9-2018

Music for AI Reports: Dual Prospects in Music Production

Achim Koh
The Graduate Center, City University of New York

MUSIC FOR AI REPORTS: DUAL PROSPECTS IN MUSIC PRODUCTION

by

ACHIM KOH

A master's thesis submitted to the Graduate Faculty in Liberal Studies in partial fulfillment of the requirements for the degree of Master of Arts, The City University of New York

2018

Music for AI Reports: Dual Prospects in Music Production

by

Achim Koh

This manuscript has been read and accepted for the Graduate Faculty in Liberal Studies in satisfaction of the thesis requirement for the degree of Master of Arts.

Date	Lev Manovich Thesis Advisor
Date	Elizabeth Macaulay-Lewis Executive Officer

THE CITY UNIVERSITY OF NEW YORK

ABSTRACT

Music for AI Reports: Dual Prospects in Music Production

by

Achim Koh

Advisor: Lev Manovich

Recent developments in artificial intelligence (AI) technology have led to industrial attempts at applying AI to music making, namely AI music. In the context of the history of music technology, AI music raises the prospect of a new phase that extends digital technology's role as central mode of music production. The computer has become an essential metamedium in contemporary cultural production, leading in the field of music to the digitization of tools and content and the digitalization of social institutions and relationships. This technological change had the dual effect of decentralizing music production while reinforcing capitalist logic in it. The rise of AI foreshadows an intensification of this dual technological potential, as projects like Google Magenta that offer new affordances demonstrate.

Contents

Introduction

Recent developments in artificial intelligence (AI) technology have led to industrial attempts at applying AI to music making, namely AI music. While AI has been extensively applied to modes of consumption of music, notably in the form of music recommendation systems, AI for music making has been a relatively underinvested field, especially in terms of industrial interest. However, advances in AI during the past decade, first and foremost in the field of deep learning, have increased the prospects of and lowered the barriers to certain applications to music production like generative composition. This technological change has been accompanied by increased activity in the field in the form of research, industrial investment and more independent individual experiments—although such categorization is not exclusive and often one type of activity depends on or bleeds into another. These new endeavors in AI music reflect a renewed mode in which corporate digital technology affects music making, and ultimately cultural production.

This thesis relies mainly on theoretical tools of media studies and science, technology and society (STS) studies in order to contextualize AI music as a socio-technological phenomenon, and analyze its social implications. Using these tools, I intend to map the actors of this fast-moving phenomenon mainly led by industrial initiatives and develop a rudimentary framework for critical engagement.

The earlier part of this thesis discusses the contexts within which AI music is situated. First, it describes the changes that occurred in the field of AI during the past decade, and its importance with regards to computational media and cultural production. In doing so, I discuss AI's reliance on data; its characteristic as a quantitative perspective on knowledge production; and its role as a capitalist engine. Furthermore, I take on a popular question about AI, namely the question of whether it will replace humans. While potentially useful, this question is prone to pitfalls that will distract from our discussion here; therefore, I clarify those pitfalls and frame the question in a more pertinent manner.

The next part of the thesis provides historical context for AI music. I rely on works on the history of music technology by Robert Strachan, Timothy D. Taylor, Paul Théberge, and more as the starting ground of my discussion of AI music. Especially, Strachan's work on the influence of digital technology in the late twentieth-century and early twenty-first-century music production will provide a foundation on which to locate AI music.[1] Digital technology in twentieth-century music production was a dual force, on one hand decentralizing-democratizing and on the other hand expanding-reinforcing the reign of capitalist logic.

In addition, I consider AI music with regards to the field of computer music, as an interdisciplinary practice between music and computer science. The practice of using computers to generate music, or more broadly the attempt to automate musical creative processes, is not a new development by itself and has a long history that at least dates back to the times of early digital computers. Nevertheless, while such endeavor has remained mostly within the domains of academia or musical experiments, recent years have seen industrial actors entering the scene. In other words, AI music is shifting from a primarily academic topic to a target of industrial pursuit, largely motivated by the advent of AI as a major technological innovation.

Turning the focus to the present, the following part illustrates recent developments in AI music, more specifically industrial efforts to apply AI technology to cultural production. This industrial interest in the computer-based generation of creative content manifests itself as capital investment—ranging from start-up companies to in-house research labs within tech giants such as Google or Spotify. These companies are producing a variety of easily accessible artifacts, from an open source tool that can create piano improvisations, to web interfaces that generate background

1. Robert Strachan, *Sonic Technologies: Popular Music, Digital Culture and the Creative Process* (New York, NY: Bloomsbury Academic, 2017).

music for videos; along with the artifacts is also produced publicity, which is no less important in our discussion. In addition to describing notable research, projects, and companies, I pay a little extra attention to Magenta, a project within Google Brain that primarily concerns machine learning methods for generating visual art and music.[2] Magenta's situation as a corporate R&D project, its deliberate commitment to open source and its collaboration both within and outside of Google makes it an especially pertinent example to this thesis.

AI music offers and lowers the barrier to new modes of affordances, not only in terms of generative composition but also in terms of new interfaces and assistive tools. On the other hand, it reinforces the network, in the Latourian sense, of a tech industry-centered knowledge production.[3] In the process, it further extends the dual potentials of digital technology with regards to music production. The resulting effects that can be observed or anticipated include a 'decentralizing' one, in the sense that skills and resources traditionally hard to attain become formalized and sometimes commodified; nevertheless, the opposite effect of centralization and concentration is also immanent in that most of these developments involve some degree of increased reliance on corporate infrastructure that constitutes said AI technology, and subsequently a more powerful capitalist logic. Ultimately, AI music is symptomatic of, and conducive to, shifts and renegotiations of power and relationships among the many actors in cultural production.

One thing I want to note is that this thesis focuses less on ontological or epistemological questions about music and artificial intelligence—e.g. how does the prospect of radically improved automation affect the way we think about the human subject and creativity; how do we

2. https://magenta.tensorflow.org/.

3. See Bruno Latour, *Reassembling the Social: An Introduction to Actor-Network-Theory* (Oxford, United Kingdom: Oxford University Press USA, 2005).

conceptualize art and music that can or cannot be quantified and automated—than sociological or anthropological ones—how such a prospect is being deployed in a social context; what are the political, economic, and cultural stakes involved. However, this is not to say that the two types of inquiries, separated by my rough deliberation, are incompatible or fully separable. Issues around the subject and creativity, musical knowledge production, economic power, and the politics of cultural visions are intertwined, as the section on the question of human obsolescence demonstrates. Nevertheless, this thesis focuses more on the social implications of AI music.

1. The Rise of Data-Centric AI

In *Software Takes Command*, Lev Manovich notes that the contemporary society has continuously moved towards the use of the computer as a universal metamedium, referring to Alan Kay's concept.[4] This universality is a key reason why the prominence of AI, described below, is the significant change that it is. The importance of computers as cultural media should be considered not only in its quantitative prominence and ubiquity, but also in its qualitative difference from pre-computer media. This difference lies in the universality of the computer as a media machine. Manovich points out that "the computer metamedium is simultaneously a set of different media and a system for generating new media tools and new types of media."[5] Here, Manovich is borrowing the term metamedium from Alan Kay, who describes the computer in a 1984 article as "a medium that can dynamically simulate the details of any other medium, including media that cannot exist physically. It is not a tool, though it can act like many tools."[6] In other words, the computer enables not only the simulation and representation of existing cultural content and practice, but entirely new ones.

Beyond the forms and modes of content and tools, computers also reconfigure human practices and social relationships within and around cultural production. Strachan points this out in *Sonic Technologies*, where he describes "the digitization of popular music practice and the digitalization of the institutions central to its production and consumption," referring to the development from the mid-twentieth century until recent years.[7] Strachan describes digitization as

---

4. Lev Manovich, *Software Takes Command* (New York, NY: Bloomsbury Academic & Professional, 2013), 101-106.

5. Manovich, 102.

6. Alan Kay, "Computer Software," *Scientific American*, no. 251 (September 1984): 52. Re-cited from Manovich, *Software Takes Command*, 105.

7. Strachan, *Sonic Technologies*, 2.

the shift of musical processes and recordings from analogue to digital media, which "fundamentally changes their nature, how they are perceived and carried out."[8] More specifically, this shift refers to "the integration of studio technologies within the personal computer and the centrality of the internet and Web 2.0 in the distribution and consumption of music."[9] On the other hand, digitalization has a more socio-cultural connotation, referring less to the technological changes than to "the ways in which the institutions (businesses, scenes and networks) of music and creative individuals have increasingly changed and adapted their central practices in the wake of digitization."[10]

Manovich states that software serves as contemporary governing force. He further details that digital media technologies "introduced *coding* as a way to store and transmit media. Simultaneously, these technologies also introduced a fundamentally new layer of media—*interface*, i.e. the ways to represent ('format') and control the signal. And this in its turn changes how media functions—its 'properties' were no longer solely contained in the data but were now also depend on the interface provided by technology manufacturers."[11] In addition to the importance of digitally encoded data and interface software as cultural media, one should keep in mind the influence of technology manufacturers as later sections discuss the history of digital music technology. Strachan also mentions his field recording practice, which was informed by four main elements: "technological possibility, an understanding of the field (in terms or institution, genre, spatiality and audience expectation), sonic affordance and wider discourses of creativity."[12]

---

8. Strachan, 2.

9. Strachan, 3.

10. Strachan, 4-5.

11. Manovich, *Software Takes Command*, 155.

12. Strachan, *Sonic Technologies*, 109.

Affordances

A brief consideration of the notion of affordance will come in handy in the following discussion as well. An important concept in fields such as media studies and human-computer interaction, affordance is commonly attributed to James J. Gibson, who described it as "what things furnish."[13] In "Theorizing Affordances: From Request to Refuse," Jenny L. Davis and James B. Chouinard provide an overview of various attempts at providing a definition of the notion.[14] Evans et al., for instance, describe it as "the variable process that mediates between properties of an artifact (features) and what subjects do with the properties of an artifact (outcomes)."[15] Faraj and Azad focus on the relational aspect, describing "the 'multifaceted relational structure' . . . between an object/technology and the use that enables or constrains potential behavioral outcomes in a particular context."[16] Drawing on previous literature, Davis and Chouinard propose an expansive notion of affordances as something relational, material, dynamic: "Affordances operate at the intersection of artifacts, actors, and situations."[17] Furthermore, they propose a framework with which to distinguish between different modes and sites of action. "Mechanisms and conditions [of

13. James Jerome Gibson, *The Senses Considered as Perceptual Systems* (Boston, MA: Houghton Mifflin, 1966), 285.

14. Jenny L. Davis and James B. Chouinard, "Theorizing Affordances: From Request to Refuse," *Bulletin of Science, Technology & Society* 36, no. 4 (December 1, 2016): 241–48.

15. Sandra K. Evans et al., "Explicating Affordances: A Conceptual Framework for Understanding Affordances in Communication Research," *Journal of Computer-Mediated Communication* 22, no. 1 (January 1, 2017): 35–52. Re-cited from Davis and Chouinard, "Theorizing Affordances," 242.

16. Samer Faraj and Bijan Azad, "The Materiality of Technology: An Affordance Perspective, 254. In Paul M. Leonardi, Bonnie A. Nardi, and Jannis Kallinikos (Eds.), *Materiality and Organizing: Social Interaction in a Technological World* (Oxford, England: OUP, 2012), 237-258. Re-cited from Davis and Chouinard, "Theorizing Affordances," 242.

17. Anthony Chemero, "An Outline of a Theory of Affordances," *Ecological Psychology* 15, no. 2 (April 1, 2003): 181–95. Re-cited from Davis and Chouinard, "Theorizing Affordances," 245.

affordances] thus create a scaffold through which artifacts *request, demand, allow, encourage, discourage*, and *refuse*, and do so through variations in *perception, dexterity*, and *cultural and institutional legitimacy*."[18]

AI in Modern Computation

This section provides some basic context about how AI came to the current prominence and what the application of AI to music making resembles. The field of artificial intelligence can roughly be described as the computer science oriented study of automated systems that make decisions and take actions. Nowadays, the popular use of the term AI often refers to applications of deep learning, a subset of machine learning methods that recently demonstrated highly superior effectiveness in tasks like natural language processing and image recognition than other computational methods. Deep learning's high performance in these areas was the combined result of improved algorithms and computational power, an increased availability of large datasets, as well as a revived academic and industrial interest in the domain. This trend initiated a new cycle of 'AI boom' in which both academia and industry, as well as media, have shown great interest in AI technology.[19] Unless otherwise noted, the term AI refers to these recent deep learning based approaches.

I mentioned above that deep learning refers to a subset of machine learning methods. While a detailed technical explanation of machine learning in general or specific deep learning methods is out of the scope of this thesis, I will attempt to provide just enough detail, not necessarily in a strictly technical language, to establish the difference between recent developments in AI and more

18. Davis and Chouinard, "Theorizing Affordances," 246.

19. AI research had, after being met with an initial enthusiasm, was subject to a long period of industrial and academic disinterest, namely the AI winter. For more on the early history of AI research, refer to Ian Goodfellow, Yoshua Bengio, and Aaron Courville, *Deep Learning* (Cambridge, MA: MIT Press, 2016).

traditional computational methods.[20]

As a subfield of computer science related to statistics, machine learning (ML) mainly concerns itself with using computers to make predictions based on patterns automatically found in data. An ML system is typically centered around a model that is trained from some data set, which includes multiple data points consisting of observations and corresponding decisions.[21]

One classic example of a data set is the Iris flower data set, which includes 150 data points of three different species of Iris flowers. In the data set introduced by Ronald Fisher, each point (or record) consists of four observations (or features), i.e. the length and width of the petals and sepals, and one decision (or class), i.e. the species.[22] Let us assume that the goal of our ML system is to distinguish between species based on the measurements. This type of task is called classification and is one of the traditional machine learning tasks along with regression, which involves predicting a continuous value as opposed to a discrete category. The model that our classification system needs is a sort of function that, given a set of four observations, outputs a decision, and does it reasonably correctly. In this case, training a model means going through the given data points in order to approach this function that we need, thus establishing a mathematical relationship between features and classes. The resulting model will ideally be able to classify an Iris flower into its correct species based on its measurements, even if that specific flower was not one of the data points used to train the model.

Note that the use of machine learning is not indispensable for the task per se. Instead of

---

20. More technically detailed resources can be found in Goodfellow, Bengio, and Courville, *Deep Learning*.

21. For a more detailed primer on machine learning, consult Gene Kogan, *Machine Learning for Artists*, http://ml4a.github.io/ml4a/.

22. Ronald. A. Fisher, “The Use of Multiple Measurements in Taxonomic Problems,” *Annals of Eugenics* 7, no. 2 (September 1, 1936): 179–88.

training a model from the data points, one could equally establish a system of human-determined rules, for example, that decides which species a set of measurements fall into. Nevertheless, in order to make such rules, one would need some prior botanical knowledge or go through a long process of trial and error. Even with expert knowledge, making a comprehensive rule might prove tricky or unfeasible; many machine learning models deal with highly complex data with much more than four features, in which the relationship between a certain feature and the corresponding decision is elusive to human interpretation. In other words, machine learning can prove very efficient when field expertise is not readily available or difficult to encode properly, or when there is a large amount of data.

The tasks of a machine learning system can extend beyond classification or regression. The predictive nature of machine learning methods allows models to be applied to generate content such as image, text and music. For example, by training a model on sequential data, the model can be used to predict the data point that should come next given certain previous data points. In the case of text, the data points could be strings of characters; in the case of music, they could be sequences of notes. This approach of applying machine learning to content generation has recently gained traction, in parallel to the rise of deep learning.

As Jean-Pierre Briot, Gaëtan Hadjeres and François Pachet note in *Deep Learning Techniques for Music Generation*, deep learning is not a precisely defined term; rather, it refers to a range of machine learning techniques that are based on artificial neural networks, using multiple layers of artificial neurons—hence the qualifier deep.[23] This type of approach, despite being largely unpursued for several decades due to the computer science community's widespread belief that artificial neural networks are essentially a dead end, nevertheless proved to be very efficient in

23. Jean-Pierre Briot, Gaëtan Hadjeres, and François Pachet, *Deep Learning Techniques for Music Generation - A Survey*, 2017, http://arxiv.org/abs/1709.01620.

various tasks involving images, text, and more following the early twenty-first century. Increased computational power, especially in the form of GPUs capable of fast calculations, and the availability of large datasets, namely the rise of Big Data as a result of digitization in many social realms, provided a material foundation for advances in machine learning techniques. Throughout the last decade or so, deep learning was further developed and widely applied, becoming a major technology synonymous in the popular context with artificial intelligence and increasingly central to not only the tech industry but also in broader social areas.

Advances in deep learning are shaping changes in music technology as well. As we have previously seen, contemporary music making relies on computational technology, from composition and production to distribution and consumption. Thus, the application of AI to music largely involves the automation of different types of musical practices that involve music in digital forms.

Tasks like music generation have been popularized to the point where start-up companies like Jukedeck and Amper are offering web-based music generation services, targeting audiences that include composers and video producers. While not strictly engaging in music generation, companies like LANDR and Neutron have developed automated audio engineering tools that control audio parameters based on the input track. These applications span many platforms and interfaces, from DAW plug-ins and standalone applications to web-based on-demand services. The availability of consumer-level products is significant in that it potentially signals a change in musical practices in a large scale.

On the other hand, tech companies like Google and Spotify have launched in-house R&D projects like Magenta, recruiting prominent scholars in the field as Jason Green reports.[24] In the

---

24. Jason Greene, “Do Androids Dream of Electric Guitars? Exploring the Future of Musical A.I.,” Pitchfork, June 12, 2017, https://pitchfork.com/features/overtones/10091-do-

case of Magenta, the project is also releasing software and datasets used in its research, allowing the broader public to use the tools in order to generate musical pieces, among other things. The many actors in AI's application to music are described in more detail in later sections. Here, suffice to note that AI involves a broad shift where cultural production increasingly relies on corporate data infrastructure as well as R&D.

A Quantitative Mode of Cultural Production

AI radically shifts the way computational technology operates, most notably through data-centric automation. As more and more social realms rely on data-based automation of decision making, the tech industry holds a position of increasingly greater power; it does so with the help of both data ownership and technological expertise. Essentially, AI as it exists in the current decade takes on the role of the next big capitalist engine, like the internet before it.

The rise of deep learning-based artificial intelligence as a popular technology is not just a matter of computer hardware and software engineering. As mentioned before, advances in deep learning were made possible partially by the advent of Big Data. The availability of massive datasets that resulted not only from the long-term digitization of contemporary societies but also a quantitatively and qualitatively radical change, notably thanks to the internet and mobile devices, in terms of means to produce and capture all sorts of data. The current AI phenomenon is best contextualized as an extension of, or a response to, Big Data. The digitalization of society, leading to big data, provided the platform for an influential industrial innovation: new and more powerful methods of automation over data, commonly referred to as AI. The recent developments in AI inherit the changes introduced by computational media throughout the twentieth century and apply

---

androids-dream-of-electric-guitars-exploring-the-future-of-musical-ai/. Google and Spotify have respectively hired experts in the field of music generation, Douglas Eck and François Pachet, each of whom were assigned in-house research labs.

large-scale automation to them, pushing social dynamics to a new phase.

As danah boyd and Kate Crawford note, it is important to keep in mind critiques of problems such as the bias inherent in the data or due to technical characteristics of systems.[25] The deployment of AI systems can also affect people disproportionately, sometimes reinforcing existing privileges and disadvantages. We can also find increasing amounts of critique of the way AI technology benefits tech companies, giving them a new scale and mode of economic power.[26] These critiques will serve as scaffolding in illustrating how AI technology potentially shifts power towards tech companies in cultural production as well.

Scientific and technological research ultimately negotiates power, even when the topic of the research is not explicitly political. In *The Postmodern Condition*, Jean-François Lyotard notes that the computerization of society comes with a reconfiguration of what counts as legitimate knowledge. To Lyotard, this reconfigured knowledge will consist of exteriorized and commodified data that is computable: "Along with the hegemony of computers comes a certain logic, and therefore a certain set of prescriptions determining which statements are accepted as 'knowledge' statements. We may thus expect a thorough exteriorization of knowledge."[27] This occurs in parallel to the shift in mode of production towards an information economy, where multinational corporations assume some of the power previously associated with nation states. The power in question is not necessarily limited to an economic one, since knowledge (science in this case) and politics are intertwined; "the right to decide what is true is not independent of the right to decide

---

25. danah boyd and Kate Crawford, "Critical Questions for Big Data," *Information, Communication & Society* 15, no. 5 (June 1, 2012): 662–79.

26. For a recent critique of the scale economy that empowers tech companies in an unprecedented level, see K. Sabeel Rahman, "The New Octopus," Logic Magazine, April 17, 2018, https://logicmag.io/04-the-new-octopus/.

27. Jean-François Lyotard, *The Postmodern Condition: A Report on Knowledge* (Manchester, United Kingdom: Manchester University Press, 1984), 4.

what is just."[28] Lyotard asks: "who decides what knowledge is, and who knows what needs to be decided? In the computer age, the question of knowledge is now more than ever a question of government."[29]

In today's AI research, in which data is the raw material and automated systems rely on models, which are quite literally exteriorized forms of knowledge, the construction and usage of data and models consequently hold socioeconomic nuances. Among these: gathering data is an important industry; availability of data is perhaps as valued as the social implications of those data; standardized datasets like those provided by Kaggle serve as de facto canons.

In their critique of Big Data, boyd and Crawford point to the lineage of data-centric thinking and knowledge production where quantification and objectivity are often confounded, a critique which also applies to AI.[30] Whereas the digitization-digitalization of the twentieth century normalized the notion of music as database, the data-centric automation of the current decade that relies on the curated and standardized artifacts that are datasets normalizes the idea of a modular and externalized model of creativity, that exists for instance in the form of encoded models. In this sense, the attempt towards a theory and/or standard practice of music stems from a fundamentally modernist quest.

While such epistemology raises many issues in broader applications, it becomes especially tricky when it comes to creative realms such as music making. Certain tasks can be measured easily through metrics such as accuracy, precision/recall rate, or speed and scale, which are commonly used to assess the performance quality of a machine learning system. When it comes to automatic generation of music, or broader AI-based automation of the music making process, the

28. Lyotard, 8.

29. Lyotard, 9.

30. boyd and Crawford, "Critical Questions for Big Data."

performance metric can be unclear. What counts as good music, or good tool, depends much on the human practice and perception—which conventional statistical metrics may have difficulties capturing. Therefore, applying quantitative methods to creative pursuits forcibly involves judgments based on subjective values from engineers and researchers. Nevertheless, this is probably not especially new in the history of music technology.

Maybe more problematic is the data itself; as Kate Crawford points out, "Data will always bear the marks of its history. That is human history, held in those data sets. So if we're going to try to use that to train a system, to make recommendations or to make autonomous decisions, we need to be deeply aware of how that history has worked."[31] While large music datasets are certainly more available than ever, many of them are limited to Western music, especially Classical music. One could not develop generative models for music with no massive dataset available; consequently, AI music research might contribute to reinforcing Western music's dominance.

Furthermore, AI music exemplifies a new trait of cultural production: an increased reliance on cloud-based computing resources and on large datasets hardly attainable through individual efforts. By contrast, individual expertise through institutionalized education, for example, somewhat loses importance. Datasets and GPU-powered computing instances act as tools that shape the concept of creativity. Access to both datasets and computing power needs to be broadened for a "democratization" of such a shift; on the other hand, a re-evaluation will be in lieu for modes of creativity that do not rely on these corporate-favoring resources.

Companies collaborate with platforms such as Kaggle and encourage participation to research areas of their interest. Meanwhile, they also employ in-house research labs that publish in academic venues. These sponsored researches are often publicly available, yet serve the

---

31. Crawford quoted in Scott Rosenberg, "Why AI Is Still Waiting For Its Ethics," *Wired*, November 1, 2017, https://www.wired.com/story/why-ai-is-still-waiting-for-its-ethics-transplant/.

companies' interest. Moreover, some of these companies (most notably Google) invite outsiders, artists and musicians in the case of Magenta, to collaborate on their research and products; these collaborators not only provide creative input that would be unavailable in traditional settings, but also 'legitimize' the research and products in question and their usage. As Crawford asks, this raises the question of "Who gets a seat at the table in the design of these systems?" These add up to a self-sustained corporate knowledge machine.

Shannon Mattern notes in her treatment of self-driving cars, quoting James Bridle, that corporate investments in AI applications bring about "*crucial social issues such as the atomization and changing nature of labor, the shift of power to corporate elites and Silicon Valley, and the quasi-religious faith in computation as the only framework for the production of truth — and hence, ethics and social justice.*"[32] This equally applies to applications in cultural production; what is at stake with the advent of AI is not simply new types of products or even creative content, but also a wide array of social renegotiations of power and knowledge.

AI in its broader sense is therefore a capitalist-cultural engine deeply intertwined with the operation of corporations. Nick Seaver draws the parallel in "Algorithms as culture": "We can replace 'corporations' with 'algorithms' . . . but we do not have to: the algorithms we care about are very often corporate products, and they seem more similar to corporations themselves—in their heterogeneity, diffuseness, and legal architecture—than to textbook algorithms like the bubble sort."[33] Just as tech companies are not entirely new capitalist entities but come with new forms of power and scale, AI exists with a continuity from previous modes of digital technology.

32. Shannon Mattern, "Mapping's Intelligent Agents," *Places Journal*, September 26, 2017, https://doi.org/10.22269/170926.

33. Nick Seaver, "Algorithms as Culture: Some Tactics for the Ethnography of Algorithmic Systems," *Big Data & Society* 4, no. 2 (December 2017): 1-12, https://doi.org/10.1177/2053951717738104.

Avoiding the Human Versus Machine Frame

An important goal of this thesis is to lay out a conceptual scaffolding that will facilitate discourse around AI music and its social implications. Considering the issues of labor and automation that are at stake with AI in general, often one can hear, deservedly, questions like: will AI replace humans? Given the extensive history of capitalism in which machines put people into precarious situations or straight out of work through the process of industrialization for the benefit of the capital, it is not difficult to see why concerns over the replacement of humans via automation are popular in the discourse about AI's social implications. Furthermore, a key goal of AI is automating tasks that are considered 'intelligent' and therefore highly correlated with human capabilities. Often in computer science literature, the benchmark for assessing the performance of an AI system is how certain humans perform given the same task. All of these contribute to the prominence of this question of human obsolescence.

Cultural production, with its complex labor relationships, is no exception to this line of questioning. This is not a new phenomenon; it has been so at least since the mid-twentieth century. Consider this quote by Hiller and Isaacson from *Experimental Music*: "almost inevitably, when the subject of our work has come up, the question has been asked: 'What is going to happen to the composer?' the implication being the composer is going to be put out of business by an 'electronic brain.'"[34]

Google's and similar companies' effort to establish themselves as a more important cultural engine takes place in parallel to their attempt to push, namely, an *AI economy*. One crucial thing in AI economy and the recent advances in deep learning that effectively herald the change in question is a renewed scale and scope of automation. This automation is enabled through decades of social

34. Lejaren A. Hiller and Leonard M. Isaacson, *Experimental Music: Composition with an Electronic Computer* (New York, NY: McGraw-Hill, 1959), 65.

digitalization culminating in the abundance of big data, academic and industrial innovations in machine learning research, and reduced cost of computational power. These preconditions are no exception to cultural production or more specifically music making, where digital objects have been the norm for decades now. Therefore, it should be no surprise to see renewed efforts towards automated music making, as recent academic and industrial trends have shown. Moreover, concerns about machines replacing humans, or the idea of *human obsolescence* as a result of AI-automation rightfully apply to cultural production as well.

However, with regards to the goal of this thesis, simply asking whether humans will be replaced by AI is at the same time too vague and too narrow a phrasing. Vague because it does not take into account the different technological implementations, human activities, and human-machine relationships; narrow because it is not only automation, but also decision making and whole infrastructures that are in stake when it comes to AI. Questioning whether AI systems will replace composers or stating that they will compose top-chart hit songs are simultaneously too vague and narrow to benefit a constructive reasoning, popular and catchy as they may be. Instead, we need to ask more detailed questions that identify relevant social actors and take their relationships into account. One such question will be: How does the incorporation of AI-powered automation in music making processes affect the labor of composers?

I try to stay clear of bold statements of technological prospect or philosophical contemplation about human replacement by automation. Essentially, such questions too often facilitate a human-centric defense of the individual or, on the opposite end of the spectrum, a techno-deterministic optimism for the quantification of everything. Both positions distract our discussion from the messy, gradual, and social nature of the human-machine relationship that cultural production embodies.

Perhaps because of this reason, many corporate messages try not to rely on the above positions, but instead to frame their product and research as pertinent to human-machine cooperation and human augmentation. While this perspective is closer to my own, I still want to point out that such a discourse still needs to be supported by social perspectives. Especially, one needs to be careful not to assume that notions like technology and creativity are somehow neutral, and instead remember that they are inherently political.

In its simple, catchy form, the human obsolescence question implies a type of imagery one could find in popular science fiction, where some generic sentient machines take over the traditional roles of people while these latter somewhat retreat from these roles. ‘Will AI replace humans?’ invites a dichotomic yes or no answer, where the choice lies between a resigned acceptance or a humanistic triumph. Such choice also usually involves an underestimation or overestimation of AI’s, and more broadly technological automation’s, potential to perform tasks currently thought of as only doable by humans.

Human-Centric Defense

For an example of underestimation, let us take a look at Douglas Hofstadter’s recent article for The Atlantic. Although this article mainly deals with text translation, it contains an argument very relevant to the discussion presented here. Consider the next paragraphs:

> The practical utility of Google Translate and similar technologies is undeniable, and probably it’s a good thing overall, but there is still something deeply lacking in the approach, which is conveyed by a single word: understanding. Machine translation has never focused on understanding language. Instead, the field has always tried to “decode”—to get away without worrying about what understanding and meaning are.
>
> Despite my negativism, Google Translate offers a service many people value highly: It effects quick-and-dirty conversions of meaningful passages written in language A into not necessarily meaningful strings of words in language B. As long as the text in language B is somewhat comprehensible, many people feel perfectly satisfied with the end product. If they can “get the basic idea” of a passage in a language they don’t know, they’re happy. This isn’t what I personally think the word “translation” means, but to some people it’s a great service, and to them it qualifies as translation. Well, I can see what they want, and I understand that they’re happy. Lucky them!
>
> Let me return to that sad image of human translators, soon outdone and outmoded, gradually turning into nothing but quality controllers and text tweakers. That’s a recipe for

> mediocrity at best. A serious artist doesn't start with a kitschy piece of error-ridden bilgewater and then patch it up here and there to produce a work of high art. That's not the nature of art. And translation is an art.[35]

In short, Hofstadter emphasizes the limitation of AI when it comes to "high art," something he argues translation ultimately is; as for the cause of this limitation, Hofstadter points to "understanding" or the lack of it.

While Hofstadter's high standards for translation and emphasis on the act of 'understanding' might be valuable to the study of the human mind and the development of an artificial general intelligence, they are of little use within the context of this thesis. The technological change that this thesis considers relevant is much broader than the exceptional and nuanced performance that some cutting-edge implementation can or cannot provide. Rather, it includes technology that may be limited to 'quick-and-dirty' performance but is adopted on a massive scale, and as a result effectuates qualitative changes in human activities. Excellence is by definition exceptional; mediocrity is how most things are done. Therefore, a simple dismissal of AI's capabilities based on its inability to do 'excellent things' relies on an arbitrarily narrow definition of the human activity in question. This can be misleading because the human activities that occur within the mediocre spectrum are not any less important; if anything, the fact that their automation is more likely and imminent means requires more attention and criticism on the process.

This does not mean that the workings and performance of current, perhaps unimpressive implementations are irrelevant; on the contrary, for critical discourse we need careful assessments with regards to how a specific AI system works and what it does or cannot do. But one must be

---

35. Douglas Hofstadter, "The Shallowness of Google Translate," The Atlantic, January 30, 2018, https://www.theatlantic.com/technology/archive/2018/01/the-shallowness-of-google-translate/551570/.

careful not to underestimate and dismiss the technological impact of AI based solely on current availability.

Consider the next statement from 2016, found at a forum of a home recording website: “Mastering really can’t be done with ‘auomatic’ [sic] settings, as anyone here will tell you - every song is different and requires different processing.”[36] Nevertheless, companies that offer automatic mastering do exist, most notably LANDR. This comment perhaps demonstrates a push/backlash against the ‘threat’ of a machine doing human work, or perhaps an adherence to the concept of human tasks.

Business-Driven Optimism

On the other hand, for examples of overestimation, one simply needs to look at headlines or business reports: “AI will replace [a type of creative job] / will be able to generate [a type of cultural artifact] by [an arbitrary year in the near- to mid-future].”[37] Both of these arguments assume a somewhat simplistic vision of technology and its relation to human activity, though in slightly different ways.

In his 2017 blog post, Rodney Brooks points out that some popular discourse around AI such as business-driven predictions tend to fuel misconceptions about the current state of technology and what can realistically be expected in the future.[38] This can be summed up by Roy Amara’s adage, “We tend to overestimate the effect of a technology in the short run and

36. “Automatic Audio Mastering Systems...,” Home Recording, accessed August 7, 2018, http://homerecording.com/bbs/general-discussions/mastering/automatic-audio-mastering-systems-391400/.

37. For an example of this type of language, see Robert Langkjær-Bain, “Five Ways Data Is Transforming Music,” *Significance* 15, no. 1 (February 1, 2018): 20–23, https://doi.org/10.1111/j.1740-9713.2018.01106.x.

38. Rodney Brooks, “The Seven Deadly Sins of Predicting the Future of AI,” accessed August 7, 2018, http://rodneybrooks.com/the-seven-deadly-sins-of-predicting-the-future-of-ai/.

underestimate the effect in the long run." Brooks further describes the common misconceptions, including a "faith-based argument" about seemingly magical (and therefore unprovable) future prospect; generalizing performance (specialized task) into competence (i.e. broader intelligence and skills); "suitcase words"; unproven assumptions that the current trend of rapid development will continue; hasty leaps into Hollywood scenarios instead of gradual changes; and an assumption of Silicon Valley-standard speed and scale in deployment, whereas real-world hardware and infrastructure is slower and more complex. Such thinking suffers from the assumption of a simplistic and dichotomous perspective on the human-machine relationship.

Acknowledging the Complexity of Human-Machine Relationship

One thing that needs to be unpacked with regards to the above predictions is the specific type and implementation of AI. There is no generic, singular AI that simply replaces everything; specific implementations matter. AI solutions are never fully *intelligent* or even solely ML-based in current real-world implementations. They are mostly combinations of various methods put together through trial and error; a gap lies between the ideals of machine music and the present implementations. Today, AI Music is being researched and developed notably in corporate-embedded research labs, consumer- or venture capital- targeting startups, and academia; in almost all cases, the research is about generating specific content like musical notes, rather than taking on the role of the composer, as we will discuss in the following paragraph. Many implementations require a combination of 'purely' generated results and manual or rule-based editing; not many are openly accessible, with the few exceptions like Google Magenta.

Another unpacking would involve formulating 'replace' with more nuance. Human-machine relationships are diverse and complex. One might try to duplicate the composer's process and replace it with a machine or automated process, making it a competitive relationship; another might try to create automated processes that assist or transforms the musician's process, building a complementary relationship. Most relationships will lie in between these two extremes. Furthermore, any human-machine replacement or the establishment of a certain relationship will take place gradually over time, if anything. Making it a binary question disregards the diverse relationships between humans and machines.

Furthermore, a wide range of human activities is simplified into a generic image of human individuals when we simply say "humans." This deviates from a constructive discussion by failing to clarify what type of human acts are being discussed; even the relatively narrow field of music composition differs wildly across genres, industries, geographic areas, and so on. Writing a classical Western symphony is a type of task very different from producing a pop song or film soundtrack, or from composing non-Western types of music—which incidentally tends to be much less researched than the former examples. Undertaking the composition from start to end is also very different from having assistive roles in certain parts of the composition process. All these are valid human activities related to music creation that are currently more or less difficult to replace using AI. Not specifying which type of human activity is being discussed, or dismissing the whole replacement scenario because it does not involve certain types of human activity, are unfortunately too often found in popular discussion.

The underestimating (or human-optimistic) statement is problematic because it ignores the wide spectrum of human activity, entangled with technology and partially mechanical or repetitive. It is not the assumption of a non-automatable human quality (which I do not intend to get into anyway) that makes this type of statement elusive; rather, the problem is in the overrepresentation of so-called human quality. Human existence and activity take extremely diverse forms, both within and across societies. A large part of this human activity, even in cultural production, is entangled with mechanical, tedious, repetitive, and often dubiously reliable tasks. This does not make such tasks non-essential to culture; on the contrary, this dull bulk of cultural labor constitutes a large part of cultural production. Therefore, the tendency to underestimate AI because it will not be able to do sophisticated tasks only humans are able to do (whether this latter assumption is correct or not) is irrelevant because one does not need to replace every single one of such sophisticated tasks in order to transform and disrupt how human cultural production functions. AI reports that throw out optimistic predictions are problematic, again not because their assessment of the technological potential is right or wrong, but in this case because they assume a trajectory that is neither well-described nor justified; while cultural projects involve various factors including consumer perception, these projections assume that developments in engineering will ensure the social acceptance of the result. They project, behaving like simple regression functions.

In other words, the ‘human obsolescence’ question disregards the multiple and gradual steps between now and then. This is problematic not because such a scenario is implausible; there is no inherent reason why automated computer agents cannot one day ‘produce music’ as well as human musicians do, nor why as music listeners we humans should not accept machine-made music as equivalent to human-made music. There is no foreseeable reason why technological development should stop at a given point, and the notion of what counts as art and music has always been one in motion throughout human history. What makes the ‘human obsolescence’ question problematic is that it almost never talks about the path between now and that future point. It is of speculative interest at best, but more often than not it distracts from more relevant questions.

What the human versus machine question does rightfully touch on, however, is the problem of labor with regards to automation. There exists, quite independently from experimental and academic artistic pursuit, a wide range of creative labor that serves diverse industries—this labor is most likely to be replaced, and the industry seems to have determined it a major goal. Small gig economies are easy targets. We can reasonably assume a commodification of assistive tools, which can have major impacts when implemented in scale.

As Timothy Taylor points out in *Music and Capitalism*, digital technologies have created some efficiencies (that are exploited for profit) and also inefficiencies (such as a too fast-paced process which harms creative decision-making), both taking more workers' time.[39] "Neoliberalism has taken advantage of new technologies to speed up the assembly line for many workers, even in those same technologies offer some others an illusion of freedom and creativity."[40] One can imagine similar ramifications with AI-powered music. As partial processes or entireties of creative labor become available for automation, the relationship between social actors will be negotiated and shifted. More content will be made, the entry barrier will be lowered, and the assembly line accelerated. Then music automation threatens much less the star musician's status (a presumably 'hand-crafted' art that caters to specific ways of consuming human narratives) than more industrialized, demand-meeting labor such as television and film scoring.

---

39. Timothy D. Taylor, *Music and Capitalism: A History of the Present* (Chicago, IL: The University of Chicago Press, 2016), 145.

40. Taylor, 153.

2. A Brief History of Digital Music Technology

The social implications of AI music can be informed from the changes in music production that digital music technology brought about. Strachan posits "three main epochs of digital music technology"[41] starting from the early days of mainframe computing from the 1950s to the 1970s. The first, "exploratory" epoch was "a period where ideas and technologies were coalescing in the somewhat esoteric and rarefied context of late Modernist art music and the 'research for research's sake' environment of academia."[42] During this period, experiments in the use of computation in music making were mainly conducted in large organizations such as Bell labs, IRCAM, MIT, Stanford. While these experiments had limited impact in terms of widespread practice, they nevertheless provided the basis for future developments.

The second "expansive" epoch from the eighties to early nineties was characterized by the "digitization of analogue instrumentation and studio technology."[43] The landmark development of this period was the MIDI standard, as well as MIDI-compatible synthesizers. As the musical processes became increasingly digitized, new common practices in composition, recording, sampling (as the widespread adoption of compositional ideas dating back to musique concrète) and sequencing ("the progressive removal of any immanent criteria for distinguishing between human and automated performance")[44] evolved. Taylor notes that "the introduction of digital technologies in the 1980s [reduced] the differences between different kinds of musical production."[45]

In the third, "convergent" epoch, "earlier patterns of digitization converge upon . . . the

---

41. Strachan, *Sonic Technologies*, 4.

42. Strachan, 4.

43. Strachan, 4.

44. Strachan, 5.

45. Taylor, *Music and Capitalism*, 33.

personal computer," most notably through the DAW as virtual environment and single integrated interface.[46] This change, backed by personal computers becoming affordable and powerful enough to run professional-level software, resulted in increased access and lowered entry barriers to studio technologies. The interoperability afforded by such a change led to a "transition in what it means to be creative. The lines between composition, production and performance have become progressively more blurred, leading to a whole new generation of practitioners whose roles are less easily placed within stratified divisions of labour that traditionally characterized the creative processes of popular music throughout the twentieth century."[47] This epoch since the late-twentieth century is further underscored by consumer digital technology becoming the central mode of musical practices. Such a change led to the simultaneous empowerment of individual musician and increased submission to capitalist logic, a trend that we continue to see being reinforced. Digitization, which brings the musical process into the realm of computation, is the precondition for data-driven AI music; the establishment of large data sets and the development of data-driven methods can be seen as a continuation of digitization. AI music as digitalization concerns the changes in the structure and practice of diverse actors, from institutions to individuals; with this regard, this thesis mainly considers the institutional layer and less the individual one, which is to yet to happen in a large scale.

As an extension of Strachan's three epochs of digital music technology, I propose that we are witnessing the beginning of a fourth epoch characterized by new, automated methods of music composition and production. The foundation that makes such practices possible were laid out through the previous digitization of music making, which brings sonic elements almost entirely

---

46. Strachan, *Sonic Technologies*, 6.

47. Strachan, 7.

within the scope of computation. The accumulation of large datasets and the explosive development of deep learning research—accompanied by a wide social interest in the field of AI in general—have produced an industrial interest in what previously was considered difficult, or remained the subject of experimental academic work: the automation of creative decision processes such as composition, arrangement, and sound engineering.

While algorithmic composition dates from earlier, and the notion that math and music are compatible is literally ancient, the idea of automatic computer composition evolved during the early days of mainframe computing. So the current epoch is where ideas and promises around automatic composition finally come forward and catch up with other development of music tech, largely thanks to the Big Data era, allowing a product-level implementation of such prospects. This industrial shift is what most distinguishes the period.

The increased availability of automated methods of music making, which results from this shift fueled by industrial interest, alters the technology's functioning in a McLuhanian sense, signaling the approach of a new phase in the history of music technology.

Decentralization

An important keyword useful for the discussion of digital media and music technology is decentralization. Media theory literature informs us that decentralizing/democratizing technologies have administered enduring and great influences throughout history. In *Strange Sounds*, Taylor notes that "those technologies that catch on are ones that lead to the decentralization of music making and listening, and more flexible ways of listening, and so MP3s or their successors are here to stay."[48] Nevertheless, 'decentralization' here does not necessarily mean a straightforward democratization or empowerment of the masses; instead, complex re-negotiations of stakes,

---

48. Timothy D. Taylor, *Strange Sounds: Music, Technology and Culture* (New York, NY: Routledge, 2014), 19.

interests and value systems take place.

Strachan points out that "Prosumer technologies such as camcorders, sound recording equipment and DAWs have provided relatively affordable tools which have been utilized in an uncountable plethora of high-quality, self-produced media."[49], an observation that also extends to the subsequent rise of smartphones. This 'democratization' of music production is exemplified by the eighties–nineties home studio phenomenon supported by MIDI technology and affordable equipment such as "synthesizers, sequencers and home recording equipment" and the post-nineties universalization of PC and laptop production, helped by cheaper consumer-level software and hardware, reduced gap between pro-level and amateur-level technologies (everything could be done on the personal computer), and the internet (especially illegally downloaded software). In return, "For the new media industry, from high-tech corporations to start-up companies, the idea that their products somehow empower consumers, enable creativity and allow a voice for individual expression has become an overriding logic of product development, as well as a key to how such products are marketed to, and understood by, the public."[50] These "twin aspects of digitalization" led to "the democratization of culture, empowerment and everyday creativity" becoming enduring concepts.[51]

This shift towards decentralization also involved an increased influence of capitalist logic in many stages of music making, where musicians are at the same time harnessing new technical potentialities and fundamentally embracing capitalist relations. Strachan notes that through this process, some traditional forms of agency are lost but also compensated through shifts in creative practice.

---

49. Strachan, *Sonic Technologies*, 20.

50. Strachan, 19.

51. Strachan, 20.

AI as data-driven automation provides reasons to posit that a new phase in the history of music technology is about to emerge. As automated processes assume part of the agencies closely tied to traditional notions of music creation, and as AI continues the late-twentieth-century trend of consumer digital technology becoming the central mode of musical production and distribution, one is faced with the question: to what end are these new technologies being used and developed? Or, in the spirit of Jacques Attali's postulation of music as harbinger of the contemporary political economy in *Noise*,[52] one might rather ask: what is the music we must aspire to, in order to realize a democratic deployment of data-driven technologies—versus a more centralized and monopolized situation?

Consumer Digital Technology as Central Mode of Production

As previously discussed, recent years have seen advances in the application of AI technology such as deep learning techniques to creative decisions in music making—both in terms of improved performance and increased access to the tools—and a concurrent rise of industrial interest in conducting such research and building products out of it. AI music signals the launch of a new phase of digital music technology; what notably characterizes this period is the data-driven automation of musical decisions in processes like composition and audio engineering.

This is not to say that data-driven automation has only begun being applied to music in the past few years. The most notable example of an already established such practice will be algorithmic recommendation systems, which apply data-driven decisions to the previously human task of music curation. In a sense, the seeds of AI music have been sowed as streaming services embraced data-driven methods in order to automate decisions. This is especially worth noting since modes of music production have always closely interacted with modes of consumption, as Taylor

52. Jacques Attali, *Noise: The Political Economy of Music*, trans. Brian Massumi (Minneapolis: Univ Of Minnesota Press, 1985).

notes in *Strange Sounds*. Nevertheless, my point here is that the expanse of automation into music production is indicative of broader and potentially more radical changes in musical institutions and individual practices.

We have seen previously that technological changes have been deeply intertwined with changes in modes of cultural production and the course of capitalism in general. This is also the case with music technology, as we will see below. However, this is not to say that technology imposes itself inevitably to cultural processes. In his analysis of digital technology's influence on musical practices, Strachan points out that "Rather than taking as wholesale the idea that developments in technology have constituted a disruptive break with previous practice, it is important to place them within the structural specificities and historical legacy of popular music [cultural] production."[53] In other words, the history of music technology should be accounted for with regards to the social aspects of music production.

In *Noise: The Political Economy of Music*, Attali addresses the relationship between music and capitalism, demarcating historic stages of music production. As Timothy Taylor summarizes in *Music and Capitalism*, Attali distinguishes among four stages of the mode of production of music: "Sacrificing" or the premodern era, "Representing" or the era of published music, "Repeating" or the era of recording, and "Composing" or the contemporary regime characterized by digital technologies.[54] The major early technological development in Western music that saw the departure from the premodern era and was the notation system, later accompanied by the printing press and movable types for music. As music came to exist as the tangible artifact that is sheet music, its mass circulation was also enabled, establishing the sheet music industry. This process of

---

53. Strachan, *Sonic Technologies*, 21.

54. Taylor, *Music and Capitalism*, 3.

commodification led to “new markets, new musical forms, genres, and techniques, and new composer-entrepreneurs In a more ontological perspective, the notation system provided a standardized way of encoding physical events into symbols. Music started becoming information as we understand it, a process that has continued to accelerate until the present era of data.

The invention and adoption of recording media initiated another big change in music production. This late nineteenth-century development along with urbanization, concert halls, apparatus of press and publicity, and the rise to hegemony of finance capitalism, extended the process of music commodification and contributed to the rise of the twentieth-century capitalist music industry. Taylor notes about the notion of exchangeability in music: “With the rise of what can be called a modern, industrialized music industry in the late nineteenth and early twentieth centuries, music produced for the purpose of exchange has become dominant.”[55]

Taylor notes on the relationship between capitalism and music production that across regimes of consumption in American culture, from late nineteenth-century to early twentieth-century advertisement industry and radio to post WWII marketing and television and neoliberalism and WWW, “music, musicians, and sound reproduction technologies were all aggressively advertised, though to varying degrees.”[56] He also notes that “campaigns [for audio recordings] were also bolstered by sympathetic journalists and other writers who covered new recording technologies positively and who helped spread a powerful ideology of the democratization of access to what was thought to be great music.”[57] As we will later see, this collaboration of developments in music technology and advertisement are almost identically found in the contemporary context of AI music.

55. Taylor, 33.

56. Taylor, 26, 35.

57. Taylor, 36.

Taylor also notes that "convincing people that to purchase recorded music was preferable to making it themselves was a slow and arduous process. . . . Consumers at first had to be assured that they weren't surrendering agency, that they were bringing their interpretive powers to the 'performance' of player piano rolls."[58] Here as well, we find parallels in AI music services that position themselves as 'assistive' tools that enhance creativity rather than replacing human roles, a message that possibly caters to the resistance of adopters.

Attali proposed, as Taylor notes, an optimistic view about his posited fourth and contemporary stage of music production, in which people will themselves compose rather than remain mere consumers of music. This proposition highly corresponds to Attali's assumption that music as superstructure predicts, or shepherds, economy. Nevertheless, the twentieth century did not play out as Attali had hoped.[59]

AI music constitutes part of the contemporary development of computational technology, which has progressed up to this point in a way that, simply put, assimilated and reconfigured everything into and around digitized forms. Now, with all media turned digital and commanded by software, data-centric computational automation takes the next step of assuming the role of creation. But this assumption is mitigated to some degree by diverse limitations, not the least of which lies in human choices about creative processes.

The field of machine learning has existed since the mid-twentieth century but was notably bolstered over the last two decades by academic advances, cheaper computing power, and the rise of massive datasets. Recently often branded as artificial intelligence (AI), the technology is

---

58. Taylor, 35.

59. For a detailed historical take on digital media's relation to music production, see Paul Théberge, *Any Sound You Can Imagine: Making Music / Consuming Technology* (Hanover, NH: Wesleyan University Press, 1997) and Strachan, *Sonic Technologies*.

drawing tremendous industrial interest, with some claiming it to be “the new electricity.”[60] The ability offered by machine learning to detect patterns from data is indeed influential in our computerized societies, where the scale and scope of data raise ontological questions.

With regards to music making, this opens up new potentials, much like how computers since the sixties and software tools influenced cultural production. The enhanced ability not only to detect patterns from large datasets and to construct models that can emulate human decisions but also to implement ‘realistic’ results is pertinent not only to automatically generated music but also in the automation of creative practice throughout the process of music production. Generative music and augmented modes of music-making question what counts as creativity.

Moreover, as Théberge notes, the late twentieth century saw consumer digital technology becoming the central mode of musical practices,[61] of which AI music can be seen as the latest development. And like with previous technologies, one can anticipate a re-negotiation of value and agency within creative practices and across different actors in the music industry, sometimes described as the democratization of certain technologies but often entailing more complex results, including an increased influence of capitalist logic over creative practice.

---

60. Shana Lynch, “Andrew Ng: Why AI Is the New Electricity,” Stanford Graduate School of Business, accessed August 7, 2018, https://www.gsb.stanford.edu/insights/andrew-ng-why-ai-new-electricity.

61. Théberge, *Any Sound You Can Imagine*, 3-6.

3. AI Music / Computer Music

"Pythagoras (around 500 B.C.) believed that music and mathematics were not separate studies," as George Papadopoulos and Geraint Wiggins point out in "AI Methods for Algorithmic Composition: A Survey, a Critical View and Future Prospects."[62] This relationship between the two fields has been an important inspiration for the application of computers to music, which started taking place in the early days of digital computation during the mid-twentieth century.

In *Experimental Music*, Hiller and Isaacson categorize what they refer to as experimental music, which in the contemporary context would be better correspond to the term computer music, into two broad approaches. One is the "experimental studies of the logic of musical composition," or algorithmic composition, and the other is the "production of musical sounds that are non-conventional," or electronic music.[63] This latter approach was pioneered by Max Matthews of Bell Laboratories; the authors of *Experimental Music* were more pertinent to the algorithmic composition, being "probably the first who used a computational model using random number generators and *Markov chains* for algorithmic composition" according to Papadopoulos and Wiggins.[64]

Since then, computers became cheaper, ubiquitous and more powerful, while software for making music and manipulating media in general have also advanced. Along with this development, computer music has also expanded as a field and become ubiquitous as a practice. As a result of the digitization of media and the digitalization of musical practices, the typical musical process in the current day comprises, or consists entirely of, the application of computers to music

62. George Papadopoulos and Geraint Wiggins, "AI Methods for Algorithmic Composition: A Survey, a Critical View and Future Prospects," in *In Aisb Symposium on Musical Creativity*, 1999, 110–117.

63. Hiller and Isaacson, *Experimental Music*, 37.

64. Papadopoulos and Wiggins, "AI Methods for Algorithmic Composition," 110.

making.

A recent development in digital computation is the advance of artificial intelligence technology based on deep learning techniques, which is based largely on the prominence of large datasets and the increased availability of computing power. As academic and industrial interest in deep learning grew, its application to music making did as well. In recent years, established machine learning conferences like NIPS and ICML included several papers and/or panels about musical applications of AI,[65] while start-up companies and corporate-embedded research labs dedicated to the field emerged. These endeavors continue and expand on earlier approaches to computer music.

The span of these recent AI music enterprises is broad and diverse, and it should come as no surprise considering the ubiquity of computer technology in all things related to music. Enumerating all different types of deep learning applications to music is out of the scope of this thesis; nonetheless, it will be helpful for further discussion to take a closer look at AI music approaches pertinent to this thesis, specifically ones that relate to music generation.

In their 2017 book, Jean-Pierre Briot, Gaëtan Hadjeres and François Pachet provide a survey of existing attempts at applying deep learning to music generation.[66] While the book is not an exhaustive survey, it provides us with a fairly recent overview of the fast-moving field, as well as a helpful framework with which to analyze technical approaches. The authors propose some criteria with which to categorize different projects: objective, representation, architecture and strategy. These are not mutually exclusive dimensions, as a choice within a specific criterion might

65. "Advances in Neural Information Processing Systems 30," Electronic Proceedings of the Neural Information Processing Systems Conference, accessed August 7, 2018, https://papers.nips.cc/book/advances-in-neural-information-processing-systems-30-2017.

66. Briot, Hadjeres, and Pachet, *Deep Learning Techniques for Music Generation - A Survey*.

dictate what the choices must be in other criteria. Objective refers to the end result that an AI system is built to generate, such as a series of musical notes, chords, harmonies, or some combination of the above.

Representation refers to the type of data used to train the system and generate content; the data could be in the form of signals, such as raw audio or processed versions of it, or in symbolic forms, such as MIDI or text. Architecture means the specific technical traits of the deep learning algorithms and how they are designed, and strategy refers to different ways of using the architectures. Approaches that fall into different combinations of these criteria allow music generation at a level considered previously impossible, for example in Papadopoulos' survey.

The practice of using computers to generate music, or more broadly the attempt to automate musical creative processes, is not a new development by itself and has a long history that at least dates back to the times of early digital computers. Nevertheless, while such endeavor has remained mostly within the domains of academia or musical experiments, recent years have seen industrial actors entering the scene. In other words, automatic music is shifting from a primarily academic topic to a target of larger-scale industrial pursuit, largely motivated by the advent of AI as a major technological innovation.

Recent Developments in AI Music

AI opens up new potentials for creation, much like how computers in the sixties and software tools influenced media art and cultural production in general. Specifically, the enhanced ability not only to detect patterns from large datasets and to construct models that can emulate human decisions but also to implement 'realistic' results is pertinent to generative art, including music. Commercial applications of automatically generated audio tracks, along with methods like data sonification and machine learning-based composition used in contemporary music, question what counts as creativity, as well as human-centric perspectives on music and other cultural fields.

Machines that do not stop at assisting but go further in taking over the whole process of creation forecast the soon-to-be-needed renegotiation of terms such as subject, agent and actor.

AI itself tends to be a vague term; the phrase points to the fluidity of the field of artificial intelligence, where things once considered human-level intelligence (and therefore important goals) become quickly part of the established computer science knowledge, somewhat losing their 'intelligent' status, as soon as someone succeeds in making a computer achieve them. Accordingly, tasks like recognizing handwritten digits or the silhouette of an object in a picture, for example, once valuable goals in the achievement of artificial intelligence, are now closer to basic foundations for more complex versions of intelligence.

Given how fluid the term AI can be, it follows that AI music also can, referring to any type of musical achievement by artificial agents that was previously considered only possible for human musicians, from singing robots to original music generated by computers. Such a broad and fluid concept would be too unfocused for this thesis. AI music as it is used in this thesis refers to a more specific phenomenon in terms of its technical implementation, its goals, and its situational context. One important technical aspect of the AI music in question as opposed to a more generic concept is the importance of machine learning and large datasets in its implementation. More specifically, the increased industrial interest stems from, and is part of, developments as well as investments in deep learning technology that recent years saw.

The fact of being centered around deep learning gives AI music certain tendencies in its goals and implementation. First, reliance on large datasets; second, applications centered around machine learning goals such as classification/regression and generation; last, reliance on cloud infrastructure. Another related characteristic of AI music is the industrial interest, as compared to academic and experimental-music based initiatives—although these fields are not mutually exclusive. This characteristic is especially notable when it comes to blurring distinctions between research and marketing, as we will see in the case of Magenta.

What I intend to designate by AI music is a recent phenomenon mostly pertinent to the tech industry. Particularly, it is a set of mainly industry-led attempts to apply deep learning research to the automation of musical creative practice. Such attempts materialize in the form of commercial or open source user-level software tools, ranging from programming toolkits to downloadable virtual instruments and web-based services and new (open or closed source) research on algorithmic methods that enable previously unachievable tasks, or improve the efficacy, in content generation or automated control of musical content. These research results are often not completely separated from the above tools, as these are built as implementations of—or at least on the basis of—researches.

The above artifacts of AI music are also accompanied by a series of narratives that are communicated as corporate marketing initiatives, media reports, and/or informal essays by researchers and developers. I consider these narratives to be not a by-product but an essential component of the phenomenon in question; the communicated 'hype' is an important mode in which corporate R&D functions, and at the same time it serves an important role in encouraging continuous corporate engagement in the field.

Content-wise, the narratives often revolve around two issues; one being whether AI can be used in a sophisticated manner enough to create human professional-level music (and thus to replace human musicians and other experts)—and the other one relating to the notion of augmented intelligence, i.e. utilizing AI in a way that assists and enhances human creative practice. The former contains a concern over human obsolescence as a result of automation; the latter is often presented as an alternative to the former, providing an optimistic view of the technological consequences. But as the history of music technology demonstrates, there are no strict boundaries between the objects of both discourses; tools developed explicitly for human automation sometimes lead to creative achievements, and well-intended creative endeavors can be nonetheless exploited for capitalist gain.

Coming back to what AI music designates, the popular use of the term refers broadly to music composed by an AI software, without additional work by a human. But it can also mean collaborative composition, assistive producing tools, customized recommendations, context-based remixes, etc.; AI is a broad term, as is music production. Nevertheless, here I focus on AI music as generative composition; specifically, I limit my scope to deep-learning based music generation. This is a new field with quite prolific technological output but relatively lower socio-cultural analysis.

AI music's reliance on in deep learning relates to its being an increasingly industry-led endeavor—although it overlaps a lot with academic research—that incorporates a field of immense industrial interest. Considering deep learning-based AI music came to prominence in part to the establishment of big musical data sets and fast GPU cloud infrastructure, large tech companies are in advantageous places to lead related researches. These researches involve automating parts of the musician's job that was previously done manually, although often it is described as something different, in the line of assistive, augmentative, or transformative.

With regards to the creative practice itself, the topics of AI music research include convenience in the sense of human labor automation; explorations of new forms of expression; and assistive tools. These topics involve the commodification of automated processes and an increased accessibility to a range of creative skills, some more focused on the former and some on the latter. Openly available software tools such as Google's Magenta provide generative methods for music elements of higher or lower levels, such as melody and polyphony, rhythm, accompaniment, score, lyrics, etc. Companies like LANDR and iZotope have products that offer AI-based automatic mastering and mixing. An increasing number of start-up companies, such as Jukedeck and Amper, offer consumer-available hybrid methods that generate full audio tracks, which can be used as background music for videos for example.

An extensive survey of available tools and active companies is outside of the scope of this thesis. Instead, later sections address one specific project by Google, Magenta. This is partly due to the fast pace in which the field is developing, where both artifacts and actors are not quite established and change quickly. In addition, Magenta's situation within Google, a tech giant heavily invested in AI, makes it more pertinent to the current discussion of corporate-led technology research and music making. Furthermore, the project's commitment to open source allows for a closer look at the tool's affordances.

The Institutions of AI Music

One notable trait of recent industrial interest in AI music relates to its institutional locus. In addition to being a research about cultural production, the research itself reflects a mode of cultural production and knowledge establishment that is pertinent to the current state of the tech industry; open source corporate R&D. Through this effort, corporate entities reposition/reinforce their products and themselves as integral components of cultural production.

The social investment in AI-based automation brings about concerns about labor replacement, social and cultural bias embedded in algorithms and datasets, and an increasing reliance on the digital infrastructure such as GPU servers and normalized datasets—leading to the issue of new centralized power in the forms of tech companies. From an STS standpoint, and using Google's Magenta project as example, I attempt to paint a snapshot of this phenomenon in action, which includes a data-driven capitalist economy, corporate-embedded research labs, open source machine learning software, academic conferences and data science competitions, as well as cloud computing infrastructure, among others—and through which tech companies like Google are essentially repositioning/reinforcing their products and themselves as integral elements of cultural production.

While Magenta is just one of many corporate-led projects around automated music making, I give it a special attention due to its institutional specificity that allows the connection with different layers of technological production. Moreover, the main narrative it pushes—that the project is not about replacing human artists and musicians, but helping them create interesting work—deserves an examination; despite the best intentions from the engineers and musicians working on the project, the technology being developed is inherently about automating, at least partially, human activities. This apparent contradiction results from an attempt to culturally legitimize an industrial engineering endeavor, but it is also reflective of the complex nature of technological change.

What differentiates current AI music projects like Google Magenta from traditional academic endeavors is the locus of research and the institutional network that supports it; one of the largest tech/advertisement company with the purpose and means to establish AI as not only its own agenda, but also a social one. This contrasts to previous types of institutions like the university and other research institutes such as IRCAM and Bell Labs, which became a refuge for composer against the establishment of capitalist logic in the music industry.[67]

Magenta and its institutional infrastructure including adjacent entities such as Google Brain, Google Art & Machine Intelligence, and Google Arts and Culture often work in collaboration. As part of Google Brain and a derivative of TensorFlow, Magenta participates in corporate promotion efforts such as Google I/O; moreover, Google's longstanding effort to push a narrative of collaboration, education and empowerment is as well in motion in the case of Magenta. Designers and UI/UX specialists, both in-house and external collaborators, allow for a quick cycle of visual

67. Taylor, *Music and Capitalism*, 33.

and interactive communication of research results, which invites more audience including developers and researchers. Indeed, this network seems eager to demonstrate the importance of HCI, which is explicitly referenced in texts from AMI, for instance. In a trend of aggressive hiring of academics by tech companies,[68] academic research itself is embracing capitalist relations, marketing and managing as mode of practice. Largely embracing open source, user participation is allowed and encouraged, but in a manner that aligns with corporate desires.

A Developing Field in Need of Legitimization

Douglas Eck's question, "can we use machine learning to create compelling art and music,"[69] although crude, serves as a sort of double manifesto, reflecting the goals of this research which include not only efficient computational methods, but also judgment over what is desirable with regards to musical creativity. François Pachet's call for a "serious" analysis,[70] while an effort to promote the album created using his software, seems to reveal an ongoing struggle to legitimize the use of AI in music through non-computer-science authority. This illustrates the not-so-simple relations among actors in the music ecosystem, where efforts from the IT industry are made to introduce this technology.

It is telling that Pachet, one of the most spotlighted researchers in the field, is calling for a musicological analysis of computer-generated music; a cultural practice needs to be legitimized by

68. Cade Metz, "Facebook Adds A.I. Labs in Seattle and Pittsburgh, Pressuring Local Universities," *The New York Times*, June 1, 2018, sec. Technology, https://www.nytimes.com/2018/05/04/technology/facebook-artificial-intelligence-researchers.html.

69. Douglas Eck, "Welcome to Magenta!," Magenta, accessed August 7, 2018, https://magenta.tensorflow.org/blog/2016/06/01/welcome-to-magenta/.

70. Francois Pachet (@francoispachet), "We Need Serious Musicologists to Analyze AI-Composed Music, in the Spirit of Chess Masters Analyzing Chess Games Played by Alpha-Zero. There Is Remarkable Stuff Going on. #HelloWorldAlbum #SKYGGE," Twitter, January 18, 2018, https://twitter.com/francoispachet/status/953980629254844416.

human perspectives, in the adjective's narrower sense. As endeavors in creative AI seem to take the form of computer scientists exploring the domain of artists and humanists, this call perhaps touches on the right notes. Nevertheless, Pachet's argument seems to be limited in that it does not necessarily call for a collaborative discourse within the building process; rather, it seems to entice a post-factum discussion of the finished pieces as a means of establishing legitimacy of new practices in AI music. Perhaps a more critically constructive discourse might be facilitated by incorporating methods from critical algorithm studies, or "the application of humanistic and social scientific approaches to algorithms . . . the domain of computer scientists" as Seaver puts it.[71]

71. Seaver, "Algorithms as Culture," 1.

4. The Case of Magenta

Magenta is a research project within Google that primarily concerns machine learning methods for generating visual art and music. While being a corporate R&D entity, the project is relatively open with regards to publishing its results and methodology, conforming to the standards of contemporary machine learning research. Magenta also refers to the project's range of software artifacts, mainly the Magenta library that extends TensorFlow, Google's flagship machine learning software library that runs on Python;[72] along with the library, the project produces and offers other artifacts such as datasets, trained models, and interactive demo applications. Hereinafter, I use Magenta to designate the research project that is embedded within Google's institutional structure; the Python library will be referred to as the Magenta library, and other software artifacts will be specified as necessary depending on context.

Magenta: Research Project

Google first publicly announced Magenta at a talk session on May 22, 2016, during Moogfest, a music, art and technology festival;[73] around this time, it set up a public Github repository for the Magenta library and soon afterward launched the Magenta website with a blog post announcing the project. The project was led by researchers at Google Brain, the company's artificial intelligence research team focused on deep learning. The lead researcher, Douglas Eck, outlines the project's goals in the inaugural blog post: to "advance the state of the art in machine intelligence for music and art generation" and to "build a community of artists, coders and machine

72. https://www.tensorflow.org/.

73. Mike Murphy, "Google Is Launching a New Research Project to See If Computers Can Be Truly Creative," Quartz, accessed August 7, 2018, https://qz.com/689887/google-is-launching-a-new-research-project-to-see-if-computers-can-be-truly-creative/.

learning researchers."[74]

Since then, Magenta has made public a series of research results via traditional venues such as academic and industrial conferences as well as open research platforms such as arxiv. It also released related software such as updates to its Python library, trained deep learning models, datasets, and interactive demos. Most of its research relates to applying techniques such as recurrent neural networks (RNNs) and variational autoencoders (VAEs) to the generation and interpolation of melodies, polyphonies, rhythmical patterns, and raw audio with regards to music, and sketches and image style transfer with regards to visual arts.

What is notable about Magenta is its close relation with other entities, both internal and external to Google. In his June 2016 blog post, along with specific research topics, Eck mentions two other Google initiatives: the Artists and Machine Intelligence (AMI) artist residency program and the Google Cultural Institute (now Google Arts & Culture). In addition, projects like Magenta have been increasingly influential in academic research as well, actively publishing in computer science journals and conferences such as IEEE, and project lead Eck taking on the role of chair in NIPS 2018.[75]

In addition to AMI and Arts & Culture, it is also worth noting that Google engages in a wide array of actions that promote cultural collaboration and education, and that often these works are supported by design and marketing resources. This is to say that the company has continuously tried to position itself in a cultural role, an effort that is further extended and perhaps more

74. Eck, "Welcome to Magenta!"

75. Douglas Eck (@douglas_eck), "I Wanted to Give a Full Tweet to the Fact That I Am the NIPS 2018 Party Chair. (Don't Hide It. I Know You're Impressed.) I'll Be Working with Others to Put Together a Night of Researcher-Generated Music for the Official Banquet. Stay Tuned for Details. Https://Nips.Cc/Conferences/2018/Committees …," Twitter, September 4, 2018, https://twitter.com/douglas_eck/status/1037137498399334401.

explicitly manifested in Magenta—where it deems itself a creative toolmaker and community builder—and related initiatives. Therefore, Magenta is more than a mere R&D project, and closer to an element of a company-wide effort to promote Google's AI technology in diverse avenues. Magenta is in a way research-as-branding.

Magenta: Software Library

By examining the affordances of the Magenta library, a mostly-digital project based on TensorFlow, the open source machine learning library maintained by Google, I attempt to draw a comprehensive spectrum of the promise and menace of automated music making, both in terms of potentialities and in its current implementations. Since the Magenta library is an extension of TensorFlow, it would be appropriate to give a brief description of Tensorflow before going into the Magenta library.

TensorFlow

In their white paper, Martín Abadi et al. explain that Google Brain developed TensorFlow as a successor to its previous large-scale deep machine learning system DistBelief, which they had worked on since 2011.[76] While they built these systems for internal use within Google, Google Brain open-sourced TensorFlow in 2015. Made public since version 0.5.0 as a C++ and Python API, the current version of TensorFlow as of this writing is 1.8.0; APIs for other languages including C, Go, and Java also exist, as well as a separately maintained JavaScript library, TensorFlow.js.[77]

Google's investment in TensorFlow must be understood within the context of the rise of

---

76. Martín Abadi et al., "TensorFlow: A System for Large-Scale Machine Learning," in *Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation* (OSDI '16, Savannah, GA: 265-283, 2016).

77. Abadi et al., 275.

deep neural network-based machine learning that was aforementioned. As the early twenty-first century proceeded tech companies embraced the renewed possibilities that deep learning offered. Google's establishment of Google Brain and its later acquisition of DeepMind well demonstrate such interest. Other giant tech companies such as Facebook, Amazon, and Apple were no exception to this trend, establishing in-house machine learning research labs and/or acquiring machine learning start-up companies.

Not only did these companies recruit machine learning talent, but they also started establishing open source frameworks for deep learning, either developing them in-house or adopting existing ones as their chosen tools. Facebook AI Research develops PyTorch; Amazon has adopted MXNet; Microsoft developed CNTK and Apple developed Core ML, although this last one is not open source and rather focused on the "easy integration of machine learning models" into Apple-compatible applications.[78] Most popular deep learning libraries as of this writing are often closely related to certain companies, with Theano being a popular independently developed exception—although Theano is no longer being actively developed since 2017, partly due to "strong industrial players [that] are backing different software stacks in a . . . competition."[79] This AI framework competition, one to establish the leading machine learning software stack, is motivated by the goal of achieving a competitive edge in the AI economy. An important thing to note here is the frequent adoption of open source by the competing companies; open source allows for more participation from users and developers, and usage is a critical factor in market dominance, which encourages companies to adopt open source as a practice.

---

78. "Machine Learning - Apple Developer," June 11, 2018, http://web.archive.org/web/20180611030548/https://developer.apple.com/machine-learning/.

79. Pascal Lamblin, "MILA and the Future of Theano," Google Groups (theano-users), accessed August 7, 2018, https://groups.google.com/forum/#!msg/theano-users/7Poq8BZutbY/rNCIfvAEAwAJ.

TensorFlow/Magenta

The Magenta library allows the user to generate MIDI data and raw audio, to interpolate between distinct patterns—generating 'in-between' spaces; it offers pre-trained models or one can train their own models using custom data. These abilities can lower requisite expertise for producing musical compositions and sound synthesis, assuming part of the agencies closely tied to traditional notions of music creation; affecting social institutions and evaluations of musicianship; and shifting the notion of creativity into new practices and experiments. Nevertheless, in the current form, the project tends to be restrictive in many ways, not the least in the amount of computational power required to train models by oneself.

This leads to the notion of creativity, which is often assumed neutral but cannot really be. Deep learning algorithms trained on Western music datasets to provide music-making affordances are no exception. Datasets and their derived models embody cultural evaluations, quite literally; from data sources to extracted features, the steps of abstraction mandatory in machine learning dictate some judgment of worthiness. Collections of classical sheet music reinforce certain types of features pertinent to Western music as desirable, which are distilled during the training cycles. Certain versions of creativity are tied to certain values and therefore to politics.

The Magenta library extends TensorFlow's functionalities to suit its purpose of training models based on (primarily musical) features and using these models to generate content. While an extensive and technical software analysis of the library is out of this thesis' scope, especially since the software is in active development, some key functionalities implemented in the current version (v0.3.8) include the following. First, the library offers tools for music information retrieval, or the analysis of musical content to extract quantified and actionable information, such as chords, chord progressions, drums, lead sheets, melodies, and piano rolls. Utility tools facilitate the information retrieval process, by parsing notation formats like ABC and MusicXML and allowing the input and output of formats such as MIDI, raw audio, and note_sequence, a format specific to TensorFlow. Key functionalities in terms of music generation are available in the form of different models such as coconet, drums_rnn, improv_rnn, melody_rnn, music_vae, nsynth, performance_rnn, pianoroll_rnn_nade, polyphony_rnn. The name of most models describe the content the model is intended to generate and the machine learning architecture used. Additional models concern music transcription and image style transfer. In terms of music, a key algorithm repeatedly used is LSTM, a type of RNN. As Kyle McDonald notes, this approach is influenced by Eck's own research on music improvisation before he joined Google.[80] For its research, Magenta relies on publicly available data that it often does not release with its models—probably for copyright reasons—or, in some cases, already-established open datasets such as the Lakh dataset.

The Narrative of Magenta

Eck's inaugural blog post on Magenta's website describes the project's goal as asking and answering the questions: "Can we use machine learning to create compelling art and music? If so, how? If not, why not?"[81] The word "compelling" here is an interesting choice. The possibility of using machine learning to create art and music per se does not seem to be a question to Eck. This is not surprising for many reasons: many contemporary artists and musicians are actively applying machine learning in their practices; the history of media art and computer music supports the possibility and validity of machine-generated content; Eck's post itself refers to several academic works that generate art and music using machine learning. Therefore, a bigger emphasis is due in the above quote to "compelling," than to "art and music." But what does "compelling" art and

---

80. Kyle McDonald, "Neural Nets for Generating Music," *Artists and Machine Intelligence* (blog), August 25, 2017, https://medium.com/artists-and-machine-intelligence/neural-nets-for-generating-music-f46dffac21c0.

81. Eck, "Welcome to Magenta!"

music refer to?

Judging the quality of an artistic work is not easy. Eck himself recognizes in the same post that "evaluating the output of generative models is deceivingly difficult," and that it will ultimately rely on the judgment of "artists and musicians" and "viewers and listeners." Nevertheless, he elaborates on specific research goals of Magenta, providing a glimpse of what would characterize "compelling art and music." Making generative models "truly generative"; taking better "advantage of user feedback"; capturing "effects like attention and surprise"; "combining generation, attention and surprise" to output a "long-term narrative arc." In other words, more automation over bigger structural elements seems to be the direction of the project. But this ambition to develop models that perform better in larger formal aspects of the generated content still sounds closer to describing what "Can we . . . create . . . art and music" means, than "compelling."

A more concrete insight into "compelling" can be found in a later discussion in the Magenta Discuss forum on Google Groups. Responding to a forum post that questions the definition of "compelling," Eck writes:

> . . .I didn't have a technical meaning for "compelling." I was just trying to challenge us to make something that holds interest long-term, not just something that is interesting because it's novel. For example, people tend to like random music if it's played on a flashy synth patch. But that kind of novelty doesn't tend to hold our interest for very long.
>
> There are a couple of ways to expand on this. First we can talk about tools that are compelling for an artist to use. I'd love for us to be part of a collaboration with a band or artist, and to have the outcome be that the work created couldn't have happened without these new tools. Second we can talk about compelling over time. Can we create models that improve gradually based on user or artist feedback and continue to hold interest.
>
> Maybe all I really mean is that we should strive to build models that people actually care about, and the people in question are both the artists themselves and those who enjoy the art.[82]

---

82. Douglas Eck, "What Does Google Mean by 'Compelling'?," Google Groups (Magenta Discuss), accessed August 7, 2018, https://groups.google.com/a/tensorflow.org/forum/#!msg/magenta-discuss/9edPR6GQkNE/roldh9EOCwAJ.

He also notes on the same thread, in a lighter tone:

> That said, I'm not against [redacted username]'s definition: "a design experience that blows out the living beejesus [sic] out of the herds of selfie/dj app inflicted youth on this planet" :)83

"Compelling" therefore seems to have to do more with external interest ("models that people actually care about") than with specific research subjects. This is not especially surprising, not the least because the appreciation of art and music, as well as the tools used to make these, is more a social process than an exclusively academic endeavor. Art and music, let alone computer-generated art and music, are social and subjective—which leaves the inquiry into methods of generating "compelling" art and music all the more relative to qualitative human assessments rather than established computational measures.

This qualitative trait is expressed in a dual idea/expectation of creativity similar to that described by Strachan with regards to electronic musicians' work. One side of this creativity aims at replicating the status quo in an automated fashion, i.e. the capability to meet certain standard qualities, and another side pursues a non-conforming and rule-breaking originality, "a completely new kind of music," as Eck testifies.

> With the right data, these networks might be able to hold on to something really artistic and really human. They might even be able to pick up on something having to do with creative intent.
>
> [. . .] When I think of Magenta, I think of us as being more like Rickenbacker or like Les Paul, than like Jimmy Hendrix. . . .The sound of failure. So much modern art is the sound of things going out of control, of a medium pushed to its limits and breaking apart. My goal with Magenta is to see someone create a completely new kind of music, a completely new kind of art, and that the Magenta work that we're doing is part of that.[84]

---

83. Eck, "What Does Google Mean by 'Compelling'?"

84. Future Of StoryTelling, *Douglas Eck – Transforming Technology into Art*, 2017, https://vimeo.com/233020628.

The demos made available by Magenta, such as nSynth, point to something publishable, demonstrable, or shareable. It's a mixture of a stylistic/producing achievement that resembles the market standards as a status quo, and something that is groundbreaking or "completely new". This dual notion/tension of creativity is also common in a musician's practice, as Strachan points out. This 'assist not replace' narrative can also be found in a comparable project led by François Pachet, Flow Machines:

> Flow Machines are cutting-edge algorithms, made to explore new ways to create.
> Flow Machines collaborate with musicians to compose the future.
> Flow Machines are AI music-making.[85]

However, such a perspective is as much, if not more, an avoidance, a shying away, a deliberate negligence of labor as a layer in musical practice, than it is a pledge of good intentions; it reiterates a modernist framing of the individual musician, without mentioning the precarity within a neoliberal economy that characterizes contemporary musicianship.

The admission that a scientific research project is operating under such indeterminate criteria and ultimate reliance on external actors makes for an interesting observation, one that relates back to the relative novelty of the field. One way to remedy lack of legitimacy, nevertheless, seems to be securing a widespread usage. Google's interest in the normalization of AI usage, and especially of TensorFlow, reflects such an attitude.[86] Furthermore, in-house collaborations such as AI Experiments carry the burden of popularizing Google's AI and projects like Magenta. As Katharine Schwab notes in her Fast Company article, "AI Experiments have two primary target audiences–the public, obviously, but also the developer community. . . .Google has a strong

---

85. "Flow Machines: AI Music-Making," Flow Machines, accessed August 7, 2018, http://www.flow-machines.com/.

86. Mark Bergen, "Google Wants to Train Other Companies to Use Its AI Tools," Bloomberg.Com, October 19, 2017, https://www.bloomberg.com/news/articles/2017-10-19/google-wants-to-train-other-companies-to-use-its-ai-tools/.

incentive to convince people that it is a responsible developer of [AI] technologies–that they can trust Google to not be evil as the company continues to cook up more and more algorithms to slip into its software."[87]

Magenta's relevance to the current discussion comes not necessarily from its direct contribution to new methods of cultural production using machine learning technology being particularly novel, although it seems to be holding quite some amount of talent and outputting potentially fields-changing research. What makes Magenta particularly significant at this moment, is that it represents an effort by Google to actively claim its role in cultural production as a toolmaker and paradigm shifter. To put it bluntly, Google wants to *be* culture.

Google's Cultural Efforts

Magenta must be understood with other cultural endeavors from Google in mind, such as the Google Arts and Culture (formerly Cultural Institute). If, as Geraldine Juárez points out, Arts and Culture serves to "bankrolling the technical infrastructure and labour needed to turn culture into data,"[88] Magenta can be said to be developing methods that utilize that data in order to create culture/cultural artifacts. As Arts and Culture in its corporate philanthropy distances itself from historical discourses and narrates a techno-centric ideology, Magenta and similar projects come with a similarly curated version of a complex history of music technology. The datasets they use is Western music, which risks a perpetuation of the current domination of Western music.

Products like Magenta (or TensorFlow, in a slightly larger scope) are both the result of

---

87. Katharine Schwab, "The Dead-Serious Strategy Behind Google's Silly AI Experiments," Fast Company, December 1, 2017, https://www.fastcompany.com/90152774/the-dead-serious-strategy-behind-googles-silly-ai-experiments.

88. Geraldine Juárez, "A Pre-Emptive History of the Google Cultural Institute," Mondothèque, accessed August 7, 2018, http://www.mondotheque.be/wiki/index.php/A_Pre-emptive_History_of_the_Google_Cultural_Institute.

corporate-embedded knowledge production process, and a means to reinforce the solidity of said process. This process is different from, for example, that of traditional institutions like IRCAM or CCRMA;[89] while both engage with novel algorithmic methods and advanced heuristics using existing methods, the Google model makes much more extensive use of design and marketing.

The power to shape knowledge is shifting towards the tech industry. The so-called AI revolution is an attempt by IT companies to gain not only economic and political leverage, but also a cultural one. In the process of developing not only means of cultural production but also modes of knowledge production itself, companies like Google and their products attain the status of creators, in addition to conduits, of culture.

89. For a detailed history of IRCAM and CCRMA, see Georgina Born, *Rationalizing Culture: IRCAM, Boulez, and the Institutionalization of the Musical Avant-Garde* (Berkeley, CA: University of California Press, 1995); Andrew J. Nelson, *The Sound of Innovation: Stanford and the Computer Music Revolution* (Cambridge, MA: The MIT Press, 2015).

Conclusion

The application of recent deep learning-based AI technologies to music making involves a range of efforts largely thrusted by corporate-embedded research labs. The history of music technology informs us of consumer digital technology becoming the central mode of musical practices since the nineties. As machine learning (now often labeled AI) is being touted as one of the latest major technological development, we can also witness an unprecedented industrial interest in musical AI applications. This interest manifests itself through an ecosystem of knowledge-making, or a network of infrastructure and practices; here I focused on Google's Magenta project, mainly because of its institutional specificity as a Google in-house lab research and the narrative it pushes through various venues, which make Magenta pertinent to the illustration of the multiple actors and context that are in play within the aforementioned cultural stack. The project's affordances point to the socio-cultural implications of the industrial push towards a more automated music making, from labor automation and increased access to music production to challenges in the notion of creativity.

Current industrial attempts to apply AI techniques to music production signal an increased reliance by the cultural realm on corporate infrastructure and knowledge production system. To the individual musician, this further extends the effect that digitalization of music technology had: some effectiveness, increased access, and exploitation by capital. In essence, Magenta is symptomatic of the commercialization of cultural innovation characteristic of the contemporary economy. It furthers the extent of neoliberal logic as a force that shapes cultural practice. This tendency, while being a continuation of neoliberal industrial-academic collaboration, takes on a new flavor in a data-driven economy empowered by the digitalization and digitization of many social aspects including cultural practices, leading to the universal convergence into digital media. In this economy, data-centric IT companies attempt to and do assume bigger cultural roles, that of instrument makers being one.

As an instance of commercialized cultural innovation, Google's Magenta project represents a dual potential of centralization-decentralization. It is centralizing in that it imposes a greater reliance on the infrastructural power that IT companies possess, ranging from computational power and machine learning ecosystems to the cultural influence to disseminate its products. It is also decentralizing in that it provides greater access to computational methods that automate and augment traditionally restrictive musical creative practices. This dual potential in turn translates to prospects of intensified exploitation of cultural labor, concentrating more power to large transnational corporates, and of more democratic cultural expressions enabled through increased access.

These possibilities are not mutually exclusive. This is not a question of 'will technology do A or B.' The question is more about what the different social actors in this context will proceed to create socially, culturally and politically. At the moment, the cultural possibilities offered by Magenta seems largely to be serving towards the benefit of Google's effort to establish TensorFlow as its next big product within the competition of AI economy. This is perhaps an inherent limitation imposed by the very manner in which the project is led, as a corporate-embedded research intertwined with academic work and marketing strategies.

Google Magenta is an example of corporate research, and more broadly the data-driven tech industry, claiming its way into (or enlarging its role within) cultural production. Such attempts, which are often hybrids of academic research, corporate R&D, and marketing, ultimately benefit the expansion of neoliberal logic in cultural production. IT companies' investment in music generation using AI is an attempt to claim a bigger role in cultural production and creativity. this brings about interesting questions about aesthetics, the politics of sound that includes it companies, both classical and pop music industries, art/music theory and more.

To paraphrase Shannon Mattern, the music is the most boring thing about AI music.[90] What is at stake in this renewed scale of automation is the reconfiguration of the notion of creativity, as well as the cultural politics surrounding this shift. What we see is a push towards tools and creativity models that rely on corporate infrastructure and large datasets.

This thesis, in its attempt to grasp a fast-changing contemporary phenomenon, leaves out many interesting avenues. Future research could, for instance, adopt ethnographic methodologies and take into account the direct perspectives of the actors involved, from engineers and research scientists to music industrials and musicians. Furthermore, a broader scope of projects need to be investigated; Magenta is a single instance in a larger field, and a wider analysis will help formalize frameworks to critique AI applications in cultural production. A continued scrutiny over how research projects like Magenta influence the diverse relationships in music production is equally needed.

The industrial investments in AI music opens up a new phase in music technology's history, to be accompanied by a re-negotiation among the actors in music production, from aesthetic judgments, changes in the notion and practice of musical creativity, and the musician's status as traditional skills become automated to the roles of IT companies and their artifacts in cultural production. This intersection of music making, AI research, and corporate interest introduce a new trait to cultural production: an increased reliance on cloud-based computing resources and on large datasets hardly attainable by individuals. Datasets and GPU-powered computing instances challenge the role of individual expertise or education with regards to creativity; a new type of infrastructural power exerts itself onto culture.

Taking into account the stakes of automation, the question emerges: does the application of

90. Mattern, "Mapping's Intelligent Agents." "Self-driving cars have sparked a 'billion dollar war over maps,' but the cars are the most boring thing about it."

AI result in a concentration of power, and how should we attempt to democratize it? In order to engage and challenge industrial efforts towards an automated, and more importantly, increasingly centralized cultural production, access to both datasets and computing power needs to be broadened for a 'democratization' of such a shift. On the other hand, this also calls for modes of creativity that do not rely on the cloud infrastructure. In the end, this is an opportunity to envision the next steps for a culture in a society where corporate AI becomes ubiquitous.

Bibliography

Abadi, Martín, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, et al. "TensorFlow: A System for Large-Scale Machine Learning." In *Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation*. Savannah, GA: 265-283, 2016.

Apple Inc. "Machine Learning - Apple Developer," Accessed June 11, 2018. http://web.archive.org/web/20180611030548/https://developer.apple.com/machine-learning/.

Attali, Jacques. *Noise: The Political Economy of Music*. Translated by Brian Massumi. Minneapolis: Univ Of Minnesota Press, 1985.

Bergen, Mark. "Google Wants to Train Other Companies to Use Its AI Tools." Bloomberg.Com. Updated October 19, 2017. https://www.bloomberg.com/news/articles/2017-10-19/google-wants-to-train-other-companies-to-use-its-ai-tools.

Born, Georgina. *Rationalizing Culture: IRCAM, Boulez, and the Institutionalization of the Musical Avant-Garde*. Berkeley, Calif.: University of California Press, 1995.

boyd, danah, and Kate Crawford. "Critical Questions for Big Data." *Information, Communication & Society* 15, no. 5 (June 1, 2012): 662–79. https://doi.org/10.1080/1369118X.2012.678878.

Briot, Jean-Pierre, Gaëtan Hadjeres, and François Pachet. *Deep Learning Techniques for Music Generation - A Survey*, 2017. http://arxiv.org/abs/1709.01620.

Brooks, Rodney. "The Seven Deadly Sins of Predicting the Future of AI." Rodney Brooks' website. Accessed August 7, 2018. http://rodneybrooks.com/the-seven-deadly-sins-of-predicting-the-future-of-ai/.

Chemero, Anthony. "An Outline of a Theory of Affordances." *Ecological Psychology* 15, no. 2 (April 1, 2003): 181–95. https://doi.org/10.1207/S15326969ECO1502_5.

Davis, Jenny L., and James B. Chouinard. "Theorizing Affordances: From Request to Refuse." *Bulletin of Science, Technology & Society* 36, no. 4 (December 1, 2016): 241–48. https://doi.org/10.1177/0270467617714944.

Eck, Douglas. "Welcome to Magenta!" Magenta. Accessed August 7, 2018. https://magenta.tensorflow.org/blog/2016/06/01/welcome-to-magenta/.

———. "What Does Google Mean by 'Compelling'?" Google Groups (Magenta Discuss). Accessed August 7, 2018. https://groups.google.com/a/tensorflow.org/forum/#!msg/magenta-discuss/9edPR6GQkNE/roldh9EOCwAJ.

——— (@douglas_eck). "I Wanted to Give a Full Tweet to the Fact That I Am the NIPS 2018 Party Chair. (Don't Hide It. I Know You're Impressed.) I'll Be Working with Others to Put Together a Night of Researcher-Generated Music for the Official Banquet. Stay Tuned for Details. Https://Nips.Cc/Conferences/2018/Committees …." Twitter, September 4, 2018. https://twitter.com/douglas_eck/status/1037137498399334401.

Evans, Sandra K., Katy E. Pearce, Jessica Vitak, and Jeffrey W. Treem. "Explicating Affordances: A Conceptual Framework for Understanding Affordances in Communication Research." *Journal of Computer-Mediated Communication* 22, no. 1 (January 1, 2017): 35–52. https://doi.org/10.1111/jcc4.12180.

Fisher, R. A. "The Use of Multiple Measurements in Taxonomic Problems." *Annals of Eugenics* 7, no. 2 (September 1, 1936): 179–88. https://doi.org/10.1111/j.1469-1809.1936.tb02137.x.

Future Of StoryTelling. *Douglas Eck – Transforming Technology into Art*, 2017. Video, 04:16. https://vimeo.com/233020628.

Gibson, James Jerome. *The Senses Considered as Perceptual Systems*. Houghton Mifflin, 1966.

Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. *Deep Learning*. Cambridge, MA: MIT Press, 2016.

Greene, Jason. "Do Androids Dream of Electric Guitars? Exploring the Future of Musical A.I." Pitchfork, June 12, 2017. https://pitchfork.com/features/overtones/10091-do-androids-dream-of-electric-guitars-exploring-the-future-of-musical-ai/.

Hiller, Lejaren A., and Leonard M. Isaacson. *Experimental Music: Composition with an Electronic Computer*. New York, NY: McGraw-Hill, 1959.

Hofstadter, Douglas. "The Shallowness of Google Translate." The Atlantic, January 30, 2018. https://www.theatlantic.com/technology/archive/2018/01/the-shallowness-of-google-translate/551570/.

Home Recording. "Automatic Audio Mastering Systems..." Home Recording. Accessed August 7, 2018. http://homerecording.com/bbs/general-discussions/mastering/automatic-audio-mastering-systems-391400/.

Juárez, Geraldine. "A Pre-Emptive History of the Google Cultural Institute." Mondothèque. Accessed August 7, 2018. http://www.mondotheque.be/wiki/index.php/A_Pre-emptive_History_of_the_Google_Cultural_Institute.

Kay, Alan. "Computer Software." *Scientific American*, no. 251 (September 1984): 52.

Kogan, Gene. "Machine Learning for Artists." Accessed August 7, 2018. http://ml4a.github.io/ml4a/.

Lamblin, Pascal. "MILA and the Future of Theano." Google Groups (theano-users). Accessed August 7, 2018. https://groups.google.com/forum/#!msg/theano-users/7Poq8BZutbY/rNCIfvAEAwAJ.

Langkjær-Bain, Robert. “Five Ways Data Is Transforming Music.” *Significance* 15, no. 1 (February 1, 2018): 20–23. https://doi.org/10.1111/j.1740-9713.2018.01106.x.

Latour, Bruno. *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford, United Kingdom: Oxford University Press USA, 2005.

Leonardi, Paul M., Bonnie A. Nardi, and Jannis Kallinikos. *Materiality and Organizing: Social Interaction in a Technological World*. Oxford, United Kingdom: Oxford University Press, 2012.

Lynch, Shana. “Andrew Ng: Why AI Is the New Electricity.” Stanford Graduate School of Business. Accessed August 7, 2018. https://www.gsb.stanford.edu/insights/andrew-ng-why-ai-new-electricity.

Lyotard, Jean-François. *The Postmodern Condition: A Report on Knowledge*. Manchester, United Kingdom: Manchester University Press, 1984.

Manovich, Lev. *Software Takes Command*. New York, NY: Bloomsbury Academic & Professional, 2013.

Mattern, Shannon. “Mapping’s Intelligent Agents.” *Places Journal*, September 26, 2017. https://doi.org/10.22269/170926.

McDonald, Kyle. “Neural Nets for Generating Music.” *Artists and Machine Intelligence* (blog), August 25, 2017. https://medium.com/artists-and-machine-intelligence/neural-nets-for-generating-music-f46dffac21c0.

Metz, Cade. “Facebook Adds A.I. Labs in Seattle and Pittsburgh, Pressuring Local Universities.” *The New York Times*, June 1, 2018, sec. Technology. https://www.nytimes.com/2018/05/04/technology/facebook-artificial-intelligence-researchers.html.

Murphy, Mike. “Google Is Launching a New Research Project to See If Computers Can Be Truly Creative.” Quartz. Accessed August 7, 2018. https://qz.com/689887/google-is-launching-a-new-research-project-to-see-if-computers-can-be-truly-creative/.

Nelson, Andrew J. *The Sound of Innovation: Stanford and the Computer Music Revolution*. Cambridge, MA: The MIT Press, 2015.

Neural Information Processing Systems Conference. “Advances in Neural Information Processing Systems 30.” Electronic Proceedings of the Neural Information Processing Systems Conference. Accessed August 7, 2018. https://papers.nips.cc/book/advances-in-neural-information-processing-systems-30-2017.

Pachet, Francois (@francoispachet). “We Need Serious Musicologists to Analyze AI-Composed Music, in the Spirit of Chess Masters Analyzing Chess Games Played by Alpha-Zero. There Is Remarkable Stuff Going on. #HelloWorldAlbum #SKYGGE.” Twitter, January 18, 2018. https://twitter.com/francoispachet/status/953980629254844416.

Papadopoulos, George, and Geraint Wiggins. "AI Methods for Algorithmic Composition: A Survey, a Critical View and Future Prospects." In *In Aisb Symposium on Musical Creativity*, 110–117, 1999.

Rahman, K. Sabeel. "The New Octopus." Logic Magazine, April 17, 2018. https://logicmag.io/04-the-new-octopus/.

Rosenberg, Scott. "Why AI Is Still Waiting For Its Ethics Transplant." *Wired*, November 1, 2017. https://www.wired.com/story/why-ai-is-still-waiting-for-its-ethics-transplant/.

Schwab, Katharine. "The Dead-Serious Strategy Behind Google's Silly AI Experiments." Fast Company, December 1, 2017. https://www.fastcompany.com/90152774/the-dead-serious-strategy-behind-googles-silly-ai-experiments.

Seaver, Nick. "Algorithms as Culture: Some Tactics for the Ethnography of Algorithmic Systems." *Big Data & Society* 4, no. 2 (December 2017): 205395171773810. https://doi.org/10.1177/2053951717738104.

Sony Computer Science Laboratories. "Flow Machines: AI Music-Making." Flow Machines. Accessed August 8, 2018. http://www.flow-machines.com/.

Strachan, Robert. *Sonic Technologies: Popular Music, Digital Culture and the Creative Process*. New York, NY: Bloomsbury Academic, 2017.

Taylor, Timothy D. *Music and Capitalism: A History of the Present*. Chicago, IL: The University of Chicago Press, 2016.

———. *Strange Sounds: Music, Technology and Culture*. New York, NY: Routledge, 2014.

Théberge, Paul. *Any Sound You Can Imagine: Making Music / Consuming Technology*. Hanover, NH: Wesleyan University Press, 1997.