SECURING EVERY LINK: EXPANDING CYBERSECURITY EDUCATION TO THE GENERAL PUBLIC

MACHINES MAKING ART: HOW CAN ARTISTS PROTECT THEIR ART FROM GENERATIVE AI?

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 Computer Science

By
Michelle Nguyen

December 1, 2023

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

Joshua Earle, Department of Engineering and Society

Introduction

As the field of cybersecurity grows, it becomes increasingly pertinent that the public learns and grows alongside it. Cyberattacks, as massive as they may be, can originate from mundane tactics. The 2013 Target data breach, in which up to 110 million customers' information was stolen, began as an email phishing attack on a smaller company ("Kill Chain", 2014). Because humans are often the weak link in cyber defense, education should be regarded as an important facet of cybersecurity itself. The question then is: How is cybersecurity education approached today and how can it be extended to the general public?

While the concept of artificial intelligence (AI) has lingered in the cultural zeitgeist as long as storytelling has existed (and the phrase has circulated since 1956), AI has recently reemerged as a hot topic and is a similarly fast-growing field (Day, 2023). It has a new face on the internet as generative AI, which has become a controversial point of discussion. Generative AI systems learn from existing material to create text, visual art, and even animations. As usage of AI content generators, specifically art generators, becomes increasingly widespread, supporters herald such systems as a democratizing, equalizing force in the entertainment industry, while artists on social media decry them. Generators such as DALL-E, Midjourney, or Stable Diffusion work by 'training' from collections of images or artwork often sourced from the internet regardless of copyright or artist consent. The LAION-5B dataset employed by Stable Diffusion, for example, consists of links to images from a myriad of sites, including DeviantArt and Pinterest (Clark, 2023). While artificial intelligence has the potential to prove helpful in some fields, the seeping of AI technology into the entertainment industry demonstrates its dangers. Regarding this, what harms do artists face from AI? What tools are in place or can be created to protect artists from these online AI art generators?

This prospectus will discuss how I intend to approach answering these questions. I begin with a brief discussion of cybersecurity education to be expanded upon in the future. The rest of the paper focuses on generative AI and will cover some background information about what AI generated art is and how it may be dangerous to artists, the different groups involved in the debate, the framework(s) I will employ, and how I plan to research. Finally, it closes with key texts related to the question.

Technical Project

As technology advances and more of our everyday lives relies on technology, the field of cybersecurity will gain increasing importance. Although much of the attention in cybersecurity is centered around large cyberattacks to be mitigated with advancing technologies (and AI intervention), the weakness in security is often as simple as human error, such as in the 2013 Target data breach. With phishing attacks against individuals being common as well, it is paramount that education and training also be considered when reviewing possible defensive cybersecurity measures to take. This paper will discuss the efficacy of cybersecurity education in its current form in schools and businesses and describe how it could be extrapolated more effectively to the general public. Relevant background will include how introductory cybersecurity classes and company training are structured. Then, advertisements, public service announcements, seminars, and games are potential vectors for learning that I will elaborate on the feasibility of.

STS Project

The displacement of workers by machines is nothing new, and has been a public concern since the advent of industrialization. Says Karl Marx, "The instrument of labour, when it takes the form of a machine, immediately becomes a competitor of the workman himself," (Marx, 1867). The question of how to introduce machines to the workplace is, then, a question fraught with ethical considerations. Nowhere is this debate more heated than in the entertainment industry. Among the demands of the actors and writers participating in the SAG-AFTRA and WGA strikes was a demand for protection against AI that has captured the public's eye. Actors fear having their likenesses captured by artificial intelligence and used ad infinitum with minimal compensation, while writers are being threatened with the elimination of their writers' rooms in favor of AI generated scripts (Coyle, 2023).

Artists on the internet, primarily digital artists, are another group threatened by generative AI. Just as AI looms on the horizon of actors and workers, it is direct competition to artists who make their income either traditionally or through the internet, especially through commission work. This is aided by the free or low price of AI art generators when compared to hiring an artist. Competition aside, proponents of AI generated art tend to describe such works as original, with the system 'learning' from previous art as humans do to create something wholly new. The veracity of this seems doubtful. While it may ultimately be a matter of semantics, artificial intelligence is not yet what the name implies, and is unable to synthesize and add thought to a work on its own. The works an AI art generator delivers are byproducts of the countless artworks it trained on, many of which were scraped from the internet without artist consent. Artists not only have to compete with cheap or free machine labor, but with machine labor that learned from

and may recreate their own art. The question, then, is: what tools are in place to protect artists on the internet from online AI art generators?

a. Relevant Social Groups

Because of their rapid proliferation across the internet, AI art generators, and more broadly, AI image generators, affect a variety of people. The wide availability of such software allows any image to be created to support any cause, even if the user has no image manipulation skills themselves. Disinformation among the public is easily bolstered by AI generated images, which can affect social currents and inform the beliefs of those who are fooled or play to their existing biases. Credible information is also adversely affected by AI images. "Not only do [deepfakes] fuel the spread of false information, they are also prone to undermine the credibility of legitimate information, creating doubts about any information encountered…" (Bontridder & Poullet, 2021, p. 4). The distrust such images foster can lead to confusion over which sources are reputable. These dangers carry over to artificially generated art as well; someone not well versed in the signs of AI generation in art may be easily tricked. With how controversial AI art usage has become, many people who are aware of the debate have a strong opinion for or against it.

The most directly affected group, however, is naturally the artists themselves. Artists, especially the digital artists most closely affected, may choose to use the internet and social media for a multitude of reasons. For many, social media is simply an environment where they can share their hobby with others who may appreciate it. However, many artists may earn some or all of their income from the internet with the advent of sites like Patreon, Etsy, and others. Social media can be used to boost an artist's portfolio, gaining attention from companies that may be looking to hire. Direct messaging also provides a point of contact from a customer or company to the artist, allowing artists to sell their art. As mentioned before, cheap or free AI art

generators now serve as competition, in addition to using artists' art without consent. Professional artist Daniel Danger worries "the images people produce with AI image generators could replace some of his more 'utilitarian' work, which includes media like book covers and illustrations for articles published online," which echoes the negative sentiment from most artists toward AI art (Metz, 2022). Some users of these art generators have also begun to pose as artists by using AI generated art to scam unsuspecting customers out of money or prompting the generator to mimic an active artist's style. Although a minority of artists encourage and welcome the use of AI art, the general community is against it or concerned about it, and have been extremely vocal on the matter online.

On the other side of the debate are the users of online AI art generators. This may include individual users who are curious about artificial intelligence, people who use AI art generators in lieu of hiring an artist, and companies who may choose to invest in such generators instead of human artists for their artistic assets. For this paper, we will not discuss the first group at length. Users creating images using AI art generators for personal use out of an interest in artificial intelligence do not threaten artists to the degree that the other groups do. For the other two groups, where they may otherwise hire an artist, they now turn to AI art generators. In the case of individuals, many may not feel able to afford artists' prices, while others may believe artists overprice their work and even hold disdain toward the costs. Companies may save money by creating artistic assets using AI art generators rather than hiring an art team or attempt to champion AI art as a legitimate art form by advertising their use of it. A recent example of this is the Netflix anime short film "The Dog & The Boy," which used AI to generate the backgrounds in each shot, citing the labor shortages in the anime industry. Instead of offering higher wages to hire more artists in a notoriously strenuous industry, Netflix employed an AI art generator and

credited the artist who worked alongside it as "Human." (Cole, 2023). This only worsens the plight of artists who could have been considered for these jobs.

Finally, the creators of DALL-E, Midjourney, and other such AI art generators represent a third party in the discussion. From the reception of AI art in the art community, one can gather that the advocates for these technologies are largely not artists themselves, but purveyors of technology. Although I doubt that these engineers aimed to harm artists, their lack of knowledge about the art community and copyright led to artists' opposition to them.

b. Methods and Frameworks

Since the parties to be discussed are relatively clear, I will employ Actor Network Theory (ANT) in analyzing the issue of AI generated art. ANT models different parties in an interaction as actors, which are not limited to humans. In this way, ANT posits that any interaction consists of said actors pushing and pulling on each other in networks (Sismondo, 2010, p. 81). For this topic, the simplest actors would be those discussed above: the artists, different users of AI art generators, and their creators, but the list of actors may also include the art generators themselves and the United States government (regarding copyright of art). The interactions of these actors and their networks will be further discussed later. My research methods will mostly consist of analyzing previous literature and public or social policy on the topic, potentially supplemented with ethnography in the form of interviewing impacted artists. I will also delve into individual case studies involving AI art.

c. Timeline

To research for this paper, I can begin by researching more specific background information about where the most popular AI art generators (DALL-E, Midjourney, etc.) source

their training images from and how they generate output pieces. Contributing to the background could be current uses of AI art in media, such as in games or movies. Once the background is solidified, I will analyze US copyright policies to find exactly what artist protections exist for traditionally created art, and compare with any policy or relevant court case on AI generated material. Finally, I will research current technological and social tools that artists may employ to protect themselves, including but not limited to AI art detection services and the University of Chicago's Glaze system.

Key Texts

 a. **Marx, K. (1887). Machinery and Modern Industry (S. Moore. and E. Aveling, Trans.). in *Capital.* Progress Publishers.**

This chapter of *Capital* discusses the tension between workers and new technologies, which will give a historical context to the AI generated art issue faced by artists. Of particular note is the section where Marx discusses workers' prior reactions to new technologies threatening their jobs.

 b. **Sismondo, S. (2010). Actor-Network Theory. in *An Introduction to Science and Technology Studies* (2nd ed.). Wiley-Blackwell.**

Sismondo's explanation of Actor-Network Theory will provide the necessary background for an ANT-based analysis of the AI art issue. He discusses its history and the assumptions it makes about science and technology. Additionally, Sismondo describes weaknesses of ANT that I should address.

 c. **Day, Kathleen. (2023). *Artificial Intelligence and Copyright Law*. CQ Researcher. https://cqpress.sagepub.com/cqresearcher/report/artificial-intelligence-and-intellectual-property-cqresrre20230421**

Day's piece is an overview of how copyright law intersects with the emergence of generative AI, and will provide suitable background for an analysis of one tool artists may have to protect themselves with. She includes a timeline of modern and historical events for context.

 d. **United States Copyright Office. (2021). *Circular 1: Copyright Basics.* https://www.copyright.gov/circs/circ01.pdf**

Circulars are files published by the U.S. Copyright Office that include updated information about copyright law. These will provide more specific information about copyright law that artists can rely on in the United States.

Additional References

Bontridder, N., & Poullet, Y. (2021). The role of artificial intelligence in disinformation. *Data & Policy, 3*, E32. doi:10.1017/dap.2021.20

Clark, N. (2023, January 17). *Artists sue AI art generators over copyright infringement.* Polygon. https://www.polygon.com/23558946/ai-art-lawsuit-stability-stable-diffusion-deviantart-midjourney

Cole, S. (2023, February 1). *Netflix Made an Anime Using AI Due to a 'Labor Shortage,' and Fans Are Pissed.* Vice. https://www.vice.com/en/article/bvmqkv/netflix-anime-dog-and-the-boy-ai-generated-art

Coyle, J. (2023, September 27). *In Hollywood writers' battle against AI, humans win (for now).* The Associated Press. https://apnews.com/article/hollywood-ai-strike-wga-artificial-intelligence-39ab72582c3a15f77510c9c30a45ffc8

Day, Kathleen. (2023). *Artificial Intelligence and Copyright Law*. CQ Researcher. https://cqpress.sagepub.com/cqresearcher/report/artificial-intelligence-and-intellectual-property-cqresrre20230421

Marx, K. (1887). Machinery and Modern Industry (S. Moore. and E. Aveling, Trans.). in *Capital.* Progress Publishers.

Metz, R. (2022, October 21). *These artists found out their work was used to train AI. Now they're furious.* CNN. https://www.cnn.com/2022/10/21/tech/artists-ai-images/index.html

Sismondo, S. (2010). Actor-Network Theory. in *An Introduction to Science and Technology Studies* (2nd ed.). Wiley-Blackwell.

United States Copyright Office. (2021). *Circular 1: Copyright Basics.*

https://www.copyright.gov/circs/circ01.pdf

United States Senate Committee on Commerce, Science, and Transportation. (2014). *A "Kill*

Chain" Analysis of the 2013 Target Data Breach.

https://www.commerce.senate.gov/services/files/24d3c229-4f2f-405d-b8db-a3a67f18388

3