musiktheater. Ab und zu verfasst er Texte zu Themen aus den Bereichen Technologie und Musikästhetik.



# musikjournalismus. bachelor und master. in dortmund.



**Kommunikation  
über Musik in Medien.**



*„Ihr Unterrichtsplan liest sich wie eine Anleitung zur Interdisziplinarität: montags Medienrecht, dienstags Gehörbildung und Harmonielehre, dann Ästhetik, mittwochs ‚Nachricht und Bericht‘, donnerstags Musikgeschichte, Klavierunterricht und die eigene Klassiksendung Terzwerk im Campusradio, freitags ‚narrative Darstellungsformen‘, danach Stimmbildung.“*

**DIE ZEIT**



Die Bachelor- und Master-Studiengänge Musikjournalismus sind eine Kooperation des Instituts für Musik und Musikwissenschaft und des Instituts für Journalistik. Die Basis bildet eine fundierte musikalische und journalistische Ausbildung in Theorie und Praxis. Auf dieser Grundlage wird erprobt und erforscht, wie »Musik in Medien« vorkommt und welche musikjournalistischen Vermittlungsformen es gibt.

Die gute Resonanz unserer Absolvent\*innen auf dem Arbeitsmarkt zeigt: es gibt einen hohen Bedarf an Musiksachverständigen, die schreiben und reden können und die Regeln des journalistischen Handwerks beherrschen.

Neugierig?  
[terzwerk.de](http://terzwerk.de)



*„Klug konzipierte, ausdifferenzierte und erfolgreiche Lehrangebote (...). Besonders im Gespräch mit den Studierenden sowie während der Visite in der Lehrredaktion war es ausgesprochen überzeugend, in welchem Maße Praxisrelevanz nicht nur postuliert, sondern tatsächlich gelebt wird.“*

**Gutachtergruppe AQAS, 2016**

**Informationen und Kontakt:**  
Technische Universität Dortmund  
Institut für Musik und Musikwissenschaft  
Emil-Figge-Str. 50 - 44227 Dortmund  
Tel. 0231-755 2957

**Studienberatung:**  
[Alexander.Gurdon@tu-dortmund.de](mailto:Alexander.Gurdon@tu-dortmund.de)

**Alle Infos unter:**  
[www.musikjournalismus.eu](http://www.musikjournalismus.eu)

