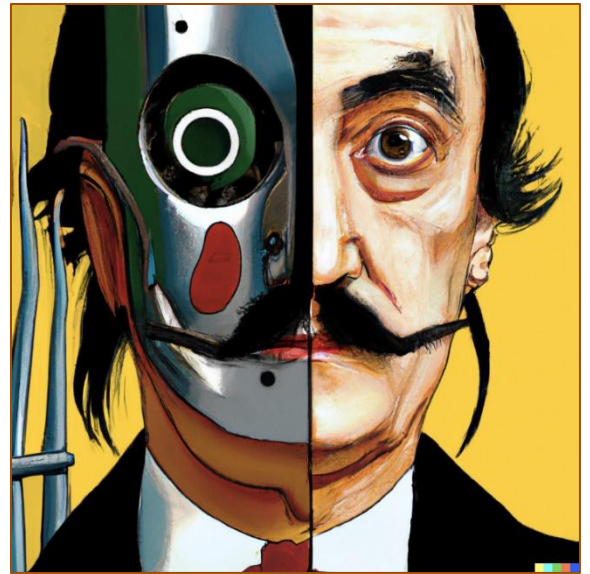


THE ETHICS OF AI ART

THE CASE OF DALL-E 2

For thousands of years, art has been an endeavor of the human race. From Rembrandt to Basquiat, from the Benin Bronzes to the new wave cinema of Hong Kong, art has been recognized as creative expressions of human intelligence. With the public release of DALL-E 2, a neural network that generates images from phrases, the definition of art might be due for reevaluation to include media produced by artificial intelligence. Generative AI, like countless other technologies emerging in the cyber-physical realm, present numerous ethical challenges.

DALL-E 2, aptly named for the combinative surrealism of Salvador Dalí and the futurism of Pixar’s WALL-E, is a machine learning model developed by research company OpenAI. The program model allows users to render images from a description in natural language. Because it relies on information databases from web servers as reference points, the illustrations are seemingly boundless. DALL-E 2 has the capability to learn and apply artistic *styles*, like “impressionist watercolor painting,” too. Realistically, the limits of its potential depend on the creativity of its input phrases.



On September 28, 2022, DALL-E 2 became universally available. With its release, renewed debate about the merits of AI art came to the fore. Specifically, artists and graphic designers began to consider how this technology can facilitate their professions. Patrick Clair, Emmy-winning main title designer, said, “It’s like working with a really willful concept artist,” (Roose, 2022). Others in the visual arts scene worry that programs like DALL-E 2 might put them out of work in the same way that automated machinery shrunk the manufacturing workforce. For instance, it isn’t hard to imagine the cover art for an electronic dance album to be designed by AI rather than a human. Why would a rock band commission Andy Warhol, say, when DALL-E 2 can generate “[rotting banana in pop art style](#)” for free?

Just as manufacturing companies defend the utility of automation, some artists don’t denounce AI in the design sector. Many retail corporations argue that computerized self-checkout stations save money by allowing employees to do other work that can’t be automated. In the same vein, AI art can maximize the efficiency of artists by sparking creativity or inspiring someone’s first step in the final project. One interior designer said, “I think there’s an element of good design that requires the empathetic touch of a human ... So I don’t feel like it will take my job away,” (Roose, 2022).

In addition to the ethics of employing AI art for commercial and professional use, the content DALL-E 2 and others produce is ripe for discussion. Dr. Eduardo Navas, an associate research professor at Pennsylvania State University who studies DALL-E 2, finds that it functions metaphorically “almost like God—all a person



has to do is to state a prompt (the word) and it *is*." Aside from some [restrictions in the algorithm](#), such as pornography and hate symbols, there are limited guardrails for what prompts can generate. While an earlier version of DALL-E filtered out all images of people, the current model allows users to render public figures in positions and settings that could be deemed offensive or simply implausible. Further, even if the average person can distinguish AI renderings from real images, some people might think a meme generated by DALL-E 2 actually happened. These artificial pictures can cause reputational damage for professionals and celebrities, and it can present national security implications for politicians and leaders.

According to Dr. Navas, the progression from DALL-E 2 static images to AI-generated video clips isn't a matter of "if, but when." This raises significant ethical concerns related to credibility and accountability. For one, falsified videos of public figures saying and doing outrageous things, or deepfakes, can erode trust in institutions. One study shows that rather than misleading people, AI deepfakes make them feel uncertain, which "reduces trust in news on social media," (Vaccari & Chadwick, 2020). It's entirely possible for bad actors to deceive the media with a false video, leading to widespread circulation even after it gets flagged for misinformation and retracted. Another AI researcher said, "If I got [an image] off the BBC website, the Guardian website, I hope they've done their homework and I could be a bit more trusting than if I got it off Twitter," (Taylor, 2022). Still, news websites regularly publish mistakes, and sometimes a redaction or correction isn't enough to curb the spread of it.

As a consequence of AI coding and algorithms, the next dilemma stems from bias and stereotypes. One journalist said, "Ask Dall-E for a nurse, and it will produce women. Ask it for a lawyer, it will produce men," (Hern, 2022). This is partly due to the web servers that provide the program with learning material. Some might argue that DALL-E 2 is still in its early phases, and the OpenAI programmers need time to work out the kinks. However, the algorithm likely can't eliminate stereotypes on the users' end. Dr. Navas, who is Latino, said that a colleague ran the prompt on DALL-E 2 to generate images of "the most beautiful woman in the world." DALL-E 2 then created images of white women (DALL-E 2 gives four initial results). Dr. Navas ran the same prompt a few days later and got a different set of images—in this case, of women who appeared ethnically ambiguous and who might be perceived as Latina. In all instances, the women in the images were portrayed in Western-style dress. It is not clear if user data was accessed for running each prompt, but problems with diversity and user privacy arise if such information played some role in the different results.

Then, there's the complicated question of ownership: who *should* own the copyright to the generated images? One could make the case for the companies behind the coding, like OpenAI. After all, if it weren't for the labor of the programmers, there wouldn't be a final product. The terms and conditions for DALL-E 2 stipulate that OpenAI holds the copyright for images rendered, but the users retain ownership of the prompts they entered manually. This might be a fair compromise, but as Dr. Navas points out, the source material for the machine learning model isn't owned by OpenAI either. With "1.5 million users generating more than two million images every day," intellectual property laws will need to be reconsidered for AI art.

To conclude, DALL-E 2 and other text-to-image AI technology showcase myriad ethical challenges, and this article in no way intends to be exhaustive. As mentioned, AI art can pose risks to human artists and their opportunities to earn money. At the same time, the programs can aid designers by sparking creativity in the initial stages of a project. On the content side of AI art, the images can be offensive or harmful, and this danger is further exacerbated with the potential for AI deepfake videos. Then, because the code relies on human input and a vast repository of reference images, the generated content can be plagued with biases. Lastly,

the realm of AI art is relatively new, so laws and policies related to ownership will require ethical reasoning to determine who or what owns the images. For better or worse, DALL-E 2 is now publicly available and gaining popularity. It's no longer a question of whether AI should generate art, but rather how ethics can guide the answers to these complicated and unique challenges.

Discussion Questions:

1. Are there ethical problems when it comes to AI generating art? Which values are in conflict in this case study?
2. If you were on the team that helped create DALL-E 2, what kinds of content, if any, would you restrict from appearing in the results? Why?
3. In regard to copyright ownership, what would be an ethical way of determining which parties, if any, deserve the right of possession? What if the art is sold commercially?
4. Should we classify AI-generated images as art? What are the qualifications for something to be considered art?

Further Information:

Hern, Alex. "TechScape: This cutting edge AI creates art on demand—why is it so contentious?" *The Guardian*, May 4, 2022. Available at:

<https://www.theguardian.com/technology/2022/may/04/techscape-openai-dall-e-2>

Robertson, Adi. "The US Copyright Office says an AI can't copyright its art." *The Verge*, February 21, 2022. Available at: <https://www.theverge.com/2022/2/21/22944335/us-copyright-office-reject-ai-generated-art-recent-entrance-to-paradise>

Roose, Kevin. "A.I.-Generated Art Is Already Transforming Creative Work." *The New York Times*, October 21, 2022. Available at: <https://www.nytimes.com/2022/10/21/technology/ai-generated-art-jobs-dall-e-2.html>

Taylor, Josh. "From Trump Nevermind babies to deep fakes: DALL-E and the ethics of AI art." *The Guardian*, June 18, 2022. Available at: <https://www.theguardian.com/technology/2022/jun/19/from-trump-nevermind-babies-to-deep-fakes-dall-e-and-the-ethics-of-ai-art>

Vaccari, C., & Chadwick, A. (2020). "Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News." *Social Media + Society*, 6(1). <https://doi.org/10.1177/2056305120903408>



Authors:

Dex Parra & Scott R. Stroud, Ph.D.
Media Ethics Initiative
Center for Media Engagement
University of Texas at Austin
December 14, 2022

Image: “Vibrant portrait painting of Salvador Dalí with a robotic half face” / OpenAI

This case was supported by funding from the John S. and James L. Knight Foundation. It can be used in unmodified PDF form in classroom or educational settings. For use in publications such as textbooks, readers, and other works, please contact the [Center for Media Engagement](#).

This work is licensed under **CC BY-NC-SA 4.0** 