

Rights of Artificial Intelligence and Neural Networks

A Research Paper submitted to the Department of Engineering and Society

Presented to the Faculty of the School of Engineering and Applied Science

University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree

Bachelor of Science, School of Engineering

Luke Benham

Spring 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

Travis Elliott, Department of Engineering and Society
Briana Morrison, Department of Computer Science

ABSTRACT

As Artificial Intelligences continue to develop creative abilities, we will have to grapple with how rights and protections are afforded to AI models and their works. The ethical principles for integrating AI into society should be established early due to the rapid and unpredictable trajectory of AI progress. The recent advances in image generation models such as DALL-E 2, Midjourney, and Stable Diffusion have demonstrated AI programs are capable of creative work previously thought to be solely in the domain of human artists. Artists and other rights holders have expressed concerns that copyrighted content has been used to train Deep Neural Networks for these image models. There is now a debate over the degree to which copyrighted content is memorized and stylistically reproduced versus being truly learned by the AI. Actor Network Theory will be used to evaluate the interactions between the artists, right holders, AI companies, and the models themselves. Image generation models are a useful case study as they are currently disrupting art and creative industries and the ethics of their development remain unclear.

INTRODUCTION

Actor-Network Theory works in the science of sociology to emphasize the ability of non-human actors to impact social networks. It works to describe social relationships as the primary form of analysis for social theory. The use of actors as a central tenet of the theory abstracts the role that groups and non-human entities play in social interactions in order to include all relevant actors into a cohesive network. By including all material and semiotic relationships ANT seeks to describe social networks entirely through internal interactions of the network without external forces.

Artificial Intelligence is a prime candidate for Actor-Network Theory analysis, as its explicit goal is to mimic the abilities of human actors as a non-human entity. By evaluating AI through the lens of ANT, the value co-creation of AI in concert with human actors can be dissected into four themes. Co-creation occurs across field advancement, supporting service providers, enabling resource integration, and supporting beneficiary well-being (Kaartemo & Hekkua, 2018, p. 5). These themes touch on the agency of the AI in a network. The ANT tenet that “humans and technologies cannot be fully separated” should be emphasized as AI systems move from being tools towards being partners in human endeavors (Bengtsson, 2018, p. 7). The building of a model using ANT will also require the ideas of both social constructivism and technological determinism on how society is shaped by technology (Matthews, 2020, p. 1). This paper will consider how AI image generation affects the behavior of actors in creative and artistic industries. While a human creator can be considered an artist, it is unclear on where the cultural and legal ownership of art will reside when art is created by a prompting user on an image model created by a separate company.

The creation of AI art has occurred since the 1970's, but the modern revolution in image generation began in 2014 with the development of generative adversarial networks. GANs work by the mutual training of a generator model which creates the image and the discriminator model which attempts to detect if the image is generated or a real image from the dataset(Goodfellow et. al, 2014). This advancement drastically improved the capabilities of image models. The next discovery that has led to the recent explosion in image quality and usage is the use of diffusion. Diffusion based models are trained by adding increasing amounts of gaussian noise to an image and then attempting to return it to the original image, sometimes with an additional text prompt(CITE). Once trained, the model can be fed a field of random noise and shape it into an image based on a text prompt. This method was another leap forward in ability and creativity, containing the latent knowledge of all types and styles of images in the dataset. The latent knowledge contained in these models has become a point of contention as diffusion models are able to replicate artistic styles, but only if the artist's images are included in the training data. These complaints arise not only from individual artists but also from large intellectual property holders.

ANALYSIS

In January of 2023 a group of artists filed a class action lawsuit against Stability AI and others for the use of copyrighted work in the training of diffusion models such as Stability Diffusion(Mattei, 2023). They claim that the training results in the creation of derivative works when the images are generated. This lawsuit is unlikely to be won by the plaintiffs as even artwork that is a direct collage is protected under Fair Use. Getty Images has also filed a lawsuit alleging that Stability AI has broken copyright laws by using copyrighted data in its training and now is competing with Getty Images(Ho, 2023). These lawsuits demonstrate the view that AI

models should be treated only as tools and not as actors. The stakeholders that include human content creators and the owners of their intellectual property have the goal to enforce human creativity as sacred as compared to art created by AI but are unlikely to be able to enforce this view legally. The AI companies desire their models to be viewed as truly learning the artistic styles and techniques, similar to how a human who has learned to mimic another artist's style is protected. This view shows the role of social constructionism on diffusion models through them absorbing the stylistic tendencies of the artists. Users have minimal ability to impact this network but would likely benefit from AI image generation being considered a tool similar to photography, in which the user of the tool would hold the copyright. The progressing technology would determine the value of human artistic contributions in this view. Another strange possibility that may seem more likely as models increase in size and complexity is that the creative work of non-human entities should be treated as similar to animals, where they are unable to hold copyright ownership so images are held by the public domain(Guadamuz, 2016).

The current equilibrium is that the for-profit companies retain ownership of the models while users have free use over the output images with the interests of artists being sidelined. Stability Diffusion and open source projects distribute the entire model in addition to their images but require that their use be limited by ethics agreements. The availability of open source models, even if they are of lower quality, redistributes power over the network from AI companies and artists towards users and consumers. This dilutes the pool of online images coming from human artists while ensuring that paywalled image models improve to stay competitive.

In order to assign rights to AI models and their works we must grapple with the fact that humans are no longer the only entities capable of complex visual creations. The various actors in

the creative industry attempt to forward their aims by analogizing diffusion models to more common entities that are easier to understand and categorize. While the actors representing previous intellectual property holders would prefer that the models be seen as just collages of their database, there is a degree to which the AI has learned and understood a portion of human artistic expression. Diffusion models no longer contain the data from the images they are trained on, only learned how an image should appear in the cultural context it was trained upon. The current copyright system was designed only to deal with human creations, and as technology advances, the human touch will diminish with each iteration. If intellectual property holders are validated that copyrighted works cannot be used for training without permission then the entire machine learning industry will be catastrophically damaged. This would lead to rights accumulating to holders similar to patent trolls, extorting those who want to innovate. If fair use remains for AI models then many artists will become obsolete as AI outpaces them in time, cost, and ability. Perhaps we should accept that humans will no longer be the most competent in every domain.

CONCLUSION

The common reaction to AI image generators is that they are incapable of producing art because it is the intention and creative expression of the artist that transforms an image into art. As new technologies are used in creating art it is usually accepted that the technology is merely a tool for the human artist, but this may change as the creative work is increasingly done by the tool. As less intention is required by the human prompter, these images will either be attributed to the expression of the AI or else lose the meaningfulness of art altogether.

Diffusion models are part of the transformation of the role of humans in creative and intellectual endeavors. The current network of creative, social, and legal systems has not updated

on another entity joining humans as creative partners. The conceptions of copyright, creativity, and art will have to change as the human touch is diminished. This does not have to mean the death of art, as if we can find beauty in nature, perhaps we can also find the meaning of art produced by an alien intelligence.

REFERENCES

- Bengtsson, S. (2018). Ethics Exists in Communication : Human-machine ethics beyond the Actor-Network. *London School of Economics and Political Science*, (pp. 1–25). <http://urn.kb.se/resolve?urn=urn:nbn:se:sh:diva-37239>
- Dafoe, A. (2015). On technological determinism. *Science, Technology, & Human Values*, 40(6), 1047–1076. <https://doi.org/10.1177/0162243915579283>
- Goodfellow, I., Pouget-Abadie, J. Mirza, M., Xu, B., Warde-Farley D., Ozair, S., Courville, A., Benigo, Y., (2020) Generative Adversarial Networks. *Communications of the ACM*, vol. 63, no. 11, pp. 139–144., <https://doi.org/10.1145/3422622>.
- Guadamuz, A. (2016). The monkey selfie: Copyright lessons for originality in photographs and internet jurisdiction. *Internet Policy Review*, 5(1). <https://doi.org/10.14763/2016.1.398>
- Ho, K. K. (2023, February 10). Getty images sues stability AI over photos used to train stable diffusion image generator. ARTnews.com. Retrieved March 17, 2023, from <https://www.artnews.com/art-news/news/getty-images-lawsuit-stability-ai-12-million-photos-copied-stable-diffusion-1234656475/>
- Hong, Y., Hwang, U., Yoo, J., & Yoon, S. (2020). How generative adversarial networks and their variants work. *ACM Computing Surveys*, 52(1), 1–43. <https://doi.org/10.1145/3301282>
- Kaartemo, V., & Helkkula, A. (2018). A systematic review of artificial intelligence and robots in value co-creation: Current status and future research avenues. *Journal of Creating Value*, 4(2), 211–228. <https://doi.org/10.1177/2394964318805625>
- Mattei, S. E.-D. (2023, January 17). Artists file class action lawsuit against AI Image Generator Giants. ARTnews.com. Retrieved March 17, 2023, from <https://www.artnews.com/art-news/news/artists-class-action-lawsuit-against-ai-image-generator-midjourney-stability-deviantart-1234653892/>
- Matthews, A. (2020). Blurring boundaries between humans and technology: Postdigital, postphenomenology and actor-network theory in Qualitative Research. *Qualitative Research in Sport, Exercise and Health*, 13(1), 26–40. <https://doi.org/10.1080/2159676x.2020.1836508>
- O’Lemmon, M. (2020). The technological singularity as the emergence of a collective consciousness: An anthropological perspective. *Bulletin of Science, Technology & Society*, 40(1-2), 15–27. <https://doi.org/10.1177/0270467620981000>
- Pinch, T. J., & Bijker, W. E. (1984). The Social Construction of Facts and Artefacts: Or How the Sociology of Science and the Sociology of Technology Might Benefit Each Other. *Social Studies of Science*, 14(3), 399–441. <http://www.jstor.org/stable/285355>

