ISSN 2637-1898

I N S Δ M

JOURNAL OF CONTEMPORARY MUSIC, ART AND TECHNOLOGY

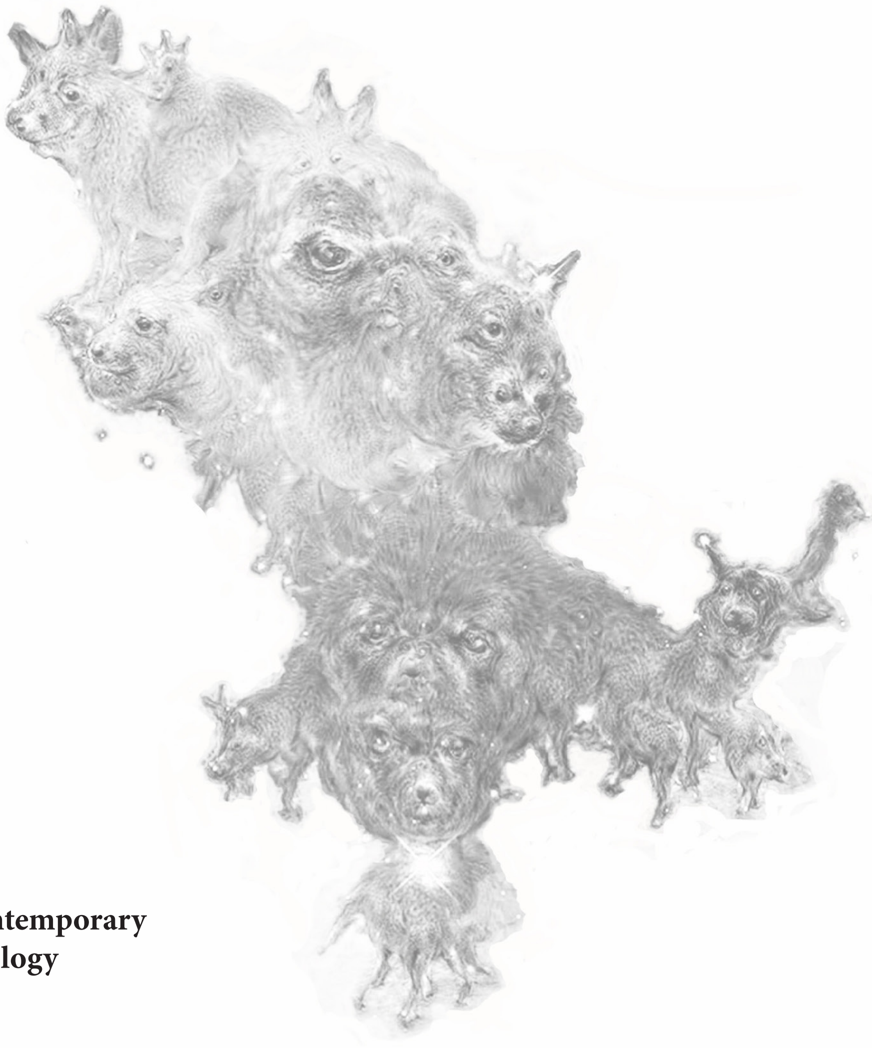
2 I/2019

Journal of INSAM
Institute for Contemporary Artistic Music

INSΔM

I N S Δ M

JOURNAL OF CONTEMPORARY MUSIC, ART AND TECHNOLOGY



INSAM Journal of Contemporary Music, Art and Technology
Issue No. 2
Sarajevo, July 2019

INSAM Journal of Contemporary Music, Art and Technology
Journal of INSAM Institute for Conteporary Artistic Music

ISSN 2637-1898 (online)
COBISS.BH-ID 26698502
UDC 78: 792

on the cover
Devine Lu Linvega / NASA, *The Puppyslug Nebula*, courtesy of NASA and Google DeepDream

design and layout
Milan Šuput, Bojana Radovanović

publisher
INSAM Institute for Contemporary Artistic Music
Podgaj 14, 71000
Sarajevo, Bosnia and Herzegovina

for the publisher
Hanan Hadžajlić

external reviewers
Amila Ramović (Bosnia and Herzegovina), Ivana Medić (Serbia), Jelena Janković Beguš (Serbia), Ksenija Stevanović (Serbia), Aleksa Milanović (Serbia), Maja Stanković (Serbia), Adriana Sabo (Serbia), Marija Maglov (Serbia)

INSAM Journal of Contemporary Music, Art and Technology is an international peer reviewed journal dealing with topical issues in contemporary art music, visual and performing arts, and technology. The journal encourages interdisciplinary and transdisciplinary approaches and innovative views on these subjects, aiming to create a platform for communicating new ideas and theoretical and artistic impulses.

INSAM Journal is published semi-annually (July and December) in open access electronic form and it can be found on the following address: https://insam-institute.com/insam-journal/

The Journal is indexed in:
CEEOL – Central and Eastern European Online Library
RILM – Répertoire International de Littérature Musicale

TABLE OF CONTENTS

(Inter)views

Main Theme:

Artificial Intelligence in Music, Arts and Theory

EDITOR'S FOREWORD

In a famous *Star Trek: The Next Generation* episode titled "The Measure of a Man," one of the most beloved androids in television history, Data, is sent to trial. The reason for this is that he (or, as Starfleet's Commander Maddox called him, *it*) and the Enterprise crew refused to allow Starfleet and Maddox to disassemble him for the purpose of studying his brain and using the technology to serve as a model for creating a new race. The emotionally intense and nerve-wrecking trial gave the viewers an overview of the many possible philosophical and moral questions regarding the treatment of a being that is not human, but is intelligent, self-aware, and perhaps even conscious to a degree.

There are many occasions in this episode where we are encouraged to reflect upon and challenge our own beliefs regarding human nature, morality, and the way we perceive and understand the technology surrounding us. However, one particular line in the episode stands out – especially in relation to the main theme of our new issue. In the very last scene, Data comes to invite Riker, who was the lead prosecutor in the case, to join everyone at the celebration. Riker refuses because, "he almost cost Data his life." But Data – being Data – rationalizes the situation: "Is it not true that had you refused to prosecute, Captain Louvois would have ruled summarily against me?" He continues: "That action injured you and saved me. I will not forget it." Riker replies: "You're a wise man, my friend," and Data's response, "Not yet, sir. But with your help, I am learning," symbolically opens up a whole new chapter in human history.

The subject of machine learning and creativity, as well as its appropriation in arts is the focus of this issue with our Main theme of – *Artificial Intelligence in Music, Arts, and Theory*. In our invitation to collaborators, we discussed our standing preoccupation with the exploration of technology in contemporary theory and artistic practice. The invitation also noted that this time we are encouraged and inspired by Catherine Malabou's new observations regarding brain plasticity and the metamorphosis of (natural and artificial) intelligence. Revising her previous stance that the difference between brain plasticity and computational architecture is not authentic and grounded, Malabou admits in her new book, *Métamorphoses de l'intelligence: Que faire de leur cerveau bleu?* (2017), that plasticity – the potential of neuron architecture to be shaped by environment, habits, and education – can also be a feature of artificial intelligence. "The future of artificial intelligence," she writes, "is biological."

We wanted to provoke a debate about what machines can learn and what we can learn from them, especially regarding contemporary art practices.

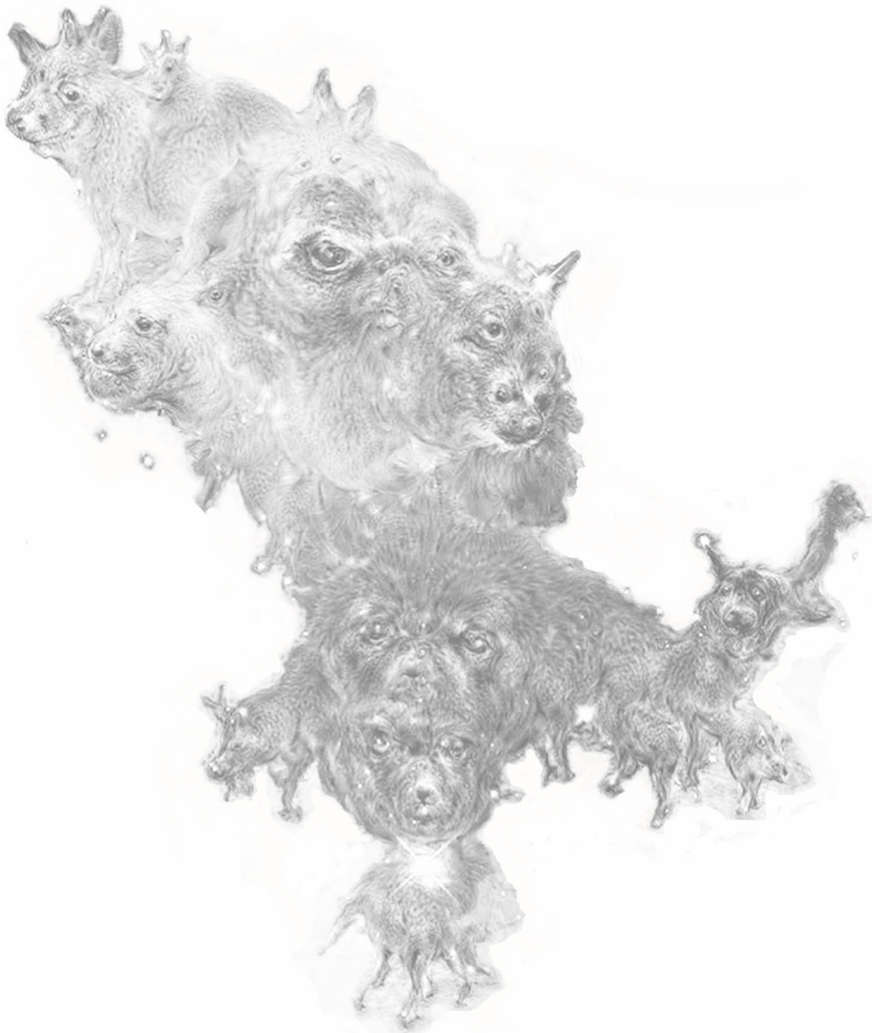
On this note, I am happy to see that our proposition has provoked intriguing and unique responses from various different disciplines including: theory of art, aesthetics of music, musicology, and media studies. The pieces in the *(Inter)view* section deal with machine and computational creativity, as well as the some of the principles of contemporary art. *Reviews* give us an insight into a couple of relevant reading points for this discussion and a retrospective of one engaging festival that also fits this theme.

Finally, I would like to welcome the new members of our Editorial Board, Ernest Ženko (Faculty of the Humanities, Koper, Slovenia) and Rifat Alihodžić (Faculty of Architecture, University of Montenegro). Likewise, a warm welcome goes out to the new members of our International Advisory Board – Miodrag Šuvaković (Serbia), Senad Kazić (B&H), Amra Bosnić (B&H), Ljubiša Jovanović (Serbia), Jelena Novak (Portugal), Daniel Becker (Italy), Olga Majcen Linn (Croatia), Sunčica Ostoić (Croatia), Omer Blentić (B&H), Michael Edward Edgerson (Republic of China), and Bill Smith (US).

On behalf of the Editorial Board and myself, I would also like to thank our peer-reviewers, and our proofreader, Hillary Sigale. Of course, special thanks goes to all the authors who have contributed to our second issue – because, as they say, the second one is the hardest one!

In Belgrade, July 1, 2019,
Bojana Radovanović
Editor-in-Chief

(INTER)VIEWS



Luba Elliott
Curator, artist and researcher
London, United Kingdom

THE NEW WAVE OF AI ART: REFLECTIONS ON ARTISTIC AND MACHINE CREATIVITY

Starting with the appearance of DeepDream's hallucinogenic aesthetic in 2015, the recent wave of art made with AI has been steadily gathering momentum. Initially an outlet for experimentation based on the latest technical developments unveiled by the AI research community, the AI art movement grew, becoming an art world trend with multiple museum exhibitions, gallery shows and media art festivals dedicated to the topic worldwide. Additionally, it had some commercial success with the memorable sale of Obvious' AI-generated artwork for $432,500 at Christie's auction.

Can machines be creative? Even though the general public may be curious to hear answers to this question, many in the AI art community are less interested

in contributing to the discussions about computational creativity, preferring instead to devote their time to exploring the applications of these AI tools across domains and datasets. Nevertheless, some of the recent technological advances such as DeepDream and Generative Adversarial Networks (GANs) do support the idea of machines being creative. Invented by the Google engineer Alex Mordvintsev in 2015, DeepDream is an algorithm that finds and emphasizes features in an image, coming out with extreme colors and puppy, slug and pagoda shapes. In this instance, you could say that the algorithm imagines new structures, colors, and creatures where they are not present in an ordinary image, and is therefore being creative. Meanwhile, GANs, particularly in their earlier years, have been creative in a different way. Tasked with generating images resembling a particular dataset, one neural network generates new images and the other determines which images are generated (fake) and which ones come from the original dataset (real). This dynamic between the two enables each neural network to improve at its individual task and the resulting images are therefore of higher quality. Back when GANs were first invented by Ian Goodfellow in 2014 and in the years shortly after, there were frequent problems with structure (human forms may have limbs at odd angles) and with counting (animals may have multiple eyes or feet). These specifics of earlier GAN models may have been considered 'problems' by the technical community – the algorithms were not perfect at completing their task of creating high-quality images resembling the original dataset, but they could be regarded as an example of algorithms being creative, precisely because they produced images that interpreted the human or animal form in ways that were different from reality.

© Google Inc.

© Google Inc.

The above examples showcase the creativity within some of the typical machine learning algorithms used in artistic practice today. However, what ultimately makes an artwork compelling is not just its aesthetics, but also the storytelling, the intent and the critical perspective that the artist contributes to the piece regardless of the medium involved. In the case of art made with AI, there is also substantial human input and decision-making involved that may not be immediately obvious: first the artist decides to incorporate AI into their work, then they pick the dataset, choose the algorithm, tweak its parameters and curate the resulting output to achieve their desired outcome. These steps are highly influential in the creation of the final artwork and leave scope for the artist's own creativity to shine through, the AI machine merely a tool.

Regardless of how you look at AI and creativity, it is clear that the arts are learning and benefiting from working in a new AI-based world. The popularity of AI across the business world and ongoing interest from the mainstream have attracted a number of artists from diverse backgrounds to work with AI, ranging from AI-researchers with an artistic streak (Mike Tyka, Alex Mordvintsev) to recent art school graduates (Jake Elwes, Anna Ridler) and established contemporary artists (Pierre Huyghe, Hito Steyerl). The engagement of all these artists with vastly different technical skills, artistic styles, conceptual ideas and social concerns have enabled many AI techniques to be tested and pushed to their limits in an artistic context. This usage of AI in art highlights not only the limitations of individual techniques and their implications for society, but also the necessity for artists to acquire basic technical skills, the need to build datasets to generate the desired images as well as the importance of individuality and artistic intent.

This new wave of AI has encouraged the art world to once again face the idea and implications of an endless art-generation machine, though this time, thanks to the ease of generating multiple variations on a theme and the high quality of the output, there have been increased concerns regarding creativity and originality. Ultimately though, what the current stage of AI art demonstrates is that the involvement of the human artist is crucial for giving meaning and context to the AI-generated or processed imagery. As our AI systems become more advanced and move towards general artificial intelligence, there may be a higher level of creativity and intent delivered by the algorithm alone. Until then, the artist is still king in AI art.

Article received: June 17, 2019
Article accepted: June 20, 2019

Andreja Andrić
Composer and programmer
Aarhus, Denmark

COMPUTATIONAL CREATIVITY: A PERSONAL APPROACH

We live in a world driven by software. Software powers and transforms the way we learn, the way we interact with each other, and the way we produce and exchange goods. Similarly, software transforms the way we create experiences and engage in art. We can talk about where this transformation has taken place in a marked degree through computational art and computational creativity.

There are two ways in which we usually understand computational creativity. One way refers to attempts to construct programs that mimic human creativity. This category contains projects such as, "I fed my system with 1000 hours of jazz and

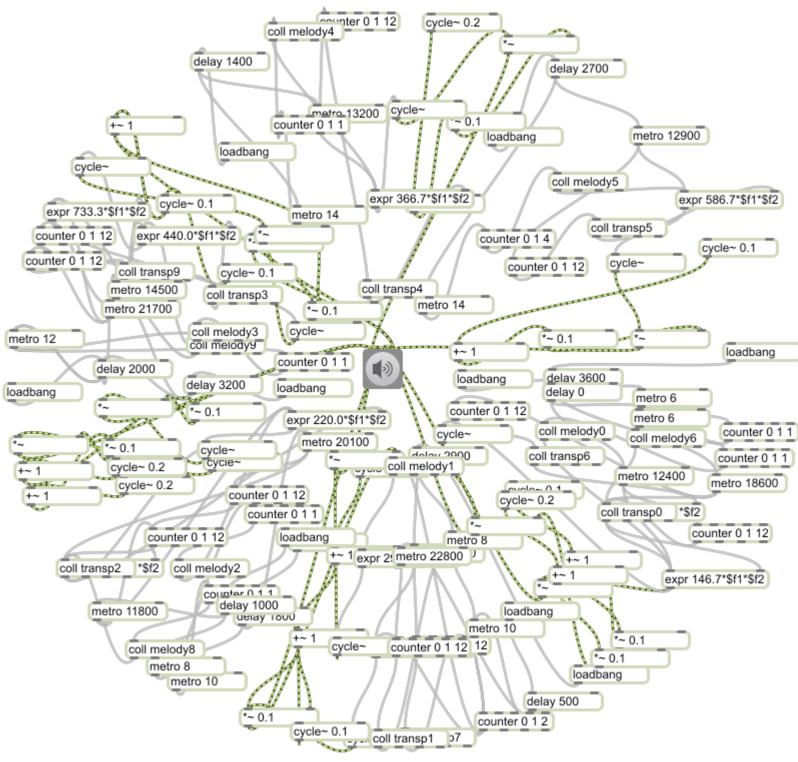
Andreja Andrić, *Portrait with Hands* © Barbara Katz now it plays jazz." Such systems have been developed since the late eighties, and have attained a degree of fame (or infamy) over the last few years with songs like *Daddy's Car* from Sony Research. The other way that we can understand computational creativity is as a different kind of creativity, that is, at least to a degree, unlike human creativity. More well-known examples of this are works of randomly generated literature, such as Nick Montfort's *World Clock*. I will focus on this second way as it is of more interest to me artistically and I find it to be an area of true artistic innovation and experimentation. I believe it offers us the possibility to extend what humans are creatively able to do, or rather, be innovative in unexpected ways, and offer new experiences that are not attainable by other means.

I was always attracted to the unexpected dream-like utterances of computational poetry. Being a musician rather than a poet, I enjoyed what similar procedures were capable of doing within the realm of sound even more. Randomized, game-like procedures for music creation have always been known to musicians, from at least the time of Mozart and his Musical Game of Dice. I was also always interested in creating musical works where the creation process executed by the computer exhibits game-like, playful attributes, while at the same time resulting in dream-like flows of computer-generated sound that resembles natural events more than human expression.

For instance, counting features prominently in a series of my works. Counting is by itself a generative procedure, as each number yields the next by the addition of one. Also, the set of natural numbers projects itself on the set of digits (or any other objects) in interesting and, at times, unexpected ways. Counting and repetition has a meditative character. It is part of many children's games, like hide-and-seek, where the seeker has to count up to a randomly decided number during which time the other players hide. The ability to find a hiding place depends on how long the counting takes, and the counting by itself also gradually increases the tension caused by the anticipation of the search that is about to take place. My piece *Effervescence* (2018), a virtuoso study for solo guitar, performed by Danish guitarist Jakob Bangsø during some of his guitar recitals, counts 4 note figures from 9 notes of an e-minor chord that spans three octaves. In this work the counting is projected onto ever different fast-paced tone patterns creating an illusion of a multiplicity of voices and melodic movement.

Besides counting, random decisions play an important part in many of my works. Many natural phenomena happen or appear to happen according to different kinds of random distribution (think of raindrops hitting the ground, autumn leaves covering the pavement, or stars seen from Earth). Many games involve throwing dice or picking cards from a deck. Divination rituals always involve a random procedure to divine the opinion of the gods. Pseudo-random functions can be easily implemented on a computer and one of the simplest implementations, called linear congruential random number generator, is also based on a procedure that somewhat resembles counting: the current number is (instead of increased by one) first multiplied by a large constant, then increased by another large constant, and finally the modulus of it is taken by a third large constant. In *Pocket Electronic Symphony #1* (2018), performed at the World Music Days 2019 in Tallinn, Estonia, the performer freely combines 15 algorithms that heavily use random functions to direct the flow of a large number of parallel sound lines. *Meditations for keyboards and string instruments* (2019), soon to be performed at the Kyoto Arts Center by the Japanese ensemble Rosetta, is a set of 200 computer-generated text-scores to be interpreted by an ensemble of human performers. The text scores are constructed out of sentences whose parts are selected at random out of a set of predetermined possibilities.

In another series of works, the computer program sets in motion ten or more circles of fifths which progress at different but constant speeds, decided randomly at the start of the performance. Here there are also counting steps that go through discrete positions on a circle. Circles also feature prominently in many folk dances, rituals, and children's games, and also reflect natural phenomena like sun rising and setting or the change of the seasons. *Spin* (2015) is a series of electronic music works which feature ten circles of fifths of sine waves each slowly moves at its own speed to reveal ever-changing tonal interrelations. The works are implemented in the graphical programming language Max/MSP, with the elements disposed in concentric circular patterns reminiscent of Ars Subtilior practice from high middle ages.



Andreja Andrić, *Spin #5* (2015)

In this text I have attempted to survey the use of computer programming in my own compositional work and show that all of it can be traced down to a few simple generative processes. All of these processes are very suitable for implementation on a computer and stand in the root of all classical programming languages. In trying to trace my own fascination with these processes, I have tried to show that these generative procedures also play a prominent part in children's games, folk dances, magic rituals, divination, and meditation practices from which they may derive their wider psychological significance.

Article received: June 10, 2019
Article accepted: June 20, 2019

Aida Adžović
Music Academy, University of Sarajevo
Sarajevo, Bosnia and Herzegovina

"THE BASIC PRINCIPLE OF CONTEMPORARY ART IS THE SYNTHESIS OF SCIENCE, ART AND TECHNOLOGY": Interview with Hanan Hadžajlić

Hanan Hadžajlić (1991) is a composer, flutist and transdisciplinary researcher. She completed DMA in flute performance at the Faculty of Music Arts in Belgrade (mentor: prof. Ljubiša Jovanović, co-mentor: Dr. Vesna Mikić). Currently she is a PhD candidate in art theory – transdisciplinary studies of contemporary art and media at the Faculty of Media and Communication in Belgrade (mentor: Dr. Andrija Filipović, co-mentor: Dr. Miodrag Šuvaković). She completed MA and BA in composition (mentor: prof. Ališer Sijarić, co-mentor: Dr. Amila Ramović) and MA and BA in flute performance (mentor: prof. Sakib Lačević, co-mentor: Dr. Ivan Čavlović) at the Music Academy of the University of Sarajevo.

Since 2018 she has been employed as a Teaching Assistant at the Department of Composition at the Music Academy of the University of Sarajevo. She is a co-founder and director of the Institute for Contemporary Artistic Music (INSAM Sarajevo, 2015). Since 2012 she is a member of the ensemble SONEMUS.

She was a scholarship holder of the Lucerne Festival Academy – Composer Seminar 2017 and Science Underground Academy 2016. Her composition *Freezing Moon* is included in the book *The 21st Century Voice: Contemporary and Traditional Extra-normal Voice* by Michael Edward Edgerton. Her music was performed in many European countries, Israel and USA.

How would you define the term 'Artificial Musical Intelligence'?

Artificial musical intelligence (hereinafter referred to as AMI) is the ability to establish a compositional process in real time and adapt the behavior/process redirection to changes in the environment (such as external information) by a modular system. In the context of my DMA research (flute performance, theme: *Flute as a 'metainterface' of modular systems in contemporary electroacoustic music*), this system refers to the *'TransFlute' Modular System* (*TFMS*).

The consistency of the process (which implies multi-processuality) is the most abstract concept of a musical composition, and in fact, it is the ultimate concept of achievement in music for composers. It refers to consistency in terms of permanence, firmness, and self-determination. Each analysis of process consistency, whether or not it is displayed numerically or verbally, is a descriptive analysis that starts from a certain thesis - the idea of relationships between individual contexts (processes) within the global context (process), where the reference is the compositional material and the possibilities of its development. The idea of mathematical concepts as an experience of nature, taken from the Antiquity and the Middle Ages, is presented in the composition by the construction of abstract models that, through mediation of proportions and analogies, allow the creation of links between completely different contexts.

The characteristic that allows a particular system to adapt its behavior to change of environment, is *the placement*, a concept that comes from cognitive science, and the reflexive reaction of the system is determined by means of its interpretation. The modularity of the concept of *placement* defines and enables the adaptation of a certain system to change of environment. Its reflex reaction, which relates to the localization of a certain occurrence (initiator of action) and the establishment of the direction of motion, determines the means of interpretation, that is, the initiator. Sound modulations are abstract self-modulating dynamic systems with the ability to adapt to the change of environment (which depends on the actions of individual modules within the general process) that arise from the physical modules connection system.

The theory of AMI first requires a summation of *basic points* of the domain of artificial and musical intelligence. The concept of artificial intelligence is based on Robert Sternberg's intelligence models, i.e. practical, creative, and analytical intelligence (component, experiential, and contextual). Intelligence is an abstract law, an entity that has the ability to solve problems in specific circumstances and use specialized intelligence in a unique way, as well as the ability to learn from its environment. The adaptive control theory, that deals with the design of machines capable of behaving in unpredictable conditions, and the control theory, which deals with the development of a particular concept of behavior of complex machines, actually determine the contexts in terms of simplified as well as unpredictable conditions and environments and accordingly define and build specific platforms enabling the potential of intelligent machine behavior. The musical intelligence base, according to Howard Gardner, refers to musical competencies, while the production or composition of music represents the highest level of musical intelligence.

The scientific-artistic potential of the field of development of general artificial intelligence synthesizes three aspects: *techne*, *poiesis*, and *mimesis*, with the intention of simulating a modular system based on a human brain model that is not deterministic; and the ontological and epistemological link represents an interface in which the two principles meet but have the possibility of separation. The concept of AMI is based on an exemplary metaphor "model and meta model of the interface," modeled on the programming of general artificial intelligence, and at the same time the point of interaction between man and machine, where man is a code maker and the machine a creator of potential artwork.

How is Artificial Musical Intelligence represented in contemporary music?

Examples of AMI are reflected in projects such as *Kurzweil Music Systems* by Ray Kurzweil, *EMI* by David Cope, or *Impromptu* by Andrew Sorensen, based on software technologies. However, although the aforementioned programs contain the basics of musical compositions achieved by specific algorithms and computer models of perception of music, they are based on the use of digital acoustic instruments and the contextualization of the musical composition as a process of algorithms that use cultural models of classical music for reference.

Each of the projects contains the concept of music competence, i.e. the ability of the program to recognize certain models and engage in further modification through the entered algorithms, resulting in a certain compositional process. In this way, in relation to the theory of interface from computer science, the programming language is a model and algorithms are the specific meta model, which would mean that the contextualization of information – both the type of sound and the initial compositional model - represents metainterface. It contains the model syntax, but it can also be viewed separately out of context, that is, as a process of algorithms. However, the AMI of these projects, based on software technology and information strictly referring to the culture of classical music, is only one of the possibilities of the application of the concept of AMI.

The field of contemporary artistic music composition strives to overcome cultural references, avoiding the use of cultural codes and finding their own models of material creation and control.

What is the 'TransFlute' Modular System, what does it consist of, how does it work, and in what way do the modular system and flute relate?

The *'TransFlute' Modular System* (hereinafter *TFMS*) is a system of physically connected specific analog modules (sound processors) whose metainterface is represented through a modular system – with modular programming of parameters and sound modulations as a compositional system.

The modular architecture of the *TFMS* means that the placement or position of modules, both physically and parametrically, plays a significant role in the predetermination of the process. Individual modules can be integrated and even isolated from the system, however, as part of the modular system, they represent the interconnection seen in the modules of one organism/system. Interdependence is reflected in the internal multi-processuality of the modular system, which implies a general process within which simultaneous multiple processes or modulation cycles occur. The cycle refers to the complementary activity of the modules, that is, the construction or reduction of the signal dimensionality, or the amplitude, frequency, phase modulation and demodulation, and self-isolation of a particular signal, single module cycle, or complete sound, i.e. all present cycles. Digital modulation can be included if there is a specific digital module in the system.

The metainterface of modular systems refers to the medium of certain architectural compositions that control the processes of all involved so – agents of the entire system. Composition architecture, that is, a composition, is an information system that is primarily partly deterministic, and its realization depends on the perception of its media by the modular system, the modulation, or the artificial music intelligence. Reflective agent metainterface, in the context of my research, refers to the flute as a medium and as an information system for the perspective of composition. The flute is the starting point of the modulation process as well as the environment of the modular system.

What about the relation between the instrument as 'external information' and the modular system? Are they equals in the composing process?

The flute as a metainterface of modular systems refers to the flute as an external initiator of the modular system process, a separate entity whose characteristics are also altered (regardless of whether the flute signal has previously passed through the digital effect). In this way, the flute and modular system relationship is bidirectional, which means that the flute influences the modulation of the analog signal – or the modular system, and the same relationship also affects the changes in the sound characteristics of the flute.

How could Artificial Musical Intelligence of the 'TransFlute' Modular System be applied to contemporary composition?

Contemporary artistic composition implies the establishment of autoreferential systems based on autoreferential material. The autoreferential system represents a separate entity, a language that functions exclusively in its own context and represents a certain law. However, compositional logic starts from the mimesis of natural laws, which means that it essentially contains an analytic aspect carried out in contextualization, though with certain references, in a musical language. Thus, according to Gardner's theory of composition as the highest instance of musical intelligence, it contains the potential of mimesis of the intelligence phenomenon, defining its own material and means of abstraction of cognitive function.

Thus, the mimesis of intelligence by compositional logic, by specific methods and techniques, becomes a composition. It is necessary to state that it is only functional in the media of its interpretation, through the analysis and musical interpretation of the symbolic system, which is the form of certain materialization of the composition. The distinction between intentional mimesis and materialization of the composition as a separate entity is its functionality in the specific circumstances defined by the medium itself. One of the axioms of this theory of artificial music intelligence is that cognitive function develops through the establishment of new synaptic connections, which means that it passes through a constant modulation process, in accordance to reactions to the specific circumstances it encounters.

TFMS has the ability to react to external information and can become a user interface for composers. With regard to the musical intelligence (process setup and ability of process redirection) of this modular system, but also to the impossibility of absolute determination of modulation by the composer, contemporary electroacoustic music for flute with the analog sound processors that make this modular system is only a hypothesis, a concept, which primary purpose is exploring the potential of intelligent music machines. Therefore, contemporary electroacoustic music for flute as the metainterface of modular systems refers primarily to transdisciplinary research of artificial music intelligence based on the synthesis of artistic and scientific perspectives.

Concerning the composition based on 'TransFlute' Modular System, is the interpreter also the composer too? Does this kind of interpretation provide the possibility of improvisation?

In the case of the integration of *TFMS* into a composition, it is not solely based on the compositional processes of the modular system, but on the interaction between the flute and the *TFMS*. Currently, Dino Rešidbegović and I are the only composers who have integrated the *TransFlute Modular System* into their compositions. The examples of this integration are compositions for flute and *TFMS*: *Concerto* for flute, processors and electronic tape (2018) and *Wreesky* (2016) by Rešidbegović and my own compositions including: *A-B-R-A-C-A-D-A-B-R-A: B-A-C-H* (2018), *A Thousand Plateause: Hommage a Deleuze & Guattari* (2018), *I Am Composition, My Name Is Politika* (2016), *Artificial Intelligence* (2016).

Rešidbegović uses his own notation systems, ARGN (Approximate Reductionist Graphic Notation) and RMC (Reductional Music Complexity), when writing the *TFMS* activity (parameters) as well as the flute part, through instructions for forming compositional processes in interaction with *TFMS*. Since Rešidbegović's compositional paradigm is based on the strict determination of a minimum of one parameter (most commonly it is a rhythmic construction or instrumental technique), he always leaves the parameter of the tone pitch system as the space for decision making by the performer.

Rešidbegović's compositional paradigm falls into the category of "determined/ structured improvisation in reference to particular composition" where the performer has the role of interpreter/composer/improviser. The interdisciplinarity of the mentioned field implies that the improviser is also a professional interpreter and composer. Compositions written for flute and *TFMS* integrate the category of "transmedia and multi-processuality/interactive process of transformation of the fundamental material", where the improviser establishes compositional processes in interaction with *TFMS* processes.

My compositions for flute and *TFMS* represent different categories of free improvisation in the compositions and their combinations: "transmedia and multi-processuality/interactive process of transformation of the fundamental material" strictly determined/structured improvisation in reference to a particular composition; partly determined/structured improvisation in reference to a particular composition. Therefore, as I have already stated, it is an interdisciplinary and interactive approach to the construction realization of compositional processes.

What is the importance and influence of science and technology for creating principles of contemporary art?

I believe that the basic principle of contemporary art is precisely the synthesis of science, art, and technology. So, I am talking about interdisciplinary, transdisciplinary/meta-disciplinary and even post-disciplinary approach. The basic principle is not contemporary, it is timeless and exists/functions from the moment when the first man entered the tertiary aspect of his being and established a creative act, that is, when he came to the creative act of discovery of natural laws.

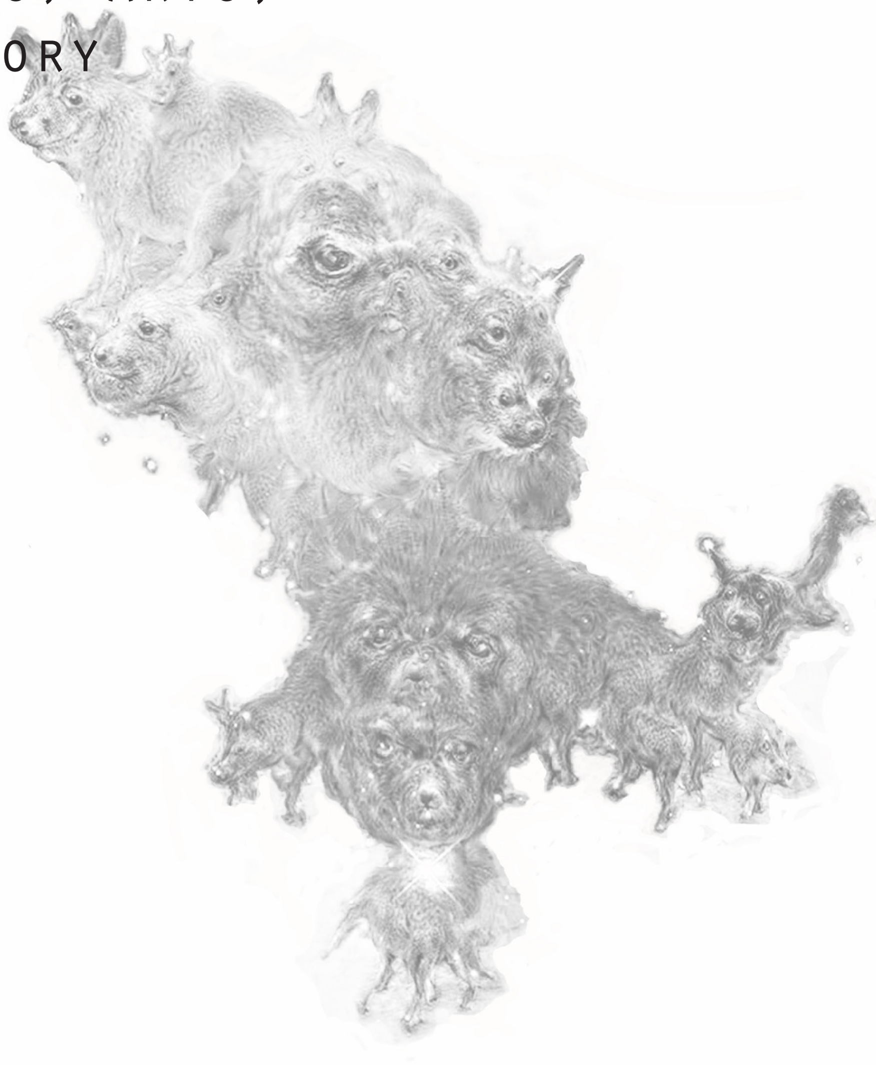
Contemporary art can be all that is today, that which is contemporary, present, and dictated by the art market; there must also be a space for different aspects of the research process and the process of becoming-of-something. However, my point is that sooner or later there will have to be a differentiation of institutionalized artistic/scientific practices that represent superficiality, brand, mainstream, instant knowledge, instant spirituality, *circus* and *kitschy* combinations of different perspectives of social, natural science, and artistic approaches, what I call *fun for spiritually poor* – from the essential search for transcultural models in nature, science, art, and technology.

The coexistence of these two *worlds* represents the image of conflict of contemporary political tendencies. The truth cannot be partial, it must cover all its aspects; but it requires a hierarchy. Therefore, when contemporary art as a thinking/action/practice/culture of living encompasses all aspects of the organization of the contemporary world, starting from economic and political, when an artist becomes someone who discovers but also produces, starts to correct his/her own mistakes, finds and invents, destroys and builds, establishes a process and provides resources for its implementation, I believe that eventually a world-wide civilization platform will be found/created. Thus, art and science offer solutions to the establishment of justice and the conditions for individual, and subsequently, the collective development and advancement of civilization. Although often negatively represented in the media, I consider the work of investor, philosopher, political activist, and philanthropist George Soros as an example of trying theoretical and practical implementation of art in the contemporary world (one example is his platform *The Open Society Foundations*).

Finally, I will go back to the initial question and to my first sentence: the basic principle of modern art is the synthesis of science, art, and technology. Art is the basic principle of living and the surviving of humanity. However, if values are expressed solely through numbers as a symbolic system of accumulated capital representation, then this world will not overcome the immoral (which might be our biggest fear today), and could even become amoral. We need to prevent this with our active work.

Article received: June 15, 2019
Article accepted: June 20, 2019

MAIN THEME:
Artificial Intelligence in Music, Arts, and Theory



UDC 78:111.852]:004.8

Miško Šuvaković*
Faculty of Media and Communications
Singidunum University
Belgrade, Serbia

ARTIFICIAL INTELLIGENCE AND AESTHETICS OF MUSIC: INTELLIGENT ANARCHY. diagramic notebook: on AI-IA *relatons*

image as noise 1–3 (audiovisual diagram)

* Author's contact information: miodragsuvakovic@gmail.com

Marx rejects the idea that *work* can ever become play.

(Marcuse, 1969, 21)

violent chanGes
Of
The
if one Tries

mind base Human

a rEcirculary
arMy

(Cage 1988)

Where thinking must stop, blueprints, schematics, and industrial standards begin. They alter (strictly following Heidegger) the relationship of Being to Man, who has no choice but to become the site of their eternal recurrence. A, Z, E, R, T . . .

(Kittler 1999, 230)

So-called Man is split up into physiology and information technology.

(Kittler 1999, 16)

Today we are calling for the rights of a new functionalism: no longer what it means, but how it works, how it functions. As if desire had nothing to say, but rather was the assemblage of tiny machines, desiringmachines, always in a particular relation with the big social machines and the
technological machines. Your particular desiring-machines: what are they? In a difficult and beautiful text, Marx called for the necessity to think human sexuality not only as a relation between the human sexes, masculine and feminine, but as a relation "between human sex and non-human sex." He was clearly not thinking of animals, but of what is non-human in human sexuality: the machines of desire. Perhaps psychoanalysis had gotten no further than an anthropomorphic idea of sexuality, even in its conception of fantasy and dreams.

(Deleuze 2004, 243)

The discourses surrounding the various forms of electronic music oscillate between two extremes: futurology and ethnography.

(During in Assche 2002, 39)

artificial general intelligence (Shanahan, 2019, 93-95)	human-level AI	possible	
artificial intelligence and intentionality	intelligence explosion AI explosion	possible project	
artificial intelligence and autonomy	superhuman-level AI	anticipation	the distribution of the sensible (Rancière 2004, 12-19)
artificial intelligence and selfhood	non-human level AI	anticipation	
artificial intelligence and consciousness	post-human level AI	SF	
artificial intelligence and new mode of existence	**AI as form of life**	**SF**	

a shift	identifying
the integration	*hegemony*
the expansion	emancipation
the live	to be

Another *Other* Beginning with AI				
present condition of the world	transformation of the present condition through world			
mechanical procedure	electronic procedure	digital procedure		
labour	act	performing	network EVENT as *epistemological break*	
political engagement with cybernetics	from ”politics of calculation” to politics of circulation	flow of artificial bodies through time and space	techno-historical event with artificial bodies	techno-de-historical event

SOUND : NOISE BASED / PRODUCED BY AI ENTITIES

building intelligence into computers and robots
human-level artificial general intelligence
computational theories of mind
neuroscience-oriented program of AI
concentrated on robotics, motor control, and sensory feedback
embodiment
the assemblagy of tiny - *esoteric* - machines
experience
computational
not closed
visual or sound layers
layered abstraction (Andersen 2018, 132)

NOISE			
continuous noise	political line	permanentn impact	
intermittent noise	political break	unstable	
impulsive noise	political impact	intensity	**subversion of the SUBJECT OBJECT relation through noise production between humans and nonhumans**
low frequency noise	under politics	mimicry	
silence as set of noises	political topography	anarchy	

SOUND impacts

forceful stimulus
core reflection
temporal instant
found something within a single sound
acoustic territories
sound between environments
temporal dynamic
the voice separated from the body/object
sound in body/object
object/body oriented mind
interiors
exteriors
as sonic impact
sound through vibrations
vibrations as non-sound but sonic impact
sonic patterns : space patterns or/and time patterns
sonic politics - opposite - to acoustic politics
politics as *hardcore* music : set of sounds between humans and nonhumans
music if it is music or antimusic is mission
mission with border lines through sounds and affective impacts
sonic wall with many gates and shadow creature
computational psychoanalytic theory
but computational is beyond psychoanalysis
most generally, a computer is any system that can output the value of a function from values of variables to which the machine is set (Glymour 2019, 240)

how their normal functioning produces the varieties of normal cognition and how damage to them produces cognitive anomalies (Glymour, 2019, 240)
social and cultural conditions of sonic impacts through computing
are you 'black box' od sounds
sonic process
image = sonic impact
for humans story on AI is a 'ghost' story
re-appropriation of medium as cyberthought
media in transition
intelligence in transition - vector
AI in transition - loop
vector and loop as affect
plunderphonics : the practice of appropriating or 'plundering' sound from a variety of sources by means of a sampler or a sequencer - or - the practice of making new music out of previously existing recordings ... (concept by John Oswald) (Reynolds in Assche 2002, 87)
the ability to perceive and comprehend interdimensional spatiality/temporality or speculative relations between object and subject: me and it
between 'black noise' and 'white spot'

the shape of SOUNDS to come
thinking outside-in NOISES
TRUTH as subversion of the postruth ideology
from AFFECT to CAUSALITY **immanence of music as truth condition**

IDEALIZATION ... MODEL ... AI potentiality ... IA ressistance

AESTHETICS OF MUSIC (Scruton, 1997)	**AESTHETICS OF EXPERIMENTAL MUSIC** (Nyman, 1999)	**AESTHETICS OF DIGITAL MUSIC** (Van Assche, 2002; Miller, 2008; Bey, 2009)	**AESTHETICS OF SOUND ART** (Voegelin, 2010; Daniels 2010, 20“11)	**AESTHETICS OF POST-DIGITAL MUSIC** (Mazierska, 2018)	**AESTHETICS OF AI MUSIC** (artificial intelligence and new mode of existence)
human body, mehanic, electronic instruments	human body mehanic, electronic instruments	digital apparatus with inputs, interfaces, commputers, outputs	space-time relations, architectural frame, sound and visual, experience, human body through sound complexity	networks, social networks	artificial body, artifical apparatus, artifical apparatus networking
voice, acoustic music, electro-acoustic music	new media music, expanded media music, open medium music	programabile media music, metamedia music	open media postmedium postmedia new media	commercial net-media music, postmedium, postmedia	music without humans : any vehicles, tools, instruments
stage, radio, TV, video, sound carriers etc.	street, garage, stage, radio, TV, video, sound carriers etc.	tape, disc, usb, online sites	gallery city architectural space natuiral space	online networks	labaratory, industry, any mode of cybernetic performing, mediation or presentation
musical culture	**alternative culture**	**software culture**	**sound culture** (Sterne 2013)	**software culture**	**AI culture**

world of music, musical institutions, commercial institutions	artworld, institutions, anti-institutins, alternative zones	world of music, musical institutions, mass and global market	art institutions, cultural institutions, architectuiral site specific	digital, online opensource or commercial networks	in radical sense: posthuman corporative institutions from labaratory to market
sound	sound, optical and haptic impact	sound, optical, haptic or/and kinematic impact	hybridity of sounds, sound and anything other mediums/medias	complex and hybrid information	the *new* artificial sensibility and sensuality by sound oriented technological impact of materialized intelligence
tone	tone, noise, silence	tone, noise, silence - any modification of the tone, noise, silence	sound	information exchange and interface presentation	potentiality of nonhuman or artificial labor, production and action
imagination	concept, propositional attitude	software	concept, affect, statement	metainterface with different platforms	generative imagination
ontology	de-ontologisation	quasi-ontology	re-ontologisation od audio and visual	processual useful ontology	onto-centrism
representation	event	impact	object, situation, event installation environment	impact-contraimpact	representing
expression (emotion)	neutrality (affect)	affect	installation	interactive affect	expression (generative emotion)
language (com-munication)	metalanguage (critical discourse)	user constructions	environment	audiovisual and symbolic communications	computational non-human thought
understanding	participation	participation	participation	poly-participation	trans-human and/or non-human social relations
tonality	dissemination	simulation	Self-Conscious collaborative	non-important	imporant, but part of the sound complexity
form	antiform, formless	any-thing	any-thing	clouds	essential in all potentials

content	set, structure, map	relation of the sets, structures, maps	relations	relations of the platforms and clouds	propositional attitudes
value	anti-value	value, anti-value	value, anti-value	value, non-value, surplus-value, capital, bitcoin, emoji...	capital affect
analysis	politization	technocratic policy	political action, aesthetic action	communication tactícts	strategy and tactics
performance	intervention	moving, flux, exchange	perceptual performance	plug-in, plug-out, networking	form of non-human life
culture	frame	condition	culture	hybridization of the apparatus	image of the new non-human world
biopolitics	anarchy (Cage 1988)	simulacrums	nomadic labor/thought	connecting, control, logistics - but - free floating	construction of technology impacts as *being*
musical medium and media	art and non art medium/media	digital equipment	any equipment or artistic tactics: appropriation, construction, performing	the social networks	state, culture, corporation, grey zones
human intention	human intention	human intention	human intention	human intention and algorithmic circumstances	post-human intention, non-human intention, AI intention

List of References

Andersen, Christian Ulrik, and Søren Bro Pold. 2018. "The Cloud Interface: Experience of Metainterface World." In *The Metainterface. The Art Platforms, Cities, and Clouds*, 121–55. Cambridge MA.: The MIT Press.

Bey, Thomas, and William Bailey. 2009. *Micro-Bionic / Radical Electronic Music and Sound Art in the 21st Century*. Bangkok: Creation Books.

Cage, John. 1988. *Anarchy*. Middletown: Wesleyan University Press.

Daniels, Dieter, and Sandra Naumann, eds. 2010. *Set this Sound: Audiovisuology Compedium. An Interdisciplinary Survey of Audiovisual Culture*. Köln: Walther König.

Daniels, Dieter, and Sandra Naumann, eds. 2011. *Set this Sound: Audiovisuology 2: Histories and Theories of Audiovisual Media and Art*. Köln: Walther König.

Deleuze, Gilles. 2004. "Your Special Desiring-Machines: What Are They." In *Desert Islands and Other Texts 1953-1974*, 242–43. Los Angeles CA.: Semiotext(e) Foreign Agents Series.

Kittler, Friedrich A. 1999. *Gramophone, Film, Typewriter*. Stanford CA: Stanford University Press.

Marcuse, Herbert. 1969. "A Biological Foundation for Socialism?." In *An Essay on Liberation*, 7–22. Boston: Beacon Press.

Mazierska, Ewa, Gillon, Les, Rigg, Tony, eds. 2018. *Popular Music in the Post-Digital Age: Politics, Economy, Culture and Technology*. London: Bloomsbury Academic.

Nyman, Michael. 1999. *Experimental Music. Cage and Beyond*. Cambridge UK.: Cambridge University Press.

Miller, Paul D., ed. 2008. *Sound Unbound. Sampling Digital Music and Culture*. Cambridge MA.: The MIT Press.

Rancière, Jacques (2004), "The Distribution of the Sensible: Politics and Aesthetics." In *The Politics of Aesthetics*, 12–19. London: Continuum.

Reck Miranda, Eduardo, ed. 2000. *Readings in Music and Arrtificial Intelligence*. London: Routledge.

Scruton, Roger . 1997. *The Aesthetic of Music*. Oxford: Clarnedon Press.

Shanahan, Murray. 2019. "Artificial Intelligence." In *The Routledge Handbook of the Computational Mind*, edited by Mark Sprevak and Mateo Colombo, 91–100. London: Routledge.

Glymour, Clark, and Ruben Sanchez-Romero. 2019. "Helmholtz's Vision. Undetermination, behavior and the brain." In *The Routledge Handbook of the Computational Mind*, edited by Mark Sprevak and Mateo Colombo, 237–46. London: Routledge.

Sterne, Jonathan (ed.). 2012. *The Sound Studies Reader*. London: Routledge.

Van Assche, Christine etc. (eds.). 2002. *Sonic Process. A New Geography of Sounds*. Paris: Centre Pompidou.

Voegelin, Salomé. 2010. *Listening to Noise and Silence. Towards a Philosophy of Sound Art*. New York: Continuum.

Article received: April 30, 2019
Article accepted: May 15, 2019
Original scientific paper

UDC 78:004.8

Asher Tobin Chodos
University of California San Diego
Department of Music
La Jolla, California, United States

WHAT DOES MUSIC MEAN TO SPOTIFY? AN ESSAY ON MUSICAL SIGNIFICANCE IN THE ERA OF DIGITAL CURATION

Abstract: The growing field of "critical algorithm studies" often addresses the cultural consequences of machine learning, but it has ignored music. The result is that we inhabit a musical culture intimately bound up with various forms of algorithmic mediation, personalization, and "surveillance capitalism" that has largely escaped critical attention. But the issue of algorithmic mediation in music should matter to us, if music matters to us at all. This article lays the groundwork for such critical attention by looking at one major musical application of machine learning: Spotify's automated music recommendation system. In particular, it takes for granted that any musical recommendation – whether made by a person or an algorithm – must necessarily imply a tacit theory of musical meaning. In the case of Spotify, we can make certain claims about that theory, but there are also limits to what we can know about it. Both things – the deductions and the limitations – prove valuable for a critique of automated music curation in general.

Keywords: music information retrieval, music recommendation, machine learning, music semantics, meaning, Spotify, digital culture

One overlooked feature of Spotify's software is that its user experience tends not to discriminate among traditional musical types. Its search box, for example, accepts virtually anything as valid input. Users can enter particular artists, albums, and songs, but they can also enter genres, moods, or other kinds of musical keywords. The resulting recommended materials are equally heterogeneous. Whether we take the "lean in" or "lean back" approach,[1] we are confronted with a mixture of

* Author's contact information: tobinchodos@gmail.com

1 Industry jargon referring to music streaming software that assumes an active ("lean-in") or passive ("lean-back") approach to what music is played.

genres, moods, playlists, or other kinds of "hubs" (Spotify's umbrella term for these variegated musical departure points) as search results. Above all of this diverse suggested material hovers the same inviting "play" button; a hub for "black history is now" is clickable in the exact same way as a "radio" station seeded by Parliament. So is the "artist" Parliament, as is their classic 1978 track, "Flashlight."

This is an important feature of Spotify's software design. This array of clickable options nurtures an impulse for instant gratification and is probably a strategy to maximize user retention.[2] It also means that Spotify is not simply a place to go to hear the music you want, but a place to *learn* about what you want as you make your way through a sea of cute icons that respond to clicks with various kinds of sonic offerings. In other words, Spotify is primarily a music discovery service.[3] As website you visit essentially to explore, Spotify communicates a certain seamless intimacy with the user. Spotify is not a machine that delivers requested goods for a fee; it is an open-ended, benevolent, and exploratory experience in which it is assumed that the data surveilled from your behavior can only enrich your relationship with the program and improve the quality of your recommended content.

It is of course natural for any profit-driven enterprise to want to project this benevolence – and, in a culture of what Shoshana Zuboff calls "surveillance capitalism" (Zuboff 2019), Spotify's practice of surveilling user behavior is an unremarkable example of what has become the dominant business model for tech companies. But it is worth pointing out that music consumption in the digital age was not always this way. Napster and MP3.com, for example, were revolutionary simply because of how much music they made easily available, not for the ingenuity with which they helped users discover new music. Today, since putting 30 million songs within reach is no longer impressive on its own, and because the excess of audio material is harder than ever to make sense of and sort through on your own, music streaming services have, increasingly, needed to become music discovery services.

It is impossible to know exactly how Spotify's music discovery engine works. This is because the system does not work in any one way at any one time for any one user.[4] Additionally, even if it were not subject to constant mutation, the actual algorithm is a carefully guarded trade secret. In spite of the limitations on what we can know about the inner workings of Spotify's music discovery engine, it nevertheless seems straightforwardly true that, no matter how Spotify's recommendations are actually made, the system must in some way be predicated on a notion, explicit or not, of musical meaning. Insofar as any recommendation, by a human or by a machine,

---

2 See Seaver (2018) for a discussion of how recommendation algorithms increasingly optimize retention over more traditional metrics for recommendation quality.

3 Spotify's gradual transition from a streaming service to a discovery service is discussed in detail below, but it is also evidenced in its recent acquisitions: the Echo Nest (acquired in 2014), Niland, Sonalytic, and MediaTv (all three acquired in 2017) are all companies acquired by Spotify that specialize in automated content curation.

4 See Seaver (2013) for an instructive reminder to remain humble about the "knowability" of recommendation algorithms.

depends on ideas of musical salience and similarity, we can say that Spotify's recommendation service represents a tacit theory of musical meaning.[5] This essay seeks to probe that theory; to make some tentative claims about what its essential contours must be (always acknowledging that the Spotify system is hidden and constantly evolving) while introducing a framework for thinking critically about it.

For good reasons, Spotify's system does not encourage this kind of critical thinking. Questions about the system's implicit theory of musical meaning can only serve to remind users that its theory is just one of many – and therefore not necessarily the best one. The success of the Spotify model depends on communicating that its catalog is both complete and effectively managed – that it has achieved a unique balance of "scale" and "care," to use the words of one of the designers of its recommendation technology (See Whitman 2012). Relativizing the theory of meaning upon which the system depends represents a disconcerting imperfection. If the technology populating my "discover weekly" playlist relies on just one way to construe musical significance, who knows what gems it might be missing, how it might be guiding my consumption habits, manipulating my moods, or shaping my personal identity.

Spotify may not go out of its way to highlight this idea, but the notion that the system is in fact predicated on such a theory can be traced back to one of the first places where Spotify's recommendation technology was laid out: the 2005 doctoral dissertation of Brian Whitman at MIT (Whitman 2005). Although it was published well before Spotify officially launched, Whitman's "Learning the Meaning of Music" introduced the basic outline of the software that would eventually power a hugely successful music intelligence company, The Echo Nest, which Spotify acquired in 2014. As I argue below, some aspects of this technology almost certainly continue to operate in present-day Spotify. And so, Whitman's doctoral dissertation forms a useful, if partial, entry point to Spotify's black box.

As is clear from the title of the dissertation, Whitman proposes this technology while engaging explicitly with the question of musical meaning. He promises that he will be "Learning the Meaning of Music" – but meaning in what sense exactly? To echo Hilary Putnam, one of the few humanistic sources cited by Whitman, what is the "meaning of meaning" in that title (Putnam 1975)? Regardless of how much of this technology is actually used for a given recommendation task by Spotify today, this article contends that a theory of musical meaning gleaned from Whitman's dissertation can be a part of the broader effort to think critically about what music "means" to Spotify. More generally, this can offer a basis for thinking critically about the consequences of the rise of automated curation in music.

This issue is analogous to questions pursued in the discipline of "critical algorithm studies." The idea of embedded bias, for example – the prospect that ostensibly

5 It is important to remember that although it is best known for its "data-driven" approach to music information and its automated personalized recommendations, Spotify actually continues to employ human curators. See Ugwu (n.d.).

objective algorithmic tools will silently encode certain assumptions – is a major theme in this field.[6] As is the related issue of "fairness," which focuses on the real-world consequences of applied machine learning, especially as it concerns social justice and inequality. These issues are clearly explained by Tal Zarsky:

> Any institutional decision that applies or allows algorithms to automatically sort, govern, and decide issues related to human actions makes two crucial assumptions: that human conduct is consistent and that with sufficient data human behavior becomes predictable (Zarsky 2016).

It makes sense that the bulk of the critical attention has, so far, been paid to machine learning applications outside of music. For example, financial institutions have begun to incorporate algorithmic recommendations into their decisions about whether to grant home loans; the question of whether those algorithms will tend to reproduce the structural injustice implicit in their ground truth data is an urgent concern for critics of digital culture. In a similar way, the algorithmic aids used in prison sentencing have been the subject of extensive reporting by, among others, the news organization *Pro Publica*. Machine bias in music, by comparison, feels less urgent. Pierre-Nicholas Schwab, an important figure who writes about fairness in machine learning, even uses music as the paradigmatic case of a place where a lack of fairness does not really matter:

> There is a big difference between a music recommendation service and a news recommendation service. What are the consequences of biased recommendations in a subscription-based service like Spotify? Getting a track recommended that you may not like and will skip. The consequences are small for the consumer (Schwab 2018).

Yet, there are other possible consequences. If we are recommended the same kind of music again and again, what does that do to our musical taste? If playlisting algorithms tend to privilege certain genres over others, do not recommendation engines represent a serious social justice concern? What, in short, are the cultural consequences of a music industry increasingly mediated by the software design decisions of a few large companies?

These considerations, and many others like them, should matter to us if we care at all about what music people are exposed to, and the manner in which our culture relates to that music. In order to investigate these questions, we need to get as strong a sense as possible of how these systems work, and then make informed decisions about how we listen to them. In this article, I look at the notion of musical meaning that is at work in the Spotify algorithm (the currently dominant music

6 See, for example, Powles and Nissenbaum (n.d), which raises the issue of embedded bias while reminding us that seeking to "fix" AI in this way actually represents a concession to its viability and inevitability.

recommendation system). I also ask whether this notion is a good one, and what is lost or gained in the transition to a culture of listening in which automated curation is the norm – a transition which, for better or worse, we are definitely making. The following discussion proceeds in three sections:

* First, I sketch a history of Spotify's development, dispelling some commonly held beliefs about it and showing how it transformed from a streaming company to a discovery company. Here, I argue that automated music recommendation services must necessarily rely on some notion of musical meaning.

* Second, I make a case for why Spotify almost certainly continues to employ some of the techniques Whitman developed in his 2005 dissertation.

* Finally, I attempt to discern Spotify's theory of musical meaning itself. I do this, first, via a close reading of the behavior of the Spotify graphical user interface (GUI) and, second, via Whitman's 2005 dissertation. In the latter case, I argue that the techniques outlined in the dissertation are novel and probably effective, but that there are interesting gray areas where Whitman addresses the question of musical meaning. In the end, I neither condemn nor endorse Spotify's system. Instead, I merely hope to show that a system like Spotify inevitably relies upon a theory of meaning; as users of that system we will benefit from paying close attention to what that theory is.

I. Spotify and the "Curatorial Turn" (2008–2018)

There is a widely held belief that when Spotify was launched in 2008, it was as a response to a music industry imperiled by the growing practice of music piracy.[7] It is true that by the time Daniel Ek and Martin Lorentzon created the startup that would eventually mature into a publicly traded corporation worth more than $20 billion, the recording industry had contracted enormously from its peak at the end of the 20th century.[8] The familiar narrative casts Spotify as a reaction to and, perhaps, a solution for the industry's financial crisis. And indeed, this sometimes seems fair: according to the International Federation of the Phonographic Industry (IFPI), for example, industry revenue in 2018 had recovered to 68.4% of that peak value, largely on the strength of a 45% growth in paid subscription streaming (IFPI 2018). As Spotify is by far the largest paid subscription service, with some 200 million active users today (87 million of whom are paying for subscriptions),[9] Spotify appears to be, from this perspective, an important driver of the industry's recovery, vindicating

7 See for example, the BBC news 2018 article "How Spotify came to be worth billions," (BBC 2018) which casts Spotify as a "response to the growing piracy problem," or Silva (n.d.), or many others that echo this idea.

8 According to Greg Kot, revenues from recorded music in America plunged from their all-time peak of $14.6 billion in 1999 to $12.6 billion in 2002, a decline of 13.7 percent. (Kot 2009, 31)

9 Apple music, though, is gaining on Spotify, with 56 million users as of time of writing (see Yoo 2019).

the altruistic posture the company occasionally projects.[10] The major record labels are frequently castigated for their repeated failures to develop viable systems of electronic distribution in the digital age. Spotify, as a kind of commercial imitation of illegal file sharing, can be seen as the music industry's belated effort to rectify that mistake. Heralded as the "solution to music piracy,"[11] Spotify is thought to restore value to the industry, connecting listeners with the music they want to hear and artists with interested audiences – and all in conformity with US copyright law. So goes, at any rate, the familiar narrative.

This narrative, however, obscures some important facts about Spotify and the relationship between music streaming and the music industry in general. First of all, it ignores the fact that Spotify has yet to turn a profit. In fact, Spotify's annual operating losses have increased sharply every single year, from €98 million in 2013 to €378 million in 2017 (Richter 2018). In 2018 and 2019, Spotify's losses have decreased, but the company remains unprofitable.[12] Although these kinds of consistent losses are not unheard of in the present investor climate,[13] Spotify's financial profile should still give pause to those who want to see it as the music industry's savior.[14] It will, after all, eventually have to turn a profit or fold. Moreover, it is important to note that these losses are not for lack of revenue or a reliable customer base, but instead point to the same old problem the music industry has always faced in the digital age: these losses are due primarily to the licensing costs paid out to the major labels, which represent Spotify's biggest operating expense by far. The fact is that customers are unwilling to pay what they used to pay for music, but major record labels remain committed to intellectual property paradigms from the 20th century, paradigms that only work with 20th century revenue streams. This has been the problem facing music sellers for the last two decades, and Spotify has not solved it. If Spotify is responding to an industry beleaguered by widespread piracy, its response fails in precisely the same way that Napster's did. The difference is that where Napster was bankrupted by aggressive litigation from the Recording Industry Association of American (RIAA), Spotify is kept from turning a profit as it funnels most of its revenue (and shares of its stock) to the major labels – which

10 As it does, for example, in Brian Whitman's lengthy 2012 blog post, "How Music Recommendation Works–and doesn't work" (Whitman 2012), discussed at length below.

11 Marsha Silva, "Spotify, the 'Solution to Music Piracy,' Is Getting Pirated by 2 Million People," published in Digital Music News, at https://www.digitalmusicnews.com/2018/03/26/spotify-piracy-hacked/, accessed May 21, 2019

12 See Spotify's publicly available financial disclosures at https://investors.spotify.com/financials/default.aspx

13 Pandora too posts losses in the hundreds of millions, and in general traditional notions of value have changed radically across the economy. As hedge fund manager David Einhorn puts it, "the market has adopted an alternative paradigm for calculating equity value." (quoted in Kim (2017))

14 Spotify's 2019 press release for investors is jubilant about its first quarter earnings while predicting another loss of €180-340 million. See https://investors.spotify.com/financials/press-release-details/2019/Spotify-Technology-SA-Announces-Financial-Results-for-First-Quarter-2019/default.aspx

constitute the controlling forces of the RIAA.

The fact that this can be said of the streaming industry's biggest player raises important questions about the financial viability of the streaming model itself; if Spotify can't make it work, one wonders, who can? Spotify has over the years shifted between various strategies for earning revenue: early on it looked to advertising, before attempting to monetize its integration with Facebook, and now it sees subscriptions as its principal revenue stream. But it would be more accurate to say that Spotify's true source of revenue has always been venture capital, which it has attracted with extraordinary success, gaining more and more money over the course of 24 funding rounds even in the face of large losses. If Spotify succeeds only in raising venture capital, growing quickly, and collecting potentially monetizable user data, it no more represents a solution for the music industry than Uber or Air B&B – both are companies that have been extraordinarily successful at raising venture capital, but which contain no special insights about the music industry.

This familiar narrative about Spotify, in which it is lumped together with other "disruptive" tech firms, also obscures another important fact: that, although it is marketed as a novel and innovative firm, it is in fact largely owned by the traditional music industry forces. Since Spotify cannot afford a market rate for the licensing fees its service requires, it has been forced to compensate the major labels, in part, with company equity rather than cash. As a result, Peter Tschmuck reports, major labels own as much as 20% of Spotify today (Tschmuck 2017, 179). This fact is perhaps the cause of the widespread concern in the music industry about so-called "playola," a word that refers to the influence major labels supposedly wield over the content of Spotify's curated playlists (not to be confused with the familiar "payola," which denotes a similar practice from radio broadcasting).[15] It is also a possible cause for the often reported homogeneity of Spotify's automated recommendations, an effect which, if authentic, would undermine Spotify's stated aims as a music discovery service.[16] In any case, it is important to remember that, although Spotify is often said to have "disrupted" the industry, it is largely owned by the major record labels, and they are the ones who benefit and receive the majority of its revenue.

The familiar narrative also overstates the relationship between Spotify and the industry as a whole. If we believe that Spotify has the potential to "rescue" the industry from the scourge of piracy, we must believe that it has a marked effect on the market itself. Yet, that may not be true at all. While Pandora has commissioned studies showing that Internet radio has positive effects on music consumption in general, there is little consensus on this point and other scholars have found quite the opposite result. Or, we may simply find that Spotify has no net effect on the music industry whatsoever. Aguiar and Waldfogel, for example, find that while Spotify does displace some lost revenue due to piracy, the new revenue is "roughly

15 For a representative complaint about playola, see, e.g. Peoples (2015).

16 Spotify's app blurb on the Google app store, for example, promises "the right music for every moment" (and, moreover, for every individual user) – not just what the major labels want to promote.

offset by revenue reductions from the sale of permanent downloads" (Aguiar and Waldfogel 2015, 22). Spotify stimulates the market in some ways while depressing it in others, and it seems impossible to know exactly how to gauge its impact on the industry as a whole. Therefore, it is not necessarily reasonable to assume that Spotify has either "rescued" or depressed the market.

Furthermore, it is not even certain that the industry's crisis in 2006 was due to piracy in the first place (the problem to which streaming is often seen as a solution). While it is true that by 2006 revenues had seen a sharp decline from their peak in the 1990s, the golden years the record industry enjoyed in the 1990s should not necessarily be seen as the norm. Instead, some have seen them as anomalous, a period of growth artificially stimulated by the advent of the CD and, therefore, inherently short-lived. Revenues had, in fact, been declining for a long time before the arrival of the CD, which gave the industry a lift largely thanks to its new ability to sell consumers CD versions of music they already owned on vinyl and tape. From this perspective, it is only reasonable to expect that this lift would be temporary – and therefore, perhaps it's inaccurate to blame the downturn on internet piracy and file sharing. The claim that piracy is responsible for the industry's downtown, though repeated constantly by the RIAA and industry insiders, is not necessarily true. As Greg Kot notes,

> It was disingenuous of the industry to blame its slump on file sharing without acknowledging the role played by rising CD prices. The average retail price of CDs had increased more than 19 percent from 1998 to 2002. Peak price was $18.99, with middlemen getting the vast majority of the split (Kot 2009, 42).

If this picture is accurate – if the industry's pains at the turn of the century were a natural regression rather than the result of disruptive new technologies or cultural shifts – then the whole idea of Spotify as the industry's savior, "restoring value" to a business struggling to accommodate new technological paradigms, is an oversimplification. Despite aligning itself with the rhetoric of disruptive innovation popular in the tech industry, in actuality Spotify is probably neither the industry's savior nor its destroyer, and, in many ways, it continues the patterns and promotes the interests of the major record labels who are among its largest shareholders. From a business perspective, Spotify is much less exceptional than it seems.

Meaning and the Curatorial Turn

But even if Spotify may not be the determining factor behind a sudden shift in the music industry, it certainly marks one. How (or whether) the streaming industry is to become self-sustaining remains a mystery; nevertheless, it is hard to imagine a future in which the music industry does not have, at its center, music streaming services. Over the last 11 years, Spotify has evolved from a music streaming company that in many ways inherited the mantle of Napster, Gnutella, and Limewire, merely seeking to provide legal access to a large catalogue of music, to a music discovery company whose most valuable properties are its recommendation engines. In this section, I trace that evolution.

In the only academic history of Spotify, Maria Eriksson et al. (2019) divide its evolution into seven periods (Eriksson et al. 2019, 43-67):

- "Beta" Period (2007–2008). Spotify released to a small circle of personal acquaintances.
- Period A (2008–2009). First public version launched in October 2008 in eight European countries. Spotify removes unlicensed music from its service. Spotify begins to sell advertising and launches ad-free Spotify Premium.
- Period B (late 2009). Global financial crisis eats into advertising revenue and venture capital. Doubts about viability of an ad-supported model leads to increased emphasis on subscription services.
- Period C (2010–2011). Spotify as a platform, emphasis on social features. Linking of Spotify and Facebook, increased practice of data extraction from users. "Related artists" function added. Spotify opens in the US.
- Period D (2011–2012). Valuation reaches $10 billion. Increased "platformization." Competition with Internet radio sites in the US (such as Pandora) leads to increased importance of recommendation and discovery.
- Period E (2013). Spotify begins to address "the abundance of choice" as a primary problem. Solution is no longer primarily social, but algorithmic. Spotify positions itself as a discovery company. Spotify acquires music recommendation company Tunigo (May 2013), which recommends music based on social activities and moods.
- Period F (2013–2015). Spotify dismantles the P2P network, opting instead to use its own servers. Spotify acquires The Echo Nest (2014), an important music information company, for $100 million.
- Period G (2015–2016). In competition with Apple Music, Spotify emphasizes its ability to create musical experiences tailored to each moment. Curation strategy combines the expertise of two acquired companies: Tunigo (expert human curation) and Echo Nest (scalable algorithmic curation). Also acquires Seed Scientific, a data science company. Summer 2015, Spotify introduces various personalized weekly playlists, such as "discover weekly."

As this timeline shows, since its founding, Spotify has nimbly adjusted to shifting market priorities and trends in startup culture, at times making dramatic adjustments to its marketing strategy and business model to accommodate these shifts. Not long after the collapse of Napster, Spotify began as a peer-to-peer sharing service that not only copied parts of Napster's technical architecture, but actually permitted the sharing of unlicensed music. When Spotify launched its first publicly available version in 2008, it removed the unlicensed music, but preserved much of the P2P architecture and kept the disruptive caché of Napster as part of its marketing strategy. After the global financial crisis cast widespread doubt on the viability of advertising for all Internet companies, Spotify recast its free tier as a marketing strategy for its subscription service, which would now become its primary revenue stream. In the wake of Facebook's monumental growth around 2010, Spotify partnered with Facebook and integrated itself into the social network giant.

Among these various adjustments, the most important one for the purposes of this paper is the so-called "curatorial turn:" the shift toward music curation as an important element in Spotify's service. Largely because of its arrival in the USA market in 2012, where it had to compete with Pandora and other Internet radio services, Spotify has increasingly positioned itself as a "music discovery service" rather than simply a music streaming service – and this remains the form Spotify takes today. Even a cursory look at Spotify's service today reveals how central recommendations are to its service. This shift can also be seen by looking at the contrast between two versions of Spotify's homepage, one from 2006 and one from Spotify's "about" section in 2019.

In 2006:

> Spotify gives you the music you want, when you want it.
> Your choice is just a search box or a friendly recommendation away.
> You'll be amazed by the speed and control you have with Spotify.[17]

And in 2019:

> With Spotify, it's easy to find the right music for every moment.
> Choose what you want to listen to, or let Spotify surprise you.
> Soundtrack your life with Spotify.[18]

The difference in tone is subtle but illustrative. In 2006, Spotify is a service that, ultimately, delivers "your choice," even if that choice can be optionally mediated by the service's recommendations (recommendations which, at the time, were probably mostly made by humans rather than machines). The leading line promises "the music you want," clearly prioritizing and emphasizing the volition of the user. This blurb also promises the user "speed and control," two features that an informed, self-directed user might value. It clearly targets a user that takes an active role in her media consumption, using what the industry terms a"lean-in" strategy.

Although it probably holds appeal for aficionados and professionals, this posture eventually became a liability,[19] and Spotify had to adjust. And this meant designing

17 Accessed via the Internet Archive, at https://web.archive.org/web/20061127231638/http://www.spotify.com/, accessed May 20, 2019.

18 Spotify Home Page, "About Us," https://www.spotify.com/us/about-us/contact/, accessed May 20, 2019.

19 In 2011, for example, Billboard published an article in which Spotify was negatively characterized

a software that had something to say about musical quality, about music qua music. In 2019, what matters is no longer the music you want, but the music that is appropriate for "every moment." The value the user might find in having control over the tool is replaced by its power to "soundtrack your life," that is, to find music that matches whatever non-musical activity you happen to be engaged in. This is a notable shift to a "lean-back" approach, a shift which has taken place with respect to the media industry in general over this decade.[20] Interestingly, this shift engenders an adjustment in Spotify's attitude toward music itself; as we lean back, music's value comes to reside primarily in its relationship to things outside of itself. A peculiar feature of the rise of curation is that the value of music is based on how it "goes with" other things rather than what it sounds like (a fact which is discussed at greater length below). This is not a posture Spotify found itself taking before the curatorial turn.

More than the size of the catalogue or the quality of the sound,[21] Spotify's current selling point is its discovery product. And although Spotify does continue to employ human curators (See Ugwu, n.d.), it probably uses more automation than any of its competitors. Spotify's service, then, is not simply to provide customers with access to an enormous database,[22] nor is it exactly to help them find music they like. Instead, what Spotify promises is to help customers find the right music for a given moment, to "soundtrack your life." On the face of it, this slogan makes a pretty bold statement: that the millions of tracks in Spotify's catalogue are "soundtrack" music. It is only made obliquely, so it is easy to miss, but it is a real consequence of the curatorial turn. Here Spotify is part of a broader trend in digital culture. As Peter Wikstrom puts it,

> In a world where information is abundant, people may not be willing to pay a premium for basic access to that information, but they are most likely willing to pay for services which help them navigate through the vast amounts of information (Wikstrom 2013, 7).

Spotify is not unique in its turn toward automated curation, but making that turn engenders certain shifts in its basic attitude toward the meaning of music. One such shift is the subtle creep of the "soundtrack," the idea that music is generally supplemental to other activities and modes of consumption.

---

as "just a huge database of songs." (cited in Eriksson et al 2019, 59).

20 For work on the rise of curation in general, see, e.g. Silberman (2015) and Gillespie (2011).

21 Even the sonic watermarks imposed by many of Spotify's music industry partners (which are noticeable) seem not to deter customers at all. See Matt Montag's blog (https://www.mattmontag.com/music/universals-audible-watermark) for a useful demonstration of those watermarks. Accessed May 16, 2019.

22 Spotify's 30 million track catalogue, while bigger than those of its competitors, is no longer really its main selling point.

II. Is Spotify Using the Echo Nest?

The rise in demand for curation services was an engineering problem that Spotify approached in more than one way. Following the broader trend of social networking after 2010, Spotify's first solution was, to use the industry's word, "social." In 2010, Spotify received $16 million in venture capital from Sean Parker, the co-founder of Napster. After Napster, Parker had gone on to become the founding president of Facebook. With his investment in Spotify, he earned a spot on its board of directors and ensured that the two companies could integrate their products smoothly. Through the integration of Spotify and Facebook, the social model of music discovery was possible: the listening habits of one's friends could be distilled and transformed into music recommendations. This strategy has the advantage of requiring relatively little engineering, and it is predicated on the intuitively reasonable assumption that people share musical tastes with their social groups. There are a number of ways in which this strategy is not particularly useful, though: first, it will never be a reliable way to expose users to music that is not already popular. Second, like all "context based" recommendation systems, it bears no formal relationship to the musical content itself. Third, it still demands the active engagement of the user, the "lean-in" attitude that Spotify had traditionally envisioned for its customers.

Automated recommendations could potentially address these shortcomings. Facing these issues, as well as competition from American Internet radio stations like Pandora, Spotify began to more aggressively develop its automated recommendation engine in 2012. It began to foreground its recommendation services, adjust its marketing strategy, and, above all, it acquired prominent companies in the music intelligence and recommendation space.

Probably the most important acquisition was The Echo Nest, which Spotify bought in 2014 for $100 million (Lunden 2014). Founded in 2005 by two graduates of the MIT Media Lab, Tristan Jehan and Brian Whitman, the company quickly grew into one of the biggest players in the music recommendation space. Its API powered the music recommendation services of major companies like MTV, Rdio, and Spotify (before the latter bought it). The technology employed by The Echo Nest is described in the academic writing of its founders (especially Whitman's dissertation), and below I will be using those texts to make some deductions about Spotify's current software. But is it reasonable to assume that Spotify is actually still using the technology it acquired in 2014? It is widely known, after all, that Silicon Valley companies regularly acquire technology without ever putting any of it to use. Additionally, 2005 was a long time ago and the technology Whitman proposed in his dissertation may well be out of date today.

There is, however, good reason to believe that Spotify does in fact use Echo Nest technology today – or, that it at least shares crucial features, with respect to its attitude towards musical meaning, with the technology Whitman developed in 2005. This can be seen by closely reading the following three documents: (1) Brian Whitman's 2005 dissertation at MIT, (2) a blog post he made detailing the Echo Nest's service in 2012, and (3) the current official documentation of Spotify's API. The similarities among these three documents, which trace a timeline as long as Spotify's own, make a compelling case for the idea that Spotify's contemporary recommendation engine shares at least some features with the software originally designed by Brian Whitman in 2005. This is important, of course, because the dissertation is in the public domain and can be read in detail. Bearing in mind the important qualifications raised by Nick Seaver (2013), and being careful about the scope of our argumentation, we can ground certain claims about Spotify and automated recommendation in a close reading of the dissertation.

In 2012 (two years before the Spotify acquisition), Brian Whitman penned a blog post (Whitman 2012) outlining the Echo Nest's general approach to music information and his own opinions on the industry as a whole. This post explicitly links the technology of the Echo Nest to the research activities of both himself and Tristan Jehan at the MIT Media Lab, and most of the features he describes in the blog also appear in his doctoral dissertation. For example, in the blog post, Whitman expresses his deeply held conviction that musical similarity derives from "cultural" meaning, not simply audio signals:

> We've shown over the years that people's expectation of "similar" – either in a playlist or a list of artists or songs – trends heavily towards the cultural side, something that no computer can get at simply by analyzing a signal (Whitman 2012).

This idea – that musical meaning resides outside of the audio signal – is the central conceptual frame for Whitman's doctoral dissertation from 2005, which positions itself unambiguously against an "absolutist" theory of musical meaning deriving "from the signal alone." It is not an overstatement, in fact, to say that this is the whole idea of the dissertation. When Whitman promises to "learn" the "meaning" of music, what he is promising above all is to capture, and render legible to machines, the difficult and unruly "cultural" information that attaches to the audio signal – and then to combine the two information streams into a single classification system into which any music can be fed. The idea that musical meaning is not in the signal alone is the single most important idea animating the dissertation *and* the 2012 blog post. Thus, we have our first clear conceptual connection between the two.

In this same post, Whitman also refers to the Echo Nest's "Audio Analysis Engine," and even provides a link to Echo Nest official documentation of this product, prepared by co-founder Tristan Jehan. This document explains how the Echo Nest's machine listening works. That is, how their system deals with the audio signal itself (as distinct from the extra-signal "cultural metadata" so central to Whitman's intervention). The Audio Analysis engine detailed in 2012 bears unmistakable similarities to the one Spotify makes available today. The 2012 document, for example, takes in an audio signal and rates it in various ways. It can evaluate it in conventional musical ways, according to its *key*, *mode*, and *tempo*. These are standard music information retrieval tasks. The 2012 document also contains more idiosyncratic measures, however, such as the abstract musical categories of *valence*, *danceability*, and *speechiness*.

Crucially, all these same categories are available today in Spotify's "Get Audio Features" API endpoint.[23] Exactly as in the Echo Nest circa 2012, Spotify today evaluates tracks for their key, mode, tempo, as well as their valence, speechiness, and danceability. Moreover, in most cases the language of the contemporary API documentation echoes verbatim the language of Tristan Jehan and Whitman in 2012. Here is Whitman characterizing The Echo Nest's machine listening tool in 2012:

> We emit song attributes such as danceability, energy, key, liveness, and speechiness, which aim to represent the aboutness of the song in single floating point scalars (Whitman 2012).

Each of these idiosyncratic metrics (danceability, energy, etc.) is outlined in the contemporary Spotify API documentation, with each one still represented as a single floating-point scalar. For more commonalities, we can look at the way these fields are defined. Here, for example, is Jehan defining mode in the 2012 documentation:

> [Mode] indicates the modality (major or minor) of a track, the type of scale from which its melodic content is derived (Jehan 2012).

And here is Spotify defining mode in the contemporary API documentation:

> Mode indicates the modality (major or minor) of a track, the type of scale from which its melodic content is derived. Major is represented by 1 and minor is 0.[24]

Similarly, in 2012, Jehan defines the key output of the Echo Nest's audio analysis tool as:

> The estimated overall key of a track. The key identifies the tonic triad, the chord, major or minor, which represents the final point of rest of a piece (Jehan 2012).

Which has been somewhat refined in Spotify's 2019 documentation:

> The estimated overall key of the track. Integers map to pitches using

23 Accessed April 1, 2019 at https://developer.spotify.com/documentation/web-api/reference/tracks/get-audio-features/

24 See "Get Audio Features" endpoint in the Spotify API, https://developer.spotify.com/documentation/web-api/reference/tracks/get-audio-features/

standard Pitch Class notation. E.g. 0 = C, 1 = C♯/D♭, 2 = D, and so on. If no key was detected, the value is -1.

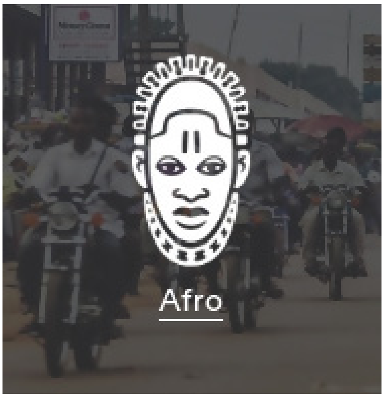
The rest of the fields exhibit the same parallelism. It seems clear that Spotify today is using the same audio feature extraction techniques that Whitman and Jehan were writing about in 2012 – which Whitman in turn explicitly connects to his own 2005 system. The grounding idea of Whitman's dissertation, moreover – that musical meaning resides not in the audio signal alone – is a prominent theme in his 2012 blog post. Reasoning from these commonalities, this article can deduce that Spotify in 2019 is still using at least some key features elaborated in 2005 by Whitman and that it is therefore likely that the "theory of musical meaning" elaborated in the one is roughly operational in the other. There is undeniably some license in this inference, and some readers may want to reject all or part of this assumption; I hope that even the most skeptical reader, however, will find the following discussion worthwhile.

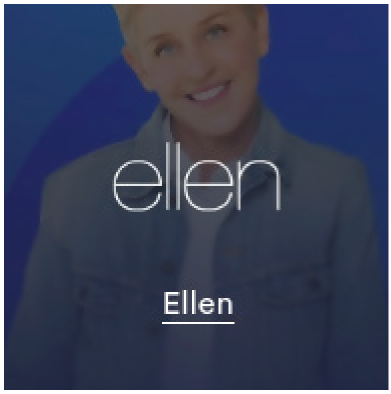
III. What does music mean to Spotify?

III (a) Spotify GUI

The theory of musical meaning I ascribe to Spotify will be principally derived from its underlying technology, which I examine mainly in the form of Whitman's dissertation (2005). Before doing that, though, it is worth taking a moment to look at Spotify's graphical user interface (GUI) to examine the notion of musical meaning implied there. Even a cursory examination of its front-end reveals some key assumptions Spotify makes about how music is meaningful to its users.

Music is grouped for Spotify users not primarily by genre or style (and certainly not by album, a concept that has grown increasingly outdated in the post-Napster world), but rather by mood, activity, and what might be termed as "musical keywords." Under the "browse" section, the user is confronted with various buttons that will lead to musical options. These are termed "hubs" in the Spotify lexicon, and they are represented by clickable square thumbnails. Hubs are distinct from the more traditional "genre" marker in that they can refer to various different kinds of musical reference. There are hubs pointing to traditional genres ("country" and "folk"), but also to activities ("party" and "chill"), as well as to politically-oriented themes ("black history is now"), sponsored content ("Spotify singles"), and, curiously, even a single hub dedicated to Ellen DeGeneres ("Ellen"). Hubs appear as thumbnail images with artwork evoking a given hub's theme (a raised fist, for example, for "black history now," a dove for "christian," and a cartoon of an African mask for the "Afro" hub). These thumbnails rework the traditional idea of an "album cover," turning it into a generic index for a given mood, more or less in the way emojis caricature human affective states. Examples of the clickable thumbnails "Ellen" and "Afro" are shown below:[25]





Although the selection of "Ellen" as a hub alongside "Afro" may seem inscrutable, the heterogeneity of the Hub themes illustrates an important feature of the kind of musical meaning the Spotify GUI seems to assume: in the curatorial phase of music streaming, music's meaning resides in its relationship to other activities or feelings. The traditional idea of *genre* is that there are certain musical properties shared among all members of a genre. The reference for a genre is, as Whitman would put it, "the signal itself." This is not true of "hubs," which are instead significant for their extra-musical references (as in the hubs "study," "sleep," "Ellen," etc.). Considered as a "hub," even the word "reggae" (apparently a *genre* word) works differently from "reggae" as a genre. Put "reggae" next to "Ellen" and you change the status of the word subtly. A "reggae" *genre* refers to the sound of the music, whereas a "reggae" *hub* is a broadly construed, fungible cultural index. Like "Ellen," it doesn't refer to a type of music so much as a musical-cultural vector.

Not coincidentally, this is exactly the kind of vector given a technical expression in Whitman's dissertation, which insists again and again that true musical meaning is informed by culture, that it is not in the audio signal alone. Of course, few people today would endorse the outmoded idea that real listening can or should take place in an idealized way, divorced entirely from extra-musical factors. Nevertheless, it is important to note that Spotify seems to have landed at the other extreme – that all music is "soundtrack." Listening on Spotify is not about attending to music but using music to evoke a desired feeling or achieve some other secondary effect. As Ellen herself puts it on a promotional web page for the "Ellen" hub, "I'm so excited to partner with Spotify on my very own music hub because music truly makes everything better. Well, music and salt."[26]

Like salt, music in the Spotify universe makes things better; presumably it

25 Screenshots taken from Spotify desktop app on May 20, 2019.
26 https://ellen.withspotify.com, accessed March 27, 2019.

also shares with salt the property of not being very good on its own. It is just one ingredient among others, one more good to consume in the effort to lead as full and happy a life as possible – something the reggae hub will help you do in a way that the reggae genre can't.

As theories of musical meaning go, this one is not crazy. The opposite extreme, where musical meaning is inherent in the idealized formal properties of a composition, is no less objectionable. It is interesting to note, however, that this is a posture Spotify arrived at mainly because it found itself having to help people discover new music; the idea of music as functional, or "relational," as Whitman sometimes puts it, is in part a byproduct of the need to make music discovery systematic and programmable. It is a music-philosophy statement arrived at because of the desperate need to accommodate a capricious market. Spotify's previous and more traditional "lean-in" posture, in which users were trusted to know what they wanted, does not rely upon any such philosophy of musical meaning. If users are finding their own music, Spotify itself is able to remain agnostic on the question of music's purpose. Users who know what they like don't want "hubs." It is only because market trends now demand a recommendation engine that Spotify has had to make choices about these questions. Its answers are visible, in part, in the user interface.

III (b) Reading Whitman, "Learning the Meaning of Music" (2005)

As Nick Seaver points out, knowing how a recommendation algorithm works is never a simple matter. Drawing specifically on his fieldwork in the music recommendation space, Seaver notes that, according to one interlocutor, there is never any single recommendation algorithm at work. Instead,

> There is not one playlisting algorithm, but five, and depending on how a user interacts with the system, her radio station is assigned to one of the five master algorithms, each of which uses a different logic to select music (Seaver 2013, 5).

When it comes to algorithms "in the wild," Seaver holds, it is never the case that they are simply a black box waiting to be opened by the right critic. The whole idea of the algorithmic black box is a red herring, a tempting fiction that tends to nourish the worst fears about algorithmic mediation. If there is a single secret code at work rather than a constantly changing and unspecifiable one, it is easy to assume the worst about it. The reality is that recommendation algorithms are far too intimately personalized, too frequently updated, and too complex for those fears to be either right or wrong in any straightforward way. This is not to say that suspicions about them are never justified, nor that the logic of a system can never be divined, but simply to remind us that we must bear in mind that our conclusions are almost always based on incomplete and possibly outdated information.

What this means in regard to Spotify, is that some types of claims are going to be more reasonable than others. We may never know how a given playlist was curated, nor, for example, what the precise proportions of "content-based" and "context-based" considerations are at work in Spotify's recommendations. But we can make empirical observations about logged recommendations, and we can think critically about the fact that every recommendation does combine the two types of signal in some way.

With these considerations in mind, one good way to approach the questions regarding the algorithm is to do a close reading of Whitman's 2005 dissertation. The technology outlined therein is distinguished above all by its ability to join two disparate subsets of music information retrieval: on the one hand, sophisticated "content-based" machine listening methods (methods that draw on the machine listening techniques alluded to above), and, on the other hand, "context-based" information culled from web crawling and other kinds of natural language processing. These two types of signal are combined into a machine learning model that, in turn, can be used to classify as-yet-unheard musical material.

Crucially, this is an approach that Whitman specifically positions against the kinds of music information retrieval techniques that derive musical meaning from the audio signal alone, which were apparently predominant in 2005. As Whitman puts it, systems that rely on the signal alone are "doomed," since they miss the essential element of human reaction. As noted above, the idea that musical meaning isn't "in the signal" is Whitman's most important theoretical commitment.

One interesting thing about Whitman's dissertation is the fact that, although it would eventually power a major corporation that many artists see as an exploitative shill for the major labels,[27] it is really an extended plea for a more nuanced treatment of musical meaning. At its heart is the kind of argument you might expect to hear from a musician: that musical meaning is hugely complex, variable, unpredictable, and contingent.[28] Whitman's language is, at times, quite personal:

> Our driving force behind this work is that fundamentally, the current approaches anger us: they don't seem right. Music is a personal force that resists 'processing,' 'packing' or 'understanding' (Whitman 2005, 91).

"Current approaches" in the above are those that take a signal-only approach (or, even worse, a context-only approach) to musical meaning. Either one, on its own, inevitably does a disservice to the true complexity of musical meaning. So far, Whitman's argument is one that probably few musicians would quarrel with. Actually, it sounds very similar to the kinds of complaints musicians frequently

27 See, e.g. (Sanchez 2018), who ranks Spotify near the bottom as one of the lowest-paying streaming services for artists, at $0.00397 per stream in 2018.

28 This is possibly because Whitman himself has performed as an avant-garde noise musician, under the stage name Blitter. According to Whitman's LinkedIn profile, Blitter's career ended in 2002. Careful not to confuse Whitman's stage name with the social network of the same name.

make of recommendation services (including Spotify): they just don't get it. But Whitman, of course, goes further than this complaint. He argues that by combining these two types of signal, one can come much closer to the true essence of musical meaning. In a basic sense, if we believe that The Echo Nest is a good system, we must agree with Whitman that he has in a non-trivial way managed to do what his title has promised: to "learn" the meaning of music. The "learn" of the title is an obvious reference to "machine learning." But what work is being done by the word "meaning?"

Newton v. Diamond and the question of musical meaning

The thesis begins by going over the well-known legal dispute between James M. Newton and the Beastie Boys over their usage of a sample from his 1978 release, *Choir*. The Beastie Boys legally licensed a few seconds of solo flute playing and looped it for their 1992 song, "Pass the Mic." The legality of the audio sample is not in dispute. Nevertheless, Newton sued for copyright infringement, arguing that the sample in question infringes upon the musical composition itself in a way not provided for by the negotiated mechanical license. Whitman uses this case to establish the central frame for his entire thesis. It proves that the true significance of music resides outside of the audio signal itself:

> When the Beastie Boys sampled his recording they took far more than the signal, even if the signal was all they took. Where can we find the rest? (Whitman 2005, 17)

After having used this case to establish the main framework for his thesis, Whitman leaves the legal questions alone. However, it is worthwhile to examine the actual facts of the case. One crucial point Whitman ignores is that the court immediately sided with the Beastie Boys. While James Newton would presumably agree with Whitman's central premise (that the Beastie Boys took more than the signal, even if it was all they took), the law does not. Strictly speaking, the only thing the case demonstrates is that James Newton *alleged* that they took more than the signal, a feeling he shares with his fellow musician Brian Whitman. Whomever we side with in the legal matter, the case does not really argue one way or another on the question of where musical meaning lives (which is Whitman's real focus in his thesis). In other words, the central frame for Whitman's "meaning," is almost off-topic.

Moreover, this case raises the issue of the "meaning" of music only in a relatively straightforward way, the same way in which almost any intellectual property dispute in music would: it points to the fact that reasonable people can sometimes disagree on what should constitute copyrightable musical property. As for the question of whether musical meaning can be convincingly derived from amalgamated reviews, Google searches, and machine learning, or whether it should be derived exclusively from the audio signal – the question to which Whitman's thesis is actually addressed – the Newton v. Diamond case bears no special relationship to it.

It is interesting to note that the case does hinge on a question of musical meaning, but one that it is different from Whitman's question. As Whitman correctly points out, the legality of the sample is not in question; the Beastie Boys obtained the rights to use the sound recording from ECM for $1000. But Newton also copyrighted the "Choir" composition, and it is this holding upon which he argues infringement has occurred. At issue, therefore, is the relationship between a sound recording and a composition, rather than, say, a listener/customer and a piece of music. The former relationship is what the legal case is about: the judges are really ruling on whether the legal instrument of a "composition" has been infringed upon by a sample deployed in a particular musical context. More specifically, what's at issue is whether or not six seconds of a sampled flute performance can constitute a vital part of the musical composition "Choir." The court upheld the verdict that, not only does the sample not constitute a vital part of the "Choir" composition, but that, even if it did, the Beastie Boys usage of it is "de minimis," that is, too insignificant to be legally actionable. As Chief Judge Mary Schroeder puts is "the dispositive question is whether the similarity goes to trivial or substantial elements."

The question is not whether the meaning can be derived from the musical stimulus but rather whether or not a small sample can infringe meaningfully upon the legal instrument known as the "composition." These are different problems. The legal case has nothing to do with the meaning of music in the broad, contextual way that Whitman will eventually construe it, that is, the sense in which music can be meaningful to a potential consumer base. Much less does it relate to the question of how that meaning can be leveraged into an effective recommendation engine. The legal case is much narrower than that, and all the argumentation connected to it remains firmly in the domain of musical form, explicitly excluding the "cultural metadata" that is so important to Whitman's work. The legal case that frames Whitman's "meaning" does tackle a problem of musical meaning, but it is not the same problem in which Whitman is interested. So, while the frame is an interesting entry to Whitman's real work, it does little to elucidate the nature of the musical "meaning" we are going to be learning about.

Whitman and Leonard B. Meyer

In spite of the critiques raised above, the case is rhetorically effective. It does seduce us into contemplating the problem of musical meaning. For Whitman, the answer is to "Learn the Meaning of Music." That is, to combine context-based (amalgamated human reactions to music) and content-based information (signal-derived) into a machine learning model that can, in turn, be used to evaluate as-yet-unheard audio signals. In sophisticated and often musically nuanced ways, ground truth data denoting the relationship of audio signal to semantic content is used to train classifiers that can determine membership of a given audio frame in a given semantic category.

At the heart of Whitman's system are machines that listen to music and, in ways informed by actual human reactions to music, determine its membership in musically useful categories. Note Whitman's usage of the idea of "meaning" in this framework:

> A model of the contextual information given a signal allows us to accurately 'understand' music (extract semantic features of link to the outside world) that hasn't even been heard yet. So what we call meaning throughout this thesis is defined as *the relationship between a signal and its interpretation*. In our work we create predictive 'machines' that analyze audio signals and extract projected community and personal reactions: these are 'meaning classifiers.' What we attempt to do here is computationally understand this extra-signal information and link it to the signal in such a way that it can be predicted for future audio (Whitman 2005, 19).

As noted above, this is a theory of musical meaning that Whitman posits in contrast to dominant intellectual trends in music information retrieval. The question of musical meaning is, of course, also dealt with in the disciplines of musicology and aesthetic philosophy, and Whitman situates his thesis in this intellectual tradition as well. Throughout the entire thesis, though, Whitman only cites one musicological source: Leonard Meyer's influential 1956 book, *Emotion and Meaning in Music*. This book serves as a humanistic counterexample to his own work, representing what Whitman terms the "absolutist view" of musical meaning:

> At the outset we should make it clear that our definition of meaning above is mostly referential, that is, it exists as the connection between two representations. This contrasts with the purely *absolutist* view discussed by Meyer, in which the meaning is encompassed purely within the composition or signal. Our approach considers both with an emphasis on referential types of meaning. Many musicologists study the absolutist view of musical meaning simply because there is no formal mechanism of analyzing the contextual information. What this thesis presents are ways of computationally representing both signal-derived and contextual music information and then ways of learning a model to link the two (Whitman 2005, 19).

Whitman's system combines digital signal processing techniques (content-based) with natural language processing techniques (context-based) to produce "meaning classifiers" – algorithms, trained on those two data sources, that can predict more "extra-signal information" for new, as-yet-unheard audio signals. It is a system for producing descriptions of music that incorporate both audio processing and large amounts of empirical, human-generated musical descriptions. Throughout this thesis, this system is associated with the words "meaning" and "understanding" (although Whitman sometimes places these words inside scare quotes).

Suppose that the system Whitman distills, a system that predicts "extra-signal information" about musical signals, is a good one. Whitman opposes it to Meyer's ideas, but how much distance does he really gain? In what follows I argue that the answer is "not much" – that is, that in spite of explicitly positioning himself against "absolutism" as encountered in his reading of Meyer, Whitman's approach actually aligns with Meyer's in most of the relevant ways.

Leonard Meyer serves Whitman in a similar way as the legal case discussed above. It is a framing conceit used to clarify his central intervention: that, contra both MIR and "many musicologists," meaning does not reside in the audio signal. For Whitman, Meyer exemplifies an approach to the question of musical meaning that attempts to derive it from the "signal" (from audio signal or representations in score, which, for Whitman, seem to be philosophically equivalent). "Many musicologists," Whitman tells us, take Meyer's approach, and they do so because "there is no formal mechanism of analyzing the contextual information." In other words, musicologists do not incorporate empirical human reactions into their theories of musical meaning because they lack any rigorous method for aggregating and processing them at scale. Whitman, of course, provides such a mechanism, and making this distinction is the beginning and the end of his engagement with Meyer and with the rest of the intellectual tradition for which he stands.

Whitman's system, however, in spite of its engagement with extra-signal materials ("cultural metadata") still has the same basic contour as Meyer's. Both address a scenario in which a signal is audited as the sole stimulus in a musical event. Meyer, availing himself of then-popular trends in psychology, characterizes music as a system of delayed gratification. Music sets us up to expect certain things and manipulates our innate desire to see those expectations fulfilled, in ways that stimulate complex affective responses.[29]

Whitman, as a software engineer, approaches the issue in a different way – but in spite of his protestations against "absolutism," not in a way that privileges the audio signal any less. Whitman produces a system that hears music and evaluates it, predicated on sophisticated audio- and language-processing techniques. Meyer sees musical affect as one case of a broader system of human affect, Whitman as a data science problem. Yet both authors see the process of musical meaning making as one in which the signal acts upon the listener (machine or human). Considered in this light, both authors agree on the signal as the primary source of musical significance.

Whitman's whole claim is that Meyer (and, it bears repeating, the entire discipline he stands for) fails to take contextual information into his account of musical meaning. But the truth is that Meyer does address it. Throughout his work, he is

29 The famous comparison from Meyer is that of the cigarette smoker whose emotions are piqued when he, craving a smoke, reaches into his pocket to find that he's out of cigarettes. Music, according to Meyer, triggers a similar affective response via a similar physiological mechanism (Meyer 1956, 14).

perfectly aware of the role that extra-signal information can play in the excitement of affect and construction of meaning. It is just that he regards this kind of information as outside his purview:

> We have found that the subjective data available, taken by themselves, provide no definite and unequivocal information about the musical stimulus, the affective response, or the relation between them (Meyer 1956, 12).

Elsewhere, he states this even more directly:

> Listeners and the objective data gathered from the observation of behavior and the study of the physiological responses to musical stimuli did not yield reliable information about the musical stimulus or the affective responses made to it (Meyer 1956, 22).

By "subjective data" (and, in a terminologically confusing choice, "listeners and the objective data gathered from [them]"), I take Meyer to be referring to listeners' reported affective states – the empirical responses of actual people reporting actual experiences to music. Thus, Meyer is here referring to more or less the concepts that Whitman terms "context" and "cultural metadata." For Meyer, this kind of "context" cannot tell us anything about the nature of the affective response itself, which is the essential substrate of musical meaning itself. This data is relevant to a conversation about musical meaning only in light of a *general* theory of affect, which is what Meyer hopes to explicate:

> This difficulty can be resolved only if the subjective data available... can be examined, sifted and studied in light of a general hypothesis as to the nature of affective experience and the process by which musical stimuli might arouse such experience. (Meyer 1956, 12)

First, Meyer says, you should postulate a general hypothesis about how meaning and affect arise. Then, and only then, can Whitman's "cultural metadata" figure meaningfully into a discussion of musical "meaning." Whitman is wrong that Meyer ignores human reaction because it's too difficult to integrate at scale. He just regards it as unimportant in a serious discussion of musical meaning. For Meyer, this discussion properly seeks to answer, "how does music work?" – not just "how has music worked for many people, and how best to use that information to synthesize future human reactions?"

In a part of Meyer's book that Whitman seems to have ignored altogether, this allows Meyer to imagine listening situations where context and conditioning do in fact play a large role in the construction of musical meaning. In this regard Meyer leaves much more space for extra-signal information than Whitman gives him credit for:

> Often music arouses affect through the mediation of conscious connotation or unconscious image process. A sight, a sound, or a fragrance evokes half-forgotten thoughts...These imaginings...are the stimuli to which the affective response is really made. In short, music may give rise to images and trains of thought which, because of their relation to the inner life of the particular individual, may eventually culminate in affect (Meyer 1956, 256).

He goes on to say:

> Neither the form nor the referential content of such experiences, however affective they may be, have any necessary relationship to the form and content of the musical work which presumably activated them. The real stimulus is not the progressive unfolding of the musical structure but the subjective content of the listener's mind. Yet...it seems probable that conscious or unconscious image processes play a role of great importance in the musical affective experience of many listeners (Meyer 1956, 258).

Note that Meyer here accepts the idea that "the real stimulus" *can* be something other than the signal itself. This is exactly the intuition animating the whole of Whitman's project, and it is one that he opposes, erroneously in my view, to Meyer's nominally "absolutist" paradigm. Again, it's not that Meyer ignores this fact of musical perception, but simply that he regards it as off-topic for an essay on musical meaning.

Whitman has created, essentially, a system for processing audio. It is one that is informed in creative ways by empirical human affective responses, but it is still a system for processing audio – that is, a system that grants the signal a kind of primacy. A signal goes in, a classifier does its work, and an output of some kind comes out. The nature of these outputs has certainly changed over the years, but the fundamental architecture of the system (audio in, evaluation out) is most likely the same. And insofar as that fundamental architecture remains in place, Whitman has gained no philosophical distance from Meyer, who also addresses the question of how a signal operates on a person. Meyer offers a psychological account rather than a data-driven one, but the philosophical approach to sound is pretty much the same. Whitman is correct that his approach, incorporating real human responses, is different from MIR techniques that derive from the audio signal alone. The intellectual intervention and technical innovation are legitimate (and, to judge from the success of the Echo Nest, practically effective); nevertheless, it would be wrong to locate Meyer and Whitman at opposite ends of the music-philosophical spectrum.

The real difference between the two authors, of course, is that Meyer is trying to understand how people relate to music and Whitman is building a machine that emulates how people relate to music. The machine's listening experience is qualitatively different from the human one; it is impossible for a machine to have the experience of the "subjective content of the listener's mind," to have the music call to mind a long forgotten experience which triggers a cascade of memories and affective states, or to experience listening in the company of friends. The machine "listens" in silence, in isolation, and without any subjective experience; in addressing itself to this scenario, there is a sense in which Whitman's system is infinitely more "absolutist" than Meyer's.

Although both these authors use the word "meaning," they are for the most part not on the same topic. Moreover, where their topics do overlap, they basically agree (they're equally "absolutist"). Whitman is not wrong that Meyer needs musical meaning to depend on the "signal," or, as Meyer calls it, the "stimulus." That is indeed the relationship under investigation for Meyer. Where Whitman is wrong is in claiming that this is not true of his own notion of musical meaning. For all his talk of musical meaning, on the mysterious relationship between signal and response Whitman is basically silent – and therefore gains no philosophical distance from Meyer. He simply writes about a different subject, namely, how best to simulate and synthesize that response. The essential, causal relationship between signal and response – the only question Meyer really targets, and a problem for countless other thinkers besides Meyer – is at once implicitly taken for granted and totally ignored in Whitman's project.

The Meaning of Meaning

What, then, is the "theory of musical meaning" employed by Spotify? Above I have sketched part of the answer: that music's meaning is functional rather than intrinsic, and that the mysterious ways in which music causes people to feel things – whatever they are (and Whitman definitely doesn't try to answer that) – will necessarily appear in a meaningful way *somewhere*, provided we gather enough data and treat it responsibly enough. In short, the "theory" of musical meaning is nothing more than the assumptions grounding the fields of machine learning and pattern recognition in general. As Zarsky puts it, the assumption is "that human conduct is consistent and that with sufficient data human behavior becomes predictable" (Zarsky 2016).

But is that really a "theory" at all? You might well answer "no," and you might be right. But what, then, do we make of Whitman's claim to have "learned the meaning of music?" And what do we make of Spotify's claim to be worth $10 a month? Are not both these claims grounded in the faith that Whitman and Spotify are at some level right about what musical meaning is? And is being basically *right* about musical meaning not *ipso facto* a kind of theorizing?

It is tempting to give Spotify a pass by declaring it a kind of "engineering" rather than "science." Very well, you might say, Spotify is wrong about meaning. So what? It's not a form of science, but just a collection of engineers trying to solve a problem and earn some money. But, as Pelillo et al (2015) argue, the era of machine learning has changed the way we should think about this traditional distinction:

> The scientist's occupation is seen today more modestly as a kind of problem-solving activity not dissimilar conceptually to that of the engineer, whereas on the other hand the work of the engineer is thought to produce a form of knowledge which is on a par with that produced by the scientist (Pelillo 2015).

Whitman himself dislikes the idea that man and machine stand in opposition. In answer to a reporter's question about The Echo Nest's potential to homogenize listening habits, Whitman defiantly responded:

> You call it algorithms but it's a lot more than that. We are obviously doing a ton of computer stuff but it's all based on what people are saying and choosing and that stuff. We hate this stupid man versus machine dichotomy.[30]

If the man-vs-machine dichotomy is "stupid," it should follow that the programmatically derived "meaning" is not just an engineering expedient, but a true statement about how music works for people in the real world. Pattern recognition and machine learning, in other words, are places where the line between science and engineering is blurred. The Spotify recommendation engine – whatever it really is – is in fact as much a theory of musical meaning, an attempt to characterize the process that causes people to like music, as it is a product designed to keep us logged in and spending money. In this regard it is not different in kind from Meyer, but rather in its approach to its own theoretical commitments. And, as I have shown, upon close inspection there are interesting deficiencies there.

In other words, Spotify sidesteps the question that should matter to it most (what does music mean?), even as it postulates a cryptic kind of answer (and keeps that answer a secret from its subscribers). The theory is that if we collect enough data, musical meaning, in all its manifold varieties, will be discerned by the system; as for specifying the nature of musical meaning itself, Whitman cites a single source as representative of hundreds of years of investigation into that topic, gives it a cursory reading, and then shrugs his shoulders because, after all, the real task is software design, not philosophizing.

This is a pretty dramatic intellectual liberty to take, one that Whitman is allowed because of a peculiar type of privilege he enjoys: the privilege deriving from the prestige of the discipline of machine learning, from the slippage inherent in that discipline between science and engineering, and from the financial promise of the system he created. But this privilege does not mean that the philosophical question

30 Emily White, "The Echo Nest CTO Brian Whitman on Spotify Deal, Man Vs. Machine, Why Pandora 'Freaks' Him Out (Q&A)", Billboard interview in 2014, https://www.billboard.com/biz/articles/news/digital-and-mobile/5944950/the-echo-nest-cto-brian-whitman-on-spotify-deal-man-vs, accessed May 20 2019.

shouldn't matter to the software designer; the inevitable fact is that the system's viability does ultimately depend on the way it construes musical meaning. If the "meaning" in Spotify is not the one customers value, or if it has not been "learned" in a way we are ready to accept, the whole Spotify enterprise is called into question. If meaning is as contingent as Whitman maintains, maybe another system would work just as well. Maybe *any* other system would work as well. Maybe there is no coherent way to measure how well such systems work the first place. Spotify seems to work pretty well, but so might a system of random recommendations. Given the capriciousness of musical affection Whitman mentions so often, that is a real possibility. A close look at Spotify's treatment of the problem of musical meaning reveals that it remains as obstinate a problem as it has been throughout its long history in aesthetic philosophy, a history that remains relevant even though Whitman dispenses with it in a brief passage or two. It is a problem as thorny and intractable as the financial crisis confronting the music industry in the 21st century, another problem that Spotify hasn't solved.

List of References

Aguiar, Luis, and Joel Waldfogel. 2015. "Streaming reaches flood stage: does Spotify Stiumlate or Depress Music Sales?" National Bureau of Economic Research Working Paper 21653. http://www.nber.org/papers/w21653.

Angwin, Julia, Jeff Larson, Surya Mattu, and Lauren Kirchner. May 23, 2016. "Machine Bias: There's software used across the country to predict future criminals. And it's biased against blacks." *Pro Publica*. Accessed January 30, 2019. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

BBC News, "Newsbeat". 2018. Online Resource, https://www.bbc.com/news/newsbeat-43240886. Accessed Jan 27, 2019.

Eriksson, Maria, Rasmus Fleischer, Anna Johansson, Pelle Snickars, and Patrick Vonderau. 2019. *Spotify Teardown: Inside the Black Box of Streaming Music*. The MIT Press.

IFPI. 2018. (International Federation of the Phonographic Industry, state of the industry) https://gmr.ifpi.org/state-of-the-industry. Accessed Jan 27, 2019.

Jehan, Tristan. 2012. "Analyzer Documentation." http://docs.echonest.com.s3-website-us-east-1.amazonaws.com/_static/AnalyzeDocumentation.pdf

Kim, Tae. 2017. "David Einhorn: Value investing may be dead and Amazon and Tesla killed it", published at cnbc.com. https://www.cnbc.com/2017/10/24/david-einhorn-value-investing-may-be-dead-and-amazon-and-tesla-killed-it.html. Accessed May 21, 2019

Kot, Greg. 2009. *Ripped: How the Wired Generation Revolutionized Music*. New York: Scribner, a division of Simon / Schuster.

Lunden, Ingrid. 2014. "Spotify Acquired Music Tech Company The Echo Nest In A $100M Deal." https://techcrunch.com/2014/03/07/spotify-echo-nest-100m/

Meyer, Leonard B. 1956. *Emotion and Meaning In Music*. Chicago: University of Chicago

Press.
Pelillo, Marcello, Teresa Scantamburlo, and Viola Schiaffonati. 2015. "Pattern recognition between science and engineering: A red herring?" Philosophical Aspects of Pattern Recognition, Pattern Recognition, Letters 64:3 –10.issn: 0167-8655. doi:https://doi.org/10.1016/j.patrec.2015.06.030. http://www.sciencedirect.com/science/article/pii/S0167865515002068.
Peoples, Glenn. 2015. "How 'Playola' Is Infiltrating Streaming Services: Pay for Play Is 'Definitely Happening'." https://www.billboard.com/articles/business/6670475/playola-promotion-streaming-services. Published at billboard.com. Accessed 01-09-2019.
Powles, Julia, and Helena Nissenbaum.n.d. "The Seductive Diversion of 'Solving' Bias in Artificial Intelligence." https://medium.com/s/story/the-seductive-diversion-of-solving-bias-in-artificial-intelligence-890df5e5ef53.
Putnam, Hillary. 1975. "The Meaning of 'Meaning'." *Minnesota Studies in the Philosophy of Science* 7:131–193.
Richter, Felix. 2018. "Spotify's Losses Widen as Royalty Costs Pile Up." Accessed March 20, 2019. https://www.statista.com/chart/4894/spotify-revenue-vs-costs/
Sanchez, Daniel. 2018. "What Streaming Music Services Pay (Updated for 2018)." Accessed February 5, 2019. https://www.digitalmusicnews.com/2018/01/16/streaming-music-services-pay-2018/.
Schwab, Pierre-Nicolas. 2018. "Workshop on fairness and ethics in recommendation systems." Accessed September 18, 2018. https://www.intotheminds.com/blog/en/workshop-on-fairness-and-ethics-in-recommendation-systems
Seaver, Nick. 2013. "Knowing Algorithms." Unpublished draft available on nickseaver.com. https://static1.squarespace.com/static/55eb004ee4b0518639d59d9b/t/55ece1bfe4b030b2e8302e1e/1441587647177/seaverMiT8.pdf.
–, 2018. "Captivating Algorithms: Recommender Systems as Traps." *Journal of Material Culture.*
Serafine, Mary Louise. 1980. "Against Music as Communication: Implications for Education." Journal of Aesthetic Education 14 (4): 85–96. issn:00218510, 15437809.
Silva, Marsha. n.d. "Spotify, the "Solution to Music Piracy," Is Getting Pirated by 2 Million People." Accessed February 4, 2019. https://www.digitalmusicnews.com/2018/03/26/spotify-piracy-hacked/.
Spotify 2019 home page. https://www.spotify.com/us/about-us/contact/.
Spotify 2006 home page, accessed Jan 30, 2019.. 2006. "Spotify Home Page, 2006." Accessed via the Internet Archive. Accessed April 17, 2019. https://web.archive.org/web/20061127231638/http://www.spotify.com/.
Tschmuck, Peter. 2017. *The Economics of Music*. Newcastle upon Tyne: Agenda Publishing.
Ugwu, Reggie. n.d. "Inside the Playlist Factory: The Unsung Heroes of the Music Streaming Boom." Published by Buzzfeed Media. https://www.buzzfeed.com/reggieugwu/the-unsung-heroes-of-the-music-streaming-boom, accessed 5-21-2019
Whitman, Brian. "How Music Recommendation Works – and doesn't work." Blog post: https://notes.variogr.am/2012/12/11/how-music-recommendation-works-and-doesnt-work/. Accessed May 5, 2019
Yoo, Noah. 2019. "Apple Music Now Has More Paid US Subscribers than Spotify: Report," Pitchfork. Online. https://pitchfork.com/news/apple-music-now-has-more-paid-

us-subscribers-than-spotify-report/. Accessed May 20, 2019
Zuboff, Shoshana. 2019. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. Public Affairs.

WHAT DOES MUSIC MEAN TO SPOTIFY? AN ESSAY ON MUSICAL SIGNIFICANCE IN THE ERA OF DIGITAL CURATION.
(Summary)

This article takes it for granted that Spotify's automated recommendation engine necessarily embeds assumptions about what is musically meaningful. Given Spotify's prominence in the 21st century music industry, the contours of that theory will have definite consequences for music culture in the digital era. This article seeks to probe the latent "theory" of musical meaning underlying Spotify's recommendation technology, proceeding in three ways: first, by narrating Spotify's transition from a streaming service to primarily a "discovery" service (the so-called "curatorial turn"). Second, by making a case for why it is useful to read Spotify against the academic dissertation of a software engineer whose company it would eventually acquire (Brian Whitman's 2005 "Learning the Meaning of Music"). Third, by performing a close reading of the Spotify graphical user interface (GUI) and the Whitman dissertation, attending to the assumptions about musical meaning embedded in both. The GUI and the dissertation turn out to go well together; both seem to see musical meaning as "relational," that is, as residing in music's relationship to things outside the audio signal itself. Nevertheless there are interesting argumentative gray areas in the dissertation on the issue of musical meaning, construed as a topic in aesthetic philosophy. By examining those gray areas, this article lays the theoretical groundwork for a quantitatively derived critique of automated music curation in the future.

Article received: April 24, 2019
Article accepted: May 15, 2019
Original scientific paper

UDC 784.9:004.934

Adriana Sabo*
Faculty of Music, University of Arts
Department of Musicology
Belgrade, Serbia

HATSUNE MIKU: WHOSE VOICE, WHOSE BODY?[1]

Abstract: This paper focuses on certain aspects of the Hatsune Miku phenomenon, a highly popular Vocaloid character from Japan. Hatsune Miku began her "life" as a software for vocal synthesis released by Crypton Future Media Inc., and has, since her first "birthday", become a virtual pop star. Despite being a fictional character, Miku takes on many of the traits a human has. She exists in a realm between human and artificial, mass media and personal space, between real and fantastic. This paper will discuss some basic information about Hatsune Miku and her large fan base, the issues of gender performativity and materiality of the body and voice, as well as the euphoric response of fans known as *moe*.

Keywords: Hatsune Miku, Vocaloid, *moe*, gender performativity, voice synthesis, materiality

Introduction

Hatsune Miku is, to put it vaguely, many things. She is a program for singing voice synthesis, developed in 2007 by Crypton Future Media using Yamaha's Vocaloid software. Hatsune Miku is a Japanese virtual idol (Guga 2015), a pop star whose official image was created by manga artist Kei (Hasse Jørgensen, Vitting-Seerup and Wallevik 2017), but also the result of the imagination and affective labor (Hasse Jørgensen, Vitting-Seerup, & Wallevik, 2017, 9) of millions of fans, an object of desire, a simulacrum. She is, essentially, a phenomenon that allows us to view the importance of technology and media in our contemporary societies. Also, she is a

*Author's contact information: adrianasabo259@gmail.com

1 This paper was written as part of a project of the Department of Musicology of the Faculty of Music, University of Arts in Belgrade, entitled "Identities of Serbian Music in the World Cultural Context", supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia, no. 1770149.

"she" in as much as her visual identity and her voice corresponds to certain gender norms we perceive as female or feminine. Given the fact that she is "not real" – as in, she is not a biological organism – there is no material body that *is* Hatsune Miku, despite this she is immensely popular, especially in Japan, so naturally the Vocaloid attracted the attention of academics. A number of academic studies were dedicated to, among other things, viewing her as a "body without organs" according to the philosophy of Deleuze and Guattari (Annett 2015; Guga 2015), focusing on the importance of fans and the audiences in creating Miku (Hasse Jørgensen, Vitting-Seerup and Wallevik 2017; Le 2013), and on her many virtual qualities (Lam 2016).

In the text that follows I intend to focus on three aspects of this phenomenon that are apparent, but whose connection and intersection are worth examining: gender, music and voice. First and foremost, Hatsune Miku is a voice, or rather, she is a Vocaloid, whose developers "digitized separable and transportable fragments of recorded voice and integrated them into a singer library" (Lam 2016, 1109–1110). She is also a software for making music and an animated female character popular with hundreds of thousands of fans. The fictional character that is Hatsune Miku does not exist if she doesn't sing, there is no Hatsune Miku without the voice. In other words, she is represented with a body, although it is not a biological body, and she has a voice that comes from *another place*, being a "nonorganic embodiment of an organic subjectivity" (Lam 2016, 1114). What we see and hear when we experience Hatsune Miku is a performance of gender – as Judith Butler understands it – and voice is an important part of that performance, happening without any explicit biological body shaping the performance and, yet, simultaneously being shaped by it. In this sense, I would like to shed light on who (or what?) is shaping her gender performance – having in mind the context of the Japanese music industry, the popularity of Vocaloids, and admiration for fictional characters in fan culture and so on. This begs the question of whose gender and what kind of gender is being performed by Hatsune Miku?

What/Who is Hatsune Miku: The Facts

The "life" of Hatsune Miku began in 2007 when she was officially released by Crypton Future Media. The name, roughly translated to English means "the first sound of the future," and it refers to a Vocaloid software voicebank that offers users the possibility of creating music for a singing voice without requiring access to a "live" singer. Miku uses Yamaha Corporation's Vocaloid 2, Vocaloid 3, and Vocaloid 4 singing synthesizing technologies, as well as Crypton Future Media's Piapro Studio, and a singing synthesizer VSTi Plugin. Other than being a voice synthesis software, she is also a fantasy character that has become something of a trademark for the software itself.

Figure 1: Official Hatsune Miku Character

Her voice was created based on the voice of Japanese actress Saki Fujita (藤田 咲), so it originated from a "live" person. The voice samples recorded by Fujita underwent substantial modifications in order to be turned into samples which can, in turn, be used to create the singing voice of Hatsune Miku:

> To develop a human-like singing voice for Miku, the Vocaloid developers digitized separable and transportable fragments of recorded voice and integrated them into a singer library. In the transmission of these fragments to Miku's voice, they were melodically modified and integrated with Miku's bodily expressions. Human voices are decomposed into binary codes and reassembled into the vocals of Hatsune Miku to mimic the naturalness of real human voices. The developers of the singing synthesis technology aim to generate the singing voice that imitates a human singer and to provide what human ones cannot do (Lam 2016, 1109–1110).

Describing the very begning of a Vocaloid's life, Daniel Black writes:

> Once extracted from the human body, this raw material can then be used to create entirely new performances independently of the body from which it was originally collected. The words and expression are provided by the owner of the computer software, who turns the generic, mass-produced raw material of the voice to whatever end he or she desires (quoted after Young 2015, 77).

After this initial modification, each software user has the opportunity to further change and adapt the voice bank sample to fit their own needs and goals as well as to match their own vision of Miku. "Basic parameters such as pitch and note length can be edited, while users can select different vocal colourings such as 'soft', 'vivid', 'sweet', 'light', and 'dark' to match the genre or mood of the song" (Prior 2018, 500). It is noteworthy that the appearance of another software that – like the Vocaloid – offers the possibility of synthesizing the singing voice, named the UTAU, was first released in March 2008 by Ameya/Ayame (Le 2013, 4). UTAU helped promote Miku further, as UTAU is free to use. The availability of softwares, as well as the fact that Miku was licensed under the Creative Commons licence, resulted in tens of thousands of songs produced and uploaded to the official Crypton Future Media website, as well as to other video sharing platforms.

Not just a singing voice synthesizer, Hatsune Miku is also a virtual character, with specific looks, movements, and gestures that are synchronized to the music. The practice of "adding" a visual image to a Vocaloid is not a new concept in Japanese culture,[2] but it seems that none of the other characters created reached the level of popularity Hatsune Miku has achieved. Other than being a voice synthesis software, Hatsune Miku is also the result of imagination brought to life thanks to 3-D animation software like Miku Miku Dance, developed by Yu Higuchi in 2008 (Le 2013, 4), which enabled users to create videos that would accompany their music, featuring Miku. "Recognizing her advertising potential, Crypton Future Media created a record label to gather and promote consumer-generated Miku tunes," which furthered her "reach," making her ever more popular (Guga 2015, 37). Her official look is that of a skinny 16-year-old girl, with big eyes, turquoise hair tied up in long pigtails, long legs, a school uniform, and a number of gadgets and accessories on her arms and head. Her figure and poses are eroticized, even more than most "biological" pop stars are.

Another important aspect of the phenomenon of Hatsune Miku is the audience, or rather, the fan base. Given the fact that she "comes to life" thanks to the use of softwares, theoretically anyone can create, recreate, and reimagine Miku. So, she is as much a product of the fans as she is a product of Crypton Future Media, and her popularity is tightly linked to the very strong culture of user-created content in

2 For example, there is an entire wikia database dedicated only to Vocaloids. See https://vocaloid.fandom.com/wiki/Vocaloid_Wiki.

Japan. Linh Le links her success to the so called *doujin* culture:

> The word "doujin" can simply be defined as self-published works that can be either original or derivative works in fields including literature, comics, software, and music (…) For Japanese pop culture, specifically for the manga and anime industries, doujin activities constitute an integral part.

In other words, collaborative works of fan communities sharing, appropriating, changing, and expanding on the artistic output of others, represents one feature of popular culture in Japan. It is no surprise, then, that the rising popularity of Hatsune Miku is often attributed to her "appearance" on the website niconico (formerly know as Nico Nico Douga / ニコニコ動画 *Niko Niko Dōga*/), dedicated to sharing videos, much like the popular western platform YouTube. Slightly different than YouTube, niconico is more like a social media platform, as it enables users to create videos in response to previously uploaded ones, create playlists based on different rankings, add comments to existing videos, change lyrics, add remixes to existing videos, and so on (Lam 2016, 1108). That is, it enables the existence of a large network of interconnected content, providing grounds for the flourishing of a "highly participatory cyber-culture" (Lam 2016, 1109). Thousands upon thousands of fans have, thus, been able to upload their own songs composed thanks to the Hatsune Miku software, accompanied by animated videos of Miku dancing, singing, and/or moving sensually to the rhythm of her music. The ever-growing community enabled the rise of a number of now-famous producers, composers, and song makers who became famous thanks to their creative use of the Hatsune Miku software. It would seem that "you cannot become rich by composing a song featuring Miku, but the brand of Hatsune Miku can make you famous enough for people to hire you" (Hasse Jørgensen, Vitting-Seerup and Wallevik 2017, 11). The importance of fan culture for Hatsune Miku will be further discussed later in the text, as we must consider their euphoric response to her, know in Japan as *moe*. For now, it is important to emphasize the fact that the phenomenon of Hatsune Miku blurred the line between producers and consumers, enabling the audience to create the very thing they intend to consume. And, "as a reward, participants are granted with a deeply personalized, satisfactory experience of disseminating parts of themselves through these artifacts they took part in making" (Guga 2015, 40).

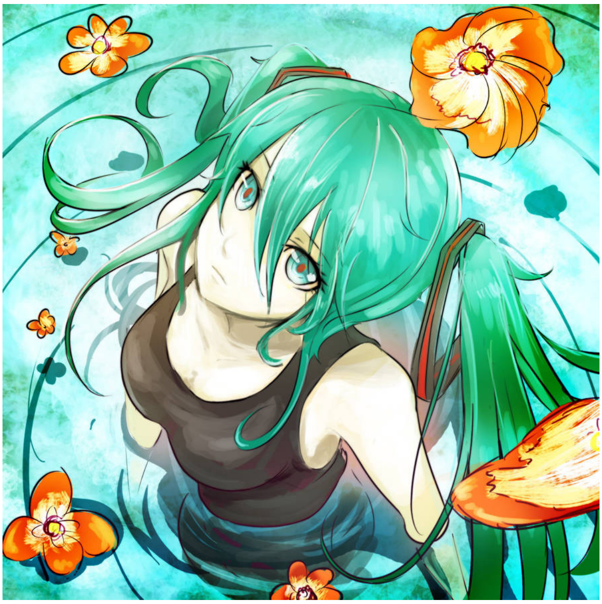


Figure 2: Hatsune Miku fanart: *End of the day 2* by beanbeancurd (available on deviantart.com)

Theory: Gender Performance, Voice, *Moe*

In the previous section, I offered some key facts about Hatsune Miku that enable us to better understand this phenomenon, as well as her immense popularity among the Japanese people. From the very beginning, I have referred to Miku as "she", implying possibly that she is, in fact, a person, a being belonging to the female gender. Yet, at the same time, it is obvious that Hatsune Miku is a computer program – albeit a very complicated, multi-layered one. Nevertheless, she is a phenomenon (which is term used in lieu of a more precise one) that is artificial and, in a sense, "non-material", for example, she can't be touched or hugged as most living creatures can be. To use terminology adapted from Judith Butler, Hatsune Miku is an artificial creation that reiterates certain gender norms that make her recognizable as a female.

In her book *Bodies that Matter. On the discursive limits of "sex"* (Butler 2011), Judith Butler further explains her understanding of gender as a performative practice. "Performativity", she writes, "must be understood not as a singular or deliberate 'act,' but, rather, as the reiterative and citational practice by which discourse produces the effects that it names" (Butler 2011, xii). She goes on to add that performativity "is not (…) the act by which a subject brings into being what she/he names, but, rather, (…) that reiterative power of discourse to produce the phenomena that it regulates and constrains" (Butler 2011, xii). In other words, Butler claims that gendered subjects are being created "within the framework of different norms; in acts of gender performativity, individuals repeat, change or 'decline' to accept those norms, and by doing so they become subjects" (Sabo 2015, 2). The collection of norms represent ideals – ideal bodies, ideal persons – that are not necessarily attainable and they themselves often change over time, but, they still represent an imagined ideal that we are "judged" against, and that governs the recognition of individuals as men or women (or rather, "normal" men and women) based on the degree to which we accept and cite the norms. Some of these (stereotypical) norms, when it comes to female gender for example, are "substantial," like heterosexuality or motherhood, some are "visual"– slender figure, long hair, big breasts, high heels, makeup – some are biological – having a uterus, ovaries, and so on – and some are related to personality traits – emotional, prone to tears, loves shopping, loves children, has natural caring abilities, and so on. Butler also emphasizes that gender performativity is not a conscious act, one cannot simply choose which gender they will be – it is just the repetition of norms that isn't (always) conscious, but which push us to accept some norms as "natural," almost as a "physical part" of our beings. One other important feature of gender performativity (elaborated in numerous publications, starting with *Gender Trouble*) is that it "happens" for others and that it is important for others to *recognize* us and accept us. It is also often defined "to the extent that one is not the other gender" (Butler 1999, 30). In other words, *the others* represent an important part of gender performativity. Butler further emphasizes the difference between concepts of gender *performance* and gender *performativity*: "the former presumes a subject, but the latter contests the very notion of the subject" (Butler 1996, 112). That is, performance implies that there is a subject – a person, an individual – following a script, so to speak, consciously repeating and interpreting it. On the other hand, performativity sheds light on a kind of "circulation" in which the subject is created by the norm it accepts and repeats, reaffirming that very norm time and again. The title of Butler's book, *Bodies that Matter*, introduces a kind of word play, derived from the two meanings of the word "matter": bodies that matter, as in, bodies that are important, and bodies that (are) matter, which implies that she focuses on the very materiality of bodies that are "produced" through different performative acts. In other words, she explores the way physical bodies, and more precisely, "sex" is constructed through performativity, wondering "how and why 'materiality' has become a sign of irreducibility, that is, how is it that the materiality of sex is understood as that which only bears cultural constructions and, therefore, cannot be a construction?" (Butler 2011, 4). In other words, she aims to question the "unquestionable" status of one's sex as a bodily trait that is "given by nature" and in a way, precedes gender. It exists "before" society and almost independently of it. In this sense, Hatsune Miku can be understood as a very good example of the way a body and its materiality are "artificially" constructed in the context of a capitalist, consumerist, technologically advanced society.[3] It has been noted many times that our bodies are influenced, formed, and of course exploited by capitalism. When

3 The interest in the ways technology changes out bodies is not in any way new and has been the cornerstone of a number of theories regarding cyborgs, post-human condition, virtual reality and so on.

Butler speaks of the way bodies materialize through acts of norm repetition, she is speaking about how biology and anatomy are shaped by society and culture. Since Miku has no material body, this process is rendered more visible, as her unique materiality is, quite literally, created by members of society. And again, she only exists as a product for consumption, as a "thing" created to be sold and bought and what we see as Hatsune Miku is a direct result of desires, ideas and ideologies of her fans, as well as the ideologies of the corporation that created her official look. And that look shows many negative, female gender stereotypes that speak volumes to how traditional and conservative values are easily incorporated with technological advancements.

In formulating her own view of performativity, Butler leans on the concept of speech acts, formulated by Austin, and later expanded on by many authors, such as Searle, Fish, Derrida etc. Yet, as Annette Schlichter notes, despite the fact that she focuses much of her attention on language – she even used the example of Aretha Franklin's performance of Carole King's "You Make Me Feel (like a Natural Woman)" to explain how gender is defined through "the restriction of gender within that binary pair" (Butler 1999, 30) – she completely omits the examination of the importance of voice for gender performativity, formulating gender performativity as a mainly visual concept. As Schlichter notes, Butler's "theory of gender performativity and the consecutive deliberations about the matter of bodies do not account for voice as sound, nor do they acknowledge the mediation of vocal acts through sound technologies" (Schlichter 2011, 32). In her attempt to expand Butler's theory to allow for the inclusion of voice, Schlichter emphasizes the fact that voice "marks a passage from the inside of bodies to the exterior, and its materiality is rather delicate, even paradoxical" (Schlichter 2011, 33), agreeing with Jacques Lacan and Mladen Dolar in claiming that the voice is "an object that emerges from the body but is neither fully defined by matter nor completely beyond it" (Schlichter 2011, 33). As Miriama Young claims that, "defining the voice is a slippery project, and one that requires a circumnavigational approach – we may only ever speak around the voice in order to get at its essence" (Young 2015, 1). Voice is, in a way, a very intimate "product" of the body, it "attracts the listener to its materiality," offering different "levels of intimacy and immediacy" (Young 2015, 2). This intimacy is in many ways complicated by the voice being mediated – whether the voice itself is changed through use of technology or whether it's transmitted to the listener via a medium like telephone, radio, TV etc. – extending the "elaborate mechanism of the human voice (…) through interaction with electronic technology" (Young 2015, 5). As will be shown later in the text, the matter of voice, gender performativity, and materiality get further complicated when applied to Hatsune Miku, given that she sings, dances, and performs without actually having a material body. The fact that a fictional character performs gender – as well as music – isn't necessarily revolutionary or strange, especially in the 21st century. What is worth noting, though, is the way that Miku's fans see her and the ways in which – within the context of the Japanese capitalist, technologically advanced society – she is treated both as a "real" human and an artificial entity from the realm of fantasy.

Japanese popular culture is booming with imaginary, animated characters that have a very unique and special place in lives of the Japanese people.[4] As was previously discussed, the fan base – people using the software, making videos, sharing, commenting, and changing them – greatly contributed to the rise of Hatsune Miku. Her immense popularity can also be attributed to the fact that "Vocaloids (…) bec(a) me part of a system of iconic anime–style characters known as *kyara*: image–beings that fans both idolize and consume" (Annett 2015, 164). In other words, she is part of a larger culture whose "members", in many different ways, consume and create their own imaginary, animated idols. Another feature of this culture that has attracted a lot of attention from psychologists as well as theoreticians, is the euphoric, almost hysterical response to a fantasy character, and the peculiar relationship people have with them, that is at the same time highly sexual, but also distant or detached.

This response is called *moe*, which is a "neologism used to describe a euphoric response to fantasy characters or representations of them…often associated with a young, media-savvy generation of *otaku*, or hardcore fans of anime, manga and videogames" as well as being "used by *fujoshi*, zealous female fans of *yaoi*, a genre of manga featuring male homosexual romance" (Galbraith 2009). Furthermore, *moe* is:

> primarily based on two-dimensional images but can also include objects that index fantasy or even people reduced to '*moe* characters' and approached as fantasy. Both *otaku* and *fujoshi* access *moe* in what they refer to as 'pure fantasy' (*junsui na fantajii*), or characters and relationships removed from context, emptied of depth and positioned outside reality (italics in original text) (Galbraith 2009).

In other words, it is precisely the fact that a character is "not real", that it is separated from any context and is completely fictional, that triggers the euphoric response in fans. This concept is primarily connected to the contemporary, technologically advanced yet strongly traditional Japanese society, in which people are mostly alienated from each other and in which human contact has become difficult to obtain. Honda Touru sees *moe* as a result of the fact that "in Japan today fulfillment as a human being can only be found inside one's own brain as a reaction to fantasy characters" (quoted after Galbraith 2009). In that sense, "a relationship with a mediated character or material representations of it is preferable to an interpersonal relationship" (Galbraith 2009). Another aspect of this "escape from reality" is that it challenges highly traditional gender norms for both men – with the imperative of working and earning money – and women – burdened with domestic work

4 This is also evident, for example, in the fact that most cities/towns in Japan have their own, animated mascot that represent them, as well as in communities of manga and anime lovers, different dating games that feature animated characters as well as the popularity of other Vocaloids and so on (Galbraith 2009).

and childbirth. "*Moe* allows men to stop performing socially sanctioned masculinity and indulge femininity, which can be very soothing (*iyasareru*)" (Galbraith 2009). Another aspect of the male response to fictional characters is the fact that the characters are highly eroticized and sexualized teenage (or even pre-teen) girls. Thus, these images are highly problematic as they, essentially, introduce young girls into the realm of sexual fantasy. Yet some psychologists view this kind of sexuality as "a sexuality deliberately separated from everyday life" (quote after Galbraith 2009), arguing that it "depends on 'fantasy contexts' (*kyokou no kontekusuto*), or (…) the 'reality of *kyara*'" (Galbraith 2009). One possible reason for this attraction to girls is their innocence, the fact that they "do not know the world" and are thus "fetishized as pure" (Galbraith 2009). In other words, there is a strong obsession with youth, which is why the school uniform as "the fetishized signifier of innocent status and character" is a necessary accessory.[5] The sexualization of girls is a manifestation of the desire to own or acquire youth in a typically masculine way: by creating and "owning" girls in a sexual way, men are given the opportunity to reach for the innocence and purity they so desire.

Figure 3: Miku by Bayeuxman (available on deviantart.com)

In conclusion, when trying to understand the phenomenon of Hatsune Miku,

5 This obsession with youth, as well as the need to detach oneself from reality has also resulted in the popularity of a character referred to as "little sister" in uniform, which enables the consumer to return to days of innocence and the time in which they could have – but didn't – enjoyed free love. This character, however "does not equate to actual incestuous desire", but can be understood "as first a longing for a time of youthful possibilities and hope (signified by the uniform) and second a desire for an uncompromising relationship not conditioned by society (the little sister)" (Galbraith 2009). Another interesting way to trigger *moe*, is to turn "cats, war machines, household appliances and even men of historical significance into beautiful little girls" (Galbraith 2009).

one can't lose sight of the fact that her popularity is also the result of the budding and normalized "attraction" of Japanese men to fictional, two-dimensional, "pure" characters who are detached from context but are, despite the crucial role that their "fakeness" plays, envisioned as "real," highly sexualized girls in school uniforms with distinctly human features. I would also argue that it is precisely this "detachment from reality" that defines not just the visual image of Hatsune Miku, but her voice and music as well. Namely, certain taboos and restrictions that are imposed on human relationships – due to the fact that human beings have emotions, can be hurt, and can be physically harmed in a number of ways – are not in play when one can design their own fantasy. In the following section, we will explore the idea that what Miku is and what her music and voice are, is thoroughly defined by the need to emphasize her artificial character, that is, to render very explicitly that *she is not a human being*.

Hatsune Miku's Body, Voice, and Music

All of the aforementioned qualities of Hatsune Miku become highly evident in situations when she is "removed" from the computer screens in her fans' homes, and put on stage in a live concert, accompanied by the "Magical Mirai" band. "The Magical Mirai (マジカルミライ), is a band that plays during live concert events featuring VOCALOID vocalists from the company Crypton Future Media, Inc." (Vocaloid fandom n.d.). There is an event that takes place around the birthday of Hatsune Miku. During these concerts, a playlist is selected from the most popular songs – some of them created by official Miku producers and songwriters – and the audience has a chance to experience Hatsune Miku "live," much like they would with a pop music diva made of flesh and blood. This is also an instance where, in the more traditional context of a rock concert, we can witness Miku's gender performance, and how it is tightly connected to her musical performance.

What we see on stage is an animated figure of Miku – initially she was shown on a screen and eventually she became a holographic image – in her school uniform enhanced by a futuristic-looking silver shirt and boots, moving and dancing to the rhythm of music and replicating the movements of human pop stars – running left and right on the stage, jumping, dancing, cupping her ear to hear the audience sing and overall, seemingly enjoying herself and "basking" in the love of her fans. Her image is utterly artificial and anatomically "impossible": for example, she has a tiny waist, very long legs and hands, and huge eyes. Her movements are also perfectly timed to the rhythm of music. In other words, it is obvious that she is not a representation of a "real" woman, she is a character that is obviously from the realm of fantasy. She repeats certain gender norms that can be understood as "female" but fails to conform to many others. While she has no physical body that is shaping gender performance and being shaped by it, there is a large group of bodies belonging to her creators and her fans, who decide how her gender performance will look. In other words, there is no biological woman who Hatsune Miku was "modeled" after, but her "subjectivity" was created according to other men's fantasies, to make her hyper-sexualized but at the same time two-dimensional and artificial. In this sense, Miku falls in line with other female characters that are often labeled as cyber-punk – one notable example in the west is Lara Croft, although she "exists" only in the realm of computer games (Flanagan 2002) – that are "primarily created and represented by men" (Flanagan 2002, 425).[6]

This delicate play of artificial, fantastic, and "natural" becomes apparent when listening to Miku sing. Specifically, it is clear that her voice isn't completely artificial, there is a real woman who offered the base for what later became an artificial voice. Her performances can be said to offer a "machine's rendering of a disembodied, often omnipresent, 'God-like' ideal" (Young 2015, 9). When it comes to Vocaloids, they are often understood as programs that enable the composer to have a voice that *sounds* human, but which also has a "mechanical" or "artificial" sound. So, in case of Hatsune Miku, it is her voice, just as much as her image, that provides the two-dimensional, detached feeling that excites her fans so much. Writing about Vocaloids, Miriama Young states that,

> these machine voices are designed as 'ideal' specimens, absent of bodily residue or the necessary signs of existence. They do not gulp, splutter, nor need to breathe. The voice is rendered in a virtuosic – albeit normalised – idealisation. The machine, then and now, enables the realm of dreams and imagination to become realised objects of desire (Young 2015, 77).

Despite the fact that it is derived from a human voice, Hatsune Miku's voice sounds completely artificial, and it is impossible to mistake her voice for that of a human singer. It is high-pitched and "sweet," as if the voice of a little girl was synthetically modified, she pronounces lyrics quickly and without taking a breath, the melody she sings is virtuosic, with many jumps and a very wide range. In other words, the voice is treated and used like any other electronic instrument – which is perfectly logical, given the fact that it is completely computer-generated. She is a simulacrum (as defined by Baudrillard), a copy without an original, and she exists mainly in relation to her fan's fantasies and imagination. Her voice, however, has a special form of materiality, given that its origin is not a human body, but a computer program. As Nick Prior puts it, her voice exists "in a different kind of materiality, one composed of diffuse digital bits spread exponentially through the circuits of corporate and peer-to-peer media" (Prior 2018, 501). This materiality is, thus, conditioned by technology, and the mediated intimacy it provides to the listeners. Technology offers the fans the possibility to have Miku all to themselves, and to adapt her voice to fit their own imagination and desires, so the materiality of her voice changes according to the carnal desires of others.

6 For more details on women's cyber-punk artifacts, see Flanagan 2002.

In other words, the entire phenomenon if Hatsune Miku – her appearance, the voice, the fan-made songs and videos, her live performances etc. – represents a kind of “assemblage” of different elements. As Nick Prior puts it,

> the terms ‘virtual’ and ‘simulated’ are only partially helpful here. ‘Assemblage’ is better at capturing how Miku’s voice is a convergent result of various human and non-human forces: silicon and carbon, corporate and grassroots, algorithmical and fleshy, local and infrastructural (Prior 2018, 502).

This multi-layered aspect of Hatsune Miku is also, interestingly enough, explored in a song titled “The End of Hatsune Miku” by cosMo that has “entered the Hall of Legend; with over five million views, it is cosMo's most popular work and one of Miku's most popular songs” (vocaloid.fandom.wiki n.d.). One of the song’s main features is the speed of the lyrics (around 240 BPM) which makes it virtually impossible for humans to sing, this, in turn, emphasizes Miku’s perfection as a virtual idol and her fantastical features. Lyrics of the song are about the uninstallation of Miku, as she *realizes* that she is a “mere imitation of humans” and that she does not exist unless she sings. She cries and looks at the sky, but she keeps singing (Vocaloid Lyrics Wiki n.d.). In other words, this is a song dedicated to something resembling an existential crisis of a computer program – or maybe the existential crisis of a fan identifying himself with Miku? – where she knows that she should feel something in the moment before being metaphorically killed by her creator, but she, in fact, doesn’t feel a thing (Vocaloid Lyrics Wiki n.d.). This song can also be said to perfectly describe how fans see Miku – she is viewed simultaneously as an artificial creation, but is at the same time approached from a very human, emotional, physical viewpoint, being a kind of artificial catalyst for (sexual) desire, love, euphoria, and excitement.

Conclusion

As was suggested throughout this text, Hatsune Miku can be understood as a symbol of contemporary (Japanese) societies, dominated and shaped by technology that is still not independent of human influence. She appears to be a virtual idol, a simulation, and a completely artificial creation, yet she is the result of the labor, imagination, and desires of biological, human organisms. Despite the fact that she herself does not have a biological “body”, there are “bodies that matter”, when it comes to Miku: bodies of the fans who create and develop her, the body and voice of the woman who initially recorded the samples, as well as the holographic or animated body of Hatsune Miku herself who, despite not being made of flesh and blood, still possesses a form of materiality that fans respond to. When it comes to the way her gender is performed, we can detect very strict, traditional norms being inscribed into her virtual body – she is the result of male desire, designed to produce moe in male fans. Her voice is sweet and tender, her body young and highly sexualized, and most of all, she is obviously a fantasy character, she is devoid of any context, she doesn't have her own "story" and is she is, in that sense, two dimensional. She exists between realms of the carnal – as a result of sexual desire, affect, strong emotion of others – and what seems to be its complete opposite – as an animated character that can never truly exist as a biological organism, as a phenomenon obviously "fake" and un-real, a result of technology. She is a subject that emerges through performative processes that "she" has no control over. Hatsune Miku can also be understood as a rather extreme and quite literal example of how social norms govern one's gender construction and performance, showing, in a radical way, that even the body, a biological entity, can be socially constructed. Lastly, Hatsune Miku is obviously the result of a highly traditional, patriarchal, conservative society which imposes strict gender norms on its people, a society whose every aspect is dominated by capitalism and consumerism and defined by the imperative of technological advancement.

List of References

Annett, Sandra. 2015. “What Can a Vocaloid Do? The Kyara as Body without Organs.” *Mechademia , Vol. 10* 163–177.

Butler, Judith. 2011. *Bodies that Matter. On the discursive limits of “sex”*. Taylor & Francis e-Library.

—. 1996. “Gender as Performance.” In *A Critical Sense: Interviews With Intellectuals*, by Peter Osborne (ed.), 109 –125. New York and London: Routledge.

—.1999. *Gender Trouble. Feminism and the Subversion of Identity.* New York, London: Routledge.

Flanagan, Mary. 2002. “Hyperbodies, hyperknowledge: women in games, women in cyberpunk, and strategies of resistance.” In *RELOAD: rethinking women + cyberculture*, by Austin Booth Mary Flanagan, 425–454. Cambridge, MA: The MIT Press.

Galbraith, Patrick W. 2009. “*Moe*. Exploring Virtual Potential in Post-Millennial Japan.” *Electronic Journal of Contemporary Japanese Studies.* Accessed April 11, 2019. http://www.japanesestudies.org.uk/articles/2009/Galbraith.html#About%20 the%20Author.

Guga, Jelena. 2015. “Virtual Idol Hatsune Miku: New Auratic Experience of the Performer as a Collaborative Platform.” Edited by Elif Ayiter, Onur Yazicigil Anthony Lewis Brooks. *ArtsIT 2014: Arts and Technology.* Cham: Springer. 36–44.

Hasse Jørgensen, Stina Marie, Sabrina Vitting-Seerup, and Katrine Wallevik. 2017. “Hatsune Miku: an uncertain image.” *Digital Creativity.*

Lam, Ka Yan. 2016. “The Hatsune Miku Phenomenon: More Than a Virtual J-Pop Diva.” *The Journal of Popular Culture Vol. 49, No. 5* 1107–1124.

Le, Linh K. 2013. “Examining the Rise of Hatsune Miku: The First International Virtual Idol.” *The UCI Undergraduate Research Journal* 1–12.

Prior, Nick. 2018. “On Vocal Assemblages: From Edison to Miku.” *ContemporaryMusic Review,* 37:5-6, 488–506.

Schlichter, Annette. 2011. “Do Voices Matter? Vocality, Materiality, Gender Performativity.”

In *Body & Society*, 17/1, 31–52.

Young, Miriama. 2015. *Singing the Body Electric. The Human Voice and Sound Technology.* Ashgate.

vocaloid.fandom.wiki n.d. *Magical Mirai (band).* Accessed April 13, 2019. https://vocaloid.fandom.com/wiki/Magical_Mirai_(band).

Vocaloid Lyrics Wiki. n.d. 初音ミクの消失 *-DEAD END- (Hatsune Miku no Shoushitsu*

-DEAD END-). Accessed April 14, 2019.

https://vocaloidlyrics.fandom.com/wiki/%E5%88%9D%E9%9F%B3%E3%83%9F%E3%82%AF%E3%81%AE%E6%B6%88%E5%A4%B1_-DEAD_END-_(Hatsune_Miku_no_Shoushitsu_-DEAD_END-).

vocaloid.fandom.wiki. n.d. 初音ミクの消失 *-DEAD END- (Hatsune Miku no Shoushitsu -DEAD END-).* Accessed April 14, 2019. https://vocaloid.fandom.com/wiki/%E5%88%9D%E9%9F%B3%E3%83%9F%E3%82%AF%E3%81%AE%E6%B6%88%E5%A4%B1_-DEAD_END-_(Hatsune_Miku_no_Shoushitsu_-DEAD_END-).

HATSUNE MIKU: WHOSE VOICE, WHOSE BODY? (Summary)

This article focuses on Hatsune Miku, a virtual idol and pop star from Japan. Miku was released in 2007 by Crypton Future Media Inc. as a voice synthesis software that uses Yamaha's Vocaloid technology. As is customary in Japan, the Vocaloid was also equipped with the fictional character of Hatsune Miku, a 16-year-old girl with a slender figure, tiny waist, long arms and legs, long turquoise hair in pigtails, and big blue-green eyes. The fictional character that is Hatsune Miku, does not exist if she doesn't sing, there is no Hatsune Miku without the voice. One feature that is very important for understanding this phenomenon, is the highly precipitative fan culture that enabled her popularity. Namely, despite the fact that her voice and image were created by Crypton Future Media Inc., her "life" continued thanks to thousands of fans who created music and videos with the software, and shared them via websites like niconico (https://www.nicovideo.jp/). Hatsune Miku is created by fans and is part of the flourishing *otaku* community of male fans who experience *moe*, a euphoric response, when seeing and hearing a fictional anime or manga character. In the *otaku* community, *moe* is achieved through the image of a fictional character that, to put is simply, has no personality, no story, and is presented as a highly sexualized and eroticized teenage girl, wearing the trademark school uniform. For a better understanding of Hatsune Miku, I turned to Judith Butler's ideas of gender performativity, especially given the fact that she, in her book *Bodies that Matter*, examines this phenomenon in connection to the materiality of the body. In case of Hatsune Miku, we cannot speak to a material body that was later turned into an animated character, yet there are different material bodies that define the performativity of her gender. For one, there is the ideal body of a schoolgirl that is fetishized by the male fans and created through an animation program, in addition to the body of the voice actress whose voice was recorded and sampled for modification in the Vocaloid software. Finally, despite the fact that Hatsune Miku doesn't have a body, her fans and creators do, and her gender and body are, in a way, a result of her fan's bodies and their own gender performativity that finds pleasure in creating and consuming fictional schoolgirls.

Article received: April 18, 2019
Article accepted: May 15, 2019
Original scientific paper

UDC 78.01:004.8

Gennady Stolyarov II*
Independent researcher
Carson City, Nevada, United States

EMPOWERING MUSICAL CREATION THROUGH MACHINES, ALGORITHMS, AND ARTIFICIAL INTELLIGENCE

Abstract: In this paper, I describe the development of my personal research on music that transcends the limitations of human ability. I begin with an exploration of my early thoughts regarding the meaning behind the creation of a musical composition according to the creator's intentions and how to philosophically conceptualize the creation of such music if one rejects the existence of abstract Platonic Forms. I then explore the transformation of my own creative process through the introduction of software capable of playing back music in exact accord with the inputs provided to it, while enabling the creation of music that remains intriguing to the human ear even though the performance of it may sometimes be beyond the ability of humans. Subsequently, I describe my forays into music generated by earlier algorithmic systems such as the *Musikalisches Würfelspiel* and narrow artificial-intelligence programs such as WolframTones and my development of variations upon artificially generated themes in essential collaboration with the systems that created them. I also discuss some of the high-profile, advanced examples of AI-human collaboration in musical creation during the contemporary era and raise possibilities for the continued role of humans in drawing out and integrating the best artificially generated musical ideas. I express the hope that the continued advancement of musical software, algorithms, and AI will amplify human creativity by narrowing and ultimately eliminating the gap between the creator's conception of a musical idea and its practical implementation.

Keywords: algorhitmic music, artificial intelligence, composition software, creativity, digital music, human-machine collaboration, musical dice games, musical technology, philosophy of music, transhumanism

* Author's contact information: gennadystolyarovii@gmail.com

The Creator's Intentions

When I began taking piano lessons more than two decades ago, I spent more time picking out and creating my own melodies than practicing the assigned pieces. Initially I would assemble musical segments from memory; whichever segments were sufficiently appealing to me were expanded upon and played through in a process of reinforcement, until I had, in my mind and memory, a finished piece. Eventually I would record the pieces on note paper, aiming to preserve them before they had fallen out of my memory. A further improvement from the standpoint of composition permanence was achieved when I obtained an electronic piano in 2001, which had the then-advanced ability to accommodate a floppy disk onto which a MIDI-format recording of one's performance could be saved. Yet, I often imagined melodies that were more complex than my playing skills allowed me to execute. For instance, in my *Variations on Alternating Marches*, Op. 15 (Stolyarov 2002), I envisioned increasingly rapid and powerful accompaniments with each variation; maintaining these accompaniments in the left hand while playing large chords with the right hand would likely only be possible for the most talented pianists – yet I had long wanted to hear the piece exactly with these kinds of features being fully implemented (Stolyarov 2002).

Even with the music I could play, I would occasionally make errors that meant I needed to start again in order to generate a satisfactory recording. There existed an ideal of the musical work in my mind, and yet the skills of a mere human piano student were not always able to reflect that ideal correctly. At the same time, my philosophical proclivities led me to contemplate key questions surrounding this endeavor. For instance, what did it mean to make a mistake in performing a piece whose "correct" version only existed in my own mind? Some philosophers, such as Plato, would have attributed music to the world of Forms, existing apart from our material reality, and containing the ideal musical pieces, of which actual human performances would be pale shadows or imitations. David Macintosh summarizes Plato's view thus:

> Plato says such Forms exist in an abstract state but independent of minds in their own realm. Considering this Idea of a perfect triangle, we might also be tempted to take pencil and paper and draw it. Our attempts will of course fall short. Plato would say that peoples' attempts to recreate the Form will end up being a pale facsimile of the perfect Idea, just as everything in this world is an imperfect representation of its perfect Form. The Idea or Form of a triangle and the drawing we come up with is a way of comparing the perfect and imperfect. How good our drawing is will depend on our ability to recognise the Form of Triangle. Although no one has ever seen a perfect triangle, for Plato this is not a problem. If we can conceive the Idea or Form of a perfect triangle in our mind, then the

> Idea of Triangle must exist. The Forms are not limited to geometry. According to Plato, for any conceivable thing or property there is a corresponding Form, a perfect example of that thing or property. The list is almost inexhaustible (Macintosh 2012).

In the realm of music, the implication of Plato's general view would be that an ideal Form of a musical piece does exist in another realm, as long as a human mind can envision it – but any attempt by a human performer to approximate that Form will always be wanting. Plato's view also implies that for any musical work that anyone can conceive of, a perfect Form will exist. Michael Bazemore notes that, "With regard to musical ontology a Platonist would hold that a work of classical music is an abstract object" (Bazemore 2015), although Plato himself held that such music ought to be put in the service of more concrete, didactic purposes. According to Mary B. Schoen-Nazzaro, Plato, "assigns four ends to music, and he sees a certain order between them: first, music moves the emotions; second, it gives pleasure; third, it disposes toward moral goodness; and fourth, it disposes toward learning" (Schoen-Nazzaro 1978, 266). Plato's ontology of music would render his posited didactic ends of music inherently incapable of full realization, as Plato also held that, "art imitates the objects and events of ordinary life. In other words, a work of art is a copy of a copy of a Form. It is even more of an illusion than is ordinary experience. On this theory, works of art are at best entertainment, and at worst a dangerous delusion" (Clowney 2018). Plato believed that music should serve uplifting moral goals, but if the music of this world is an imitation of an imitation of a Form and therefore even farther removed from the ideal world than is everyday life, then, in Plato's framework, music ultimately could not fully actualize those goals in practice – creating an unsatisfactory conclusion.

Yet, I was not and am not a Platonist, and even the didactic ends posited by Plato do not require, in my view, that music exist in a realm of Forms apart from the physical world. A more this-worldly variant of musical Platonism can be found in the theoretical writings of composer Ferruccio Busoni (1866–1924). Erinn Elizabeth Knyt explains that,

> Unlike Plato, Busoni does not believe an *Idee* [a fundamental musical idea] to be an ideal metaphysical 'type' that the phenomenal object merely represents. He considers the Idee to be a tangible image in the mind of the composer. It is not related to specific compositions, musical tones, or rhythms. But, it is an image formed in the psyche, something drawn from the human experience. (Knyt 2010, 116)

However, Busoni's view, while more appropriately locating the source of musical ideas within the mind of the composer, still requires music's core idea or essence to exist apart from the music itself – rather, being found in the external experiences that the composer brings to the creative process. This position also raises the dilemma of how to understand and define a given work of music if the listener is not extensively familiar with the source experiences and history of the composer.

My view is more closely aligned with that of Aristotle, who posited that the essence of any object or phenomenon is present in that object or phenomenon, existing in this world and not any other superior world of ideal Forms. In Book VII of his *Metaphysics*, Aristotle writes, "The essence of a thing is what it is said to be in virtue of itself" (Mattey 1998). But then, from an Aristotelian standpoint, could there be an essence of the *correct* version of a musical work when that work had never yet been correctly instantiated into the world via a flawless performance? The notation I wrote down could be considered my intention as the creator of the music – but is the music in this intended form only a concept, an aspiration until some sufficiently capable performer were able to play it flawlessly according to the notation?

While creating music, I also wanted to find ways to accelerate the process – to reduce the distance between the conception of a piece and its implementation in practice. Historically, the practice of composition has involved extensive study of music theory, learning to play one or more instruments, a process for writing down individual notes to record the creator's intentions for the work, and then seeking out musicians to perform the work. To the extent that technology could aid in reducing the effort involved in each of those processes, I hypothesized that this would place the real-world instantiation of music closer to the creator's original envisioned nature of that music and the thoughts that generated it. The logical conclusion of this process, in my mind, would be a technological future so advanced, that a creator's thoughts could become translated into music almost instantaneously. While humankind has not yet reached such a future, it has come significantly closer in the years since I imagined the possibility.

The questions surrounding the extent to which a composer can control a musical work and the precision with which the composer's intentions can be realized have been explored in prior eras, and varying conclusions have been reached. For instance, Karlheinz Stockhausen (1928–2007), in his article "… How Time Passes…", originally published in 1957 and republished in English in 1959, undertakes a systematic discussion of the variability that occurs in human performances of musical works and suggests that the difficulty of achieving an accurate performance is increased if the composer specifies more complex elements to which the musician must adhere. Stockhausen observed, for example, that, "In some recent scores, the notation of duration-relationships has become extremely differentiated. The result has been that, with an increase of metric-rhythmic complexity, the degree of precision in playing correspondingly decreased. To put it differently, the more complicated the way in which a time-value was indicated, the less sure the performer was about when it should begin and end" (Stockhausen 1959, 30). Stockhausen's thinking regarding this dilemma ultimately led in different directions from my own, in that Stockhausen attempted to accommodate some variability within the performance of his work but sought to have such variability, "precisely specified in its statistical parameters, i.e. as deliberately composed Variable form (even though the parameters like exact onset time, duration, precise pitch, etc. of the individual events themselves, by definition, even the number of events in a collection are not specified by the composed of the variable form)" (Koenigsberg 1991, 29). For me, however – even though variability in performances may be difficult to avoid due to human error, uncertainty, or difficulty in carrying out the composer's intentions (even when the performer is the composer) – the variable form of composition – which, to some extent, involves a loss of the definitiveness of the composer's intentions – is not a solution to this issue. In my view, both then and now, it remains important for the vision of the composer (which is often a singular vision) to find some actualization in the external reality, if only the proper means for such actualization could be deployed. Furthermore, in my view, some kinds of variability are more consistent with the composer's intentions than others – for instance, variations in dynamics and tempo per some (reasonably limited) discretion of the performer could potentially preserve the underlying essence of the work, but adherence to the composer's prescribed pitch and meter is crucial when the composer has provided singular specifications. Yet if human performers could not implement the composer's design in a manner that expressed the specific aspirations in the composer's mind, then what else could overcome those limitations? I wished to have a reliable pathway toward hearing a sufficiently developed instance of one of my own works, and being able to say regarding it, "This is indeed what I had intended."

Software as Solution

The narrative in this section relates the manner in which many of the previously articulated dilemmas were resolved for me personally through the course of my poietic and musical creation enabled by computer software.

In 2007 I was first introduced to Anvil Studio, a free program for input of musical notation and its playback via computerized instruments, with the ability to export MIDI file versions of the resulting works.[2] In Anvil Studio, it is possible to assemble music, note by note, and hear it played back in the exact manner in which it was input. Subsequently, it is possible to import the resulting MIDI file into other composition programs, such as Finale, which include more realistic simulations of instruments and even a "Human Playback" feature, which attempts to emulate various potential styles of human performance. It was then possible to record the resulting enhanced version of the music in a widely shareable and playable file format, such as MP3. Here I will not attempt to make comparisons of quality between a computerized performance and a human one; rather, I posit that the arrival of composition software

2 Anvil Studio. Free music composition, notation & MIDI-creation software. https://www.anvilstudio.com/

enabled a resolution to the philosophical dilemma that I previously faced. Provided that I entered the notation exactly as I intended it to be, the playback of the notation by the computer would be a definitive instantiation of those intentions – and, if the music did not sound as I had envisioned it, then the proper recourse would be to alter the inputs given to the computer. If a human musician were to subsequently perform the piece by following the same notation, it is likely that the human musician would impart certain nuances and performance styles that computers, as of yet, are not able to emulate. However, any such aspects brought by the human musician would be *additions to* or *interpretations of* the definitive version of the music, but the definitive version created through the use of software is able to be replayed in the exact same manner as many times as one might wish to hear it (as could, potentially, the human musician's interpretation of it, if it were recorded and shared).

Computer playback therefore enables the Aristotelian essence of a work of music to come into being and be reliably instantiated within the physical world of sounds. The essence of the music is ultimately determined by its creator – the person who decides upon and inputs the notes – but it is only able to be rendered concrete by means of machines which (provided they function as configured) will not commit the kinds of errors to which all humans are vulnerable to some degree. Through the precision of musical software, ideas which might have previously only existed in the mind or on paper can become auditory realities at the push of a button. For some such works it may be necessary to supplement traditional Western notation with other instructions to fully convey the composer's intentions. Or, if the composer intends the work to solely be performed by means of computers and other electronic devices, it may, in some cases, not even be necessary to have an official score for the work; rather, instructions given to the electronic device to enable it to perform the music may be sufficient, with the output of those instructions constituting the definitive musical work.

Another major advance that musical software allowed me to realize fully was the ability to create musical works which were too technically challenging to be played by humans at all – but, because of the ability to be played back by the computer as intended, could nonetheless have their musical essences established in this world. In 2008, I was finally able to hear my *Variations on Alternating Marches* in full after inputting it into Anvil Studio. I also ventured into experiments with multi-instrumental music where some aspect of a given instrument's part would pose challenges to a human performer. For example, the Fibonacci Rondo, Op. 54 (Stolyarov 2008a) is based on the famous Fibonacci sequence of numbers, which begins as 1, 1, 2, 3, 5, 8. If one assigns the value 1 to the note C, then one can assign the following values to other notes in relation to it: 2 = D, 3 = E, 5 = G, 8 = C one octave above the "1" note. Then, using two eighth-notes, one can represent the numbers being added, while the following quarter note represents their result. So, two eighth-note C's will be followed by a quarter-note D to represent "1 + 1 = 2." Then the eighth-notes C and D, followed by a quarter-note E, represent "1 + 2 = 3."

Then the eighth-notes D and E, followed by a quarter-note G, represent "2 + 3 = 5." Then the eighth-notes E and G, followed by a quarter-note C from the next octave, represent "3 + 5 = 8." Thereafter, the same pattern is applied to other harmonies – both major and minor – to ensure a melodic progression. This composition is written for a piano, two string sections, and timpani. It probably could not be played by a human orchestra, as the 32nd-notes in one of the string sections are simply too fast for human players to produce consistently.

A subsequent experimental work of mine from 2008, titled *Man's Struggle Against Death, Op. 57*, aims to depict the sequential overcoming of the seven major types of cellular and molecular damage involved in biological aging or senescence (Stolyarov 2008b). This composition – written for organ, two pianos, harpsichord, timpani, a brass section, and a strings section – consists of seven variations on the same theme – with the theme representing the consistent, unyielding human effort to defeat death and achieve indefinite longevity. Every time that a variation on the theme is played, this represents one of the causes of senescence finally being overcome by human ingenuity. Accordingly, the melody becomes more jubilant and determined as the composition progresses, because there are fewer perils awaiting man and the amount of tasks remaining is reduced. Once the seven variations are complete (which corresponds to the attainment of indefinite life), the coda of the work is meant to evoke the last line of John Donne's sonnet, *Death, Be Not Proud*: "And death shall be no more; death, thou shalt die." Like the quest to attain indefinite life, the performance of this work would be beyond the capabilities of humans alone – but aided by machines that can replicate extremely rapid, complex note progressions, success comes within reach.

Indeed, computer-aided musical creation can be considered a fundamentally transhumanist endeavor in that it extends the capabilities of humans to bring into being music that is not constrained by the limitations of the biological human organism. A specific celebration of this premise can be found in my *Transhumanist March*, Op. 78 (Stolyarov 2014). This march for piano depicts the accelerating improvement of the human condition and the overcoming of human limitations through technological progress. An ambitious and benevolent melody intensifies throughout the piece, pushing onward to champion the radical improvement of the human condition through lifting of age-old barriers and the conquest of both space and time. While the initial theme could be played by a human, the subsequent variations on it become increasingly challenging. The march concludes with an extremely complex variation for two pianos, which play identical parts, but always staggered by one 64th-note – an effect which adds depth to the sound but could not be consistently sustained by two human pianists. Subsequently, in January 2018, I integrated parts of the Transhumanist March (as an orchestral version) into the Fourth Movement of my *Symphony No. 1*, Op. 86 (Stolyarov 2018). The themes from the *Transhumanist March*, occurring toward the end of the symphony, point toward the prospects for a brighter future of humankind, if humans can preserve all of the past gains of civilization and extrapolate upon them, creating an era where our capabilities could be greatly expanded through technology.

Early Methods of Automated Playback

Even prior to the advent of computers, however, automated playback was available in a more limited form, due to technological progress in the design of self-playing instruments. Precursors to the player piano, or pianola, were developed as early as the 1840s, but this instrument came into widespread prominence during the first decade of the 20th century and enabled the automated playback of musical works through inserting perforated paper rolls that convey instructions to the piano. In the 20th century, several composers developed works of immense complexity for the player piano, which would be impossible for a human performer to play unaided. Conlon Nancarrow (1912–1997) composed works directly for player piano, beginning in, "the mid-1930's, when he found pianists unable to play works like the *Toccata* for Violin and Piano and the *Prelude and Blues* (both composed in 1935) at the speeds or with the clarity that he demanded" (Kozinn 1997). Indeed, Nancarrow's, "frustrations with the limitations of live performance technique led him to compose almost exclusively for mechanical player pianos" (Kozinn 1997) – a similar sentiment to the one that led me to explore the creation and playback of musical works via computer software seven decades later. Nancarrow harnessed the player piano to innovate with tempo relationships – for instance, $\sqrt{(2)}/2$ in his Player Piano Study No. 33 (1965-1969) and e/π in his "Transcendental" Player Piano Study No. 40 (1969–1977) that no human would be able to reproduce with exactitude (Gann 1997). Indeed, even Nancarrow had to approximate, as even player pianos are only capable of a certain level of precision; for instance, "for his Study no. 33, Nancarrow approximated $\sqrt{(2)}/2$ within 99.97 percent as the ratio 41:29" (Pesic 2017, 204). However, the capabilities of the player piano still vastly expanded both the melodic complexity and the tempo varieties available for composers to deploy. Works for the player piano continue to be created by contemporary composers, including Marc-André Hamelin (b. 1961), who wrote *Circus Galop* in 1991–1994 and the *Solfegietto a cinque* in 1999 – a work which greatly expands upon C.P.E. Bach's 1766 *Solfegietto* in terms of length, complexity, and ornateness. The advantages today's computer software have over the player piano include a greater range of possible instruments to replicate, greater affordability for the composer, and greater portability of the software and its products (digital files) as compared to the large, heavy player pianos and the paper rolls that they utilize. Accordingly, access to automated playback has become greatly expanded in the age of personal computers.

Algorithms and Humans Together

As I continued using software to create musical experiments, I pondered the possibility of harnessing the massive processing capabilities of computers to generate new musical ideas. My subsequent research into this subject suggested that this concept was not novel; indeed, endeavors in algorithmic musical composition have been pursued since the Age of Enlightenment. The Musikalisches Würfelspiel, or musical dice game, was a common creation of musicians in the mid-to-late 18th century, where each roll of the dice was mapped to a particular pre-composed measure of music, and the measures would be assembled into melodies depending on the outcomes of the dice rolls. The genre began with Johann Philipp Kirnberger's publication (1721–1783) of *Der allezeit fertige Menuetten- und Polonaisencomponist* ("The Ever-Ready Minuet and Polonaise Composer") in 1757. Kirnberger's rule set enabled the generation of 1,679,616 possible musical combinations. The Musikalisches Würfelspiel genre culminated in perhaps the most famous musical dice game which was attributed to Wolfgang Amadeus Mozart (1756–1791). This work was published posthumously in 1793 and enabled the generation of minuets and trios of 16 bars each. In this musical dice game, one would roll two six-sided dice 16 times to create the minuet and roll one six-sided die 16 times to create the trio, allowing for 6616 = 129,629,238,163,050,258,624,287,932,416 unique minuet/trio combinations.

Between 1967 and 1969, John Cage (1912–1992) and Lejaren Hiller (1924–1994) utilized components of Mozart's Musikalisches Würfelspiel to develop the multimedia composition/performance called HPSCHD, which included seven harpsichord solo pieces along with 52 tapes generated by a computer. The harpsichord solos utilized 64 out of the 176 measures from the *Musikalisches Würfelspiel*, supplemented in some of the solos by measures from six other works of Mozart, as well as excerpts from the compositions of Ludwig van Beethoven (1770–1827), Frédéric Chopin (1810–1849), Robert Schumann (1810–1856), Louis Moreau Gottschalk (1829–1869), Ferruccio Busoni, John Cage, and Lejaren Hiller. Hiller and Cage created two programs, DICEGAME, to compose the harpsichord solos, and HPSCHD, to compose the music for the computer tapes (Di Nunzio 2014).

Since at least the 1990s, independent programmers have created downloadable software instantiations of both the Kirnberger and Mozart musical dice games, automating the process and allowing finished musical combinations to be heard and downloaded in MIDI format, creating the potential for subsequent editing and development.[3]

In 2015, conducting an experiment that explored the intersection of human and

3 A program that allows the generation of music from both Kirnberger's and Mozart's dice games was developed by Peter Baumann in 2006-2007 and can be accessed and downloaded for free at http://www.combib.de/programme/musikalischewuerfelspiele.html.

algorithmic creativity, I randomly generated a minuet and trio from the Mozart *Würfelspiel* and then composed four variations on it for piano and harpsichord, resulting in a piece of nearly five minutes in length (Stolyarov 2015a). When creating the variations, I was cognizant of the near-certainty that, given the vast number of unique minuet/trio combinations attainable, this particular iteration of the Musikalisches Würfelspiel had never been heard before – and that by creating the variations and publishing an augmented version that was definitively played back by a computer program, I was concretizing into existence and imparting unique significance onto what had previously been merely one potentiality among over 129 octillion. Even though the original measures were (possibly) composed by Mozart and assembled together in a specific manner by an algorithm, it was still my decision as a creator to emphasize and build upon that particular combination, and the original minuet/trio and variations thereon became their own definitive piece with an original version that everyone could hear and which represented the real-world essence of that piece.

Algorithmic composition has advanced significantly since the early musical dice games aided by the advent of computing in the mid-20th century. Between 1955 and 1957, Lejaren Hiller and Leonard Isaacson (b. 1925) used the ILLIAC I computer at the University of Illinois at Urbana-Champaign to create the Illiac Suite, "the first musical composition for traditional instruments that was made through computer-assisted composition" – one which was informed by Hiller's view that, "the process of musical composition can be characterized as involving a series of choices of musical elements from an essentially limitless variety of musical raw materials" and that such choices could be automated by applying, "a model that allow the computer to make organizational decisions respect to musical composition features. The model adopted was the Monte Carlo method, an algorithm which uses the generation of random numbers" (Di Nunzio 2011).

In 1963 composer, architect, mathematician, and musical theorist Iannis Xenakis (1922–2001) published *Musiques formelles*, which became expanded and translated into English in 1971 under the title of *Formalized Music: Thought and Mathematics in Composition*. Xenakis articulated the idea that computers are tools to extend our ability to compose music, but that they also do not represent a fundamental disconnect from previous musical frameworks and methods of composition, which also rely on mathematical principles and mechanisms of choice within formalized frameworks. Xenakis explains that, "Computers resolve logical problems by heuristic methods. But computers are not really responsible for the introduction of mathematics into music; rather it is mathematics that makes use of the computer in composition" (Xenakis 1971, 132). Xenakis then proceeds to express six principles that provide a bridge into computerized composition:

> 1. The creative thought of man gives birth to mental mechanisms, which, in the last analysis, are merely sets of constraints and choices. This process

> takes place in all realms of thought, including the arts.
> 2. Some of these mechanisms can be expressed in mathematical terms.
> 3. Some of them are physically realizable: the wheel, motors, bombs, digital computers, analogue computers, etc.
> 4. Certain mental mechanisms may correspond to certain mechanisms of nature.
> 5. Certain mechanizable aspects of artistic creation may be simulated by certain physical mechanisms or machines which exist or may be created.
> 6. It happens that computers can be useful in certain ways (Xenakis 1971, 132–133).

Xenakis himself developed computer systems such as UPIC and GENDY to provide him with content to use in his music, although in Xenakis's process, "what the computer was outputting was not the composition itself but material with which Xenakis could compose" (Maurer 1999).

In 1965, then-17-year-old inventor Ray Kurzweil (b. 1948) appeared on the American television show "I've Got a Secret" to demonstrate an excerpt from a composition generated by a computer he had custom-built. Kurzweil later described it as his, "first pattern recognition project. I built a computer, programmed it to recognize the melodies of the music I would feed into it and then write original music using the same kinds of patterns. So it would write music, recognizable as Mozart, Bach or Chopin" (Ray Kurzweil, quoted in Candela 2004). Kurzweil went on to develop major innovations and inventions in the musical realm, including the development of the first synthesizer capable of replicating the grand piano and orchestral instruments – the Kurzweil K250 in 1984, "whose 'piano mode' was indistinguishable from a grand piano when played for musicians in listening tests" (Kane 2016). Indeed, much of my own later musical creation only became possible because of the advances in computerized instruments that built upon the principles in Kurzweil's work.

David Cope (b. 1941), a composer and former professor of music at the University of California, Santa Cruz, developed programs such as EMI (Experiments in Musical Intelligence) and the later Emily Howell, which have created thousands of pieces of music based on the styles of historical composers. These programs have also created thousands of experimental compositions that could possibly be attributed to the "style" that each program has attained through the novel recombinations of the vast assortment of musical elements and techniques with which it has been provided. Cope explains that he began his foray into the use of artificial intelligence in composition as a way to enhance his own creativity:

> I began Experiments in Musical Intelligence in 1981 as the result of a composer's block. My initial idea involved creating a computer program which would have a sense of my overall musical style and the ability to track the ideas of a current work such that at any given

> point I could request a next note, next measure, next ten measures, and so on. My hope was that this new music would not just be interesting but relevant to my style and to my current work. (Cope n.d.)

Cope describes his approach as based on the premise that,

> every work of music contains a set of instructions for creating different but highly related replications of itself. These instructions, interpreted correctly, can lead to interesting discoveries about musical structure as well as, hopefully, create new instances of stylistically-faithful music. My rationale for discovering such instructions was based, in part, on the concept of recombinancy. Recombinancy can be defined simply as a method for producing new music by recombining extant music into new logical successions. (Cope n.d.)

The recombination of known elements and techniques is not a barrier to creativity but, rather, it has been a foundation of creativity from the onset. Cope observes that, "most of the great works of Western art music exist as recombinations of the twelve pitches of the equal-tempered scale and their octave equivalents. The secret lies not in the invention of new letters or notes but in the subtlety and elegance of their recombination" (Cope n.d.). At the same time, Cope has continued to be involved in the curation of the pieces produced by EMI and Emily Howell – in the sense that it remains his decision, as a human well-versed in the history and theory of music, to identify which pieces are reasonable facsimiles of the styles of particular composers, and also to determine which of the "avant-garde" melodies generated by the programs are inherently interesting and worthy of featuring and designating as definitive works. Cope emphasizes the essential continuity between this process and the creative efforts of prior composers: "Ultimately, the computer is just a tool with which we extend our minds. The music our algorithms compose [is] just as much ours as the music created by the greatest of our personal human inspirations" (Cope n.d.).

WolframTones, developed based on the work of mathematician and computer scientist Stephen Wolfram (b. 1959), is another free music-generation engine, created in 2005. It is described as using, "various Wolfram Language algorithms to form music out of cellular automaton patterns" and attempting to, "search the universe of possible rules for ones that have relevant kinds of complex behavior" (WolframTones 2005). Some of the randomly generated 15-second musical patterns from WolframTones sound intriguing and intricate, whereas others seem more noise-like, or simplistically repetitive, or end abruptly before the 15-second time interval elapses. However, due to the ease of generation of new patterns, it again becomes the province of the human interacting with WolframTones to select the most promising patterns to download and preserve, and potentially build upon. Another experiment of mine from 2015, the *Variations on a Theme* by WolframTones, Op. 80, began with a 15-second theme generated by WolframTones for a harp and a string section, upon which I developed ten variations for a string section, three harps, and two pianos (Stolyarov 2015b). Out of the initial melodic segment, I was able to craft two distinct alternating themes that evolved with each variation and complemented one another in mood. Several listeners of the piece have commented that they enjoyed the variations more than the original WolframTones theme, and indeed, the variations were more deliberately crafted, more melodic, and were specifically designed to appeal to the human ear. However, the fact that they were inspired by and derived from a randomly generated musical segment is a testament to how the products of algorithms are able to stimulate human creativity and the development of music that clearly reflects human intentionality.

In February 2019 the Chinese technology company Huawei released a complete recording of the finished version of Franz Schubert's Symphony No. 8 – Schubert (1797–1828) had only composed the first two movements of the *Unfinished Symphony No. 8* during his lifetime. The finished symphony was performed live at the Cadogan Hall in London, United Kingdom, on February 4, 2019.[4]

The final two movements were composed by an artificial intelligence (AI) system running on a Huawei Mate 20 pro smartphone. Per Huawei's description, "The Mate 20 Pro smartphone listened to the first two movements of Schubert's Symphony No. 8, analysed the key musical elements that make it so incredible, then generated the melody for the missing third and fourth movement from its analysis" (Huawei 2019). The orchestral score for the finished symphony was arranged by composer Lucas Cantor, who assembled the melodies provided by the Huawei smartphone into coherent Third and Fourth Movements of the symphony (Davis 2019). However, Cantor did make intentional selections regarding *which* of the melodies generated by the smartphone to integrate and also regarding *how* to orchestrate them to enable the live performance to take place. Through the collaboration with Cantor, the Huawei smartphone AI's capabilities were greatly extended – enabling the completed symphony to be created – but Cantor's own capabilities were likewise enhanced in that he could work with the numerous melodic ideas supplied by the AI. Cantor has stated that this experience was, "like having a collaborator who never gets tired, never runs out of ideas" (Davis 2019) and that it, "proves that technology offers incredible possibilities and has a significant and positive impact on modern culture" (Huawei 2019).

While the Huawei-generated Third and Fourth Movements of Schubert's Symphony No. 8 have been criticized for not being sufficiently reminiscent of Schubert's own style (Richter 2019),[5] from the standpoint of sheer technical accomplishment and the enjoyment that the movements confer on their own

4 Excerpts of performance of the Huawei Schubert Symphony No. 8 were captured in this video recording by Chris J. Kenworthy: https://www.youtube.com/watch?v=9FV75jDzse0.

5 For instance, Goetz Richter has written that, to him, "these movements sound only a little like Schubert and a lot like film music." (Richter 2019, Goetz 2019)

terms, these additions to the symphony are of significant merit. Were a human composer to create anything comparable, the result would feature prominently in that composer's oeuvre. In my impression, it is indeed the case that the Third and Fourth Movements resemble more of a late 19th-century style – perhaps evoking the music of Gustav Mahler, especially given the triumphant, epic ending. However, this does not necessarily mean that the music is "not like Schubert;" rather, it may answer a counterfactual question that arose from Schubert's tragically short life. Schubert died in 1828 at the age of 31, and, were it not for the unfortunate illness which killed him, could have potentially survived, with some luck, into the 1880s. Could his composing style have evolved to generate works similar to those last two movements? This is at least plausible. The emergence of artificially intelligent systems, drawing upon a vast database of Schubert's compositions but also capable of innovating beyond it and guided by humans who choose which of these potentially innovative pathways are explored further, allows us to at least more vividly imagine what an older Schubert's music might have sounded like.

The Promise of Amplifying the Creator through Artificial Intelligence

While humans cannot yet compose music with a mere thought, the tools available for innovative musical creation have greatly expanded since I first wished for this possibility. Composition software and computer playback of music on increasingly realistic digital instruments allow for definitive versions – the concretized essences – of music, as intended by its creator, to come into being in the physical world. Algorithmic composition, based on both random (or pseudorandom) and preprogrammed processes, has existed for over 260 years, but now can be carried out in much more efficient and diverse ways through an expanding array of computerized tools. Presently, as prominently demonstrated by the work of David Cope and the Huawei/Cantor collaboration in completing Schubert's Symphony No. 8, narrow AI systems are becoming increasingly competent at both emulating historical composition styles and expanding them in new directions. Narrowing the gap between conception and performance also suggests various possibilities for instantiating the work of music. Composers can use software to hear their works both in the intermediate stages of creation and in their final versions. If they wish for the works to eventually be performed by human musicians, they are still able to develop scores for the musicians to follow – but it is also possible for the first recorded and published performances of the work to antedate the creation of such scores (since the computer software would already have the instructions for performing a work entered via its interface).

The future which these developments point to promises to be even more intriguing, as artificially intelligent systems become more versatile and even begin to acquire capabilities across domains, approaching the status of an artificial general intelligence. It is already possible for an AI to generate a plethora of musical ideas for a human composer to integrate into a coherent work or use to develop variations that add a human touch. At the outer edge of this frontier would be the generation of entire complete musical pieces by AI systems, where human involvement would be only in the form of a curator who decides what to publish, what to feature, and what to draw people's attention to. On the other hand, the emergence of this capability is not intended to suggest a diminished role for human composers – in fact, it is quite the contrary. The participation of algorithms and artificial intelligence does not in any way compromise the ability of music to be precisely developed to match the creator's intentions. Just as historical composers have drawn upon folk melodies for inspiration, or written variations on themes authored by other composers, so the musical creators of the present and future would be able to intentionally decide which elements from algorithms and AI systems to build into their own works. Once the decision is made and instantiated, the resulting work – no matter which proportions of its content were generated by a human as compared to an algorithm – becomes just as intentional and just as definitive as a more traditionally composed piece of music would be. The Aristotelian essence of a work of music is in the music itself, rather than in the specific pathway to its creation – so the sound of the music can encapsulate its essence as long as it accurately conveys the intention of the composer, as derived from potentially any source or process of the composer's choosing. It is possible, even, for the composer to hear an algorithmically produced musical work for the first time and subsequently, either accept that work as aligning with the composer's intentions or modify it to make it align, thereby bringing it into that composer's *oeuvre*.

The same tools that allow increasingly intricate works to be created by machines, could also broaden the possibilities and heighten the quality of what can be produced by the human mind. Already experimental technologies have been developed to detect the transmission of human brain waves and, to some extent, interpret their content. One example of such experimentation is the electroencephalogram (EEG) technology developed in 2017 at the Toyohashi University of Technology, Japan, which is able to, "recognize the numbers 0 to 9 with 90% accuracy using brain waves […] while [a human is] uttering the numbers" and which, "has also realized the recognition of 18 types of Japanese monosyllables from EEG signals with 60% accuracy, which shows the possibility of an EEG-activated typewriter in the near future" (Toyohashi University of Technology, Japan, Committee for Public Relations 2017). If future innovations extrapolate upon this technology, it is not beyond the realm of possibility, within the proximate several decades, for the capability to emerge to invent devices which could record the brain signals corresponding to a melody envisioned within a human mind and import it into a musical program that would translate it into notation and then play it back. Artificially intelligent features within the musical program could develop variations and creative orchestrations of the melody, with the human creator capable of providing input, adjusting the parameters within which these variations and orchestrations are generated, and curating the output using ever more convenient and intuitive user interfaces. Using augmented-reality glasses or holographic projections, the human creator might be able to see and work with the resulting music from any location, without even accessing a desktop or smartphone screen. The already-existing ability to hear instant playback of one's music – instead of needing to hire musicians and arrange for them to perform, as many past composers needed to do – shortens the time period between having the initial idea and its implementation, but the improving convenience of every step in the composition process will allow more time and energy to be spent on the actual development of and experimentation with musical ideas and their integrations into works of greater intricacy and ambition. The confluence of these developments will also greatly lower the practical barriers to entry involved in the creation and distribution of music. One day, hopefully within the lifetimes of the readers of this paper, any human who thinks of an interesting melody will be able to seamlessly develop it into a full-fledged, beautifully arranged work of music for the world to hear. What were once only thoughts, or potentialities, or pale reflections of the creator's wishes would come into full, vibrant reality as man and machine extend their creative symbiosis.

List of References

Amaranth Publishing. "Musikalisches Würfelspiel (Musical Dice Game)." Accessed April 13, 2019. http://www.amaranthpublishing.com/MozartDiceGame.htm.

Anvil Studio. Free music composition, notation & MIDI-creation software. https://www.anvilstudio.com/.

Baumann, Peter. 2006–2007. "Musikalische Würfelspiele." ComBib. http://www.combib.de/programme/musikalischewuerfelspiele.html.

Bazemore, Michael. 2015. "The Aesthetics of Classical Music." *Internet Encyclopedia of Philosophy*. https://www.iep.utm.edu/aest-c-l/.

Candela, Tony. 2014. "Interview with Ray Kurzweil, Part 3." *American Foundation for the Blind*. November 12, 2004. https://www.afb.org/blindness-and-low-vision/using-technology/interviews-technology-pioneers/ray-kurzweil/part-3-4.

Clowney, David W. 2018. "Plato". Rowan University. http://users.rowan.edu/~clowney/Aesthetics/philos_artists_onart/plato.htm.

Cope, David. "Experiments in Musical Intelligence", accessed April 13, 2019, http://artsites.ucsc.edu/faculty/Cope/experiments.htm.

Davis, Elizabeth. 2019. "Schubert's 'Unfinished' Symphony completed by artificial intelligence." *Classic FM*. February 6, 2019. https://www.classicfm.com/composers/schubert/unfinished-symphony-completed-by-ai/.

Di Nunzio, Alex. 2011. "Iliac Suite." *Musica Informatica*. Last modified April 17, 2014. https://www.musicainformatica.org/topics/illiac-suite.php.

Di Nunzio, Alex. 2014. "HPSCHD." *Musica Informatica*. Last modified December 3, 2014. https://www.musicainformatica.org/topics/hpschd.php.

Gann, Kyle. 1997. "Conlon Nancarrow: An Annotated List of Works." KyleGann.com. https://www.kylegann.com/cnworks.html.

Hamelin, Marc-André. 1994. "Circus Galop - Marc-André Hamelin." YouTube. Video, 5:10. https://www.youtube.com/watch?v=Jv4epBitpic.

Hamelin, Marc-André. 1999. "Marc-André Hamelin – Solfeggietto a cinque for Player Piano." YouTube. Video, 3:59. https://www.youtube.com/watch?v=hzFOkGfqaeE.

Huawei. 2019. "Huawei Presents Unifnished Symphony, Powered by Huawei AI." https://consumer.huawei.com/uk/campaign/unfinishedsymphony/.

Kane, Sean. 2016. "How a pianist became the world's most famous futurist." *Business Insider*. Last modified March 19, 2016. https://www.businessinsider.com/how-ray-kurzweil-became-famous-2016-3.

Kenworthy, Chris J. 2019. "Huawei's Unfinished Symphony - Cadogan Hall, London (Mon 4 Feb 2019)." YouTube. Filmed February 4, 2019, at Cadogan Hall, London. Video, 12:55. https://www.youtube.com/watch?v=9FV75jDzse0.

Kirnberger, Johann Philipp. 1757. *Der allezeit fertige Polonoisen- und Menuettencomponist*. Berlin: George Ludewig Winter, and IMSLP, April 20, 2015. Accessed April 13, 2019, http://ks4.imslp.info/files/imglnks/usimg/e/ea/IMSLP374746-PMLP243537-kirnberger_allezeit_fertiger_usw.pdf.

Knyt, Erinn Elizabeth. 2010. *Ferruccio Busoni and the Ontology of the Musical Work: Permutations and Possibilities*. Stanford University. https://stacks.stanford.edu/file/druid:ck155rf0207/Dissertation%20R-augmented.pdf.

Koenigsberg, Christopher K. 1991. "Karlheinz Stockhausen's New Morphology of Musical Time." Mills College. https://issuu.com/markusbreuss1/docs/karlheinz_stockhausen_s_new_morphol.

Kozinn, Allan. 1997. "Conlon Nancarrow Dies at 84; Composed for the Player Piano." *New York Times*. August 12, 1997. https://www.nytimes.com/1997/08/12/arts/conlon-nancarrow-dies-at-84-composed-for-the-player-piano.html.

Kurzweil Technologies, Inc. 1965. "Ray Kurzweil on 'I've Got a Secret.'" Vimeo and YouTube. Video, 3:20. https://vimeo.com/270066141 and https://www.youtube.com/watch?v=X4Neivqp2K4.

Macintosh, David. 2012. "Plato: A Theory of Forms." *Philosophy Now*, Issue 90 (May/June 2012). https://philosophynow.org/issues/90/Plato_A_Theory_of_Forms.

Mattey, G.J. 1998. "Instructor's Notes: Aristotle's Metaphysics, Books I, IV, VII, XII." University of California, Davis. http://hume.ucdavis.edu/mattey/phi001/metanote.htm.

Maurer IV, John A. 1999. *A Brief History of Algorithmic Composition*. Stanford University. https://ccrma.stanford.edu/~blackrse/algorithm.html.

Mozart, Wolfgang Amadeus. 1793. *Musikalisches Würfelspiel, K.516f*. Bonn: N. Simrock, Plate 48, and IMSLP, August 6, 2008. Accessed April 13, 2019. http://ks4.imslp.info/files/imglnks/usimg/b/bc/IMSLP20432-PMLP47543-mozart_-_dice_waltz.pdf.

Pesic, Peter. 2017. *Polyphonic Minds: Music of the Hemispheres*. Cambridge: The MIT Press.

Richter, Goetz. 2019. "Composers are under no threat from AI, if Huawei's finished Schubert symphony is a guide." The Conversation, February 2019. http://theconversation.com/composers-are-under-no-threat-from-ai-if-huaweis-finished-schubert-symphony-is-a-guide-111630.

Schoen-Nazarro, Mary B. 1978. "Plato and Aristotle on the Ends of Music." *Laval théologique et philosophique*, 34, (3), 261–273. https://www.erudit.org/fr/revues/ltp/1978-v34-n3-ltp3386/705684ar.pdf.

Stockhausen, Karlheinz. 1959. "… How Time Passes …" *Die Reihe*. Vol. 3: 10–40. https://www.artesonoro.net/artesonoroglobal/HOW%20TIME%20PASSES%20BY.PDF.

Stolyarov II, Gennady. 2002. *Variations on Alternating Marches, Op. 15 (2002)*. YouTube: https://www.youtube.com/watch?v=rVjTL3QzAGY and The Rational Argumentator (MP3): http://rationalargumentator.com/music_stolyarov/Stolyarov_Variations_on_Alternating_Marches_Piano.mp3.

Stolyarov II, Gennady. 2008a. *Fibonacci Rondo, Op. 54 (2008)*. YouTube: https://www.youtube.com/watch?v=HK6MITRdjOU and The Rational Argumentator (MP3): http://rationalargumentator.com/music_stolyarov/Stolyarov_Fibonacci_Rondo_Remastered.mp3.

Stolyarov II, Gennady. 2008b. *Man's Struggle Against Death, Op. 57 (2008)*. YouTube: https://www.youtube.com/watch?v=y_i_iZMdhjA and The Rational Argumentator (MP3): http://rationalargumentator.com/music_stolyarov/Stolyarov_Man%27s_Struggle_Against_Death_Remastered.mp3.

Stolyarov II, Gennady. 2014. *Transhumanist March (March #12), Op. 78 (2014)*. YouTube: https://www.youtube.com/watch?v=rzs4Fk1DzEI, and The Rational Argumentator (MP3): http://rationalargumentator.com/music_stolyarov/Stolyarov_March_12.mp3.

Stolyarov II, Gennady. 2015a. *Variations on a Randomly Generated Minuet and Trio by W.A. Mozart, Op. 81 (2015)*. YouTube: https://www.youtube.com/watch?v=tDOWjsgXoOI and The Rational Argumentator (MP3): http://rationalargumentator.com/music_stolyarov/Stolyarov_Variations_Minuet_Mozart.mp3.

Stolyarov II, Gennady. 2015b. *Variations on a Theme by WolframTones, Op. 80 (2015)*. YouTube: https://www.youtube.com/watch?v=Tm-4tu0shI8, and The Rational Argumentator (MP3): http://rationalargumentator.com/music_stolyarov/Stolyarov_Variations_WolframTones.mp3.

Stolyarov II, Gennady. 2018. *Symphony No. 1, Op. 86 – Movement 4 – 'Preservation' (2017–2018)*. YouTube: https://www.youtube.com/watch?v=ZwvHq23R4o4 and The Rational Argumentator (MP3): http://rationalargumentator.com/music_stolyarov/Stolyarov_Symphony_1_Movt_4.mp3.

Stolyarov II, Gennady. 2017–2018. *Symphony No. 1, Op. 86 – 'The Contemporary World' (2017–2018)*. YouTube. https://www.youtube.com/watch?v=6V42kdj7DCc.

SunSITE Austria at the University of Vienna. 2008. "Musikalisches Würfelspiel rules for Minuet / Trio composition." Last modified September 21, 2008. Accessed April 13, 2019. http://sunsite.univie.ac.at/Mozart/dice/rules.html.

Toyohashi University of Technology, Japan, Committee for Public Relations. 2017. "Success in Recognizing Digits and Monosyllables with High Performance from Brain Activity Measurement." Last modified April 17, 2017. https://www.tut.ac.jp/english/introduction/docs/pr20170417_nitta.pdf.

WolframTones. 2005. Accessed April 13, 2019. http://tones.wolfram.com/.

WolframTones. 2005. "How It Works: Scientific Foundations." Accessed April 13, 2019. http://tones.wolfram.com/about/how-it-works.

EMPOWERING MUSICAL CREATION THROUGH MACHINES, ALGORITHMS, AND ARTIFICIAL INTELLIGENCE (summary)

My early thinking on the philosophy of music and the essence of a musical work was motivated by my imagination of melodies which were more complex than my playing skills could execute. I asked questions such as, "What did it mean to make a mistake in performing a piece whose 'correct' version was only formed in my own mind?" and "From an Aristotelian standpoint, could there be an essence of the correct musical work when that work had never yet been correctly instantiated into the world via a flawless performance?" While creating music, I also wished for ways to accelerate the process – to reduce the distance between the conception of a piece and its implementation in practice. Some of these dilemmas were resolved via accessible software that enabled input of musical notation and its playback via computerized instruments. Such playback would be a definitive instantiation of those intentions – and if the music did not sound as I had envisioned it, then the proper recourse would be to alter the inputs given to the computer. Computer playback therefore enables the Aristotelian essence of a musical work to be reliably instantiated within the physical world of sounds. The essence of the music is ultimately determined by its creator – the person who decides upon and inputs the notes – but it is only able to be rendered concrete by means of machines which will not commit the kinds of errors to which all humans are vulnerable to some degree. Another major advance that musical software enabled was the ability to create musical works which were too technically challenging to be played by humans at all – but, because of the ability to be played back by the computer as intended, could nonetheless have their musical essences established in this world. Indeed, computer-aided musical creation can be considered a fundamentally transhumanist endeavor in that it extends the capabilities of humans to bring into being music that is not constrained by the limitations of the biological human organism. The idea of algorithmic composition is not novel; indeed, it was implemented in musical dice games extending as far back as Johann Philipp Kirnberger's *Der allezeit fertige Menuetten- und Polonaisencomponist* in 1757. My experiments involved creating variations on randomly generated music from Wolfgang Amadeus Mozart's *Musikalisches Würfelspiel* and a 15-second theme from WolframTones. This paper also explores the innovations in artificial intelligence and human-machine collaboration, including the works of Ray Kurzweil, David Cope, and Lucas Cantor – the composer who curated the completed version of Schubert's *Unfinished Symphony* using musical ideas generated by a Huawei smartphone AI. While humans cannot yet compose music with a mere thought, the tools available for innovative musical creation have greatly expanded since I first wished for their creation. The future which these developments point to promises to be even more intriguing, as the same tools that allow increasingly intricate works to be created by machines, could also broaden the possibilities and heighten the quality of what can be produced by the human mind.

Article received: April 13, 2019
Article accepted: May 15, 2019
Review article

UDC 78.02:004.8
78.09:004.8
78:37]:004.8

Harun Zulić*
Canada Music Academy
Toronto, Canada

HOW AI CAN CHANGE/ IMPROVE/INFLUENCE MUSIC COMPOSITION, PERFORMANCE AND EDUCATION: THREE CASE STUDIES

Abstract: The use of artificial intelligence in science is happening more and more frequently, and often artificial intelligence can be seen in different approaches to creating music and art. In this paper, I will present some of the research that has been carried out, which involve the use of artificial intelligence in the field of composition, performance, and music education. The main focus in the field of composition will be on AIVA – the first virtual composer created with artificial intelligence, which is registered with an author's rights society. In the field of performance, we'll mostly talk about Yamaha's experiment where the world-renowned dancer Kaiji Moriyama controls a piano with his body movements, and in the context of education, this paper reviews some of the possibilities in a variety of artificial intelligence approaches to music education. Lastly, I will conclude the paper by presenting the direction of and possible future for the use of artificial intelligence in music.

Keywords: artificial intelligence, composition, performance, music education, human

Introduction

In the middle of the last century, several scholars from different fields (engineering, psychology, mathematics, economics and political science) began

* Author's contact information: harun92zulic@hotmail.com

to think about creating an artificial brain. The official beginning, or its date of recognition within the academic community, occurred in 1956. The name artificial intelligence was coined in 1955 by John McCarthy.[2] "Artificial Intelligence (AI) is the part of computer science concerned with designing intelligent computer systems that exhibit the characteristics we associate with intelligence in human behaviour—understanding language, learning, reasoning, solving problems, and so on" (Barr and Feigenbaum 1981, 3). Since 1956, the development of technology has rapidly grown. Today, in the 21st century, AI is all around us. In the 63 years since AI was officially born, it's been used in a variety of fields, some of which are: finance, marketing, healthcare, medical diagnosis, robotics, automation, optical character recognition, nonlinear control, semantic webs, education, transportation, music, artificial life, game theory, computational creativity, speech recognition, bio-inspired computing, face recognition, hybrid intelligent system, etc. Considering the fact that the field of AI is very diverse and complex, in the rest of this paper we will only talk about the influence of AI on music.

Connecting artificial intelligence with music began in the mid-1960s, and it relates to research that focuses on music as a cognitive process or as a set of activities modelled with the aid of computer programs (Berz and Bowman 1995, 20). The first published paper on algorithmic music composition using the "Ural-1" computer was "An algorithmic description of process of music composition" by R. Kh. Zaripov, in 1960. Cognitive psychology had a significant role in the development of this field. One of the first significant discoveries in this field was the Ray Kurzweil computer that was programmed to compose music (New York, 1948). Kurzweil is one of the world's leading inventors, thinkers, and futurists, who is involved in a diverse number of fields including artificial intelligence, entrepreneurship, exponential organizations, future forecasting, optical character recognition (OCR), text-to-speech synthesis, speech recognition technology, and electronic keyboard instruments. At the age of 12, he became fascinated by the possibilities of the computer, and by age 15, he had written his first computer program. In one of his interviews, Kurzweil mentioned that his first project involved the computer recognizing patterns in music. Melodies composed by this computer can be compared with melodies of each composer from whom it learned to compose. This computer program led to Kurzweil winning first prize in the International Science Fair and the Westinghouse Science Talent Search.[3]

2 John McCarthy (Boston, 1927 – Stanford, 2011) was an American computer and cognitive scientist. His "contributions to computer science and artificial intelligence are legendary. He revolutionized the use of computers with his innovations in timesharing; he invented Lisp, one of the longest-lived computer languages in use; he made substantial contributions to early work in the mathematical theory of computation; he was one of the founders of the field of artificial intelligence; and he foresaw the need for knowledge representation before the field of AI was even properly born" (Hayes and Morgenstern 2007, 93).

3 "I went on I've Got a Secret. I went on and played a piece of music and then whispered in Steve Allen's ear, 'I built my own computer.' And he said, 'Well, that's impressive. What's that have to do with that piece of music you just played?' And I said, 'Well, the computer composed the piece of music.' And then Bess Myerson, who was a former Miss America, was stumped, but then Henry Morgan,

Another very significant project of Kurzweil's was the Kurzweil Music Systems and to create this program Kurweil collaborated with multi-musician Stevie Wonder and software developer Bruce Cichowlas. They demonstrated an engineering prototype of the Kurzweil 250 (K250) in 1983 and introduced it commercially the next year. "The K250 is considered to be the first electronic musical instrument to successfully emulate the complex sound response of a grand piano and virtually all other orchestral instruments" (Kurzweil Technologies n.d.). This patent has greatly contributed to the overall development of the music industry. Years after this program was created many additional discoveries were made using artificial intelligence including: "intelligent instruments; deeper, multifaceted representations for scores and sounds; intelligent musical data bases; singing and talking input with singing and talking output; a better understanding of human musical cognition and musical universals; new musical machines with capabilities beyond those of a single performer; more intelligent sound-analysis systems; performance systems capable of intelligent response to musical sound; and new and interesting compositional rule structures," (Roads 1980, 23).

Computational art eludes to the futuristic possibilities of artificial intelligence. Despite the opinions of many that a machine is not capable of creating art, current developments and examples in computational art have presented the world with a new medium of art. Computer scientist Donald E. Knuth (1995) suggests that "science is what we understand well enough to explain to a computer. Art is everything else we do. (...) and Science advances whenever an Art becomes a Science." Kurt continues this thread of thinking saying that, "by concerning this symbiotic relationship between art and technology, computational creativity is a field that explores the evolving correlation between human intelligence and machine intelligence" (Kurt 2018, iii).

Another significant moment in the history of combining artificial intelligence and music was the first International Computer Music Conference (ICMC) held in 1974 at Michigan State University in the United States.

> "The International Computer Music Conference — The ICMA[4] cosponsors the annual ICMC, which, since its inception in 1974, has become the preeminent yearly gathering of computer music practitioners from around the world. The ICMC's unique interleaving

who was a film star, actually guessed it, which was pretty insightful. Computers were not that well known at that time" (American Academy of Achievement 2018).

4 "The International Computer Music Association is an international affiliation of individuals and institutions involved in the technical, creative, and performance aspects of computer music. It serves composers, computer software and hardware developers, researchers, and musicians who are interested in the integration of music and technology. ICMA functions include: Presenting the annual International Computer Music Conference; Professional Networking; Publication of the newsletter array; Specialized Publications, Recordings and Projects; Sponsored Research; ICMA Commissions and Awards; Membership directory; Searchable database of all musical works performed at ICMC and/or commissioned by ICMA" (ICMA 2007).

> of professional paper presentations and concerts of new computer music compositions—refereed by ICMA-approved international panels— creates a vital synthesis of science, technology, and the art of music" (ICMA 2007).

In the rest of this paper, I will introduce artificial intelligence as an artistic phenomenon in relation to musical composition, musical performance, and music education.

Artificial Intelligence in Composition

Historically speaking, composing is the creative act of a human being acting as a composer by implementing his inspiration or inner ideas in a written form, forming a complete composition. But, is this still a valid definition in the 21st century – or even for the second half of the 20th century? The moment that the first computer that was able to compose music appeared, this definition became the subject of the questioning. Many wondered whether music composed by artificial intelligence could really be considered as originating from a creative process since the source of the composition is not human creativity, but instead software creativity.[5] However, people could overcome this barrier if the process of creativity was understood as a new way of behaving, which would include a piece of software (or one of its parts) that goes beyond the physical details of the program (Colton et al. 2015, 5). Since the origination of the first programs capable of composing music, no artificial intelligence system that composes music has become a part of the broader community of music, but now, in the 21st century, this is changing. With the rapid development of technology, artificial intelligence has enabled a faster flow of information, and thus faster ways of solving the problems we face in the digital world. Thus, the possibilities for developing newly advanced composer-software are much greater, as are the possibilities of its dissemination within the digital world. Some of the companies that are using artificial intelligence to produce music are: Popgun (2017), AIVA (2016), Melodrive (2016), Flow Machines (2016), IBM Watson Music(2016), ORB Composer (2015), Amper Music (2014), Humtap (2013), Jukedeck (2012), Ludwig 3.0 (2011), Chordpunch (2011), Google Brain: Magenta (2010), Google Experiments: Music and AI (2009), The Echo Nest (2005), Brain.FM (2003). The development of music produced by AI and the number of companies that participate in this development are significantly increasing. Numerous companies, ranging from small start-ups to key global players, have

5 "There is, of course, much progress still to be made technically, so that software can be creative and be seen to be creative, in order for consumers to be provided with valuable artefacts and enjoyable creative experiences. In addition to the technological hurdles faced, it is clear that certain sociological issues stand in the way of progress. That is, people naturally tend towards thinking that nuts-and-bolts, bits-and-bytes machines will never have a creative spark, and different sets of people instantiate this tendency in different ways" (Colton et al. 2015, 4).

seen the future in artificial intelligence. The platforms listed above are some of the most important developments for creating music and algorithms for automating the music composition. Considering the large number of platforms, the next sections will just focus on the AIVA platform that became the first virtual artist to have its creations registered with an author's rights society.

Artificial Intelligence Virtual Artist or AIVA, "is an AI capable of composing emotional soundtracks for films, video games, commercials and any type of entertainment content" (AIVA 2016). A project of Aiva Technologies, AIVA was founded in February 2016 by Pierre Barreau,[6] Denis Shtefan, Arnaud Decker and Vincent Barreau. The source of its information is a rich history of about 30,000 scores of musical composition written by composers such as Beethoven, Mozart, Bach, etc. Learning from the significant contributors to musical history helped AIVA capture the concepts of music theory and understand the art of music composition. Also, that helped AIVA, "to create a mathematical model representation of what music is. This model is then used by Aiva to write completely unique music" (Ibid.). Along with the reinforcement of deep learning algorithm techniques found in TensorFlow, AIVA uses CUDA, TITAN X Pascal GPUs, and cuDNN. But, AIVA is still only able to compose for piano, the orchestration, arrangement, and production of the music require human skills. Also, it is important to emphasize that AIVA uses GPU computing, which has created a plagiarism checker, which is able to understand whether a created track partly or fully plagiarized from the database AIVA learned from. Besides that, several Turing tests completed with music professional participants have confirmed that the compositions of AIVA can't be differentiated as human or AI creation. The registration of its creations in an author's rights society, SACEM,[7] the author's rights society for France and Luxembourg, made it the first software to be recognized for creating unique works both formally and officially. Never before has there been a case where artificial intelligence is recognized by a rights society as a composer. Creators of AIVA say that it, "is able to write beautiful and emotional music, a deed that is considered to be deeply human" (Medium 2016).

The process of composing which AIVA uses is quite different from the process followed by human composers. It uses deep neural networks to look for patterns and rules in compositions and uses this information to learn the basics of style and music. While composing AIVA predicts what should come next in the track. After

6 Pierre Barreau is an entrepreneur, computer scientist, composer, director and chief executive officer as well. The idea of creating a virtual composer from him comes after watching the science fiction movie "Her" (The film is about a super intelligent form of AI that cannot take physical form) in which the AI composer composes a musical piece. After that, he founded AIVA with few people close to him and with similar interests.

7 "Since 1851 SACEM (Society of Authors, Composers and Publishers of Music) is a non-profit non-trading entity owned and managed by its members according to the business model of a co-operative. To guarantee authors' intellectual property rights over their works, SACEM collects and distributes royalties, thereby playing a crucial economic role to preserve musical creation" (SACEM n.d.).

this happens, she creates a set of mathematical rules for a particular style of music, and then, she is ready to compose (Aiva Technologies 2017). However, this raises the question of how advanced this type of technology is and if it can fully replace the living composer.

> “Storytellers, film producers and the whole entertainment industry rely on music to turn moments into magic. But interactive content such as video games have hundreds of hours of gameplay, and typically only two hours of music, since human composers are limited by physical constraints. And rather than looping the same tunes over and over again, AI gives an alternative to augment human creative abilities, so that games may finally have hundreds of hours of original music” (Ibid.).

So, it seems that it will take a very long time for learning algorithms to be able to replace the human composer.

Immediately after the software was founded, AIVA composed its first composition on February 8, 2016, called *Genesis*, which is also the name for the first album that was produced with its compositions. All compositions from the album were produced by CEO Pierre Barreau. Composition Genesis is formally constructed in a variational form with epic character, which was achieved in the orchestral version of this piece. One can recognize the remarkable similarities between this composition and those created in the epic style of German composer and producer Hans Zimmer, whose music exudes masculinity, heroism, and military sound, most often expressed through the use of a great orchestra, as well as a large number of percussive instruments.

Figure 1: First page score of the composition Symphonic Fantasy in A minor, op. 21, *Genesis*

The composition consists of 38 bars divided into four parts, of which the first part (A) presents the main theme played by the bassoon and Horn in the key of F, while the other three parts (B, C and D) represent its variations, with the main thematic material being predominantly played by the string section.

After classical music, AIVA expanded its compositions to the entirety of World music (Tango, Middle Eastern, Chinese, etc.) as well as Rock music. One of AIVA's World Style albums is 艾 娲 (ài wā),[8] inspired by the music of China.

“Much like film music has a single theme reprised several times throughout the soundtrack, we wanted to apply this idea of theme variation to our Chinese Album and have some repetition of themes present in different tracks. To achieve this, we applied techniques of deep learning to create musical variations of a theme in the Chinese album presented above. For this, we give a musical theme in MIDI format to our variation module, that then iteratively applies changes to the harmony, melody and rhythm of the piece, while staying faithful to the original theme, but moving away further from it with each iteration” (Medium 2018).

AIVA's second album was released in 2018 and contains 24 compositions from various genres.

Although the achievements of artificial intelligence in the field of composition have been impressive so far, there is still much to do if we want artistic intelligence to live up to the same abilities as a human composer. At this point, artificial intelligence still needs help while composing, mostly in orchestration and music production. But what is a very important development, is that in the field of music production artificial intelligence has reduced the time spent on repetitive tasks, especially in games. It is important to emphasize that artificial intelligence still lacks that human creative element which is very important in art. Artificial intelligence is currently just a means by which composers can find new ways of composing. Its role is currently solely dependent on the human or composer.

Artificial Intelligence in Performance

As noted earlier, artificial intelligence has played an important role in the development of computer generated music (although most efforts have been on compositional and improvisational systems). But, little effort has been devoted to using AI in performance systems. This may be due to the fact that performing, improvisation, and music composition are seen as a creative and autonomous manifestations that are not feasible for a machine. However, we cannot resist the fact that artificial intelligence is slowly becoming a nucleus of computer music. Every day through the development of new algorithms, it offers us new approaches and new knowledge needed for solving music problems.

> “When thinking about issues relating to performance (in particular) it is important to think about the way that the audience interacts with the performer and the performance, enabling the audience to fully

8 “AIVA’s Chinese name ‘艾娲’ (ài wā), also the title of this album, has a specific meaning in Mandarin. The character ‘艾’ (ài) means "elder", and here represents the extensive knowledge that our AI has learned from the greatest composers in history. The character ‘娲’ (wā) is from Nüwa ‘女娲’ - the mother goddess of Chinese mythology” (Aiva 2018).

> understand the systems that are being used, could offer performers, designers and technologists a whole range of new and exciting performative possibilities that go beyond the tradition performer – consumer paradigm and offer new ways in which the audience can effect and interpret the system used" (Chamberlain 2017).

Every day, interest in having computers perform musical tasks, by playing music, processing music, or creating music is increasing. So far, computers have only been following the commands of a human to generate and perform the compositions. In live music performance, a computer and a man collaborate, "as a versatile tape machine—playing its part (which could include multiple instruments and lines) with admirable virtuosity but in blissful ignorance of the circumstance of the performance and its fellow musicians" (Baird, Blevnis and Zahler 1993, 73). The problem with the collaboration between computer and human performers is that the computer is not able to monitor abrupt tempo changes, dynamics, articulation, or basically cooperate with the human on a high level. The computing performer should be able to recognize the mistakes of other performers and make alterations in the performance to compensate for the mistakes of others (Ibid.). But, besides these current limitations, the computer is a good accompanist for a human performer, as long as it is not the leader.

One of the last (incredible) moves in the application of artificial intelligence in musical performance is Yamaha's project, where the world-renowned dancer Kaiji Moriyama controls a piano using dance movements. The concert presentation of the project entitled "Mai Hi Ten Yu" was in Japan on November 22, 2017, and it was

Figure 2: Live performance, Mai Hi Ten Yu, 2017. © Ayane Shindo

sponsored by Tokyo University of the Arts and Tokyo University of the Arts COI.[9] The emphasis in the performing process is on the collaboration between artificial intelligence and dancers, more specifically in this case, Kaiji Moriyama. Throughout the performance, he was accompanied by the Berlin Philharmonic Orchestra Scharoun Ensemble as musical support. The original system that Yamaha uses, turns human movements into musical expression through AI technology, as a technical collaboration for performance.

> "The AI adopted in the system, which is now under development, can identify a dancer's movement in real time by analyzing signals from four types of sensors attached to a dancer's body. This system has an original database that links melody and movements, and, with this database, the AI on the system creates suitable melody data (MIDI) from the dancer's movements instantly. The system then sends the MIDI data to a Yamaha Disklavier™[10] player piano, and it is translated into music. To convert dance movements into musical expression, the Yamaha Disklavier™ is indispensable because it can reproduce a rich range of sounds with extreme accuracy through very slight changes in piano touch. Moreover, we use a special Disklavier in the concert which was configured based on Yamaha flagship model CFX concert grand piano to express fully and completely the performance of the talented dancer Moriyama" (Yamaha 2018).

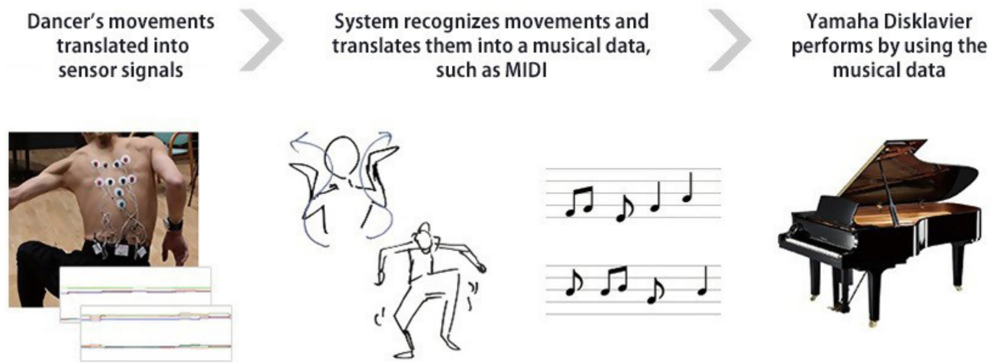


Figure 3: : Graphic representation of converting dance movements into musical expression

With this technology, dancers no longer have to adapt or adjust to music, instead music would adjust itself to the dancer's movement. Although this project was presented to the public, Yamaha's technology is not fully ready to be commercialized

9 See https://www.youtube.com/watch?v=tLFe2AzCodk, Accessed March 9, 2019.

10 "An original system developed by Yamaha, the Disklavier™ is a hybrid player piano that performs automatically and can reproduce the movements of the keyboard and pedal with extreme accuracy. The latest model, the Disklavier™ ENSPIRE, includes recordings of the original sound sources of world-famous artists. This feature enhances the enjoyment of musical performances that users can experience at their leisure" (Yamaha 2018).

for massive practical use or as a new form of artistic expression. But it is certain that once this kind of performance software is released to the public, it will bring significant changes to the field of music and dance as well create new dimensions of artistic process.

Artificial Intelligence in Music Education

The use of computer technology in education began in the 1960s and 1970s, shortly after the emergence of artificial intelligence, with the development of what was called Intelligent Computer-Assisted Learning (ICAI) by Carbonell (1970). Soon after that, in 1982, Sleeman D. H. and Brown J. S. coined the term "Intelligent Tutoring Systems"[11] in their volume of the same title (Nkambou, Bourdeau and Mizoguchi 2010, 2). Today, AI has entered into almost all spheres of human learning, and the possibilities of its application for educational purposes continue to increase. Simon Holland lists several definitions from various authors regarding the role of artificial intelligence in education (AI-ED). The first one is that AI in Education is any application of AI techniques or methodologies to educational systems. A second, more narrow definition says that AI-ED is any computer-based learning system which has some degree of autonomous decision-making with respect to some aspect of its interaction with its users. The third definition of AI-ED that Holland points out is the use of AI methodologies and AI ways of thinking being applied to the discovery of insights and methods for use in education, whether or not AI programs are actually involved at the point of delivery (Nkambou, Bourdeau and Mizoguchi 2010, 2).

Each of these definitions refers to hypermedia i.e., multimedia contents that are interconnected. "Historically, the first use of computers in teaching music or teaching any other subject for that matter, was associated with programmed learning, derived originally from behaviorism" (Holland 1989, 22). AI in music education has long been present in schools and we are often not even aware of how much it is used. Music is a challenging field for Artificial Intelligence in Education because it requires creativity and problem solving from the students and teachers. However, the effectiveness of hypermedia in teaching is greater than the effectiveness of a classical teaching method where the teacher talks and the students listen. Many studies have confirmed that a person is able to remember about 20% of information disseminated if they only heard it, 40% if they saw and heard it, and 75% if they saw it, heard it, and actively used it. By introducing a hypermedia paradigm for creating a learning program, the motivation of students to learn is greater because of the use of different non-textual media.

Artificial intelligence in music education is at its nascent stages. Its presence

11 "Intelligent Tutoring Systems (ITSs) are complex computer programs that manage various heterogeneous types of knowledge, ranging from domain to pedagogical knowledge. Building such a system is thus not an easy task" (Nkambou, Bourdeau and Psyché 2010, 361).

in schools is still lagging behind the use of AI in other subjects and this is also the case for the teaching of musical culture or instruments in general. The reasons for this can be economic or political, but also generally social. Unfortunately, society, especially in underdeveloped countries, refuses to accept AI as a part of our lives. However, there are many useful ways to apply AI to music education. Today there are thousands of good quality programs / apps that the teacher is able to use while teaching Music, Music theory, Solfeggio, Harmony, Counterpoint, etc. Many music programs that were developed to make Music class more interesting and productive are often only applied on an individual basis. Some of educational hypermedia music programs whose purpose is to train the recognition, or reproduction of intervals, chords, or melody scales, modes, durations, tunings, and flexible melodic dictation are Ear Trainer, Interval, Listen, MiBAC Music Lessons, Seventh Heaven, Perceive, Practica Musica, and MacGAMUT. A lot of these applications have clear educational goals, but they are still not being used in schools. One program which offers tuition in basic music theory is Practice Room. In addition to these virtual programs many generally useful musical computer tools are also applicable to education: analysis tools, innovative musician interfaces, music editors, sequencers, computer - aided composition tools and multimedia reference CD-ROMs on masterworks (Holland 2000, 2). Music Teachers should be aware of the presence of these programs and use them as aids for teaching. Of course, it would help if schools were equipped with computers or electronic equipment to make classes faster and more contemporary. It seems as if it will take some time for this approach to become the typical method of teaching music. It is important to note that for many applications, especially those that study the field of harmony and counterpoint, there is scope for improvement of existing works. Therefore, teachers are bound to recognize the limitations of the software, and be aware of whether they are appropriate for use. Nevertheless, at the moment, artificial intelligence is certainly a good means for self-learning music.

Conclusion

Since the beginning of the 21st century, we have witnessed an enormous increase in the use of artificial intelligence in science and art. Its development has led to change in human thought, due to an understanding of communication technology that allows us to ignore the confines of geography, while compressing space and time constituting a virtual reality. With the more frequent use of artificial intelligence in the field of music, some musicians feel threatened, while others remain amazed at the opportunities that it uncovers. The fact is that artificial intelligence has become an inevitable factor in our future. Today, many artists collaborate with artificial intelligence in creative endeavors and in the future, this link between artists and artificial intelligence will become even more recognizable. All the strides in artificial intelligence discussed in this paper are just the beginning of a new era, which will truly begin in the next decade of this century. Virtual artists like AIVA will become the reality of our everyday lives. One of the expected improvements that will surely happen is in the field of musical performance and of dance, as we have seen from the collaboration of Japanese dancer Kaiji Moriyama with AI technology. If we are still wondering whether artificial intelligence can create art that emulates human emotions, the answer is yes. If we have not convinced ourselves yet, surely the future of this technology will prove to us that this is possible.

List of references

Aiva Technologies. 2017. "Press release." Accessed February 28, 2019. https://www.luxinnovation.lu/wp-content/uploads/sites/3/2018/01/fundraising-en-aiva.pdf.

Baird, Beidget, Blevnis Donald, and Zahler, Noel. 1993. "Artificial Intelligence and Music: Implementing an Interactive Computer Performer." *Computer Music Journal*, No. 2, Vol. 17: 73–79.

Barr, Avron, and Feigenbaum A. Edward. 1981. *The handbook of artificial intelligence*. Stanford: William Kaufmann.

Berz, William L., and Bowman, Judith. 1995. "An Historical Perspective on Research Cycles in Music Computer-Based Technology." *Bulletin of the Council for Research in Music Education*, No. 126 (Fall 1995): 15–28.

Chamberlain, Alan. 2017. "Are the Robots Coming? Designing for Autonomy & Control in Musical Creativity & Performance." Audio Mostly 2017: Augmented and Participatory Sound/Music Experiences, 23-26 August. Queen Mary University of London (London, UK) ACM - DOI https://doi.org/10.1145/3123514.3123568 UoN ePrints

Colton, Simon, Pease, Alison, Corneli, Joseph, Cook, Michael, Hepworth, Rose, and Ventura, Dan. 2015. "Stakeholder Groups in Computational Creativity Research and Practice." In *Computational Creativity Research: Towards Creative Machines*, edited by Tarek R. Besold, Marco Schorlemmer and Alan Smaill, 3–36. Paris: Atlantis Press.

Hayes J. Patrick, and Morgenstern, Leora. 2007. "On John McCarthy's 80th Birthday, in Honor of His Contributions". *AI Magazine*, No. 4, Vol. 28: 93–102.

Holland, D. Simon. 1989. "Artificial Intelligence, Education and Music: The Use of Artificial Intelligence To Encourage and Facilitate Music Composition by Novices." PhD diss., The Open University.

Holland, D. Simon. 2000. "Artificial Intelligence in Music Education: A Critical Review." In *Readings in Music and Artificial Intelligence*, edited by Eduardo Reck Miranda, Vol. 20: 239-274. https://pdfs.semanticscholar.org/85a6/1705375edb5e2cc6090563168c46e77f215a.pdf?_ga=2.158404858.1159493817.1553204540-18698389.1553204540.

Kurt, E. Deniz. "Artistic Creativity in Artificial Intelligence." Master degree, Radboud University.

Nkambou, Roger, Bourdeau, Jacqueline and Psyché, Valéry. 2010. "Building Intelligent Tutoring Systems: An Overview." In *Advances in Intelligent Tutoring Systems*, edited by Roger Nkambou, Riichiroa Mizoguchi and Jacqueline Bourdeau, 361–375. Berlin-Heidelberg: Springer-Verlag. https://www.researchgate.net/

publication/226301831_Building_Intelligent_Tutoring_Systems_An_Overview.
Roads, C. 1980. "Artificial Intelligence and Music." *Computer Music Journal*, No. 2, Vol. 4: 13–25.
Nkambou, Roger & Bourdeau, Jacqueline & Mizoguchi, Riichiro. 2010. Introduction: What Are Intelligent Tutoring Systems, and Why This Book?. In *Advances in Intelligent Tutoring Systems,* 1–12. Heidelberg: Springer-Verlag Berlin Heidelberg.

Website content

AIVA. 2016. "AIVA - Artificial Intelligence Virtual Artist." Accessed February 26, 2019. https://www.aiva.ai/about#about.
Aiva. 2018. "AIVA - 艾娲 (ài wā) - 30 min playlist of AI generated music." Accessed March 5, 2019. https://www.youtube.com/watch?v=Bd9p2oi760g.
American Academy of Achievement. 2018. "Ray Kurzweil – Inventor and Futurist." Last modified June 8, 2018. http://www.achievement.org/achiever/ray-kurzwell/#biography.
Kurzweil Technologies. n.d. "Kurzweil Music Systems." Accessed February 25, 2019. http://www.kurzweiltech.com/kms.html.
Medium. 2016. "Composing the music of the future." Accessed February 28, 2019. https://medium.com/@aivatech/composing-the-music-of-the-future-4af560603988.
Medium. 2018. "How we used our Music Engine to create the first AI-generated album of Chinese Music." Accessed March 5, 2019. https://medium.com/@aivatech/how-we-used-our-music-engine-to-create-the-first-ai-generated-album-of-chinese-music-9d6fa984b4e8.
SACEM. n.d. "SACEM – In a nutshell." Accessed February 26, 2019. https://societe.sacem.fr/en.
Roberts, Jacob. 2016. "Thinking Machines: The Search for Artificial Intelligence. Does history explain why today's smart machines can seem so dumb?" Science History Institute. Accessed February 21, 2019. https://www.sciencehistory.org/distillations/magazine/thinking-machines-the-search-for-artificial-intelligence.
The International Computer Music Association. 2007. "About ICMA." Accessed February 26, 2019. http://www.computermusic.org/page/3/.
Yamaha. 2018. "Yamaha Artificial Intelligence (AI) Transforms a Dancer into a Pianist. – Dancer: Kaiji Moriyama, Accompanied by the Berlin Philharmonic Orchestra Scharoun Ensemble." Accessed March 9, 2019.
Yamaha Corporation. 2018. "Yamaha Artificial Intelligence (AI) Transforms a Dancer into a Pianist - Full version." YouTube video, March 9, 2019. https://www.youtube.com/watch?v=tLFe2AzCodk.

HOW AI CAN CHANGE/IMPROVE/INFLUENCE MUSIC COMPOSITION, PERFORMANCE AND EDUCATION: THREE CASE STUDIES
(Summary)

This paper shows the connection between artificial intelligence and its influence on composition, musical performance, and musical education. These three case studies have been discussed through the lens of the previous research carried out in this field and the research done specifically for each new development. Additionally, there is a unique case for the author's free choice. The first case is about Artificial Intelligence Virtual Artist or AIVA, the first virtual composer to have its creations registered with an author's rights society (SACEM). In the aforementioned case, the beginnings of AIVA were touched on, as well as its creations, past achievements, and the future goals for AIVA to compose direct orchestral music. The second case is about Yamaha's project where the world-renowned dancer Kaiji Moriyama controls a piano using his dance movements. This section talks about the process of the project itself and its important role for future research, considering it was the first collaboration between AI, music, and movement of its kind. The third part focuses on the importance of music programs and applications in education. Finally, the conclusion speaks to the importance of artificial intelligence for the future of humanity, primarily in the field of music, and some assumptions on which artificial intelligence will be based on in the future.

Article received: April 13, 2019
Article accepted: May 15, 2019
Review article

UDC 7.036:004.8

Daniel Becker*
Bibliotheca Hertziana – Max Planck Institute for Art History
Rome, Italy

THE INTELLIGENT WORK OF ART: HOW THE NARCISSUS THEME ECHOES IN NEW MEDIA ART

Abstract: This paper will deal with the mythological figure of Narcissus in new media art. In visual arts in general, this myth is usually used to reflect on the relationship between the artist and his actual work. There are countless examples of artists from antiquity to the present age that deal with subjectivity in their work by recurring the Narcissus theme. But different to those adaptations, works of the New Media Art since the 1970s reflect more about the technology and subjectivity of the observer through the theme of Narcissus. The use of time-based media allows the artists to address the observer immediately through interaction and let him become a part of the work and therefore become a part of the cognitive process. The argument of this paper is that only through the use of time-based art could the self-awareness of the observer be discussed instead of only a reflecting on the work itself and the reception-process. Against this backdrop, the paper will focus on the use of AI as a 'material' in contemporary art and how it extends this cognitive process. In addition to other works from the history of new media art the work *Narciss* (2018) by the German art collective Waltz Binaire will be in the center of this discussion about AI *in* and *as* artistic practice.

Keywords: artificial intelligence, consciousness, media-awareness, Narcissus, new media art, self-awareness, Waltz Binaire

*Author's contact information: danielbecker2000@gmx.de

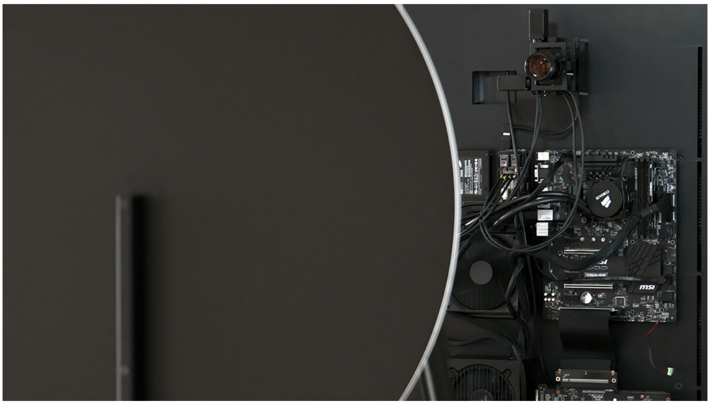


Figure 1: Waltz Binaire, *Narciss*, 2018, installationshot.
© The artists.

Figure 2: Waltz Binaire, *Narciss*, 2018, installationshot.
© The artists.

Narcissus as a theme in art

"[…] the project Narciss uses this scientific milestone to raise a question at the core of human experience: What do we always look at, but never fully understand? Ourselves"
(Waltz Binaire 2019)

Like the claw in a vending machine this work is looking for something. The robot-arm goes up and down, to the left and to the right, at its end a camera lens is looking into a mirror. It is whirring when it does so, and it seems like it is seeking something. This work described above is called *Narciss* (2018) by the German art collective Waltz Binaire. It is an AI based installation that tries to find itself (Fig.1 and 2).

The *Narciss* installation consists of common computer parts and a movable camera lens mounted on a vertical rectangle. A screen is located on the back of this construction. Opposite to the bare computer equipment, connected by a metal frame, is a circular mirror. It cannot be denied that the two opposing forms can remind one of a one (1) and zero (0), as the artists themselves write: "Narciss is a robot, built to analyse its own physical embodiment. Its design uses a reduced visual language with a high density of intended associations. The two opposite components, a circular mirror (O) and a computer (I), are designed to symbolize the duality of input and output (I/O)" (Waltz Binaire 2019).

Visual minimalism, the unveiled display of the hardware, as well as the reference to the binary of 1 and 0, make clear the intention that this should be a technological archetype. Here, attention should not be drawn to a design, nothing should be hidden, but the viewer should concentrate on what Narciss is actually doing: looking into the opposite mirror with the camera eye, interpreting this image of itself using an image recognition software, and showing this interpretation in the form of text on a screen. "By constantly panning and zooming, Narciss receives a feed of different perspectives and sub-regions of its hardware. This restless choreography resembles an urge of intention, a never-ending curiosity, and the self-looping nature of narcissism" (Waltz Binaire 2019).

In this way *Narciss* deals with fundamental questions about consciousness and self-awareness in the context of humanity and AI. Therefore, by choosing this title for their work the artists place their work into a widely traditional category of art, as the Narcissus theme is central to western culture concerning consciousness and self-awareness. Fundamentally, it is based on a depiction of a man written by Ovid in the third book of his *Metamorphoses* (Ovid 1922, 337–434). Although the story is fairly well-known, it seems reasonable to recall it briefly here.

The myth tells the tragic story of the young Narcissus, son of the nymph Liriope and the river god Cephisos, whose appearance was admired by everyone. However, Narcissus spurns his countless admirers and harshly rejects the love of the nymph Echo. Disappointed and angry, the rejected nymph curses him: he shall be destined to the same fate as her, he shall fall in love just as eternally and this love will remain unrequited. Nemesis, the goddess of revenge, fulfills this wish by making Narcissus fall in love with his own reflection, which he sees in a pond while hunting. Instead of recognizing it as his reflection in the surface of the water, Narcissus considers the object of his desire to be someone else. Only when the image becomes blurred when Narcissus attempts to embrace this stranger by touching the water does he realize that it is his own image and that his love can never be reciprocated. In mourning and disappointed at this fate, his strength leaves him, and he dies on the shore of the pond. His body disappears and at the place where he had kneeled in front of the pond, a crocus-colored flower grows instead, this flower is then called the narcissus.

There are a lot works that explore the meaning of Narcissus in Western culture, but I like to concentrate on its impact on contemporary new media art like *Narciss* by Walt Binaire.[1] To do so, one has to understand that there are two major works of

1 For the impact by the Narcissus theme of art and culture in general see for example: Barolsky, Paul.

art which defined the discourse surrounding Narcissus until the later 20th century: one created by Michelangelo Merisi, known as Caravaggio, and one created by Salvador Dalí.[2]

In *Metamorphosis of Narcissus* (1937) Dalí makes it possible to experience the transformation from man to flower in the two-dimensionality of an image. The left half of the picture schematically shows a naked man kneeling in the water with his head lowered. Like the rocky landscape in the background, he is kept in earth tones and only stands out through illuminated features. The right side of the picture, on the other hand, is dominated by a stone-like hand holding, in its fingers, an egg from which a narcissus grows. Its form and contour are similar to that of the kneeling figure; this repetition of form occurs a third time in the rock formation found in the background of the upper righthand corner of the picture. The psychoanalytic symbolism of the egg, the ants, or the carrion-eating dog is evident and widely-known (Lomas 2011, 27 et seq.). In this context it is more important to examine how Dalí unites the different stages of the metamorphosis in the pictorial plane. Here, the stages of the metamorphosis are depicted simultaneously within the pictorial space but cannot be perceived at the same time. Only from a distance can the figure on the left be recognized as a kneeling person; the figure on the right, however, can only be recognized as a hand and not just as a rock formation at a close range. The third repetition of the form can only be seen by a concentrated view of the background, which limits the field of vision accordingly and ignores the other two forms. Therefore, Dalí breaks with the perspective axiom of the eye point by applying a process that takes into account the spatial situation of the observer, as can be also found in his other optical illusions and stereoscopic images (Lomas 2011, 30 et seq.).

Dalí illustrates the metamorphosis by transferring the process of transformation to the viewer. The process of observation is thus no longer that of Narcissus, who observes his mirror image, but the process in which the recipient observes Narcissus.

Louise Vinge also understands Dalí's representation of narcissus in a classical iconographic tradition. For Vinge, the kneeling pose, the reddish hair, and the bent arm posture can be traced back to Caravaggio's *Narcissus* (1594–96) (Vinge 1966, 44). This reference certainly makes sense, since Caravaggio's *Narcissus* was often

1955. "Very Brief History of Art from Narcissus to Picasso". *The Classical Journal*, Vol. 90, 255–259. Cole, Michael. 2002. "The Demonic Arts and the Origin of the Medium". *The Art Bulletin*, Vol. 84, 621–640. Sennett, Richard. 1977. "Narcissism and Modern Culture". *October*, Vol. 4, 70–79.

2 "This Narcissus [by Caravaggio, D.B.], which for many years received virtually no attention, represents the best-known pictorial representation of the myth and the one that is most germane to our understanding of the theme today – with the probable exception of Metamorphosis of Narcissus [by DalÌ, D.B. …], which was created 350 years later. Caravaggio was considered to have based the details of his depiction 'too closely' on Ovid's account. At the time, no-one was interested in emulating the intense portrayal of the connection between Narcissus and his reflection, whilst his external appearance was rejected as being too coarse. Only when Freud's theories of narcissism began to be recognized did people realize that in this painting Caravaggio had shown himself to be far ahead of his time, and that his Narcissus, like modern humanity, was seeking self-knowledge" (Welsch 2012, 23).

understood as a homoerotic motif and corresponds to such psychoanalytic concepts as drive and repression which were frequent themes dealt with in Surrealist Art.

At the same time, Caravaggio emphasizes the boundaries between Narcissus and his mirror image like no other work of art has. The water edge of the spring forms the horizontal central axis of the picture, almost cutting it into two symmetrical halves. Narcissus and his mirror image are also separated by color. While the mirror image is only a shadow image that almost disappears in the darkness of the water, Narcissus stands out from the unspecific black background through the use of light accents found in Caravaggio's chiaroscuro painting. His bright robe, his illuminated arms, and his protruding knee are only partially reflected by the water.

Caravaggio holds Narcissus in suspense, he leaves him undefined, because it is not clear whether Narcissus bends forward, closes his eyes completely, and approaches the surface of the water for a kiss, or whether his left hand touches the water making waves and tearing Narcissus from the illusion causing him to wrench open his half-closed eyes in horror. In Caravaggio, Margrit Brehm, too, does not see a juxtaposition of different points in time, but rather understands Narcissus as in a 'floating' moment, which becomes comprehensible through the "tense posture in which the ambivalence between the pull exerted by the mirror image and the powerful support of oneself, i.e. the counter-movement, becomes apparent. This magical attraction is also emphasized by the closed circular form shaped by the real and mirrored arms in the compositional scheme" (Brehm 2001, 338).[3]

Starting from this ambivalent composition, Christiane Kruse discusses Caravaggio's *Narcissus* with regard to the difference between the mediator of representation and the theme of representation. In this way, she describes the artist as a media theorist and reception aesthetician at the same time. She understands the circular figure formed by the arm position of Narcissus and his mirror image as "the imprisonment of the ignorant boy in his own illusion" (Kruse 2003, 342). The void that arises within this circular figure is dominated by the "oversized phallus" that "pushes itself like a barrier between the youth and his image," Kruse continues (Kruse 2003, 342). However, a space opens up between the picture and the image, but it does not fall into an illusionistic depth but into a plane of undefined darkness. This can be understood through Kruse's description of the ambivalence of painting and spatial illusion as the thematization of the medium and its reception, and according to her, the "cool transparency of the mirror image and the opacity of the canvas soaked in impenetrable black [function as] polar guiding metaphors of the medium that thematizes itself" (Kruse 2003, 343). At the same time, Caravaggio involves the viewer in the sense of the *rilievo*, in that the young man stands out from the flat black background through the virtuoso *ciaroscuro*. It is not the moment of recognition nor the tragedy of the myth that is emphasized here, for here all the matter of touching the water is negated. Instead, the desire is emphasized as the boy's body and knees stand out, which is also supported by the common biographical-

3 This and the following german quotes are translated by the author.

homoerotic reading of this subject.

Contrary to this more stereotypical reading, Kruse states that Narcissus believes he looks at someone else; unlike the viewer, he does not recognize that it is his mirror image. She understands the Narcissus myth as a process of media knowledge. The first stage of this process is ignoring of the medium, in which Narcissus fails by perceiving the illusion as real; in the second stage, Narcissus recognizes his mirror image as a mirror image, but against his better judgment he allows himself to continue to be deceived and thus moves only in an aesthetic world; in the third stage, Narcissus recognizes his mirror image but he no longer allows himself to be deceived, instead he questions it with regard to its medial qualities and reflects on its function as mediator between the real and aesthetic world. Whether Kruse believes that Narcissus actually reaches this third stage of reception is not clarified because she denies that he has this level of insight at the beginning: "He [Narcissus, D.B.] not only knows no mirrors, no mirror images, he knows nothing of the laws of catoptrics, he also knows no visual media. [...] In other words: Narcissus lacks what I would like to call media-awareness" (Kruse 2003, 309–310).

Self-awareness and media-awareness

Marshall McLuhan, one of the great media theorists of the 20th century, finds that the figure of Narcissus has importance in the context of media, too:

> "The Greek myth of Narcissus is directly concerned with a fact of human experience. As the word Narcissus indicates, it is from the Greek word narcosis, or numbness. The youth Narcissus mistook his own reflection in the water for another person. This extension of himself by mirror numbed his perceptions until he became the servomechanism of his own extended or repeated image. The nymph Echo tried to win his love with fragments of his own speech, but in vain. He was numb. He had adapted to his extension of himself and had become a closed system" (McLuhan 2001 [1964], 45).

Like Kruse, McLuhan describes a similar imperfection and lack of media-awareness held by the figure of Narcissus. The inability to recognize himself as a mirror image on the surface of the water leads Narcissus to becoming trapped in a cycle causing him to appear, at least from the outside, as narcotized, numb, or unconscious. In this respect, self-awareness and media-awareness are mutually dependent. Only by being aware that the surface of the water is a reflection could Narcissus gain awareness that he is looking at an image of himself. Such a reading is precisely the opposite of the psychoanalytic understanding of narcissism that one is trapped in admiration for oneself. McLuhan also enables the persistence of negative interpretations with his reference to narcosis and numbness. However, his contextualization of the narcissist within media theory allows a way out of this stereotypical interpretation that appears to be primarily understood by artists.

Since the 1970s, artists have employed this theme in their work, especially with regard to new technologies. The relationship between the self and medialization is particularly virulent in the close-circuit installations like the ones created by Nam Jun Paik and Dan Graham. In the context of the raising popularity and usage of video and media art since this time, Rosalind Krauss speaks about video as an "Aesthetic of Narcissism" (Krauss 1976).

In contrast to the "reflection" of material art, such as paintings, which deal with the relationship between image and image carrier and thus move at the threshold of two entities, video art is more often characterized by its "reflexiveness" (Krauss 1976, 56). Therefore, she understands technical-material components of the video (camera, monitor, etc.) not as actual objects, but as a constellation. In this constellation the artist is not reflecting about the technical equipment itself, but he is instead constantly experiencing reflexiveness of his own image, because this new media gives him endless feedback:

> "Unlike the other visual arts, video is capable of recording and transmitting at the same time – producing instant feedback. The body is therefore as it were centered between two machines that are the opening and closing of a parenthesis. The first of these is the camera; second is the monitor, which re-projects the performer's image with the immediacy of a mirror" (Krauss 1976, 52).

The technical closed-circuit situation in which performance and recording actually happen is, for Krauss, the analogy of Narcissus' mythological viewing situation; the situation in which Narcissus is trapped and from which he is only able to escape with his death corresponds to the 'closed circle' of the video installations, in the endless loop of media in which the performer finds himself. Although Krauss mentioned Lacan and the viewing situation she is just focusing on the aspect of production and not of the actual exhibition of video art. The double-bind of the mirror stage – to oneself as a whole but at the same time as some else – which Krauss assumes to be the starting point for the "Aesthetic of Narcissism," therefore cannot be applied to the medium of video in general but just to the special case of a close-circuit performance, in which an actual encounter with a media counterpart takes place. After all, the actual figure of Narcissus does not play a prominent role in this context because it lacks to consider the viewer. Moreover the focus lies on the psychoanalytical dimension of narcissism and its relevance concerning the artist. But instead of Narcissus as subject, Narcissus as observer has become an increasingly evident theme in art since the 1990s and at a time when digitality determines the discourse about new media art.

In Nicolas Anatol Baginsky's work *Public Narcissism* (1999–2001), for example, visitors were filmed on the escalator of the VW Group's World of Experience in Wolfsburg, from which a face-finding software extracts individual faces and makes them into portraits. These portraits are then shown on one of the displays next to the escalator, so that the visitor is suddenly confronted with their own portrait. Next, these portraits are assigned to classes and superimposed so that they create an oversized image of a chimera which appears on another big screen. The fact that the work focuses on Narcissism – and not Narcissus – makes sense insofar as the visitors, who see their own portrait on the screens, recognize themselves and, unlike the mythological youth, do not mistake themselves for someone else. In this way, the work deals more with the psychological component of reception, evoking emotions such as surprise, ecstasy, shame, and also fear of surveillance in the visitors. However, in the superimposed portrait, the visitor is confronted with something different. Based on the classification, the software generates a portrait that theoretically contains parts of each visitor at that time. Because this picture no longer shows a clear face, the viewer becomes aware that the digital calculation does not correspond to the human demand for a face, but to an ideal of beauty determined by the AI. The media-awareness comes precisely from the fact that the viewer is aware of himself and can thus distinguish himself from the digitally calculated model.

A similar effect can be found in the work *Liquid Views – The Virtual Mirror of Narcissus* (1992–1993/2008) by the German artists Monika Fleischmann and Wolfgang Strauss, but on a different, more sensual level. In this installation, the recipient sees themselves placed in the actual role of Narcissus by viewing and touching their own mirror image on a screen. Through the technology of the touchscreen, the installation registers the touch and animates the virtual mirror image with ripples. Like a "tactile shock" (Becker 2016, 95), the user now understands that the work is not a simple close-circuit-installation, but that his portrait is filmed, rendered, and animated – without a temporal delay being perceived, it all happens synchronously to the moment of touch – making the viewer aware of the technical constellation. This form of human-machine-interaction was by no means familiar at the beginning of the 1990s, since at that time a large computer in a separate room had to be borrowed for the technical implementation. The narrative setting with the narcissistic subject and the animation of the familiar image of oneself therefore served to familiarize people with the new and strange technology. In this case, it is the shock or eureka effect, in relation to Jacques Lacan's mirror stage, which shows the user the condition and abilities of technical media in the digital age.

In these works, where the recipient is integrated in and basically slips into the role of Narcissus, self-awareness and media-awareness are mutually dependent, and maybe even more so, as only through the consciousness of self-reflection do the media conditions become conscious.

Me, myself, and AI

Narciss by Waltz Binaire is different, because the recipient is not a user who participates in or interacts with the work, but rather a viewer who watches the process of an observation. And yet, the work differs from classical art, such as the piece by Caravaggio, where one 'only' looks at the represented Narcissus. More than any other media in art, electronic media, through its time-based nature, allows processes to be thematized. Not just a moment, even if the moment can be embedded in an anticipatable narrative, but the transformation through the before-after can be represented, as well as the process of knowledge itself. With *Narciss*, the viewer can simultaneously observe how the camera eye searches the mirror image of its uncovered hardware and how it interprets what it sees using its AI. These interpretations, which can be very different, are shown to the viewer on the display, for example: a toaster oven, a video game controller, a city by night, a bicycle is parked in front of a television, a pair of scissor sitting on a table. At the same time there are also interpretations that come very close to the real situation: a man looking at a laptop, a person's reflection in the mirror, a person is taking a picture of their reflection in a mirror.

The AI of the installation is controlled by an application based on openFrameworks. The snapshots are analyzed by the im2txt caption generator in Google's Tensorflow framework and textualized into descriptions of what can be seen in the picture. This could be seen as an entertaining gimmick, because the work neither recognizes itself as what it is, i.e. a work of art, nor can it go beyond the status of an animal that does not know or recognize its mirror image. But this is not the intention of the piece either, simply because the text output is descriptive (a ...) and not reflexive (I am a ...), this means that the point of the exhibit is not about the self-knowledge of the AI. The artists are much more concerned with the concept of the viewer's self-confidence in their own self-awareness and how they project it onto the robotic Narcissus: "The project Narciss aims to question our self-righteous model of self-awareness, the quality of our subjective findings while investigating ourselves and the resulting unequal distribution of dignity" (Waltz Binaire 2019).

Waltz Binaire themselves write that they were inspired by Jacques Lacan's concept of the mirror stage and Gordon Gallup Jr.'s mirror self-recognition test. Lacan introduced the mirror stage as a psychological phenomenon in the 1950s, manifesting the assumption that self-awareness is one of the fundamental characteristics for intelligence in regard to social participation. In psychology, the mirror-test originally created by Gordon G. Gallup in 1970 and later also done by Beulah Amsterdam in 1972 examined the self-recognition and -awareness by animals as well as infants. Here, the ancient myth of the adolescent Narcissus serves as blueprint not only for pathological self-indulgence but also for the conception of the self in general. But how relevant is this question of self-awareness and self-knowledge in the work of Waltz Binaire? This question only makes sense if one does not relate it to the autonomous work, but to the contemplation of the piece by the viewer – do we see ourselves in the installation that searches for itself? In regards to the role of the observer, "[t]he human observer is excluded from this internal cycle, yet invited to participate as a superior judge" (Waltz Binaire 2019). The viewer is a judge who judges whether the object is intelligent enough to belong to a social group and whether it can satisfactorily imitate intelligent human behavior, such as in the 'Turing Test' or 'Lovelace 2.0 Test of Artificial Creativity and Intelligence.'

To understand viewers as passive judges gives away a certain potential of the work. Thus, only the question of body-spirit dualism is continued and runs from René Descartes to Otto Rössler's Endophysics all the way through Western intellectual history and precisely does not do justice to the transition from the body to the disembodiment in the Narcissus myth. Narcissus is about touching, seeing, and, in relation to the nymph Echo, hearing. A poly-sensual experience, which also takes place in *Narciss*, occurs when one understands the installation as a reflection of one's own technical infatuation. In fact, the American philosopher Shaun Gallagher defines the minimal nature of a self as the "immediate subject of experience, unextended in time" (Gallagher 2000, 15). The experience of a now and here is therefore the difference between a classic piece of timeless art and an idealistic nowhere of seeing everlasting archetypes. That is why Krauss' "Aesthetic of Narcissism," where she also refers to Lacan, is not expedient here. Narcissism deals with a moment of awareness and the mirror-gazing of the artist, whereas the work by Waltz Binaire focuses on becoming aware of the difference between the networked-based, almost humanoid AI and the actual, bodily human condition. In a certain way *Narciss* reacts here to a social condition based on an enduring acceleration in culture by automation.

Narciss is not a temporary installation, it is partly a work-in-progress dealing with machine learning and seeing. It is also against the psychological position to preserve the status in sterile tests. To understand the work narcissistically means to see oneself in it, a performative, searching, unfinished self. It not so much raises the question if the robotic installation is intelligent in anyway, but moreover what is the concept and idea of human intelligence and consciousness itself. The opposite would be to touch the installation, to capture the camera lens, and to freeze the image. But that would be an action that would have the consequence of allowing one to cling to the present, without self-awareness, and without media-awareness. But at least there is a flower.

List of References

Becker, Daniel. 2016. "What you see is what you get: Das Porträt im Angesicht des Digitalen". *Hermeneutik des Gesichts. Das Bildnis im Blick aktueller Forschung*, edited by Uwe Fleckner and Titia Hensel, 81–100. Berlin: De Gruyter.

Gallagher, Shaun. 2000. "Philosophical Conceptions of the Self: Implications for Cognitive Science". *Trends in Cognitive Sciences*, Vol. 4, 14–21.

Krauss, Rosalind. 1976. *Video. The Aesthetics of Narcissism*. October. Vol. 1, 50–64.

Kruse, Christiane. 2003. *Wozu Menschen malen: Historische Begründung eines Bildmediums*. München: Wilhelm Fink.

Lomas, David. 2011. *Narcissus Reflected: The Narcissus Myth in Surrealist and Contemporary Art*. Exhibtion catalogue. The Fruitmarket Gallery, Edinburgh. London: Reaktion Books.

McLuhan, Marshall. 2001 [1945]. *Understanding Media*. Abingdon: Routledge.

Ovid. 1922. *Metamorphoses*. Translated by Brookes More. Boston: Cornhill Publishing Co.

Posèq, Avigdor. 1991. "The allegorical content of Caravaggio's Narcissus". Source. Note in the History of Art. Vol. 10, 21–31.

Vinge, Louise. 1967. *The Narcissus Theme in Western European Literature up to the early 19th century*. Lund: Gleerups.

Waltz Binaire. 2019. http://waltzbinaire.com/work/narc/, ac. 22.4.2019

Welsch, Maren. 2012. "The Mirror of Narcissus. Mythology and Zeitgeist". In *The Mirror of Narcissus. From Mythological Demigod to Mass Phenomenon*, edited by Beate Ermacora and Maren Welsch, 17–35. Exhibtion catalogue. Galerie im Taxispalais, Innsbruck. Köln: Snoeck.

THE INTELLIGENT WORK OF ART: HOW THE NARCISSUS THEME ECHOES IN NEW MEDIA ART
(Summary)

The paper has shown how the classic Narcissus myth influences the arts and, in particular, new media art. Starting from early milestones in art history focused on this context such as works by Caravaggio and Salvador Dalì, we discussed to what extent the self-awareness associated with Narcissus always includes media-awareness. It is emphasized that this subject always functions on a double level: on the one hand it functions on the representation of the image, and on the other hand it functions on the reception of the image. Following up with this aesthetic dimension, the paper focuses on new media art especially in the context of digitization. It shows to what extent the media conditions of new media art such as interactivity or time-baseness are suitable for more deeply illuminating the relationship between medium and reception. The Narcissus myth was revived in the 1990s, because new and foreign possibilities of media experience were made possible and allowed for a new examination and understanding of the myth. Just as the literary Narcissus experiences the medium of the mirror image in the narrative of the myth, the recipient experiences the media conditions through his performative action. Against this background, the focus shifts to the work *Narciss* by the German artist collective Waltz Binaire, which takes a more

current look at the Narcissus theme.

The reason why this work is central here – and is important throughout this paper – is that it uses AI to represent the Narcissus myth. At the same time, it ties in with classical art historical works in which the viewer merely observes the scenery. Through this combination, the work creates a reference to newer approaches in neuroscience. In particular, the execution of the work by AI raises the question of whether a Narcissus is observed in the actual sense, or whether the entire installation situation, including the viewer, represents the myth and the robot-esque component is only the mirror image of him? Finally, this is theme is seen against the backdrop of the philosophical dualism of body and mind. Here the concepts of self- and media-awareness refresh a general and continuous view of Western culture and history.

Article received: April 23, 2019
Article accepted: May 15, 2019
Original scientific paper

UDC 612.821.89:004.8

Robert Max Williams*
University of Louisville
Louisville, Kentucky, United States

Roman V. Yampolskiy**
University of Louisville
Louisville, Kentucky, United States

OPTICAL ILLUSION IMAGES DATASET

Abstract: Human vision is capable of performing many tasks not optimized for during its long evolution. Reading text and identifying artificial objects such as road signs are both tasks that mammalian brains never encountered in the wild but are very easy for us to perform. However, humans have discovered many very specific tricks or illusions that cause us to misjudge the color, size, alignment, and movement of what we are looking at. A better understanding of these phenomenon could reveal insights into how human perception achieves these extraordinary feats. In this paper we present a dataset of 6,725 illusion images gathered from two websites, and a smaller dataset of 500 hand-picked images. We will discuss the process of collecting this data, models trained on the data, and the work that needs to be done to make this information of value to computer vision researchers.

Keywords: Computer Vision, Optical Illusions, Human Vision, Machine Learning, Neural Networks, Cognition

1. Motivation

Being able to understand and intentionally create illusions is currently only possible for humans. The ability to accurately recognize illusory patterns using a computer, and to generate novel illusion images, would represent a huge advancement in computer vision. Current systems are capable of predicting the effect of specific classes of illusions, such as color consistency illusions (Robinson, Hammon, and Sa, 2007) and length illusions (Garcia-Garibay and Lafuente, 2015; Bertulis and Bulatov, 2001). A reinforcement learning system learned to perceive

* Author's contact information: robertmaxwilliams@gmail.com
** Author's contact information: roman.yampolskiy@louisville.edu

color consistency illusions after training to predict color values where half of the image was covered in a tinted film, showing that perception of an illusion can emerge from the demands of seeing in a complicated world (Shibata and Kurizaki, 2012). It is also important to consider whether making a perceptual mistake similar to the mistakes of human perception constitutes having a visual experience similar to humans (Yampolskiy, 2017).

Recent work on generative adversarial networks (GANs) (Karras et al., 2017) has shown that high resolution images of faces can be created using a large dataset of 30,000 images. This quantity and quality of images is not available for optical illusions; as discussed below, naively applying their methods to this dataset does not have the same results.

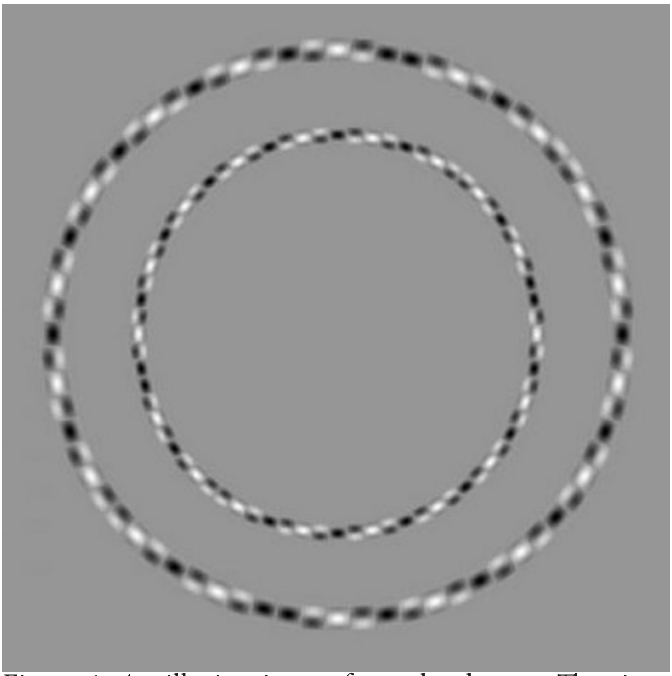


Figure 1: An illusion image from the dataset. The rings are circular and concentric, but the patterns and changes in contrast make them appear to be warped. © viperlib.york.ac.uk

The number of static optical illusion images available are in the low thousands, and the number of unique kinds of illusions is certainly very low, perhaps only a few dozen (for example, the Scintillating Grid illusion, Cafe Wall Illusion and other known categories). Creating a model capable of learning from such a small and limited dataset would represent a huge leap in generative models and our understanding of human vision.

2. Related Works

Research into biologically plausible models makes it possible to learn about visual phenomenon by conducting experiments on proxies for the real human vision system. Elsayed et al. found that by selecting the right models, adversarial examples for these models were also effective on time-limited humans (Elsayed et al. 2018). In their experiment, they created adversarial images for an ensemble of image classification neural networks designed to be similar to human vision. The adversarial images cause the machine learning classifier to classify them incorrectly by only making subtle changes to the image pixels, and they were testing whether these subtle changes would also cause humans to incorrectly classify the altered images. To make the neural networks similar to human vision, they preprocessed their input images to mimic some aspects of human vision, such as higher resolution in the center and lower resolution on the outside. Participants were shown an image in one of two classes, for example, an image of a snake or spider. The images were only shown for 63 milliseconds, meaning that there was not enough time to look at multiple places in the image or reason about its contents on a semantic level. Only the first few "layers" of human vision can work in that short a time span. Their result was that images with subtle changes that could fool an ensemble of neural networks also caused a significant decrease in accuracy for the time-limited humans. This means that current models learned using convolutional neural networks are internally similar to the simplest parts of human vision, and attacks on these neural networks transfer to the visual abilities of time-limited humans. The adversarial examples they created constitute a new class of optical illusions, which can fool the eye into making a mistake when first glancing at an image.

The Brain-Score metric measures internal and behavioral similarity between computer and primate image recognition (Schrimpf et al. 2018). As this metric is developed and models with higher scores are created, those models may be capable of experiencing additional kinds of optical illusions that are otherwise only experienced by primates.

To our knowledge, no dataset of this kind has been created before.

3. Data Collection

3.1. Image Sources

Twelve different websites that collect and display optical illusions (such as the one shown in Figure 1) were considered for inclusion in the dataset. Most proved to be too small or did not contain the right content. For instance, the site "Visual Phenomena & Optical Illusions" contains many interesting and visually powerful demonstrations of optical illusions, but very few still images that by themselves contain a visual effect ("Visual Phenomena & Optical Illusions" 2018). In the end, "Mighty Optical Illusions" (Mighty Optical Illusions, 2018) and "ViperLib" (Thompson and Stone 2018) proved to be the best sources of illusion images, both containing labeled, almost exclusively static images.

Mighty Optical Illusions is a blog-style website, with pages in chronological order labeled as different kinds of illusions and miscellaneous categories. These categories are used as training labels for the classification models. Most of the content on the site are static images, with only a few animations, meaning most of the data could be used.

ViperLib has image pages organized into exclusive categories, but many of them are animations which do not create an illusion when viewed as static images.

The "Illusions of the Year" contest also seemed to be a good source of images, but they only post the winning results publicly (Neural Correlate Society 2018). Emails to the website owner requesting all of the submissions were not answered.

3.2. Data Collection Results

We created a web scraper to go through each page of Mighty Optical Illusions and download the images on the page (source is available at Williams 2018). In total, 6,436 images were obtained, along with their metadata such as categories and page titles. ViperLib was scraped in a similar manner, obtaining 1,454 images also organized into categories and with page titles.

Each image from the Mighty Optical Illusions dataset has one or more tags describing it. Tags such as "anamorphosis" or "impossible objects" were associated with specific kinds of illusory effects, while other tags such as "murals" or "animals" describe the medium or contents of the images. To simplify the training of the classifier, a folder was created for each tag and all images using that tag were placed in its folder. This means that many images were duplicated across categories. In the multi-label classification experiment (Section 4.3), images are included in the datasets based on having or not having a particular tag, so no duplication occurs.

A subset of the data, referred to in Williams paper (2018) as "illusions-filtered," was selected manually as the highest quality illusion images. These images were selected based on having an immediate visual effect without needing any context, such as apparent motion illusions. This hand-picked subset of the data represents the classes of illusions that can be understood solely based on visual stimuli and seemed like the most likely candidates for illusions that a computer could experience, create, and discover. With the current state of machine learning, I expected that identifying and generating pattern-based illusions, such as motion illusions, would be a much easier task than understanding real world objects well enough to identify perspective illusions or Escher-like impossible objects.

3.3. Qualitative Analysis

To determine the feasibility of learning the dataset, we considered how meaningful the classes are and if the images were representative of their class enough to be learned.

The Mighty Optical Illusions dataset was used for the classification experiments (in section 4.1) because it had more images and diversity. However, the labels seem somewhat arbitrary and are difficult for a human to understand from the images alone. Looking at Figure 2, it is not immediately obvious that images in each column belong together.

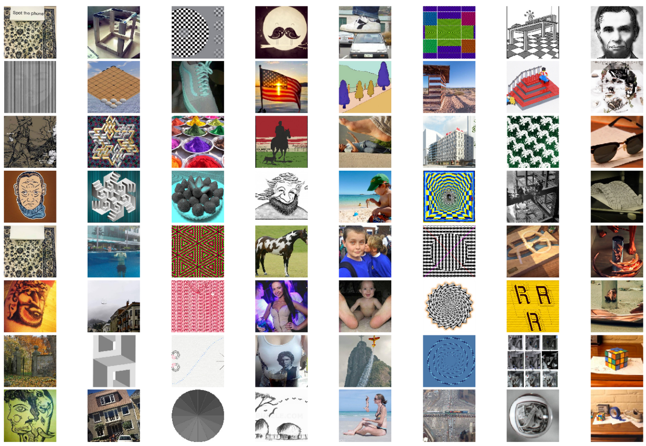


Figure 2: Images from the dataset. All images in the same column have a label in common. Labels are Spot The Object, Impossible Objects, Color Adapting, Multiple Meanings, Relative Sizes, Seemingly Bent, Escher Style, and Anamorphosis, from left to right.

The first class is "Spot the Object," where images contain something that is hard to find but is easy to see after it's been pointed out. Classifying whether or not an image contains a hidden object is a very difficult task, since some of these illusions can require minutes of searching to find the object. This means that to confirm that something does not have a hidden object, you would need to search for as long as the expected time needed to find a hidden object. The "Impossible Objects" category contained images or illustrations of perspective illusions and Escher-like geometries. Given a geometric scene, careful spatial reasoning is required to tell if there is impossible geometry. There are also images based in perspective illusions, with various combinations of clouds, refraction, reflective water, and slanted landscapes that create impossible seeming scenes. "Color Adapting" refers to the eye's ability to adapt to changes in lighting and illusions that are created by taking advantage of this ability, but this category includes a wider variety of images, and it seems that anything color related received this tag. The "Multiple Meanings" category contains images which have more than one appearance depending on how you look at them. Some are very subtle, so this category is difficult for the same reason as finding a hidden object. It overlaps heavily with "Impossible Objects," since many impossible objects appear as being two thing simultaneously which cannot exist at the same time.

"Relative Sizes" contains familiar objects in contexts that make them seem far larger or smaller than they really are. A mismatch between the apparent size of an object in an image and your commonsense knowledge about the actual size of objects is easy for humans to identify, but it seems like a task that would be very difficult for a neural network to learn without being specifically designed for this task. The category "Seemingly Bent" contains a large amount of illusions that are immediately apparent without additional context or knowledge, so I expected this category to be one of the easiest to identify with machine learning. "Escher Style" is the same as "Impossible Objects" but limited to impossible geometries and often the images are in the pencil-sketch style of Escher's artworks. "Anamorphosis" refers to images where viewing them with a specific perspective or lens changes their appearance. For instance, images of sculptures which produce an image when a cylindrical mirror is placed in the center and images that appear different when viewed up close or far away.

Within each class, many of the images do not contain illusions or are only meant as references. For instance, in many "Spot the Object" illusions, a second image with the hidden objects highlighted or circled is provided. How to make use of these images in a machine learning model is unclear. Many "Impossible Object" images also include images of how the object was constructed, none of which contain an actual illusion.

Many confounding factors make using this dataset in a traditional machine learning workflow difficult. This difficulty distracts from the key question: can machines perceive optical illusions as humans do? Expensive hand-sorting of the data could solve this problem, by isolating exactly the images of interest and putting them in consistent and meaningful classes of illusions. Illusions of restricted kinds could be automatically generated, such as variations on motion illusions based on known patterns. Overall, the dataset contains a large portion of images which clearly demonstrate illusions, but many non-illusions are present which makes learning difficult.

4. Machine Learning Results

Three different kinds of models were tested on subsets of the data. Two classifiers were trained to test how visually distinguishable the given classes are, and a generative model was trained to see if new instances of known illusions could be created by naively applying existing methods for image generation.

4.1. Single-label Classifier Results

A pretrained "bottleneck" model (TensorFlow 2018) was used to classify images from Mighty Optical Illusions. Only the last few layers had to be retrained, making use of transfer learning from a much larger dataset to learn how to classify images in general. In this case, the pre-trained model was "Inception v3", trained on over 14 million images in the ImageNet dataset. The pre-trained model converts the very high dimensional image data into a lower dimensional "feature space" vector. This means that the image, a vector of around two hundred thousand values, is compressed into a vector of around two thousand value which contain enough information to accurately classify an image.

Learning new image classes from this feature space representation requires significantly less data and computation time, meaning that these experiments can be run on a normal PC in a few minutes, instead of the days or weeks on a GPU that was required to train the original Inception v3 model.

Each image in the training data may belong to multiple classes, which was not accounted for in the model. In Section 4.3, a multi-label classifier was created for a different subset of the data. The results of training can be seen in Figure 3.

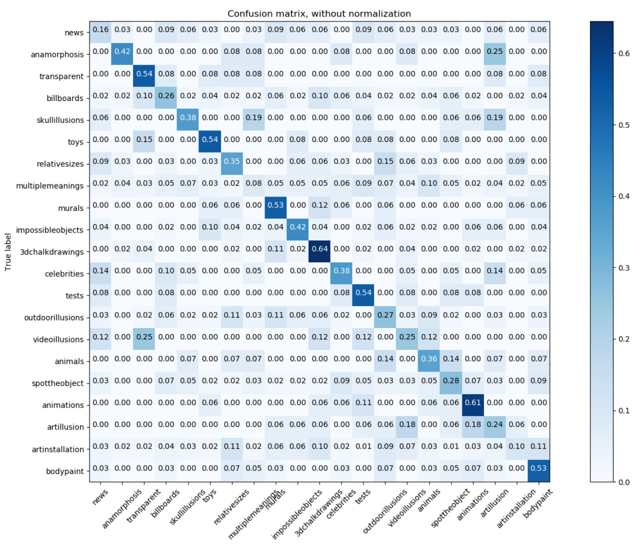


Figure 3: Confusion matrix for a classifier trained on the Mighty Optical Illusions data.

The model performed significantly better than random, meaning that the given classes are meaningful in a way that can be detected using a model trained on normal classes of images. Out of the 21 classes, most classes were predicted with 40-60% accuracy. The very poor accuracy on news, multiple meanings, and art illusions is explained by the lack of defining features for these classes. For example, the "art illusions" category overlaps evenly with most other categories.

An interpretation study could reveal more about how the neural network is able to distinguish these classes, such as the methods used in Zhang, Wu, and Zhu (2017) that show which areas of the image are important to classification and what the key features of each class are.

4.2. Generative Adversarial Network

A trial run using a generative adversarial network was attempted. Using HyperGAN (Martyn, 2017) on a hand-picked subset of the data with no hyperparameter optimization, nothing of value was created after 7 hours of training on an Nvidia Tesla K80. The training progression is shown in Figure 4.

When trained with homogeneous data (such as only using images of faces), GANs are able to create varied and convincing imagery. However, when applied to a varied, multimodal dataset, performance degrades and the generator only learns to generate a single type of image, a problem known as mode collapse (Barnett 2018).

The output produced by the GAN subjectively resembles some sort of scene or objects. It has learned many underlying patterns in the dataset, such as high contrast edges, varied shading, and spatially confined objects. On a small-scale visual level, the generated images appear to be a plausible photographic scene. However, on a larger scale it fails to recreate anything resembling the images in the dataset.

The GAN could be pretrained on a larger dataset to overcome the issue of having such a small dataset. Dataset expansion techniques, such as rotating, cropping, and scaling images, could also be applied to increase the amount of data available for training. The GAN was run with default parameters, which are likely far from ideal for this dataset. Tweaking the parameters to better suit the specifics of this dataset may prevent mode collapse and increase the quality of training.

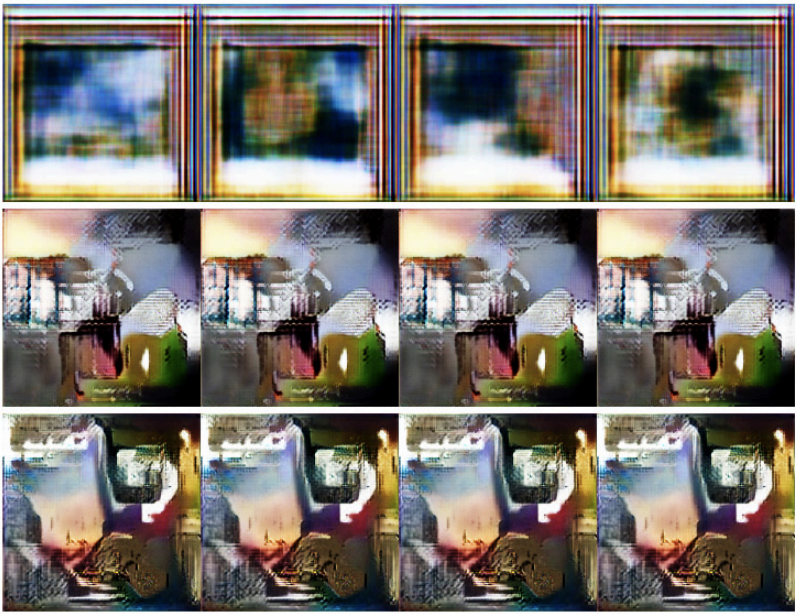


Figure 4: Failure of GAN to generate imagery similar to the dataset. Top to bottom is the progression from start to finish. Images in the same row are from the same training step but with different random input vectors (Goodfellow et al., 2014). The lack of variety is abnormal and may lead to insights into how to correct the problem.

4.3. Multi-label Classifier

In multi-label classification, each image can be in more than one class, and the classifier outputs true or false for each label. Most of the images in the Mighty Optical Illusions dataset have more than one label, so this technique is appropriate for the dataset.

4.3.1. Model

For this experiment, ResNet50 with pre-trained weights was used (Simonyan et al. 2015). The final classification layers were removed and replaced with a densely connected layer with ReLU activations and a prediction layer with sigmoid activation on a single output. This is the same bottleneck training technique used in Section 4.1. An instance of this model was created for each of the target classes and trained separately. To obtain a vector representing all of the labels, the output of all of the models is concatenated together.

4.3.2. Data

Labels that occur less than 70 times in the dataset were removed, as well as labels that do not indicate the content of the illusions, such as "Animals" or "Murals." Only 8 of the 42 labels met these specifications: Spot The Object, Impossible Objects, Color Adapting, Multiple Meanings, Relative Sizes, Seemingly Bent, Escher Style, and Anamorphosis. In an initial testing run, images with none of these labels were left in the dataset to provide negative samples. This leaves a large majority of the images with no label, meaning that a model that always predicts 0's for each class will be largely accurate, and the model failed to learn to classify on any of the labels significantly better than random. To better evaluate the model, the data was made into even splits for each label: 50% images which have the label the model is being trained on, and 50% with any other label. For example, the split for "Color Adapting" would consist of 50% images that have the "Color Adapting" label (and possible other labels) and the other 50% would consist of images randomly sampled from the rest of the dataset, contained any label except "Color Adapting". This is repeated for every label. Another similar dataset was made, but with a third category of images with no labels to provide negative examples. In that dataset, the data for each label was split with 50% having the label the model is being trained on, 25% having any other labels, and 25% with no labels at all.

4.3.3. Results

The model trained on the original, biased data only learned to predict false for every label and fit some of the training data. The models trained on balanced data, however, were able to generalize to the held-out validation set with some success. The validation accuracies for the two models trained on balanced data are shown in Figure 5. Both models failed to learn some classes, and scored very well on some labels, such as Color Adapting. Inspecting images in each class shows that there are many surface characteristics, such as the high contrast shapes in Color Adapting, that make it easy to recognize the label without being able to identify the presence of a color adapting illusion. This dataset allows the model to "cheat" and identify illusions by the way they are presented instead of imitating the human visual system and identifying them as illusions. A means to deactivate illusions without changing surface characteristics would enable a more rigorous test of the model. For instance, in "Skye's Oblique Grating," if the odd check marks are not rotated 90 degrees from the even check marks, the illusion disappears. A model that memorized the appearance of the Skye's Oblique Grating might miscategorize the deactivated version as being an illusion.

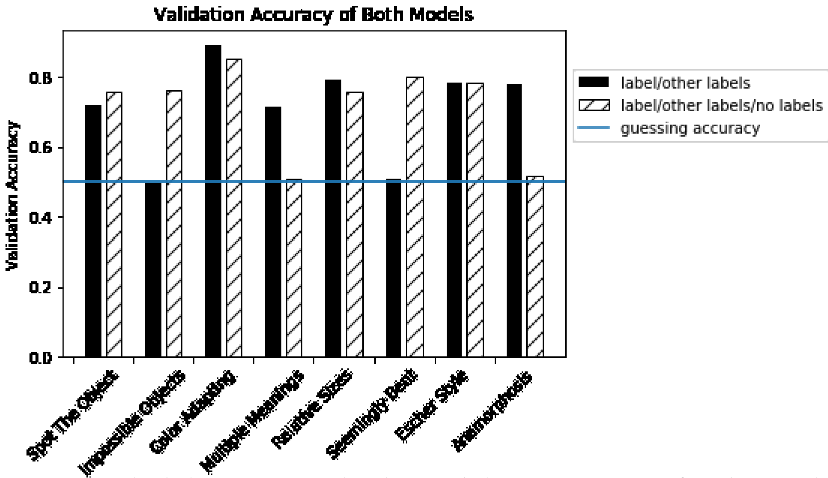


Figure 5: Black bars are single class validation accuracy for the model trained on data split between having that label and not having that label but having a different label. Striped bars are accuracy for the model trained on data split between having that label, having any other label, and having no label. The horizontal line shows the baseline accuracy for a model that always guesses the most likely class, 50%.

5. Conclusion and Future Work

The only optical illusions known to humans have either been created by evolution (for instance, eye patterns in butterfly wings) or by human artists. Both artistic designers of illusion images and the glacial process of evolution have access to active vision systems to verify their work against. An illusion artist can make an attempt at creating an illusion, observe its effect on their own eyes, and add or remove elements to try to create a more powerful illusion. In an evolutionary process, every agent has a physical appearance and a vision system, allowing for patterns to be verified in their environment constantly. A GAN trained on existing illusions would have none of these advantages, and it seems unlikely that it could learn to trick human vision without being able to understand the principles behind the illusions. Because of these limitations, it seems that a dataset of illusion images might not be sufficient to create new illusions and a deeper understanding of human vision would need to be obtained by the network somehow. This could be done by having a human giving feedback as the network learned, or by learning an accurate proxy for human vision and trying to deceive the proxy as in Elsayed et al. (2018).

Appendix A.

Downloading the Dataset Images are currently hosted on the machine learning cloud platform "Floydhub."

- https://www.floydhub.com/robertmax/datasets/illusions-jpg
- This contains all images that were downloaded, using the same numbering scheme as the metadata on the linked github repository.
- https://www.floydhub.com/robertmax/datasets/illusions-filtered
- This folder contains images hand picked for having obvious visual effects without having to follow special instructions.

List of References

Barnett, Samuel A. 2018. "Convergence Problems with Generative Adversarial Networks (GANs)." CoRR abs/1806.11382.

Bertulis, Algis, and Aleksandr Bulatov. 2001. "Distortions of length perception in human vision." *Biomedicine* 1 (1): 3–23.

Elsayed, Gamaleldin F., Shreya Shankar, Brian Cheung, Nicolas Papernot, Alex Kurakin, Ian Goodfellow, and Jascha Sohl-Dickstein. 2018. "Adversarial Examples that Fool both Computer Vision and Time-Limited Humans." eprint: arXiv:1802.08195.

Garcia-Garibay, Otto B., and Victor de Lafuente. 2015. "The Muller-Lyer illusion as seen by an artificial neural network." 25745398[pmid], Front Comput Neurosci 9 (February): 21. issn: 1662-5188. doi:10.3389/fncom.2015.00021. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4333816/. 5

Goodfellow, Ian J., Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. *Generative Adversarial Networks*. eprint: arXiv:1406.2661.

Martyn, Mikkel et al. 2017. "HyperGAN." https://github.com/255BITS/HyperGAN.

Mighty Optical Illusions. 2018. "Mighty Optical Illusions." Accessed January 15, 2018. https://www.moillusions.com/.

Karras, Tero, Timo Aila, Samuli Laine, and Jaakko Lehtinen. 2017. "Progressive growing of gans for improved quality, stability, and variation." arXiv preprint arXiv:1710.10196.

Robinson, Alan E., Paul S. Hammon, and Virginia R. de Sa. 2007. "Explaining brightness illusions using spatial filtering and local response normalization." Vision Research 47 (12): 1631–1644. issn: 0042-6989. doi:https://doi.org/10.1016/j.visres.2007.02.017. http://www.sciencedirect.com/science/article/pii/S0042698907000648.

Schrimpf, Martin, Jonas Kubilius, Ha Hong, Najib J. Majaj, Rishi Rajalingham, Elias B. Issa, Kohitij Kar, et al. 2018. "Brain-Score: Which Artificial Neural Network for Object Recognition is most Brain-Like?" bioRxiv preprint.

Shibata, K., and S. Kurizaki. 2012. "Emergence of color constancy illusion through reinforcement learning with a neural network." In 2012 IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL), 1–6. November. doi:10.1109/DevLrn.2012.6400580.

Simonyan, Karen, and Zisserman, Andrew. 2015. "Very Deep Convolutional Networks For

Large-scale Image Recognition" In 2015 ICLR. abs/1409.1556. arXiv: 1409.1556. http://arxiv.org/abs/1409.1556.

Neural Correlate Society. 2018. "Illusions of the Year." Accessed January 15, 2018. http://illusionoftheyear.com.

TensorFlow. 2018. "How to Retrain an Image Classifier for New Categories." Accessed April 10, 2018. https://www.tensorflow.org/hub/tutorials/image_retraining.

Thompson, Peter, and Rob Stone. 2018. "ViperLib." Accessed January 15, 2018. http://viperlib.york.ac.uk. "Visual Phenomena & Optical Illusions." 2018. http://www.michaelbach.de/ot/.

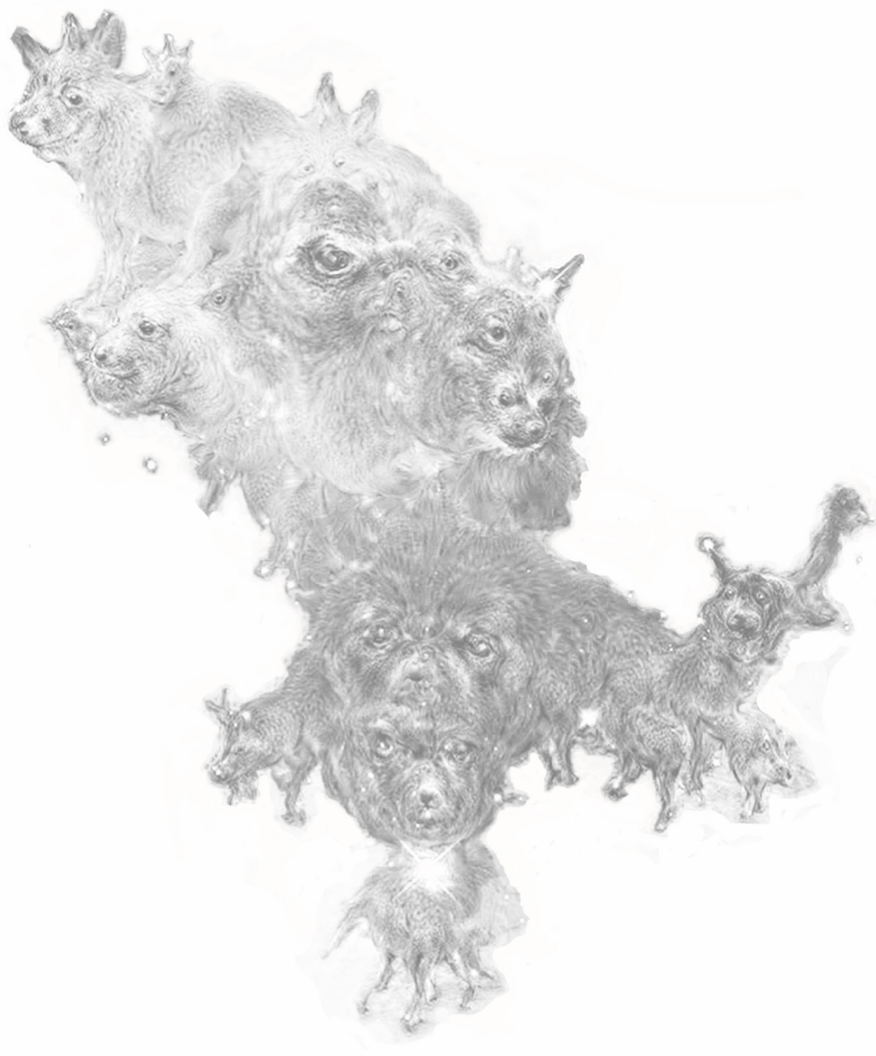
Williams, Robert Max. 2018. "optical-illusion-dataset." Accessed September 21, 2018. https://github.com/robertmaxwilliams/optical-illusion-dataset.

Yampolskiy, Roman V. 2017. "Detecting Qualia in Natural and Artificial Agents." CoRR abs/1712.04020. arXiv: 1712.04020. http://arxiv.org/abs/1712.04020.

Zhang, Quanshi, Ying Nian Wu, and Song-Chun Zhu. 2017. "Interpretable Convolutional Neural Networks." eprint: arXiv:1710.00935. 6

Article received: February 25, 2019
Article accepted: May 15, 2019
Original scientific paper

REVIEWS



Klara Petrović and Luja Šimunović*
KONTEJNER | bureau of contemporary art praxis
Zagreb, Croatia

Device_art 6.018: Machines are Not Alone

The sixth edition of the triennial festival Device_art took place from December 2018 to January 2019 in Croatia, with the main exhibition hosted by the Museum of Contemporary Art in Zagreb. In addition to the exhibition, the festival program included a conference, performances, lectures, and workshops in Zagreb, Split, and Rijeka.

KONTEJNER | bureau of contemporary art praxis was established in 2002 and is the organization behind Device_art, an international art festival that has taken place every three years since 2004. Device_art works to investigate the role of technology in art and society. During past editions, the festival has focused on a comparative approach, with a main partner from a specific country (i.e., USA, Canada, Japan, Czech Republic). Both countries showcasing their investigations into the use of technological devices, machines, gadgets, and robots as an artistic medium.

In 2018, Device_art 6.018 delivered its audience a conceptual turn by inviting renowned curator ZHANG Ga to co-create a thematic exhibition named: *Machines are Not Alone*. The theoretical basis of the concept was derived from works by Gilles Deleuze, Felix Guattari, Brian Massumi, Erich Hörl, Gilbert Simondon and Thomas Lamarre, whose selected texts in the festival catalogue support a new, open-plane approach to our understanding of what constitutes a machine or being a machine, agency of subject and object, as well as envisioning a broader, fluid approach to what we consider to be part of nature, and the ecological system that surrounds us.

Should we search for a supporting thread of thought as we ascend the steps of the Museum and begin the exhibition, we should look no further than the opening sentence of the curatorial text by ZHANG Ga: "The world is machinic." Perhaps this thought is an abbreviation of Deleuze and Guattari's writings in *Anti-Oedipus, Capitalism and Schizophrenia* (1983): "Everywhere it is machines – real ones, not

*Authors' contact information: klara.petrovic1@gmail.com and luja.kontejner@gmail.com.

figurative ones: machines driving other machines, machines being driven by other machines with all the necessary couplings and connections." Asking "who are these machines," and in a conceptual circuit envisioning a world that transcends categories of nature and technology may be one of the overarching themes one will find while visiting this exhibit. To find the answer(s) to this question it may be best to take a look at works supporting this view.

The works of Marnix de Nijs, Saša Spačal and Mirjan Švagelj, Ida Hiršenfelder, and Tin Dožić all deal with a more somber and anxiety-ridden side of this concept in which the *machinic* works as a proposed "way out" for the artist's concerns of excess and global pollution and the ways we deal with inevitable (self)destruction. In Dožić's *GoldRush* we encounter obsolete technologies, that were once an inherent part of the Earth, transformed to serve humanity and in the process became toxic to the source they came from. This inherent chemical interconnectedness reveals itself in the possibility of gold extraction from used up processors and graphic cards proposing the question of how we value gold in a depleted world.

Spačal and Hiršenfelder, on the other hand, deal with the depletion of the Earth through the idea of the oil shortage. Their *Sonoseismic Earth* tackles the task of visually representing the seemingly minute, but realistically unnatural, tectonic changes in the Earth's crust caused by fossil fuel extraction. The pale globe sits helplessly on a pedestal filled with bones and polluted water which ooze an unpleasant odor into the surroundings, the globe shivers as humans approach it. It fears further abuse. The peril of Earth is contrasted with the idea of human lunacy in De Nij's work, situated so it can be viewed from the gallery where the *Sonoseismic Earth* trembles in silence. De Nij imagines a world where humans have brought their need for natural luxuries to a level of absurdity. The *Autonomous Oil Reserve (AOR-200)* is a consumer system that protects a personal oil reserve and self-destructs if someone enters the surrounding space where the barrels are held. The system works so that it feeds its own alarm, burning itself out until it ultimately destroys itself so that nobody can have the precious, stored liquid.

Not only humans will suffer the consequences of their own deeds: the species often forgotten at the "top of the pyramid," the slimy, the delicate, and the endangered organisms living in ocean shallows are put under a magnifying glass in Gail Wight's *Pool*. We must think of the natural balance in nature that we are destabilizing, and Wight urges us to look at it head on.

How will we ever get out of this "postmordial" swamp? Can the subjectivity of the machines help drag us out of the inevitable? A lone crane stands limply on a sand colored wooden base. Topped with a decorative plastic leaf, it stands as a deserted island amidst the other works in the space. Titled *I will be back sometime*, Dorial Gaudin's work evokes empathy with the structure's glitchy motions, a giant waiting to be used once more, as if it had lost its life purpose in the absence of functionality. Artists such as Dorian Gaudin, Ralf Baecker, Adam Donovan, and Katrin Hochschuh all deal with attaching emotionality to the physical behaviors of artificial intelligence. In Adam Donovan's and Katrin Hohschuh's *Empathy Swarm* the robots try to understand us and emotionally reflect our behavior by following, gathering, and running away. Ralf Baecker's *Interface I* strays from the need for humans as it is autonomous. In the encapsulated space, stripped of anything relating to our everyday surrounding, the strings, playing a very sophisticated game of seemingly random tug-of-war, serve as a perfectly austere structure investigating complex interactions in general.

Among these works concerned with ecology, and those with an almost decidedly pessimistic outlook reflecting the Anthropocene, and the artists creating potentially autonomous, random machines and robots we also encounter works that use technology as a medium to relay history, intimacy, and the mystical, spiritual qualities of physics. Such works are those by Mirjana Vodopija, Navid Navab and Martin Howse.

Vodopija's installation *Vibrating Landscape* serves as a visualization of a mind trying to grasp conflicting needs, a game of shifting from open spaces to those most intimate thoughts, trapped in our memories. *Test Execution Host* by Martin Howse tells a converging tale of a man's history and its relationship to geology. It is a conceptual machine, one conceived by Alan Turing, with tubes dripping cyanide on rocks and books, a screen monitors it all, relaying a story of a person's death, but also the story of the decay of the world in an ecological sense. Aptly named *tangibleFlux φ plenumorphic ∴ chaosmosis*, as the title itself evokes an alchemic, mystical condition, Navid Navab's installation comprises of three pedestals, a triptych of altars each carrying a small metallic ball levitating, turning, jumping from one side to the other, dependent on the magnetic field installed below. A simple physical phenomenon delivered in a sacral, exhilarating way.

"Everything is machinic" is the poignant phrase that served as the motto of Device_art 6.018. The solidarity, autonomy, companionship, and relationships of machines and life was deeply thought upon, researched, and discussed among artists and curators so a series of possible outlooks on the present and future place of the machines' in the world could be presented to the viewer. In the words of KONTEJNER's curators: "The artists showing their work (...) reveal to us in various ways this machinic nature of the world and reality; the machine that surrounds us and that is in us, the universal machine and the machinic connections that determine our very existence."

Beside man's need for destruction, there is a force more powerful which ties together the strings of thought generated throughout the exhibition, and that is creation. From human ingenuity comes the machine, as a predetermined idea in the human conscience, consisting of man-made physical components, as well as the essence of the human being as its creator. The machines are not alone; they are intrinsically a part of a greater mechanism in which everything is inherently *machinic*.

Article received: May 25, 2019
Article accepted: June 1, 2019

Jelena Čupić*
University Archive of University of Music and Performing Arts
Graz, Austria

Jorge Solis and Kia Ng (Eds.), *Musical Robots and Interactive Multimodal Systems* (Springer Tracts in Advanced Robotics 74), Berlin Heidelberg: Springer Verlag 2011, ISBN 978-3-642-22291-7

In the era of technology and artificial intelligence, it is not a question of if, but, rather, when will robots become part of our everyday life. Between dystopian and utopian ways of seeing this robotic future, it is definitely more appealing to choose to side with the utopian perspective. Bruno Siciliano, the editor of the *Springer Tracts in Advanced Robotics (STAR)* stresses in his Foreword that *STAR* is devoted to bringing the most recent advances in the robotics field to the entire research community. The collection of scientific papers published in this volume, *Musical Robots and Interactive Multimodal Systems,* edited by Jorge Solis and Kia Ng are the first in the series to cover the subject of musical robotics, a new emerging field of human-robot interaction. The volume consists of fifteen chapters divided into three parts: one introductory chapter[1] and two sections with seven chapters each.

The preface states that the fourteen carefully selected scientific contributions should "[…] highlight cutting edge research related to […] exploring musical activities, interactive multimodal systems and their interactions with robots to further enhance musical understanding, interpretation, performance, education and enjoyment." The chapters are thematically organized, with the first section being *Understanding Elements of Musical Performance and Expression*.[2] This section

*Author's contact information: jelenacupic@gmx.at

1 Jorge Solis and Kia Ng, "Musical Robots and Interactive Multimodal Systems: An Introduction", pp. 1–12.

2 The studies in this section are: Rolf Inge Godøy, "Sound-Action Chunks in Music", pp. 13–26, Fabrizio Agenti, Paolo Nesi and Gianni Pantaleo, "Automatic Music Transcription: From Monophonic to Polyphonic", pp. 27–46, Antonio Camurri and Gualtiero Volpe, "Multimodal Analysis of Expressive Gesture in Music Performance", pp. 47–66, Joseph Malloch, Stephen Sinclair, Avrum Hollinger and Marcelo M. Wanderley, "Input Devices and Music Interaction", pp. 67–84, Diana S. Young, "Capturing

concentrates on the development of multimodal systems to provide untaught and efficient human-machine interaction with the goal of conceiving more advanced methods for the analysis, modeling, and understanding of musical performance and innovative interfaces for musical expression. The second section, entitled *Musical Robots and Automated Instruments*,[3] focuses on the advancement of automated instruments and anthropomorphic robots designed to study human motor control from an engineering point of view and it aims to propose new ways of musical expression. Although the singular chapters are, as stated above, connected on a general thematic level, every one of them focuses on one specific subject and/or experiment and by doing so, they aim to cover all of the relevant aspects that make up the "bigger" section theme, in as much detail as possible, at least as far as the current technology allows.

The seven chapters in the first section are all about musical performance, what makes it unique, why it is important, how it can be better understood and improved, and what the current technology can offer to musical performance. So, the exact focus in this section varies. In his study, "Sound-Action Chunks in Music," Rolf Inge Godøy explores the connections between sound features and action features, i.e. musical features and body movement, and brings those features together, focusing on the micro- and meso-levels of the sound-action links, which, according to his proposal, manifest at the timescale of a chunk, meaning in excerpts in approximately the 0.5 to 5 seconds range, forming one sound-action chunk. The chapter "Automatic Music Transcription: From Monophonic to Polyphonic" focuses on audio analysis and transcription. Fabrizio Agenti, Paolo Nesi, and Gianni Pantaleo explain automatic music transcription as the process of analyzing a musically recorded signal, or a musical performance, and converting it into a symbolic notation or any equivalent representation. Their aim is to investigate the evolution of algorithms that help us understand music, as well as the evolution of the two models – from monophonic to polyphonic, based on the most used techniques in the recent literature, such as: Nonnegative Matrix Factorisation, Hidden Markov Models, Bayesian models, generative harmonic models, and the use of jointed frequency and time information. The authors credited MIREX (Music Information Retrieval Evaluation Exchange) as one remarkable achievement in unifying the existing approaches. The rising

Bowing Gesture: Interpreting Individual Technique", pp. 85–104, Kia Ng, "Interactive Multimedia for Technology-Enhanced Learning with Multimodal Feedback", pp. 105–126, Frédéric Bevilacqua, Norbert Schnell, Nicolas Rosamimanana, Bruno Zamborlin and Fabrice Guédy, "Online Gesture Analysis and Control of Audio Processing", pp. 127–142.

3 The studies in this section are: Eiji Hayashi, "Automated Piano: Techniques for Accurate Expression of Piano Playing", pp. 143–164, Roger B. Dannenberg, H. Ben Brown and Ron Lupish, "McBlare: A Robotic Bagpipe Player", pp. 165–178, Koji Shibuya, "Violin Playing Robot and *Kansei*", pp. 179–194, Jorge Solis and Atsuo Takanishi, "Wind Instrument Playing Humanoid Robots", pp. 195 – 214, Ajay Kapur, "Multimodal Techniques for Human/Robot Interaction", pp. 215–232, Guy Hoffman and Gil Weinberg, "Interactive Improvisation with a Robotic Marimba Player", pp. 233–252, Jorge Solis, Klaus Petersen and Atsuo Takanishi, "Interactive Musical System for Multimodal Musician-Humanoid Interaction", pp. 253–268. The studies are followed by an Author Index.

significance of the expressive gesture for music performance is the focus of the fourth chapter: "Multimodal Analysis of Expressive Gesture in Music Performance." In this paper, the research of Antonio Camurri and Gualtiero Volpe joins findings from numerous disciplines: psychology, biomechanics, computer and social sciences, and performing arts, the result is an automatic system that can classify gestures according to basic emotion categories and simple dimensional approaches, which is exceptionally important for group playing. The topics included are interaction between the performers, between the performers and conductor, as well as the interaction between performers and audience. In the joint paper, "Impute Devices and Music Interaction," by Joseph Malloch, Stephen Sinclair, Avrum Hollinger, and Marcelo M. Wanderley the understanding of playing an instrument is discussed. The authors state, that impute devices are extremely relevant to our ability to understand the playing of an instrument. The focus of the study is the design and conceptualization of the digital musical instruments (DMIs) and approaches to instrument design are presented in three different contexts: application to new music performance, use within specialized medical imaging environments, and interaction with virtual instruments. Chapter six: "Capturing Bowing Gesture: Interpreting Individual Technique" is dedicated to the string players, and more specifically, to the myriad of ways in which they control their bow. Diana S. Young focuses on the significance of the bowing parameters and introduces a measurement system for violin to accurately capture bowing techniques during realistic playing conditions. The implication being, that by capturing the individual bowing techniques, the understanding of the physical elements of performance will be improved. Right after this paper understanding the individual techniques for string players, Kia Ng introduces in his paper an interactive multimedia system for music education. The study "Interactive Multimedia for Technology-Enhanced Learning with Multimodal Feedback" focuses on the i-Maestro 3D Augmented Mirror (AMIR) which can offer offline and online feedback for technology-enhanced learning of strings. Since playing an instrument is a physical activity, the idea behind this project is to capture a performance in detail and be able to use the video capture in order to improve posture. This kind of musical robot can assist both teachers and students. The last chapter of the first section "Online Gesture Analysis and Control of Audio Processing" presents the notion of temporal mapping. In this paper, Frédéric Bevilacqua, Norbert Schnell, Nicolas Rosamimanana, Bruno Zamborlin, and Fabrice Guédy show the importance of the temporal aspects of the relationship between gesture, sound, and musical structures.

The second part of the volume concentrates on the development of anthropomorphic robots and automated musical instruments that allow us to study human motor control, facilitate the human-robot interaction from a musical point of view, and propose innovative ways of musical expression. Every chapter of this section is concentrated on a specific difficulty that instrument players have, whether it is a technical or even a biological issue. Combining these studies and observing them together, one can get a fairly clear picture of how the human body functions. As for the technical difficulties, such as those involved in “touch”, a superior automatic piano is designed to produces soft tones. The development of this piano and the analysis of its actions are presented by Eiji Hayashi in the paper "Automated Piano: Techniques for Accurate Expression of Piano Playing." The next two papers, "McBlare: A robotic Bagpipe Player" by Roger B. Dannenberg, H. Ben Brown and Ron Lupish and "Wind Instrument Playing Humanoid Robots" by Jorge Solis and Atsuo Takanishi focus on human anatomy. McBlare aims to exceed the physical limits of a human bagpipe player and give composers new possibilities, and Humanoid Robots – such as flute- and saxophone-playing robots – serve to study human motor control. For these purposes, some human organs (lungs, lips, tongue, arms and fingers) were mechanically reproduced. On an emotional level, Koji Shibuya introduces a violin-playing robot and the idea of *kansei* (Japanese word similar to “feeling” or “mood”), the purpose being to develop a musical robot that can perform expressive musical sounds. However, expressive playing is not the only goal of this chapter. It is rather important, according to the author, to develop a robot that understands and can express human *kansei* to facilitate smooth human-robot communication. With the introduced topic of human-robot communication it is almost ‘natural’ for the volume to culminate with research on human-robot interactions. In his paper "Multimodal Techniques for Human-Robot Interaction¸" Ajay Kapur concentrates on fusing together musical gesture extraction, musical robotics, and machine musicianship, stating that by using multimodal systems for machine perception of human interaction, robots can be trained to generate a mechanical response. The studies are implemented in India to promote Indian culture and show how technology can be useful in making new music. The last two chapters of this section: "Interactive Improvisation with a Robotic Marimba Player" by Guy Hoffman and Gil Weinberg, and "Interactive Musical System for Multimodal Musician-Humanoid" Interaction by Jorge Solis, Klaus Petersen and Atsuo Takanishi focus on human-robot communication. In their paper, Guy Hoffman and Gil Weinberg describe “Shimon,” an interactive robotic improvisation system for marimba-playing. All the techniques used in order to achieve the requirements of a performing robotic musician are described. Shimon is a pioneer project and the first of its kind, which uses anticipatory gesture-based methods to music viewing. The results emerged from a number of human-subject studies, testing the effect of robotic presence on the synchronization of musicians, as well as the audience’s perception of the duo.

Rounding up all the separate subjects covered in the volume, the last chapter brings into focus the concept and implementation of an interactive musical system for multimodal musician-humanoid interaction. It is explained that, in order for the machines to communicate with humans, they must be able to emulate two of human’s most important perceptual organs: eyes and ears. When it comes to musical interaction, a great part of the performance is still based on improvisation. The future prospects for this project are nonetheless very encouraging. The researchers are working on implementing a feature that would allow musical robots to improve their own performance by listening and analyzing the sound they produce. Still, the biggest challenge lies in the research regarding the ability of musical robots to emulate emotion and recognition through musical sounds, but the technologies are promising and the researchers are optimistic.

Musical Robots and Interactive Multimodal Systems is a very intriguing collection of papers, which shows how the limits of the human body can be easily exceeded thanks to robotics and at the same time underlines that the limits of robotics are further than we could ever imagine. The fourteen chapters give but a glimpse into the world of technology and one should stay eager to see the future developments from the respective authors, or to see what kind of new developments their work will inspire by other researchers in their respective fields.

Article received: May 30, 2019
Article accepted: June 1, 2019

Boris Filipić*
Nezavisni istraživač
Mostar, Bosna i Hercegovina

Osvrt na knjigu *Preobražaji inteligencije: Šta da radimo sa njihovim plavim mozgom?* Catherine Malabou

(*Preobražaji inteligencije: šta da radimo sa njihovim plavim mozgom?*, Catherine Malabou, Biblioteka Philoxenia, Fakultet za medije i komunikacije, 2018, Beograd, 141 str. / sa francuskog preveli Srđa Janković i Suzana Bojović. Naslov originala: Métamorphoses de l'intelligence – Que faire de leur cerveau bleu?, PUF, 2017)

O čemu govorimo kad spominjemo inteligenciju? Ako tvrdimo da rapidnim razvojem novih tehnologija mašine postaju umjetno inteligentne, na osnovu koje definicije prirodne inteligencije povlačimo ovu paralelu/analogiju? Catherine Malabou, prominentna suvremena francuska filozofkinja, hegelijanka, učenica Jacquesa Derride, posvetila je posljednje decenije svoga rada izučavanju neuroznanosti, posebice uloge mozga – organa na osnovu čijih operacija se u svjetlu novih uzbudljivih neuroznanstvenih otkrića razumijevamo ljudima kao - inteligentnim subjektima. Upravo su karakteristike mozga, točnije njegova plastičnost tj. sposobnost promjene i readaptacije neurosinaptičkih spona koje ga čine funkcionalnim uslijed izvanjskih utjecaja okoline, kulture ili traume, ono što nas kao inteligentnu vrstu, prema autoričinoj knjizi *Što da radimo sa svojim mozgom?* iz 2004. godine, značajno distingvira od sintetičkih inteligentnih entiteta u povoju. Plastičnost je koncept kojim se Malabou bavi tokom cijele svoje akademske karijere, preuzimajući nasljeđe iz Hegelove *Fenomenologije duha*. Međutim, samo godinu dana nakon što je u pomenutoj knjizi iznijela ovu, po čovjeka spram mašine emancipatornu tezu, otkriće neurosinaptičkog čipa tvrke IBM obara njene temeljne pretpostavke. Ovaj novi precizni sintetički simulator moždane plastičnosti proširio je ionako već široko polje implikacija i pitanja. Dok je filozofsko pitanje iz knjige *Što da radimo sa svojim mozgom?* imalo izvjesnu humanističku i političku elaboraciju, pitanje Malabouine najnovije knjige *Preobražaji inteligencije: Što da*

* Author's contact information: kukoljpromet@gmail.com

radimo s njihovim plavim mozgom? djeluje kao dramatična promjena mjesta traženja odgovora na ljudsku posebnost.

Koncept inteligencije, nesagledljiv i apstraktan, te, kako se ispostavilo kroz svoju naučnu istoriju, krajnje štetno razumijevan, Malabou razmatra kroz tri paradigmatske smjene koje ujedno i strukturiraju knjigu.

Početke istraživanja inteligencije pronalazimo krajem 19. stoljeća, napose kroz prizmu novonastalih disciplina fiziološke i eksperimentalne psihologije. Eugenička istraživanja Francisa Galtona, te prve psihometrijske skale Theodorea Simonea i Alfreda Binneta otvaraju stoljeće puno malignih aproprijacija i smrtonosnih interpretacija inicijalnih naučnih ideja. Iako obe škole polaze iz različitih, inicijalno dobrohotnih ideala i namjera, prema Malabou, obje konvergiraju i kulminiraju u nehumanim državnim politikama sterilizacije ogromnog broja ljudi, bjelačkim rasističkim politikama, te naposljetku nacističkim koncentracijskim logorima. Eugenika Francisa Galtona otvorila je imaginarij zapadnog svijeta o pospješivanju ljudske vrste selektivnim razmnožavanjem 'genijalnih ljudi', konstruiran po redukcionističkom shvaćanju Darwinove teorije evolucije i prirodne selekcije. Kroz 20. stoljeće primarno razumijevanje inteligencije (genijalnosti, faktora G, generalne mentalne sposobnosti) je hereditarno i urođeno. Psihometrijske skale utabale su put brojnim, i danas postojećim, testovima inteligencije koji su mahom korišteni za marginalizaciju širokih slojeva čovječanstva, pretežito po rasnoj i društvenoj osnovi. Potraga za 'genom inteligentnosti' princip je bihevioralne genetike čije su nasljedne pretpostavke oborene konačnim sekvenciranjem i mapiranjem samog ljudskog genoma 2003. godine, uslijed kojeg se ispostavilo da takav gen - ne postoji.

Važna je i linija otpora koja kroz 20. vijek prati psihologiju, te na koju se Malabou kroz cijelu knjigu referira. Filozofska formacija kornjače, Malabou se koristi ilustrativnim primjerom borbenog stava Rimljana, predstavlja niz filozofa (Od Heideggera do Foucaulta, Agambena i Deleuzea) koji, vođeni Bergsonovim centurionskim idejama o instinktu nasuprot inteligencije, zauzimaju niz označitelja za psihologiju i njene krake kao nizove štitova koji im omogućuju distanciranje od ozbiljnog suočavanja s problemom. Tako je psihologija, a s njom i njen pojam inteligencije – policijski normativ, aparat za poslušnost, priznanje i disciplinu, naposljetku i jednaka – gluposti. Pitanje koje Malabou otvara je, što će ova formacija kornjače poduzeti kad se suoči s nečim mnogo neizbježnijim nakon rapidnog razvoja neuroznanosti i kibernetike, ne tajeći da je u svom ranijem radu i sama bila dijelom formacije kornjače. Zašto ne uzeti u obzir i instinkt i intelekt i inteligenciju suprotstavljenu samoj sebi, zajedno sa njenom glupošću? Više puta proklamiran slogan Malabou: Život je jedan.

20. stoljeće završilo je kao bezuspješan pokušaj kvantifikacije i evaluacije ideološki shvaćenog, ali naposljetku krajnje eluzivnog koncepta inteligencije. Početkom 21. stoljeća u svjetlu otkrića molekularne biologije i neuroznanosti počinje epigenetička paradigma inteligencije, u kojoj se još uvijek nalazimo, a koja razlikuje genotip (genetski zapis) i fenotip (fiziološke karakteristike svakog pojedinca), te sve što

djeluje na njihov odnos. Napredak u neuroznanosti pokazao je da kompleksnost mozga, tj. trilijunskog broja neurosinaptičkih poveznica u njemu, daleko nadmašuje ljudski genom koji čini tek 30 000 gena. Epigenetički razvoj inteligencije, točnije uzajaman odnos genotipa i fenotipa, prenatalno, natalno, te u nešto manjoj mjeri i tokom ostatka života, prema Malabou, koncept inteligencije opisuje kao metod, a ne entitet, kao fleksibilnu stabilizirajuću putanju grešaka i učenja, a ne fiksirano svojstvo. U ovome joj naročito pomažu ranija istraživanja Pierrea Piageta. Na stanovit način inteligencija se poistovjećuje sa samim konceptom plasticiteta. Dijalektički odnos epigeneze čine genetske predispozicije, te u većoj mjeri iskustvo, učenje i kultura predstavljaju značajniji determinizam od onog koji je u 20. stoljeću pripisivan genetskom. Epigenetički uslovljen razvoj mozga Malabou pronalazi u idejama habitusa – ishodišnog mjesta inteligencije ispreplitanja biološkog i društvenog alata, tijela i mozga i njegove uslovljivosti – prirodnih sposobnosti ostvarivanja neprirodnih sposobnosti, iz rada društvenog teoretičara Pierrea Bourdieua. Prema Bourdieu svaka inteligencija je prije svega vještačka.

Iako je mozak rezultat organske evolucije, epigenetika naglašava prostor ontogenetičke individualnosti mozga – otkriće IBM neurosinaptičkog čipa, te naročito čip TrueNorth omogućilo je pothvate sintetiziranja ljudskog mozga u SAD-u i Europskoj Uniji. Blue Brain Project nastoji modelirati mozak koji će uz plastičnost sve jače neurosinaptičke arhitekture otvoriti prostor za nadilaženje ljudske inteligencije.

Ako je slaba umjetna inteligencija automatska tehnologija koju svakodnevno koristimo, primjerice Googleovi samovozeći automobil, a srednja ona koja će uspješno položiti i Turingov test, tj. bit će nam ravnopravna, eksponencijalni rast kapaciteta komputacijske tehnologije u (skoroj) budućnosti će nas suočiti s umjetnom superinteligencijom, entitetom koji će imati mogućnost kontrole nad vlastitom modifikacijom, tj. plastičnošću. Pesimizam i katastrofična predskazivanja česti su pokušaji mišljenja o ovoj, trećoj smjeni paradigme inteligencije.

Malabou kroz cijelu knjigu, iako s razumijevanjem prihvaćajući upozorenja Stephena Hawkinga, primjerice, ipak nastoji otkloniti potrebu za tehnofobnim formiranjem 'kornjača' u budućnosti fleksibilnih automatizama vještačke inteligencije.

Budućnost svakako donosi ljudske manipulacije genetikom. Modificiranje (ne)željenih gena već je opcija u našoj post-genomskoj eri (genetički skalpel crispr-cas9 napravljen je 2012. godine), ali još uvijek uz prisutnost opasnosti duha ondašnje Galtonove eugenike, no mašine će u skoroj budućnosti biti sposobne epigenetički se automodificirati i automanipulirati svojim vlastitim kodom. Plastičnost postaje tako raspoloživa (samo)programiranju prema računalnoj arhitekturi. Mapiranje sintetičkog mozga novi je projekt sekvenciranja ljudskog genoma. Ostaje otvoreno pitanje podložnosti ovih novih polja društvenoj i političkoj manipulaciji. Upravo Bourdieu artikulira ovaj problem. Kao što tijela preuzimaju naviku unutar društvenog poretka, tako fizički i moždana shema biva podložna automatizaciji. Očekivanja i strahovanja od zloupotrebe znanosti srodne onoj iz perioda prve paradigme ogledaju se u očitom potencijalu sprege političkog poretka i programiranja neuroplastičnosti. Nove mogućnosti uniformiranja psihe i tijela prisutne su već u kalkulacijama vremena i mjestima izlaganja pojedinca marketingu, no za Malabou je današnja kognitivna era izazov za kojeg moramo odbaciti ranije dihotomije intelekta i inteligencije, čovjeka i mašine. Potrebna je izvjesna smjena zastarjelih diskursa. Epigenetički potencijal treba tražiti u dijalektici prirodnog i mehaničkog. Zajedničko posmatranje automatizma i spontanosti otvara ovo polje.

Prateći filozofiju tehnologije Johna Deweya, Malabou ističe socijalni karakter inteligencije, nužnost komunikacije unutar partikularne lokalne zajednice u želji za adekvatnijem rješavanju problema. U tome je zapravo i fundament demokracije. Plastičnost kao definirajuća karakteristika mozga je po sebi nejasna, podložna je i za ostvarivanje slobode i za nametanje dominacije, kao i za pasivizaciju. Automatizam mašine se burdijeovski čini kao učinkovita karakteristika pri naumu za nametanjem političkog poretka nad društvom i pojedincima. Malabou ovdje uvodi drugu, vrlo važnu karakteristiku inteligentne mašine – sposobnost greške. Prema Alanu Turingu, mašina koja je nepogrešiva ne može biti inteligentna, odstupanje od automatiziranog i discipliniranog ponašanja je ono što inteligenciju čini inteligentnom. Automatizam jednom prekinut greškom iziskuje nagli skok u kreativnost readaptacije. Čak i kod razvoja djece postoje fazni skokovi inteligencije razdvojeni naglim prekidima, primjerice Piaget formulira niz prelaznih točki kod prilagođavanja prilikom odrastanja gdje su upravo greške uslov daljnjeg razvoja inteligencije pojedinca.

Greška, prekid, smetnja u naviknutom/automatiziranom funkcioniranju preduslov je iznalaženja novih modusa popravke i regeneracije sistema uslijed novih problema. To je ono što inteligencija u suštini jeste. Obuhvatno razmatranje svih ovih aspekata donosi zaključak o autonomnosti unutar mehaničkog automatizma inteligentne mašine. Na stanovit način, senzibilitet za vanjski svijet koji je potreban umjetnoj inteligenciji za učenje na greškama prije upozorava na ljudsku podložnost automatizaciji, nego vice versa.

Slike Jacksona Pollocka su primjer automativne umjetnosti. Ni namjerni, ni nasumični potezi pri slikanju, nego nešto između to dvoje, ono je što Malabou, slijedeći dijalog dvojice naučnika iz filma *Ex Machina* (Alex Garland, 2014) o umjetnoj inteligenciji i automatizmu, smatra točkom istinskog automatizma – onog između kontingencije i nužnosti. U filmu mašina savršeno precizno 'simulira' Pollocka.

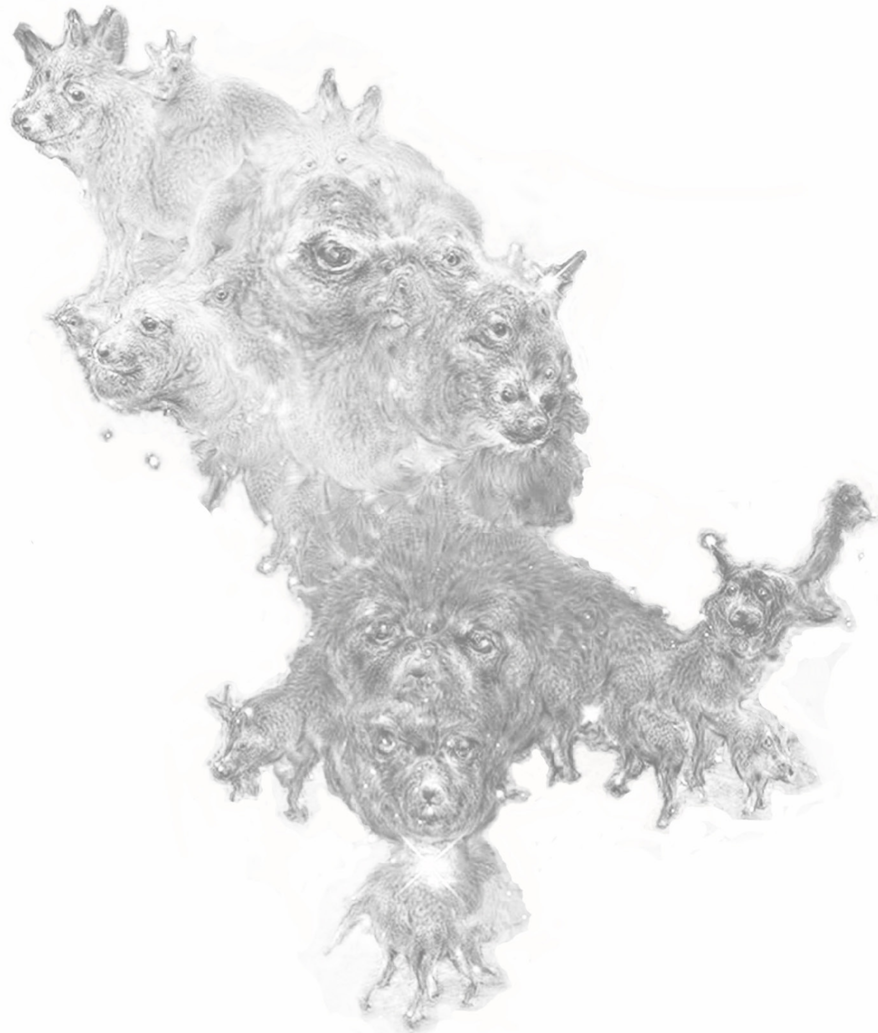
Programiranje plasticiteta, dakle, također biva transformirano. Razvoj neuroznanosti učinio je ranije pomenutu formaciju kornjače, a primarno biopolitičku kritiku izlišnom u susret novoj smjeni paradigme. Filozofski štitovi padaju jer se sve više čini da se i mnogo toga tradicionalno shvaćeno kao transcendentalno može mapirati neuronima. U guranju granica prilikom novih horizonata koji nas čekaju, Malabou predviđa razvoj humanističkih nauka ka svojim neuro predznacima. Neurobiologija već iskazuje svoje normative tumačeći simboličko i transcendentalno kroz termine neuronske prijemčivosti. Zanimljiva je rasprava o tonalnoj i atonalnoj muzici u kojoj neurobiolozi kratkovidno tumače tonalni sistem kao prirodno više stimulirajući po ljudski mozak, te tim vrijedniji od atonalnog. Prema neurobiolozima, ljudski mozak je skloniji figurativnoj umjetnosti, nego apstraktnoj. Upravo izazivanje takvih mijopičnih esencijalizama zadatak je novih pedagogija kroz neurolingvistiku, neuroekonomiju – neurohumanističke nauke.

Poroznost granica između neuroznanosti i humanističkih nauka, prateći Foucaultovo čitanje Kanta, te njegov imperativ guranja granica ljudskog uma ka transgresiji transcendentalnog kako bi se utvrdile nove granice, otvaraju prostor za kritičku ontologiju nas samih. Nove granice trebaju nam pomoći ocrtati upravo (neuro)humanističke nauke nove paradigme. Mozak je slobodan, nije predviđen za normative, podložan je prekidima i readaptaciji, stoga je nužno kritički pristupiti automatizmima normalizacije bilo koje vrste, uvijek otvarajući prostor novim načinima čitanja, slušanja, mišljenja. Manuel Castells predvidio je današnju sve više upražnjavanu praksu učenja na daljinu. Fleksibilniji pristup edukaciji kroz online alate omogućio je učenicima lakše balansiranje društvenog, porodičnog i edukacijskog vremena. Učenici biraju što, kad i gdje će učiti. Demokratizacija nove virtualne sfere učenja ogleda se u izmjeni modela ex cathedra, posljedično i shvaćanju inteligencije. Informacija je najprije dematerijalizirana, da bi naposljetku sama postala nova materijalnost. Cyberspace je otvorio novu univerzalnost, lišenu centraliziranog značenja i hijerarhije. Razvoj inteligentnih sustava koji se mijenjaju prema ranijim restabilizacijama uslijed eventualnih grešaka u pravi plan stavio je sinergiju vještina, znanja, sjećanja učenika. Novi edukativni digitalni alati omogućuju različitim ljudima i različitim 'inteligencijama' kooperaciju, optimizaciju mašte i intelektualne energije bez prepreka karakterističnih za sve ranije epohe. Mi smo dio novog digitalnog ekosistema koliko je on dio nas. Naša kreativnost, kapaciteti za suradnju i otvorenost unutar nastajućih inteligentnih sustava izmještenog učenja također su alati za nove pedagogije. Naposljetku, Malabou zaokružujući knjigu u duh filozofskog imperativa hvatanja u koštac s novim inteligencijama, živim i neživim, podsjeća na grčki koncept *metis* – lukaviju inteligenciju, sposobnost navigacije kroz nove probleme, iznalaženja novih metoda, mudrost, iskusnost. Prije nego li je ustoličen Platonov logos, metis je označavao inteligenciju manifestiranu u krajnje neizvjesnim situacijama. Ostaje nam otkriti i upregnuti metis za novo doba.

Preobraženja inteligencije srčan su i investiran rad Catherine Malabou u svjetlu presijecanja novih disciplina i njihovih novih implikacija po čovjeka i društvo. Pogrešno rezonovanje iz knjige *Šta da radimo sa svojim mozgom* iz 2004. godine zajedno s navedenim, odskočna su daska u uzbudljivo novo naučno polje koje se pred nama otvara. Ono što možda ostaje nerazmotreno u ovoj dosta podrobnoj i detaljnoj analizi sprege filozofije, neurokibernetike i novih pedagogija je ekonomski okoliš u kojem do nje dolazi. Knjiga naposljetku govori s dosta optimizma spram novih horizonata, dok su pesimistična predviđanja divljeg neoliberalnog okvira unutar kojeg same države već kompetitivno traže 'singularnost', tj. umjetnu superinteligenciju za militarne svrhe tek ovlaš dotaknuta. Ipak, kompletnost Malabouinog rada ogleda se upravo u perspektivi i razumijevanju otvorenosti naše neuroplastične kibernetičke budućnosti. Suština pojma inteligencije je da naposljetku samo ona može razriješiti svoj problem. Postoji samo jedan život.

Article received: May 30, 2019
Article accepted: June 1, 2019

Contributors' Biographies



Miško Šuvaković (1954) is Dean of Faculty of Media and Communications, Singidunum University, Belgrade, Serbia, and professor of theory of art and media in the Ph.D. program of trans-disciplinary humanities and theory of art. He works on theories of avant-gardes, neo-avant-gardes and contemporary art. He has published around 50 books in the Serbian, Slovenian and English languages.

Asher Tobin Chodos is a composer, pianist and musicologist. He has been named a fellow of the Brubeck Institute, the Asian Cultural Council, and the Ucross Foundation. He has performed to wide acclaim around the world, and his compositions have been commissioned by many noted contemporary ensembles.

Adriana Sabo (1989) is a Junior researcher at the Department of Musicology, Faculty of Music in Belgrade. She is a PhD candidate at the same department, with master degrees in musicology (2012) and gender studies (2015). She is a recipient of the scholarship given by the Ministry of Education, Science and Cultural Development of the Republic of Serbia, member of the Serbian Musicological Society and a contributor for the Center for Popular Music Research (Belgrade). Her research mainly focuses on the issues of gender and music, popular music, otherness and connections between music and politics.

Gennady Stolyarov II, FSA, ACAS, MAAA, CPCU, ARe, ARC, API, AIS, AIE, AIAF, is the Chairman of the United States Transhumanist Party / Transhuman Party, an actuary, independent philosophical essayist, science-fiction novelist, poet, amateur mathematician, and creator of 88 technologically enabled musical works. In December 2013, Mr. Stolyarov published *Death is Wrong*, an ambitious children's book on life extension. Mr. Stolyarov is the Editor-in-Chief of *The Rational Argumentator*, an online magazine championing the principles of reason, rights, and progress, which has been published since August 2002. Mr. Stolyarov regularly produces YouTube Videos discussing life extension, politics, philosophy, and related subjects. Mr. Stolyarov can be contacted at gennadystolyarovii@gmail.com.

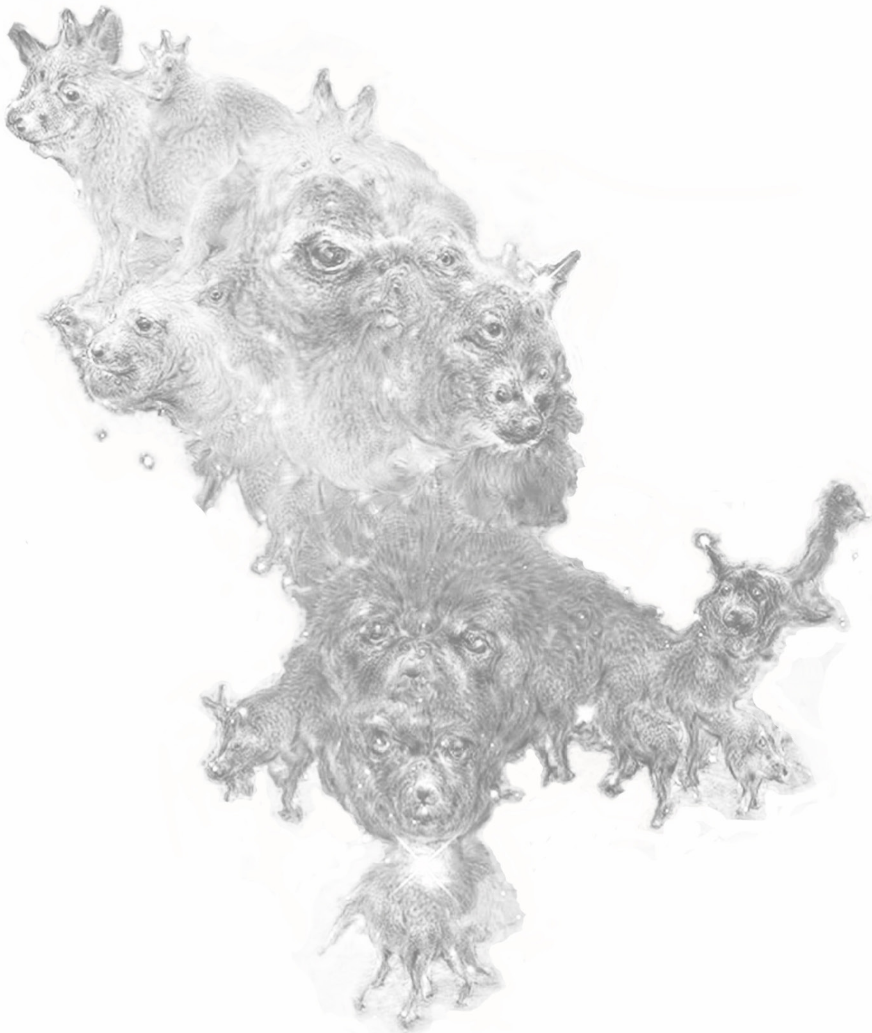
Harun Zulić (1992), is a professional musician who accesses music through writing, teaching, composing, performance as well as the event organizing activities. He is born in B&H and is currently based in Toronto (Canada) where he is employed at the Canada Music Academy and Music at Home. As an adventurous researcher of music, Zulić has participated as a lecturer at several international symposia in B&H and Serbia, and has published articles in several proceedings and journals such as "Didactic guidepost – Journal for Teaching theory and practice", "Scientific Meeting Vlado Milošević: Tradition as an Inspiration – Thematic collection of the scientific conference in 2017.", "Balkan Art Forum – Proceedings of scientific conference in 2017", etc. In his published works, his interest includes the area of Bosnian-Herzegovinian music, as well as world music problems in the area of classical music. He holds a Master's Degree in Music Theory and Education – Musical Forms and Styles from the University of Sarajevo (Forming in instrumental works of Arvo Pärt composed in tintinnabuli technique, supervised by Ivan Čavlović).

Daniel Becker studied art history and German language and literature in Hamburg, where he wrote his master's thesis on *Narcissus in Media Art*. His focus is on contemporary art, digital aesthetics and the history of art and technology. From 2012–2014 he worked on the digitization project of the Hamburger Kunsthalle. From 2014–2018 he was a scholarship holder at the International Doctoral Program *Mimesis* of the Ludwig-Maximilians-University Munich and is doing his doctorate there on the topic of *Schnitt-Stellen. Aesthetic implications of interfaces*. Since February 2019 he is a research assistant in the department of Tristan Weddigen. Here he is pursuing a project on *The Aquatic and Fluids as a Sculptural Element in Italian Art*.

Robert Max Williams is a computer science student and machine learning researcher. Currently he is completing his computer science undergraduate at the University of Louisville. His fields of interest include software design, esoteric programming languages and neural networks. He also enjoys traditional dance and creating electronic music.

Roman V. Yampolskiy holds a PhD degree from the Department of Computer Science and Engineering at the University at Buffalo. There he was a recipient of a four year NSF (National Science Foundation) IGERT (Integrative Graduate Education and Research Traineeship) fellowship. Before beginning his doctoral studies Dr. Yampolskiy received a BS/MS (High Honors) combined degree in Computer Science from Rochester Institute of Technology, NY, USA. After completing his PhD dissertation Dr. Yampolskiy held a position of an Affiliate Academic at the Center for Advanced Spatial Analysis, University of London, College of London. In 2008 Dr. Yampolskiy accepted an assistant professor position at the Speed School of Engineering, University of Louisville, KY. He had previously conducted research at the Laboratory for Applied Computing (currently known as Center for Advancing the Study of Infrastructure) at the Rochester Institute of Technology and at the Center for Unified Biometrics and Sensors at the University at Buffalo. Dr. Yampolskiy is also an alumnus of Singularity University (GSP2012) and a visiting fellow of the Singularity Institute. Dr. Yampolskiy's main areas of interest are behavioral biometrics, digital forensics, pattern recognition, genetic algorithms, neural networks, artificial intelligence and games. Dr. Yampolskiy is an author of over 100 publications including multiple journal articles and books. His research has been cited by numerous scientists and profiled in popular magazines both American and foreign (New Scientist, Poker Magazine, Science World Magazine), dozens of websites (BBC, MSNBC, Yahoo! News) and on radio (German National Radio, Alex Jones Show). Reports about his work have attracted international attention and have been translated into many languages including Czech, Danish, Dutch, French, German, Hungarian, Italian, Polish, Romanian, and Spanish.

Guidelines for Authors



Guidelines for authors

Authors must submit original, unpublished articles.

All the manuscripts should be accompanied by author's name, affiliation, e-mail address, and a short biography (up to 150 words per author). Articles can be submitted in English (preferably) and Bosnian.

Manuscripts should be written in .doc or .docx format, in Times New Roman font, font size 12 with 1.5 line-spacing.

Original scholarly paper intended for sections The Main Theme and Beyond the Main Theme should include a short abstract (100-200 words), 5-10 keywords, as well as the summary (500 words). For articles in Bosnian, summary must be written in English. Do not include citations in the abstract. Keywords must be chosen appropriately in order to be relevant to the subject and content of the paper.

Regarding the citations, authors should use the author-date system with the separate bibliography, following the guidelines given in Chicago Manual of Style (Chicago: University of Chicago Press, 2010; http://www.chicagomanualofstyle.org/tools_citationguide.html). Please note that the list of references (bibliography) given at the end of the article must only include works that are cited in text.

Book, conference, and festival reviews should bring to attention relevant and valuable contributions or events that are in interest scope of our Journal. Reviews must contain a dose of critical appraisal instead of being written merely as summary. The title of the book review should include necessary information regarding the volume, as in following example:

- William Myers, *Bio Art – Altered Realities*. London: Thames and Hudson, 2015, 256 pp., ISBN 9780500239322
- *Margins, Futures and Tasks of Aesthetics*, Conference of the IAA, Helsinki, Finland, July 5–7, 2018.
- Sonemus Fest, Sarajevo, Bosnia and Herzegovina, April 16–21, 2018.

Manuscripts can be equipped with photos, illustrations, drawings, and tables. These should be of good quality (resolution higher than 300 dpi), in .jpg or .tiff formats, and submitted as files separate from the text. All visual materials must have permission for publishing from the author, photographer or the respected owner of the rights.

Word count:

- Original scholarly papers (Main Theme and Beyond the Main Theme sections) – 3000-6000 words
- Book, conference, and festival reviews – 1000-1500 words
- Interviews – 1000-1500 words