In the age of generative AI, does a meaningful distinction remain between human crafted and technologically generated content?

Berenike Melchior, March 12, 2024

In the wake of recent ground-breaking advances made in the field of generative AI, concerns about the reproducibility of art, dating back to the invention of the printing press, are more present than ever. In his article on AI creation and copyright, Stefano Quintarelli poses a series of questions regarding the content generated by these novel AI systems.[1] At the core of these questions lies the theme of human versus AI authorship and the concern about the loss of what it means to be human. I will approach these in the following essay, evaluating technological, artistic, and economic considerations.

The most powerful image generation models are based on some of the largest neural network architectures today. Trained on datasets containing billions of images, such as LAION-5B, even the creators of these models are still decoding their capabilities. The complexity is so great, that the main methods of understanding the scope of their abilities are empirical in nature. The mid-2023 paper "Diffusion Art or Digital Forgery? Investigating Data Replication in Diffusion Models" is one of such studies and may provide some crucial insights into the question of the originality of generative models. The research team set out to discover whether those models at the heart of state-of-the-art image generation are creating "unique works of art, or […] replicating content directly from their training sets" They found not only that Stable Diffusion is able to reproduce training data, but also that generated objects were sometimes semantically equivalent reconstructions of objects found in the training data[2]. While these reconstructions make up a small subset of generated images, the proof that they do indeed occur says a great deal about the purported originality of generated images and should greatly impact the outcomes of related copyright lawsuits, such as the one filed by Getty Images against StabilityAI, alleging copyright abuses.

Among AI company CEOs, such as Midjourney's David Holz, generative models are largely promoted as enhancers of creative production: "An engine for the imagination"[3]. Undeniably, generative AI has introduced a novel and extremely accessible tool kit with which to produce media, be it visual, textual, or musical. What artist duo Holly Herndon and Mat Dryhurst term as *spawning*: "creating new artworks from a training data corpus"[4], may simply extend the common practice of appropriation and sampling in art-making. From this lens, generative AI acts as an enhancer for creativity. However, the question of authorship remains. Are these models merely new tools, freeing us from menial tasks and expanding our choices, or do they possess the capacity to create their own art? Creativity is a defining feature in the process of art-making and may be considered a key human trait. Writer Arthur I. Miller defines creativity as "The production of new knowledge from existing knowledge, achieved through the process of problem solving.",

[1] Stefano Quintarelli, *AI Creation and Copyright: Unraveling the debate on originality, ownership* (Euractiv 2023).

[2] Gowthami Somepalli et al., Diffusion Art or Digital Forgery? Investigating Data Replication in Diffusion Models (IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019), 6048.

[3] James Vincent, *'An engine for the imagination': the rise of AI image generators* (The Verge, 2022).

[4] Chris Stokel-Walker, *This couple is launching an organization to protect artists in the AI era* (Input, 2022).

characterised in particular by certain markers of creativity, such as "inspiration, intuition […] and to be "out there" gaining worldly experiences […]."[5] While generative AI has achieved something akin to "the production of new knowledge" with outcomes remarkably indistinguishable from human produced content, the experience of being "out there" remains beyond AI's reach. Walter Benjamin posits that "The uniqueness of the art work is identical with its being embedded within the context of tradition."[6] Considering tradition as the intersection of human history and culture, one can deduce from Benjamin's writing that an inherent trait of art itself, and thereby the creativity required to produce it, is indeed its situatedness within worldly experiences. While the debate continues, I pose that this argumentation constitutes sufficient evidence as to continue to differentiate between human authored art and AI generated content.

Something greater may be at risk if we don't explicitly distinguish between machine and human capabilities. The commodification of data by technology companies predates the dawn of generative AI. Writer Shoshana Zuboff has studied the transformation of user behaviour data into "prediction products", making those private companies collecting and trading in this freely available, human-produced data immensely wealthy[7]. Within our economic framework, where humans receive monetary rewards for their labour, it may put those already precarious workers at risk to assign the possibility of authorship to machines, who themselves don't stand to gain much, yet whose creators stand to gain a lot. As Stefano Quintarelli states in his concluding paragraph "There are significant economic interests involved, including the future exploitation of the wealth of information produced to date."[1] In a world where data is commodified and sold for the benefit of a select few, is not each act of production of data a form of labour, which should be fairly rewarded within our economy? While tech CEO's make a case for voluntary accountability in order to mitigate the throttling of innovation and creativity[8], we may stand to lose much more than innovation, if we continue to ignore the commodification of our work and our private lives.

If, as James Betker, co-creator of DALL-E 3 says, "the "it" in AI models is the dataset."[9] then one way forward may be the independent maintenance of datasets by entities that don't stand to profit, and a non-negotiable requirement of greater transparency, both for safe but also for a true fair use.

---

[5] Arthur I. Miller, *Kreativität und Künstliche Intelligenz* (KUNSTFORUM International Bd. 278, 2021), 93.

[6] Walter Benjamin, *Das Kunstwerk im Zeitalter seiner technischen Reproduzierbarkeit* (3rd Edition, 1935), 7.

[7] Shoshana Zuboff, *The Age of Surveillance Capitalism* (2019).

[8] StabilityAI, *Response to United States Copyright Office Inquiry into Artificial Intelligence and Copyright* (2023), 15.

[9] James Betker, *The "it" in AI models is the dataset*. (Non_interactive - Software & ML, 2023).