

Is AI Art Theft?

Trystan S. Goetze, Harvard University

✉ tgoetze@fas.harvard.edu | 🐼 [@ErrantCanadian@zirk.us](https://twitter.com/ErrantCanadian)

Job Talk, California Polytechnic State University, San Luis Obispo, 16 February 2023

0. Introduction

In August 2022, an AI-generated image won an art contest.

Professional artists immediately decried the winner as a cheat, and soon launched a protest movement against AI image generators.

Artists make two ethical claims against AI image generators:

- (1) The machine learning techniques that enable these generators *steal creative work* from human artists.
- (2) AI image generators *pose an existential threat* to the vocation of visual art itself.

I am mainly concerned with (1), but (2) will come up.

Here is the plan:

1. Brief analysis of art theft.
2. Outline of how AI image generation works.
3. John Locke's account of labour and property, applied to art.
4. Decolonial theory and critical data studies on extractive AI.
5. Implications for AI development

1. What is Art Theft?

- **Heist:** Unauthorized removal of physical artworks.
- **Plagiarism:** Claiming another's creative work as yours.
- **Style Theft:** Mimicking the distinctive qualities of a particular artist to a high degree.
- **Labour Theft:** Artist's labour is used for your benefit, but artist isn't compensated.

2. AI Image Generation

The recent prevalence of AI image generators is in part due to the success of a new set of machine learning techniques, exemplified by OpenAI's DALL•E 2 application.

Algorithms are then trained on these image-annotation pairs. The

Gault (2022).

Protest actions include the proliferation of hashtags such as [#CreateDontScrape](#) and [#NoAI](#), a surge of posts on the art sharing site ArtStation containing an image with "AI" crossed out (Weatherbed, 2022), and a lawsuit filed against the creators of the Stable Diffusion AI image generator (Butterick, 2023).

There is a kind of irony in juxtaposing Lockean and decolonial theory, given Locke's personal involvement in American colonization and the transatlantic slave trade, and the tensions this creates within his philosophy. See Arneil (1996), Lewis (2003), and Mills (2022).

See Ramesh et al. (2021, 2022) for details on DALL•E 2. The prize-winning piece above used Midjourney.

Radford et al. (2021).

result is a mathematical model of patterns in these pairs. The original use case was to generate text annotations for images not in the training set.

Image generation AI basically runs in reverse. Given a novel text annotation, it creates an image that corresponds to the input prompt.

Heist? No: AI image generators can't take physical works.

Plagiarism? Sometimes: models can "memorize" images from the training data and return copies at generation time.

Style Theft? Yes, and this is sometimes hyped as a feature.

Labour Theft?

3. Locke, Labour, and Theft

James Moor argues that new technologies create new possibilities for action, and this opens gaps in our norms.

It's philosophy's job to fill these gaps. But what kind of philosophy?

I follow Nancy McHugh's feminist pragmatist approach to philosophical inquiry into the impacts of science & technology.

Core principle: centre the lived experiences of people resisting the negative impacts of technologies.

Artists articulate their anti-AI position in terms of *labour theft*. How might we substantiate it?

Locke argues that we own what we produce from unclaimed natural materials because: we own ourselves, so we own our labour, and so we own the products of *mixing our labour* with those materials.

Locke claims that because we are self-owners, even in the state of nature we have a right to self-defence. This extends to a right to defend the products of our labour.

A Lockean account of theft of physical property:

1. A farmer claims a plot of land and produces grain on it.
2. The farmer's labour is mixed with the grain, so they own it.
3. No one else has a right to that grain, even if they mix additional labour with it (e.g. baking a cake).

Lockean accounts of intellectual property note that the labour of the mind also transforms unclaimed natural materials.

Concepts, stories, artistic processes, math, natural laws, cultural

These differ from earlier methods, such as generative adversarial networks (Goodfellow et al., 2014).

Carlini, Hayes, Nasr, et al. (2023).

Knight (2022).

Moor (1985).

McHugh (2015) draws on the feminist tradition of situated knowing and Deweyan pragmatist epistemology.

Locke (1690).

Locke also claims that seizure of property is sometimes permissible. In particular, he argues that Europeans were justified in seizing lands from indigenous nations in the Americas. This infamous conclusion depends on some racist assumptions Locke makes about both the people of the Americas and what kind of labour is valuable.

Cf. Diderot (2002).

works in the public domain, etc. are building-blocks of creativity. These are not just unclaimed, but *unclaimable*.

A Lockean argument that AI art is theft:

1. Artists take building-blocks of creativity and mix them with their intellectual labour to create new works.
2. Because artists' intellectual labour is mixed with their works, they own those works.
3. Thus, AI developers do not have a right to human-made art.

Developers and users of AI image generators do not have a right to images produced thereby, either.

Objection: Artists draw on the works of others when creating their works. Some even refer to this as a permissible form of *stealing*!

One answer: the processes involved in AI image generation are fundamentally different from the human creative process.

A different answer: the *speed* and *scale* of how AI image generation uses human artists' work is a morally significant difference from the ways human artists draw on one another's works.

4. Colonialism and Extraction

To see how the speed and scale of AI image generation matters, consider the difference between:

- (a) a small community cutting timber for themselves, and,
- (b) a large corporation clear-cutting a forest.

AI development more closely resembles (b). One artist "stealing" aspects of another's work for inspiration more closely resembles (a).

Conquest pattern: powerful entity appropriates a resource, declares a right to it, extracts economic value, accrues profit, all of which poses an existential threat to the original owners of the resource.

Data colonialism applies the conquest pattern to information.

Until recently, data extraction has been focused on "rendering the self" from data trails left by our unreflective behaviour.

Now, the extraction seeks to mine our creative self-expression.

Note that this argument does not depend on any particular *legal* framework of intellectual property rights.

Cf. Kleon (2012).

Birhane (2021), argues that generative AI cannot be creative because computational systems are deterministic; Sato & McKinney (2022) argue that AI cannot be creative unless it is an embodied agent.

Zuboff (2019). See also Birhane (2020), Couldry & Mejias (2019), Crawford (2021), D'Ignazio & Klein (2020), Thatcher et al. (2016).

5. Implications for AI Development

Other kinds of generative AI are based on similar processes of scraping huge datasets for model training.

If AI image generators involve theft of creative labour, then the same argument applies to these cases.

Furthermore, it isn't just generative AI that uses massive datasets. Language understanding models, for example, may also be theft.

Some possible recommendations for responsible AI development:

- Opt-out as a default.
- More conscious curation of datasets.
- Compensation for data subjects.
- AI techniques that do not require massive datasets.
- Serious consideration of the social impact of automating creativity.

Examples: GPT (text generation), Copilot (code generation), Magenta (music generation).

Other reasons for making these changes: trauma inflicted on content reviewers who remove harmful content from datasets (Perrigo, 2023) environmental and social justice impacts (Bender, Gebru, et al. 2021).

Cf. Danaher (2019), who argues that we should automate *work* but not creativity.

Works Cited

- Arneil, B. (1996). *John Locke and America*. Oxford University Press.
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? *ACM Conference on Fairness, Accountability, and Transparency*. FAccT '21, Virtual, Canada. <https://doi.org/10.1145/3442188.3445922>
- Birhane, A. (2020). Algorithmic Colonization of Africa. *SCRIPTed*, 17(2), 389–409. <https://doi.org/10.2966/scrip.170220.389>
- Birhane, A. (2021). The Impossibility of Automating Ambiguity. *Artificial Life*, 27, 44–61. https://doi.org/10.1162/artLa_00336
- Butterick, M. (2023, January 13). *Stable Diffusion Litigation*. <https://stablediffusionlitigation.com/>
- Cartini, N., Hayes, J., Nasr, M., Jagielski, M., Sehwag, V., Tramèr, F., Balle, B., Ippolito, D., & Wallace, E. (2023). *Extracting Training Data from Diffusion Models* (arXiv:2301.13188). arXiv. <https://doi.org/10.48550/arXiv.2301.13188>
- Couldry, N., & Mejias, U. (2019). *The Costs of Connection: How Data Is Colonizing Human Life and Appropriating It for Capitalism*. Stanford University Press.
- Crawford, K. (2021). *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale University Press.
- Danaher, J. (2019). *Automation and Utopia: Human Flourishing in a World without Work*. Harvard University Press.
- Diderot, D. (2002). Letter on the book trade (A. Goldhammer, Trans.). *Daedalus*, 131(2), 48–56.
- D'Ignazio, C., & Klein, L. F. (2020). *Data Feminism*. MIT Press. <http://datafeminism.io/>
- Gault, M. (2022, August 31). An AI-Generated Artwork Won First Place at a State Fair Fine Arts Competition, and Artists Are Pissed. *Vice*. <https://www.vice.com/en/article/bvmvqm/an-ai-generated-artwork-won-first-place-at-a-state-fair-fine-arts-competition-and-artists-are-pissed>
- Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). *Generative Adversarial Networks* (arXiv:1406.2661). arXiv. <https://doi.org/10.48550/arXiv.1406.2661>
- Kleon, A. (2012). *Steal Like an Artist*. Workman Publishing Company.
- Knight, W. (2022). Algorithms Can Now Mimic Any Artist. Some Artists Hate It. *Wired*. <https://www.wired.com/story/artists-rage-against-machines-that-mimic-their-work/>
- Lewis, D. (2003). Locke and the Problem of Slavery. *Teaching Philosophy*, 26(3), 261–282. <https://doi.org/10.5840/teachphil200326328>
- Locke, J. (1690). *Second Treatise of Government* (C. B. Macpherson, Ed.). Hackett.
- McHugh, N. (2015). *The Limits of Knowledge: Generating Pragmatist Feminist Cases for Situated Knowing*. SUNY Press. <https://sunypress.edu/Books/T/The-Limits-of-Knowledge>
- Mills, C. W. (2022). Locke on Slavery. In J. Gordon-Roth & S. Weinberg (Eds.), *The Lockean Mind* (pp. 487–497). Routledge.
- Moor, J. H. (1985). What is Computer Ethics? *Metaphilosophy*, 16(4), 266–275.
- Perrigo, B. (2023, January 18). Exclusive: OpenAI Used Kenyan Workers on Less Than \$2 Per Hour to Make ChatGPT Less Toxic. *TIME*. <https://time.com/6247678/openai-chatgpt-kenya-workers>
- Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G., & Sutskever, I. (2021). *Learning Transferable Visual Models From Natural Language Supervision* (arXiv:2103.00020). arXiv. <https://doi.org/10.48550/arXiv.2103.00020>
- Ramesh, A., Dhariwal, P., Nichol, A., Chu, C., & Chen, M. (2022). *Hierarchical Text-Conditional Image Generation with CLIP Latents* (2204.06125). arXiv. <https://doi.org/10.48550/ARXIV.2204.06125>
- Ramesh, A., Pavlov, M., Goh, G., Gray, S., Voss, C., Radford, A., Chen, M., & Sutskever, I. (2021). *Zero-Shot Text-to-Image Generation* (arXiv:2102.12092). arXiv. <https://doi.org/10.48550/arXiv.2102.12092>
- Sato, M., & McKinney, J. (2022). The Enactive and Interactive Dimensions of AI: Ingenuity and Imagination Through the Lens of Art and Music. *Artificial Life*, 28, 310–321. https://doi.org/10.1162/artLa_00376
- Schuhmann, C., Vencu, R., Beaumont, R., Coombes, T., Gordon, C., Katta, A., Kaczmarczyk, R., & Jitsev, J. (2022, December 12). *LAION-5B: A new era of open large-scale multi-modal datasets*. LAION. <https://laion.ai/blog/laion-5b>
- Thatcher, J., O'Sullivan, D., & Mahmoudi, D. (2016). Data colonialism through accumulation by dispossession: New metaphors for daily data. *Environment and Planning D: Society and Space*, 34(6), 990–1006. <https://doi.org/10.1177/0263775816633195>
- Weatherbed, J. (2022, December 23). *ArtStation is hiding images that protest against AI art*. The Verge. <https://www.theverge.com/2022/12/23/23523864/artstation-removing-anti-ai-protest-artwork-censorship>
- Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. Public Affairs.