WHIPLASH DRUM DEVICE

THE CREATION OF ARTISTIC MEDIA BY ARTIFICAL INTELLIGENCE AND ITS EFFECTS ON HOW HUMANS PERCEIVE CREATIVITY

A Thesis Prospectus
In STS 4500
Presented to
The Faculty of the
School of Engineering and Applied Science
University of Virginia
In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science in Electrical Engineering

By
Leonardo Anselmo

December 2, 2022

Technical Team Members:
Leonardo Anselmo, Uriel Gomez Ibarra, Max McCullough, John Lilly, Davis Lydon

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

ADVISORS

Rider Foley, Department of Engineering and Society

Harry Powell, Department of Electrical and Computer Engineering

Introduction

One of the largest problems in the realm of music is finding the most effective way to practice. Preparing for a rehearsal can be a stressful and frustrating time if one does not have the proper materials or infrastructure to play. Of course, one's practice style also depends on the type of music they are playing. This paper will focus on musicians that wish to play together, whether it be in a band or for recreation. Some popular methods of practice are to play alongside digital audio or playing with other musicians, though both methods come with their own issues. Playing along with digital audio does not provide the same necessary feedback as playing along with other physical instruments. Furthermore, while playing in the presence of other musicians is one of the best methods of practicing music (Nielsen, 1999), not everyone has the time or capability to meet in person to practice. Playing music live with others involves subtleties and cues that are difficult to replicate in any other format. A method must exist, then, where players can practice with a physical instrument while working on their individual parts.

Practicing an instrument can be a challenging task if one does not know where to begin. Searching for educational musical equipment is even more difficult, as each developer has their own idea for the quickest improvement in musical skills. There is currently a disconnect between music technology and music education. There are musical apps and websites that beginners can use to learn the basics of an instrument (Crazy Ootka Software AB, 2022). There are also machines that are more capable of playing instruments than any human who has ever lived (Polyend, 2015). The problem arises when we try to find a product in the center of both extremes – something physical and capable of assisting players using methods that are known to increase productivity.

Devices advanced enough to replicate the subtleties and nuances of human creativity pose an obvious problem to human musicians. With the introduction of neural networks, research has already produced several AI capable of generating digital artwork indistinguishable from human creation (DreamStudio, 2022). People have even won awards for AI-generated pieces, passing them as their own creations (Metz, 2022). Below in **Figure 1** is an award-winning AI-generated piece. Although AI currently performs best in the world of digital art, the technology will eventually spread to other regions of creativity.

Figure 1: Award Winning AI Generated Piece Titled "Théâtre D'opéra Spatial" (Metz, 2022)

Our technical project, titled Whiplash, after the movie of the same name (Chazelle, 2014), aims to assist musicians with practicing music when other human musicians cannot be present. While our project intends to achieve a net-positive outcome for education, it has many implications on the effects of technology on the arts. I intend to explore how technology is changing the way we perceive creativity.

Technical Topic: Whiplash Drum Device

There currently exist several working models of robotic drum-playing machines. One of the most prominent robotic-humanoid drummers is “Stickboy” (Barnes, 2007). Stickboy is a six-armed robot that has been on tour since 2007. Stickboy must be programmed to play, as it has no internal algorithm to play autonomously to a specific song. This means that while Stickboy may be much more capable of playing than any drummer, it is also incapable of creating its own music. On the other end of the spectrum, the Polyend Perc (Polyend, 2015) is an entirely functional device capable of playing an entire drum set. The device involves aluminum enclosures that contain small cylinders to “beat” the drum, rather than strike it. The Perc also has no algorithm to automatically track music and play along, as it must be entirely programmed to the user’s specifications. Below in **Figure 2** are the two devices.

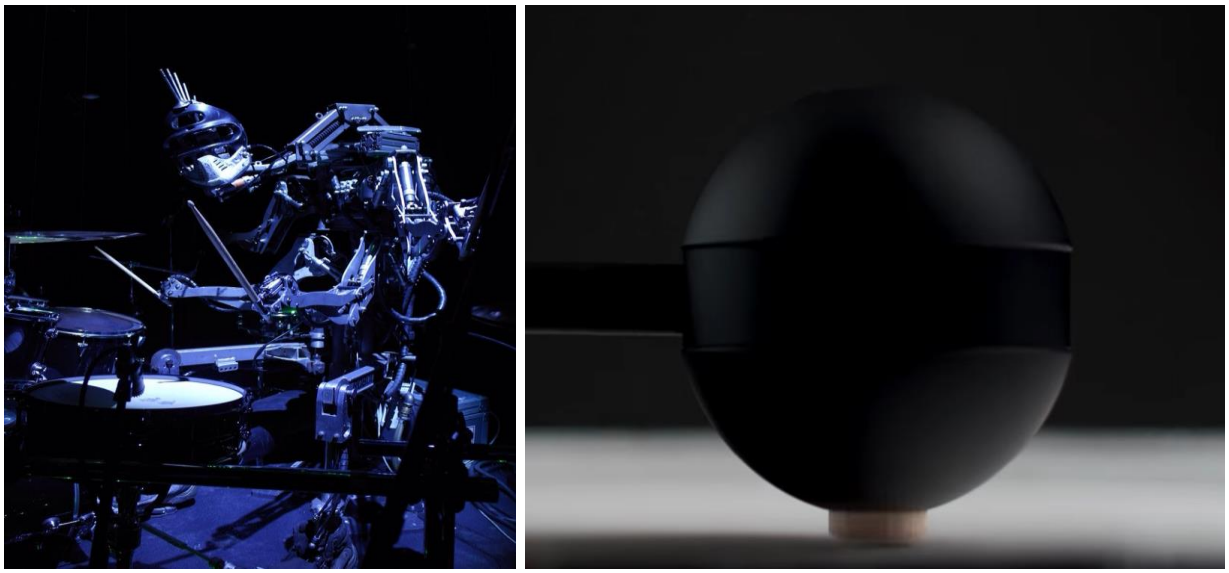

Figure 2: [From Left to Right] Stickboy (Barnes, 2007) and Polyend Perc (Polyend, 2015)

While both Stickboy and Polyend Perc fill their niche of being able to display drumming capability, neither can play without being programmed to a song's specifications. Our device, Whiplash, utilizes a digital signal processing (DSP) element to autonomously track the beat of a song and play a physical drum to that beat. Rather than hard code each drum hit as the current technology does, we can focus on making our project generalizable to any musical audio using a temporal algorithm. Temporal algorithms, or software that focuses on time precision, have been used previously to train robotic musicians to play music (Fryen et al., 2020), though Whiplash does not utilize any reinforcement training.

The chipset we have chosen is a Texas Instruments (TI) mixed signal processing 432 (MSP432) chip. We are also using a TI CC3220SF launchpad (Texas Instruments, 2018) that enables us to access each input/output of the chip easily. The user plays audio from their preferred device (phone, computer, etc.) and the signal is passed to the microcontroller. The microcontroller then performs the DSP to parse out the song's beat. The beat signals are then transformed internally into pulse-width modulated (PWM) signals that can be read by servos, specifically HS-805BB servos (HiTEC, 2007). The generated PWM signals are sent to a servo

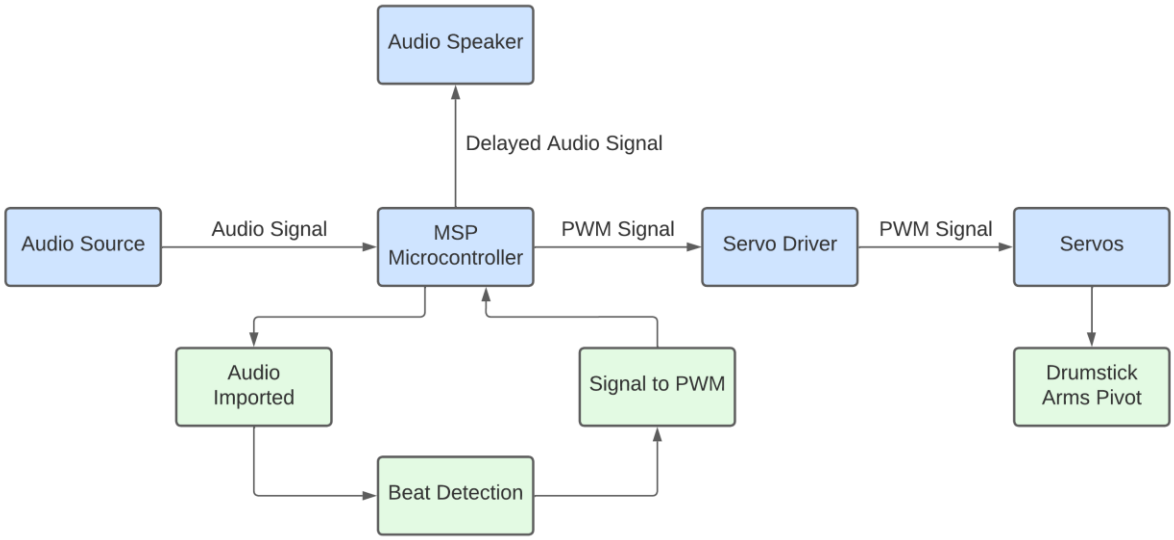

Figure 3: Whiplash Block Diagram (Anselmo, 2022)

driver printed circuit board (PCB) created to protect the MSP432 and servos from overcurrent. The servo driver amplifies the signals and sends them off to the servos, enabling them to spin the drumsticks and hit the drum. **Figure 3** above shows the process in a block diagram.

The device will have several other features. The strength of the drumstick hit will depend on how loud the song is playing at the moment of that hit. The device will also have an output to a speaker, so the user is able to hear the music live. The output speaker will have a delay when playing audio so the device can hit the drum exactly when it is expected to. The physical body of the device was designed using FreeCAD (Riegel, 2002) and printed using a 3D printer. The body has mounts for the servos, while the drumsticks will be screwed into a mounting plate on the servos themselves. A model of the 3D design is provided in **Figure 4** below.

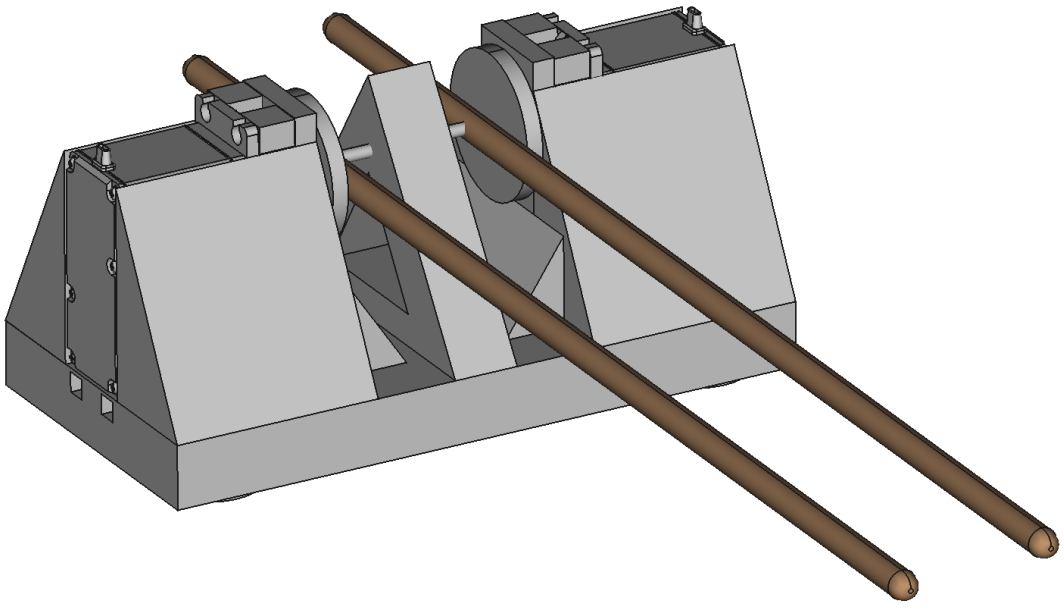

Figure 4: Whiplash 3D Model (Anselmo, 2022)

Whiplash helps in solving the previously stated problem by providing a musical companion to practice with. The user will receive the benefits of physical feedback while being able to practice with a partner that cannot tire. The user's only responsibility is to choose their desired song and Whiplash will do the rest. One may argue that Whiplash is no different than a metronome. While Whiplash only plays to the beat of a song, it also adapts to any shift in tempo and beat that may occur during the song. A metronome is only capable of playing a constant beat, so Whiplash enables a musician to seamlessly practice an entire piece with no pauses to adapt the device to each part of a song.

While the device provides an extra musician where there may not be one, it also provides context to the human and social dimensions of technology in general. A common side-effect of research is that it typically leads to revelations that are later applied to wildly separate fields (Cetinic & She, 2022). For example, our beat detection algorithm enables humans to explore the digital bits of a song and further our understanding of music. Whiplash is a bridge technology that may eventually expand to devices with onboard AI, capable of generating their own algorithmic compositions. The device currently is nowhere near capable of fully replacing a human musician; however, if the project were to be expanded and researched further, it may very well be. The device can be reformatted to parse out individual drums in a song or "taught" by machine learning algorithms to generate its own drum track. This would allow the device to interface with a full-scale drum set and create its own music as a human would. This issue then expands into technology's effect on musical creativity and general creativity.

STS Topic: The Creation of Artistic Media by Artificial Intelligence and its Effects on How Humans Perceive Creativity

AI is a double-edged sword because it both makes our lives easier while replacing the jobs of others. Technology can be described as a human means of interfacing with nature, while AI is described as interfacing human learning with machine learning technology (Grba, 2022). Technology is created to assist humans with everyday living, whether it makes life more convenient, more luxurious, or simply allows us to survive. The problem arises when we view technology in terms of who or what it can (and may) replace.

Technology, so far, has been advantageous in the development of creativity. Artists and consumers can much more easily access, use, and create media. Currently, the analog world is digitizing, allowing computers and devices to perform the tasks that humans once did. Software engineers can exhibit their creativity by writing algorithms to generate artistic media. Humans have become more connected through the metamorphosis from analog to digital media.

The primary concern of AI is the replacement of human beings by technology. Such an occurrence has already taken place in industries such as manufacturing, automotive, and sales (Flynn, 2022), while the creative industry is a new focus. Traditional artists cannot compete with the vast processing speed of current technology and are therefore losing the spotlight. Software engineers interested in machine learning and neural networks can thrive in the current environment of automated creativity.

The STS framework I will be using is technological momentum. The primary social groups affected in my research are musicians, artists, software engineers, and consumers of both art and music. In the book “The Evolution of Large Technological Systems” Hughes reinforces the previously stated idea that technology is a means of interfacing with nature. He states,

"Technological systems solve problems or fulfill goals using whatever means are available…the problems have to do mostly with reordering the physical world in ways considered useful…" (Hughes, 1987). Technology in the creative industry has evolved and gained momentum to perform this task.

In 1957, the first machine learning algorithm was invented to play checkers. Over years of research and development of more robust hardware, innovators designed generative adversarial networks (GANs) in 2014 (Poltronieri, 2019). GANs are the AIs responsible for creating art, built from the infrastructure of algorithms that once served an entirely different purpose. Such is an example of Hughes' innovation in technological momentum. Further investigation into the expansion of GANs also reveals the development and invention stages of such an AI. Fixing this framework to my research will allow me to pinpoint the niches in which AI fits itself in to the point where the social construction of the technology becomes clear. The continued invention of AI would be described as "radical" by Hughes as it has spread itself into many regions of human innovation.

New forms of technological entrepreneurship have also arisen in the art world. For example, Spotify commercialized streaming technology by paying artists to host their discography on their own servers. In 2019, streaming accounted for 83% of all recorded music revenue, as shown in **Figure 5** to the right (Friedlander, 2022). Hughes explains that the growth of technology is a result of a drive for high

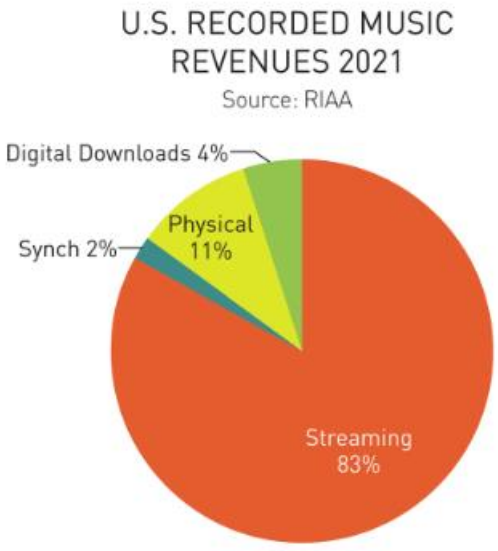

Figure 5: U.S. Recorded Music Revenues 2021 (Friedlander, 2022)

diversity and competition in a field. Such an idea is reinforced by a framework on technological entrepreneurship, stating, “Technological entrepreneurship…can be defined as the setting up of new enterprises by individuals…to exploit technological innovation” (Siyanbola, 2011). Spotify, YouTube, Amazon, and many others compete today to dominate the media streaming market by exploiting algorithms to provide humans with recommended content. Similarly, the AI-generated art market is expanding as well. DreamStudio AI, Dall-E Mini (Dayma, 2021), MidJourney (Holz, 2022), and several other GANs compete in a new market of AI media. While GANs like DreamStudio have utilized a paid system, others stay entirely free due to advertisements. Diving deeper into AI art through the lens of technological entrepreneurship will provide further insight into the incentives and payoffs of producing such technology.

To display the differences in several GANs, I input the prompt “A Vision of a Parallel Dimension” into DreamStudio, Dall-E Mini, and MidJourney, as shown in **Figure 6** below.

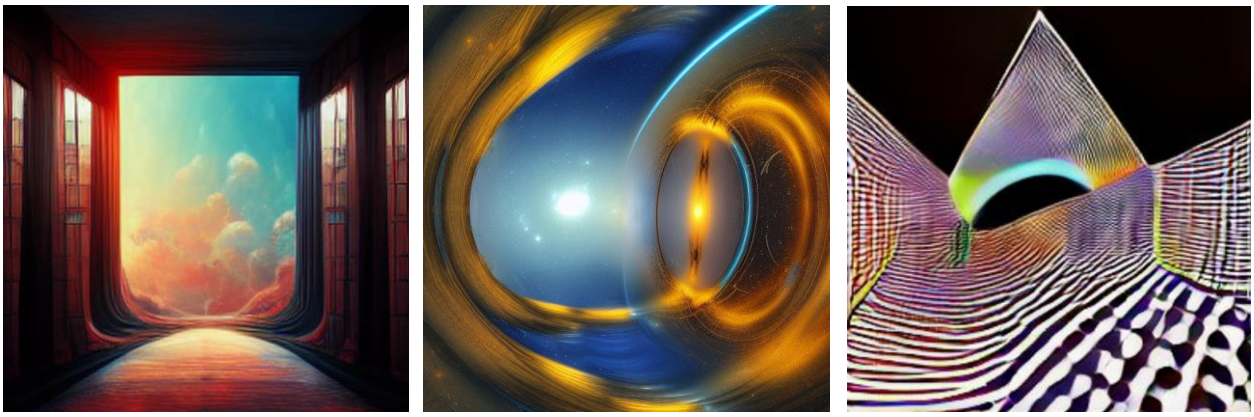

Figure 6: [From Left to Right] "A Vision of a Parallel Dimension" Prompt using MidJourney (Holz, 2022), DreamStudio (DreamStudio, 2022), and Dall-E Mini (Dayma, 2021)

Research Question

My research question is as follows: How has technology affected the way we perceive creativity? For most of human history, art was created solely by humans; however, through AI-generation, new methods of artistic expression have been introduced, challenging the current definition of what "art" is. As the inventors of technology that threatens to replace human creativity, we must assess the level of this threat. By gathering human perspectives about AI-generated media, we can learn more about the position in which humans are situating themselves in preparation for the future of AI.

To analyze the question, I will conduct a survey that asks questions regarding AI-generated and human-generated art. In addition, I will interview students in art majors, software majors, and majors unrelated to the two. The idea of the survey is to gather numerical data that can be statistically analyzed. The primary idea of the interviews is to gather the perspectives of the primarily affected social groups while having unrelated social groups as a control. The interview portion will also provide a qualitative rather than a quantitative view, as in the survey.

The survey will involve several sections: art, music, and research questions. The art section will present a piece from a random pool of AI-generated and human-generated art. The user will rate a piece from 1 to 10 based on specific aspects of each composition. The music section will be the same, on a 1 to 10 scale. The survey idea was pulled from a paper about who/what should receive credit for AI-generated art (Epstein et al., 2020), therefore the last section will involve related research questions. The questions will ask what people think of AI in the art world, who should obtain the credit for AI generated media, and whether there should be certain restrictions. The survey will also gather metrics such as age, major/occupation, etc.

Once the data is collected, I will run a statistical analysis on each individual part to find if there is a statistically significant view for or against AI art. I can discover if people prefer pieces generated by AI or humans and whether age or occupation influences that preference. I can also gather a consensus on whether people think AI-generated media is a net benefit or detriment to how we perceive creativity. Of course, the larger the sample size, the better the data; however, some biases may occur due to the population I will survey.

Conclusion

The expansion of technology has changed how we produce and consume creative media. While specific problems have been solved within the expansion of such technology, newer problems arise. Our drum device, Whiplash, solves the problem of practicing music when other musicians are not around to play along with. Our solution is a bridge technology that contributes to the expansion of automated creative technological systems. Creative media has begun to be produced from start to finish by AI, thereby removing the human element. I hope to learn from my research how people perceive such a shift in our culture, and whether people believe it leads to a net positive or negative for human creativity.

Bibliography

Barnes, F. (2007). *Stickboy—Robocross Machines* [Portfolio]. The Drummer. https://robocross-machines.com/robots/stickboy/

Cetinic, E., & She, J. (2022). Understanding and Creating Art with AI: Review and Outlook. *ACM Transactions on Multimedia Computing, Communications, and Applications*, *18*(2), 1–22. https://doi.org/10.1145/3475799

Chazelle, D. (Director). (2014). *Whiplash* [Drama, Music]. https://letterboxd.com/film/whiplash-2014/genres/

Crazy Ootka Software AB. (2022). *Perfect Ear: Music & Rhythm* (Version 1) [Swift; IOS]. Crazy Ootka Software AB. https://play.google.com/store/apps/details?id=com.evilduck.musiciankit&gl=US

Dayma, B. (2021). *DALL-E* (Version 2) [Python; Windows]. OpenAI. https://www.craiyon.com/

DreamStudio. (2022). *DreamStudio* (Version 1) [Ubuntu]. DreamStudio. https://beta.dreamstudio.ai/dream

Epstein, Z., Levine, S., Rand, D. G., & Rahwan, I. (2020). Who Gets Credit for AI-Generated Art? *IScience*, *23*(9), 101515. https://doi.org/10.1016/j.isci.2020.101515

Flynn, J. (2022, September 19). 36+ Alarming Automation & Job Loss Statistics [Careers]. *36+ Alarming Automation & Job Loss Statistics*. https://www.zippia.com/advice/automation-and-job-loss-statistics/

Friedlander, J. P. (2022). *YEAR-END 2021 RIAA REVENUE STATISTICS* (Year-End Music Industry Revenue Report, p. 3) [Sales & Revenue]. RIAA. https://www.riaa.com/reports/2021-year-end-music-industry-revenue-report-riaa/

Fryen, T., Eppe, M., Nguyen, P. D. H., Gerkmann, T., & Wermter, S. (2020). *Reinforcement Learning with Time-dependent Goals for Robotic Musicians* (arXiv:2011.05715). arXiv. http://arxiv.org/abs/2011.05715

Grba, D. (2022). Deep Else: A Critical Framework for AI Art. *Digital*, *2*(1), 1–32. https://doi.org/10.3390/digital2010001

HiTEC. (2007, February 24). *HS-805BB Mega Giant Scale Servo* [Hardware]. HiTEC. https://hitecrcd.com/products/servos/giant-servos/analog-giant-servos/hs-805bb/product

Holz, D. (2022). *MidJourney* (Version 3) [Python; Windows, macOS, Linux]. https://www.midjourney.com/home/

Hughes, T. (1989) The Evolution of Large Technological Systems. In: Bijker, W.E., Hughes, T.P. E Pinch, T.J. The Social Construction of Technological Systems. New Directions in the Sociology and History of Technology. (pp. 51-82) Massachusetts: MIT Press

Metz, R. (2022, August 3). AI won an art contest, and artists are furious [News]. *CNN Business*. https://www.cnn.com/2022/09/03/tech/ai-art-fair-winner-controversy/index.html

Nielsen, S. G. (1999). Learning strategies in instrumental music practice. *British Journal of Music Education*, *16*(3), 275–291. https://doi.org/10.1017/S0265051799000364

Poltronieri, F. A. (2019). Technical Images and Visual Art in the Era of Artificial Intelligence: From GOFAI to GANs. *Proceedings of the 9th International Conference on Digital and Interactive Arts*, 1–8. https://doi.org/10.1145/3359852

Polyend. (2015). *Polyend Perc* [Music]. Polyend. https://polyend.com/legacy/polyend-perc/

Riegel, J. (2002). *FreeCAD* (0.20.1) [C++, Python; Windows, Linux, macOS]. FreeCAD. https://www.freecadweb.org/

Siyanbola, W. (2011). Framework for Technological Entrepreneurship Development: Key Issues and Policy Directions. *American Journal of Industrial and Business Management*, *01*, 10–19. https://doi.org/10.4236/ajibm.2011.11002

Texas Instruments. (2018, December 2). *SimpleLink™ Wi-Fi® CC3220SF wireless microcontroller LaunchPad™ development kit* [Electronics]. Texas Instruments. https://www.ti.com/tool/CC3220SF-LAUNCHXL