Katrina Geddes
Joint Postdoctoral Fellow
NYU Law/Cornell Tech
kgg283@nyu.edu

September 1, 2023

To the Copyright Office,

Please find below my comments in response to your inquiry regarding the effect of artificial intelligence on copyright law.

AI-Generated Artwork and Fair Use

A. How Text-to-Image Generation Works

Popular generative image models, such as Stable Diffusion, Imagen, Midjourney, and DALL-E 2, are able to generate high-quality synthetic images from text prompts entered by users.[1] They are known as "diffusion models" because they gradually add ("forward diffusion") and remove ("reverse diffusion") Gaussian noise from training images in order to generate synthetic images.[2] These probabilistic models are trained to predict a de-noised variant of input x_t, where x_t is a noisy version of the original input x.[3] They learn to iteratively de-noise x_t using guidance from text encodings.[4] So-called *latent* diffusion models are particularly efficient at this task because they work on a compressed image representation (rather than the image itself) in low-dimensional *latent space*, rather than high-dimensional pixel space.[5] This reduces the computational demands of training diffusion models, in terms of both the time and energy required.[6]

Text-to-image diffusion models are relevant for copyright scholars because they are trained on extremely large (and minimally curated) datasets, which contain large

[1] Google's text-to-image diffusion model, Imagen, uses large language models to encode text for image synthesis, with impressive photorealistic results. Large language models are trained on vast text-only datasets, which means that they are exposed to a very rich and wide distribution of text, in comparison to datasets of image-text pairings. See Chitwan Saharia, et al., Photorealistic Text-to-Image Diffusion Models with Deep Language Understanding, arXiv:2205.11487v1 (2022).
[2] Louis Bouchard, How Stable Diffusion Works? Latent Diffusion Models Explained (August 27, 2022), <https://www.louisbouchard.ai/latent-diffusion-models/>
[3] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, Björn Ommer, High-Resolution Image Synthesis with Latent Diffusion Models, arXiv:2112.10752v2 (2022).
[4] Jay Alammar, The Illustrated Stable Diffusion, https://jalammar.github.io/illustrated-stable-diffusion/ (2022)
[5] Louis Bouchard, How Stable Diffusion Works? Latent Diffusion Models Explained (August 27, 2022), <https://www.louisbouchard.ai/latent-diffusion-models/>; Lilian Weng, What are diffusion models? (July 11, 2021), <https://lilianweng.github.io/posts/2021-07-11-diffusion-models/>.
[6] Robin Rombach, *supra* note 3.

amounts of copyrighted content.[7] In addition, the synthetic images generated by diffusion models are often similar to the copyrighted images on which they were trained. There is evidence that some diffusion models (including Stable Diffusion and Imagen) "memorize" and regenerate near-identical replicas of training images.[8] Computer scientists from one study described memorization as "pervasive" in large diffusion models, and unlikely to be alleviated by data deduplication.[9] This raises important legal and ethical questions about the originality of diffusion outputs.[10]

B. Are Generative Art Models Engaged in Copyright Infringement?

An artist could bring a claim against the developers of a generative model if the model was trained on the artist's (registered) copyrighted work, and the model produced output that was substantially similar to the copyrighted work.[11] The artist would need to prove that the defendants violated at least one of the artist's exclusive rights under 17 USC 106, such as the right of reproduction.[12] The model must have reproduced *protectable* aspects of the original work, not unprotectable aspects, such as the artist's distinctive style. An artist's signature style is not protected by copyright law; another artist may use the same subject and style provided they do not substantially copy the first artist's specific expression of their idea.[13]

There are a few caveats to this. First, generative art models frequently struggle to separate an artist's distinctive style from the type of subject to which that style was ordinarily applied. For example, if a user prompts DALL-E 2 to create a pointillist painting of a wedding in the style of Seurat, the model will produce an image that could credibly pass as a Seurat original.[14] However, if DALL-E 2 is asked to produce a pointillist painting of Times Square, it will produce "something unstructured and primitive-seeming, as helpless as Seurat would have been at this task."[15] Adam Gopnik argues that this is a reminder that "Seurat *is* his people, as van Gogh is his cypresses."[16] Similarly, asking the model to produce a Thiebaud painting of a battle scene would generate "a gibbering nightmare of unrelated form, vaguely and nightmarishly evocative of soldiers and tanks," because battles are "not a variant of

---

[7] Shawn Shan, et al., GLAZE: Protecting Artists from Style Mimicry by Text-to-Image Models, arXiv:2302.04222 (2023).
[8] Nicholas Carlini, et al., Extracting Training Data From Diffusion Models, arXiv:2301.13188v1 (2023) (here the authors define "memorization" using image similarity metrics such as the distance function l_2.)
[9] Nicholas Carlini, et al., Extracting Training Data From Diffusion Models, arXiv:2301.13188v1 (2023).
[10] Gowthami Somepalli, et al., Diffusion Art or Digital Forgery? Investigating Data Replication in Diffusion Models, arXiv:2212.03860v3 (2022).
[11] See Fourth Estate Public Benefit Corporation v. Wall-Street.com, LLC, 139 S.Ct. 881 (2019).
[12] The reproduction right is generally construed so broadly that in most cases it encompasses the work that would have been done by 106(2) (the right to prepare derivative works).
[13] Dave Grossman Designs, Inc. v. Bortin, 347 F.Supp. 1150, 1156–57 (N.D.Ill.1972) ("For example, Picasso may be entitled to a copyright on his portrait of three women painted in his Cubist motif. Any artist, however, may paint a picture of any subject in the Cubist motif, including a portrait of three women, and not violate Picasso's copyright so long as the second artist does not *substantially copy* Picasso's specific expression of his idea").
[14] Adam Gopnik, "What Can A.I. Art Teach Us About the Real Thing?", The New Yorker, March 1, 2023.
[15] Ibid.
[16] Ibid.

a Thiebaud theme but an absence within Thiebaud-world."[17] This inability of generative models to successfully separate an artist's distinctive style from the application of that style to particular expression may increase the model's vulnerability to copyright infringement liability.[18]

Secondly, although an artist's distinctive style is not copyrightable, the application of that style to a new work may represent a violation of the artist's moral rights, specifically, the right of attribution. For example, an artist whose brand and reputation has been built on pacifist imagery may object to the application of their signature style to military propaganda, where this leads the public to erroneously attribute the propaganda to the artist. In the U.S., moral rights only apply to works of visual art, and are limited to rights of attribution and integrity.[19] Thirdly, although copyright law does not protect an artist's distinct aesthetic sensibility, the reproduction of such sensibility in a secondary work may be probative of copying, and support a claim that the secondary work was not created independently.[20]

If a generative model violates the right of reproduction by reproducing protectable aspects of an artist's work, the artist will need to prove the following three elements of infringement: (a) copying; (b) what the model created is a "copy"; and (c) the copying rises to the level of "improper appropriation." The plaintiff can satisfy the first element (copying) by proving access and probative similarity (similarities between the works that are probative of copying). Access is satisfied by showing that the artist's copyrighted work was contained in the model's training data. Probative similarity is satisfied by demonstrating similarities between the artist's work and the model's output. The similarities need not be extensive, nor involve protected elements of the artist's work; they simply need to be "similarities that one would not expect to arise if the two works had been created independently."[21]

Litigation surrounding generative models may ultimately challenge the relatively narrow definition of "access" within copyright jurisprudence. If a generative model produces songs that sound substantially similar to Taylor Swift's music (because it was trained on a corpus of pop country music) but it was not trained on any of Swift's songs, can Swift bring a copyright infringement claim against the model's developers? In general, a plaintiff trying to prove access must show that the defendant "had a reasonable opportunity to view [or hear] the copyrighted work."[22] A "bare possibility" of access is not sufficient, nor is a theory of access "based on speculation and conjecture."[23] A plaintiff can demonstrate a reasonable opportunity for access by

---

[17] Ibid.
[18] However, future iterations of these generative models will likely be better at separating the artist's style from particular forms of expression.
[19] 17 U.S. Code § 106A.
[20] See, e.g., Steinberg v. Columbia Pictures Industries, Inc., 663 F.Supp. 706 (1987).
[21] Laureyssens v. Idea Group, Inc., 964 F.2d 131, 140 (2d Cir. 1992); 4 Nimmer on Copyright § 13.01[B]; Rentmeester v. Nike 883 F.3d 1111 (2018).
[22] Batiste v. Lewis, 976 F.3d 493 (2020).
[23] Batiste v. Lewis, 976 F.3d 493 (2020).

showing that their copyrighted work "was so widely disseminated that the defendant can be presumed to have seen or heard it."[24] Taylor Swift could establish an inference of access (and therefore copying) by providing evidence of her music's popularity. However, this inference could be disproved by showing that the model's training set does not contain any of Swift's songs.[25] Artists whose livelihoods are disrupted by AI song generators might find this to be an unsatisfactory outcome, and advocate for a more generous interpretation of "access" in the context of generative models.

To prove the second element of an infringement claim (what the model created is a "copy"), a plaintiff would need to show that the model's output is tangible, fixed, and intelligible.[26] Any transitory copies of the artist's work made in the process of training the generative model would be unlikely to meet this definition.[27] To satisfy the third element ("improper appropriation"), an artist would need to prove comprehensive nonliteral similarity ("substantial similarity") between the model's output and the artist's copyrighted work. Several different tests have been developed to determine whether there is "substantial similarity" between an original copyrighted work and an allegedly infringing work. These include: the totality approach; the "ordinary observer test";[28] the "more discerning observer" test (where the plaintiff's work incorporates public domain elements); the extrinsic/intrinsic test; and the Altai abstraction filtration test for nonliteral similarity of software. The extrinsic/intrinsic test developed by the Ninth Circuit evaluates the "objective" similarities between the protectable elements of the two works, after filtering out the unprotectable elements, and then engages in a holistic, subjective comparison of the two works based on total concept and feel.[29]

The vulnerability of generative models to copyright infringement liability has prompted a variety of scholars to offer technical solutions. Computer scientists from the Harvard School of Engineering and Applied Sciences (SEAS) have proposed a method of modifying a generative model learning algorithm (with minimal performance degradation) into one that "protects against violations" of copyright law by lowering the probability of it generating infringing output (works that are similar to copyrighted works on which it was trained) so that it resembles the infringement probability of a "safe" generative model that was *not* trained on these copyrighted

---

[24] For this purpose, courts typically consider "the degree of a work's commercial success and ... its distribution through radio, television, and other relevant mediums": Design Basics, LLC v. Lexington Homes, Inc., 858 F.3d 1093 (2017); Cholvin v. B. & F. Music Co., 253 F.2d 102, 103–04 (7th Cir. 1958); Loomis v. Cornish, 836 F.3d 991, 995 (9th Cir. 2016). See also ABKCO Music, Inc. v. Harrisongs Music, Ltd., 722 F.2d 988, 998 (2d Cir. 1983); Cholvin v. B. & F. Music Co., 253 F.2d 102, 103–04 (7th Cir. 1958).

[25] Ty, Inc. v. GMA Accessories, Inc., 132 F.3d 1167 (1997) (an inference of access "can be rebutted by disproving access or otherwise showing independent creation).

[26] 17 USC 101

[27] See Cartoon Network LP, LLLP v. CSC Holdings, Inc. 536 F.3d 121 (2008) ("Given that the data reside in no buffer for more than 1.2 seconds before being automatically overwritten, and in the absence of compelling arguments to the contrary, we believe that the copyrighted works here are not "embodied" in the buffers for a period of more than transitory duration, and are therefore not "fixed" in the buffers").

[28] Whether an ordinary observer, unless they set out to detect the disparities, would be disposed to overlook them, and regard the aesthetic appeal as the same. See Mannion v. Coors Brewing Co, 377 F.Supp.2d 444 (2005).

[29] See Rentmeester v. Nike 883 F.3d 1111 (2018).

works (the "access-free" model).[30] Since the probability of a model that was *not* trained on copyrighted work *C* generating output that is substantially similar to *C* is relatively low, the authors argue that reducing the infringing probability of the original model (which *did* have access to *C)* to resemble the (very low) infringing probability of the access-free model should shield the original model from liability.[31]

While this represents an interesting and novel technical solution, it also reflects a fundamental misunderstanding of copyright law. Courts are not interested in the *probability* of a model generating infringing output; they are interested in specific instances of infringement, i.e., specific infringing *outputs*. If a model trained on copyrighted content does *in fact* produce a specific output that is substantially similar to copyrighted material on which it was trained, that specific event may constitute infringement, even if the overall probability of the model generating such output is relatively low. For this reason, the authors' claim that their technical solution offers "a rigorous guarantee of no substantive similarity" is misleading. Substantially lowering the probability of a model generating infringing output is a positive development, but as long as the probability is not *zero*, even one instance of infringing output will suffice to trigger a copyright infringement claim.

Furthermore, the authors define "similarity" between input and output data in terms of the "probability distributions of generative models rather than on particular outputs themselves." This quantitative measure of similarity is very different from, and arguably incompatible with, how substantial similarity is interpreted by copyright jurists. In copyright law, "substantial similarity" refers to the degree of similarity between the protectable elements of a copyrighted work and the protectable elements of an allegedly infringing work, and several different legal tests have been developed to measure it, as aforementioned.

The authors claim that their technical solution will "reduce the task of determining if some types of infringements have occurred to well-defined, quantitative questions."[32] Again, this statement reflects a misunderstanding of copyright law. A court will not "reduc[e] the task of determining a copyright infringement to a quantitative question of the acceptable value of *k*," *k* being the closeness between the infringing probability of a model that was trained on copyrighted work *C*, and the infringing probability of a model that was *not* trained on *C* (the "access-free model"). Accepting a quantitative measure of this kind would eliminate the judicial discretion associated with the application of fair use. It would transform fair use from an open-ended legal standard to a predetermined threshold probability that would bind judges in infringement suits moving forward.

---

[30] Nikhil Vyas, Sham Kakade, Boaz Barak, "Provable Copyright Protection for Generative Models," arXiv:2302.10870 (February 2023).
[31] Nikhil Vyas, Sham Kakade, Boaz Barak, "Provable Copyright Protection for Generative Models," arXiv:2302.10870 (February 2023).
[32] Ibid.

C. Would Generative Art Be Protected By Fair Use?

If an artist is able to prove the various elements of a copyright infringement claim (copying, a copy, and improper appropriation), the next question is whether the defendant can establish a defense to infringement, such as fair use. Courts determine whether a particular use of a copyrighted work is fair based on (1) the purpose and character of the use; (2) the nature of the copyrighted work; (3) the amount and substantiality of the portion used, in relation to the copyrighted work as a whole; and (4) the effect of the use upon the potential market for, or value of, the copyrighted work.[33] Whether the first factor weighs against a finding of fair use depends on the purpose to which a generative model is put. If a generative model is trained on the *expressive* features of a copyrighted work for a *non-expressive purpose* (for example, learning the distinctive features of Yayoi Kusama's artworks in order to identify infringing content), then such use might be considered fair. In contrast, if a model is trained on the expressive features of Kusama's artworks for an *expressive* purpose (for example, to generate similar-looking artwork without commissioning Kusama), then this use is presumptively unfair because it encroaches upon the market for the original artwork.[34] Although it is transformative,[35] courts have historically been less tolerant of such uses where they embody an expressive (and substitutive) purpose,[36] rather than a non-expressive purpose.[37]

The second fair use factor (nature of the copyrighted work) would weigh *against* a finding of fair use to the extent that a generative model was trained on highly creative (rather than factual) works. The Supreme Court's generous interpretation of "transformative" fair use in *Google v Oracle* contrasts significantly with the Second Circuit's much narrower interpretation of transformativeness in *Warhol v Goldsmith*, illustrating that in the area of visual art, courts are much more concerned about preserving an artist's market for derivative works. Courts require secondary visual artworks to embody a distinct and different purpose whereas Google was permitted to reproduce Oracle's declaring code for the *same* purpose: calling up specific task-implementing programs. The Supreme Court's opinion in *Google v Oracle* specifically noted that "copyright's protection may be stronger where the copyrighted material ... serves an artistic rather than a utilitarian function." This differential treatment of copyrightable works in the technology space can be attributed to the relatively thin

---

[33] 17 USC 107

[34] Mark Lemley and Bryan Casey, Fair Learning, 99 Texas Law Review 743 (2021).

[35] Campbell v. Acuff-Rose Music, Inc., 114 S.Ct. 1164 (1994).

[36] See, e.g. Andy Warhol Foundation for Visual Arts, Inc. v. Goldsmith, 11 F.4th 26 (2021) (lower courts should not interpret *Cariou* as meaning that any secondary work that adds a new aesthetic or new expression to its source material is necessarily transformative; overly generous interpretations of "transformativeness" risks eliminating a copyright owner's entire market for derivative works, many of which add new expression, meaning or message).

[37] See, e.g. Sony Corp of America v. Universal City Studios, Inc., 464 U.S. 417 (1984); Google LLC v. Oracle America, Inc., 141 S.Ct. 1183 (2021) (reproducing code to reimplement a user interface for the same purpose (to enable programmers to call up implementing programs to accomplish specific tasks) but in a different computing environment (smartphones) is a sufficiently transformative fair use).

protection afforded to largely functional works, and it maintains a longstanding thread of judicial deference to technological innovation.[38]

The third fair use factor (amount taken) would also likely weigh against a finding of fair use if a generative model was training on the expressive (protectable) features of artistic works in order to generate similar-looking content. Finally, the fourth factor (market effects) would also likely weigh against a finding of fair use, if a generative model produces images that are substantially similar to copyrighted images on which it was trained. The production of such output would clearly undermine the market for the original copyrighted works; users of generative models could simply use text prompts to obtain images in the style of their favorite artists without commissioning the artists themselves. This rapid and affordable means of generating artistic works would significantly diminish the livelihoods of human artists.

D. Ongoing Litigation

In January 2023, three artists launched a class action lawsuit against Stability AI, DeviantArt, and Midjourney for copyright infringement related to their development and use of an AI text-to-image generator known as Stable Diffusion. The class action complaint alleges that the images generated by Stable Diffusion are "based entirely" on copyrighted training images (for which the defendants did not seek a license or otherwise offer compensation), and represent unauthorized derivative works.[39] The complaint alleges that the defendants infringed the plaintiffs' right of reproduction, right to prepare derivative works, and rights of distribution, performance, and display. The complaint also alleges that the AI-generated images compete with the original artworks, because they allow users to obtain images "in the style of" a particular artist, without commissioning or licensing artwork from the artist themselves. Each of the plaintiffs owns a copyright interest in works used as training data for the AI model. The complaint incorrectly asserts that the copyrighted training images are "stored at and incorporated into Stable Diffusion as compressed copies." The model does not store copies of the original works; it derives patterns, rules, and relationships from the training images and stores them in the form of abstract mathematical parameters.

In February 2023, Getty Images filed a complaint against Stability AI, alleging that the defendant had copied more than 12 million photographs owned or licensed by Getty Images, along with associated captions and metadata, in order to train its text-to-image generator, Stable Diffusion.[40] Getty alleges that Stability AI used its copyrighted images to train Stable Diffusion to produce images (in response to text

---

[38] See, e.g. Sony Corp of America v. Universal City Studios, Inc., 464 U.S. 417 (1984); Sega Enterprises Ltd. v. Accolade, Inc., 977 F.2d 1510 (1992); Authors Guild v. Google, Inc., 804 F.3d 202 (2015).

[39] Sarah Andersen, Kelly McKernan, and Karla Ortiz v. Stability AI Ltd., Stability AI, Inc., Midjourney, Inc., and DeviantArt, Inc. (Class Action Complaint, filed January 13, 2023).

[40] Getty Images (US), Inc. v. Stability AI, Inc. (Complaint filed February 3, 2023).

prompts) that are "highly similar to and derivative of" Getty's own images. The reproduction of a modified version of the Getty Images watermark in images generated by Stable Diffusion underscores the "clear link between the copyrighted images that Stability AI copied without permission and the output its model delivers." Getty claims that Stability AI specifically chose to train its model on Getty's images due to their high quality and rich metadata, thereby taking advantage of the investments required to generate its collection of high-quality images with detailed image-text pairing. Getty claims infringement of its right of reproduction, and its right to create derivative works.

In addition to its claims for relief regarding copyright infringement, Getty's complaint also alleges trademark infringement. Some of the images generated by Stable Diffusion contain a modified version of the distinctive Getty Images watermark, and this creates confusion as to the source of the image and falsely implies an association with Getty Images. Getty alleges that Stability AI's incorporation of its trademarks into "low quality, unappealing, or offensive images" thereby dilutes those marks, which have historically been associated with "premium quality visual content." The incorporation of a distorted version of the Getty Images watermark on "bizarre or grotesque synthetic imagery" ultimately tarnishes Getty's hard-earned reputation for high-quality visual content.

E. Should AI-Generated Artwork Receive Copyright Protection?

The answer to this question is inextricably bound up with the fair use status of diffusion outputs. If generative models are found to engage in copyright infringement (by generating output that is substantially similar to copyrighted images on which they have been trained), then the answer is simple: an infringing derivative work receives no copyright protection.[41] If, however, a generative model is trained on public domain works or otherwise licensed copyrighted works, and it produces original synthetic images, should those images receive copyright protection? And if so, who should be regarded as the author of those images: the programmer, the user who entered the text prompts, or the model itself? These are difficult normative questions to answer.

From a utilitarian perspective, some scholars argue that AI-generated art *should* receive copyright protection because this will incentivize further innovation, which serves the public interest.[42] There are a few assumptions to unpack here. Firstly, this argument assumes that the *non-copyright incentives* for AI innovation are insufficient to support ongoing investment in AI technology. This assumption is false; technology firms have strong *non-copyright* incentives to invest in generative models, including

[41] 17 U.S. Code § 103(a).
[42] Ryan Abbott, and Elizabeth Rothman, Disrupting Creativity: Copyright Law in the Age of Generative Artificial Intelligence, Florida Law Review (forthcoming) (2022).

the licensing and subscription fees they charge users for access to their services.[43] Moreover, once these models exist, the marginal cost of generating additional creative works is extremely low. In other words, generative models will continue to be built in the absence of copyright protection for their outputs.[44]

Secondly, the argument that AI-generated art unambiguously serves the public interest finds strong opposition within the artistic community.[45] If one views human-authored art as an indispensable source of political and cultural expression, social commentary, first-person self-consciousness, and community dialogue, then the diminishment of this expression by generative models represents a loss of social welfare.[46] If, over time, a larger fraction of new cultural works is generated by AI models rather than human authors, this will significantly reduce "the quality of the epistemological and cultural signals that current generations send to future ones."[47] In other words, the social practice of human authorship will be impoverished by the encroachment of generative models.[48]

At the same time, generative models are *more likely* to displace human authors if their outputs are freely available to the public; a consumer is unlikely to commission a human artist for a specific work if a model can generate similar work for free. For this reason, some scholars have argued that extending copyright protection to AI-generated output would "level the playing field" between human authors and generative models. However, there are alternative methods of erecting cost barriers to AI-generated art (for example, charging licensing and subscription fees) that do *not* involve extending copyright protection to model outputs. Furthermore, if generative models are trained only on public domain works (to avoid copyright infringement), then they are unlikely to be able to generate works in the style of living artists, thereby preserving the commissionable market for these works. This is particularly the case if consumers want their commissioned artworks to enjoy copyright protection, but AI-generated works are not eligible for such protection.[49]

---

[43] This is in addition to the copyright protection that the underlying software receives.
[44] See, e.g., Robert Yu, The Machine Author: What Level of Copyright Protection Is Appropriate for Fully Independent Computer-Generated Works?, 165 U. PA. L. REV. 1245, 1264 (2017); Daniel J. Gervais, The Machine as Author, 105 Iowa Law Review, 2053 (2020); Pamela Samuelson, Allocating Ownership Rights in Computer-Generated Works, 47 U. PITT. L. REV. 1185 (1986) (copyright law "has allocated rights only to humans for a very good reason: it simply does not make any sense to allocate intellectual property rights to machines because they do not need to be given incentives to generate output"); Ralph D. Clifford, Intellectual Property in the Era of the Creative Computer Program: Will the True Creator Please Stand Up?, 71 TUL. L. REV. 1675 (1997).
[45] See, e.g., Robert Denicola, Ex Machina: Copyright Protection for Computer-Generated Works, 69 RUTGERS L. REV. 251 (2016) (arguing that computer-generated works provide the same benefits to the public as human-authored works, and that the grant of copyright is not primarily for the benefit of the author, but for the benefit of the public). See also Fogarty v. Fantasy, Inc., 510 U.S. 517, 524 (1994) ("The primary objective of the Copyright Act is to encourage the production of original literary, artistic, and musical expression for the good of the public").
[46] See, e.g., Daniel J. Gervais, The Machine as Author, 105 Iowa Law Review, 2053 (2020); Carys J. Craig & Ian R. Kerr, The Death of the AI Author, 52 OTTAWA L. REV. 31, 42 (2019).
[47] Daniel J. Gervais, The Machine as Author, 105 Iowa Law Review, 2053 (2020).
[48] Carys J. Craig & Ian R. Kerr, The Death of the AI Author, 52 OTTAWA L. REV. 31, 42 (2019).
[49] Ryan Abbott, and Elizabeth Rothman, Disrupting Creativity: Copyright Law in the Age of Generative Artificial Intelligence, Florida Law Review (forthcoming) (2022).

F. Attributing Authorship of AI-Generated Art

(i) Programmers As Authors

Theoretically, if AI art were to receive copyright protection, a secondary question is who should be regarded as the author for copyright purposes. One possibility is to bestow authorship upon the developers of the generative model.[50] Creating a high-quality generative model is intellectually and creatively demanding, in addition to being time-consuming and expensive.[51] However, there are many ways to reward programmers for their labor, without attributing to them authorship of generated outputs.[52] Additionally, given the minimal role played by programmers in generating specific outputs, attributing authorship to them would strain existing definitions of authorship under copyright law. It would over-reward programmers who are "no more able to anticipate the output than anyone else."[53] Vesting ownership of all generated output in the programmer may also discourage users from using the model to generate (and disseminate) new works.[54]

Jane Ginsburg and Luke Budiardjo argue that the designer of a generative model should be considered the sole author of its outputs only where the model is "fully generative" in the sense that it requires *no further input or intervention* in order to execute its given functions, that is, to generate creative expression.[55] Ginsburg and Budiardjo compare this to a wildlife photographer setting up a camera in a specific location, at a specific time, with specific camera settings, and waiting for the result.[56] Even if the precise composition and content of the final photo is unexpected, the photographer is responsible for both its conception *and* its execution, by virtue of setting up "a set of deliberate executional steps."[57] The photographer's intellectual conception of the work need not precisely foresee the final result; an artist's conception of their work often evolves in tandem with its execution, and copyright law protects works that result from accidental creativity.[58]

Analogously with the wildlife photographer, where the designer of a generative model is wholly responsible for *conceiving* and *executing* creative expression (because the model, once designed, requires no further instructions), then the designer is the author of the model's outputs, no matter how unexpected.[59] Like the wildlife

---

[50] Ryan Abbott, and Elizabeth Rothman, Disrupting Creativity: Copyright Law in the Age of Generative Artificial Intelligence, Florida Law Review (forthcoming) (2022).
[51] Pamela Samuelson, Allocating Ownership Rights in Computer-Generated Works, 47 U. PITT. L. REV. 1185 (1986).
[52] Tech firms can charge licensing and/or subscription fees to users of their generative models.
[53] Pamela Samuelson, *supra* note 51.
[54] Robert Denicola, Ex Machina: Copyright Protection for Computer-Generated Works, 69 RUTGERS L. REV. 251 (2016).
[55] Jane C. Ginsburg and Luke A. Budiardjo, Authors and Machines, 34 BERKELEY TECH. L. J. 343 (2019).
[56] Ibid.
[57] Ibid. This continues to be the case even if a downstream actor is involved (for example, an assistant clicks a button on the camera), provided that the photographer has sufficiently bound the assistant's choices such that they are limited to fulfilling a set role within the photographer's creative plan.
[58] Jane C. Ginsburg and Luke A. Budiardjo, *supra* note 55.
[59] Ibid ("as long as those designers, by designing the tool's algorithms, or training a "learning" generative model to produce outputs, control the inner workings of the system, they have also executed the resulting works").

photographer setting up their camera, when the model designer "sets up a process which will lead to the creation of the work without the contribution of any other creative forces, the designer will be the author of the end result."[60] This is *not* the case, however, for contemporary generative models, such as Stable Diffusion and Midjourney, whose output is dependent on (and influenced by) text prompts entered by downstream users. These models are not "fully generative" (as Ginsburg and Budiardjo use the term) because they cannot create expressive works without the contributions of users.[61] For these models, the user's creative choices *disrupt* the designer's claim of authorship over all images generated by the model.[62] The designer cannot anticipate how a given user will "complete" the work, and therefore cannot claim sole authorship of the result, because it is not the product of the designer's creative plan, which is incomplete without the contribution of the user.[63] For this reason, the designer of a generative model such as Midjourney cannot claim sole authorship over the images the model generates.

(ii) Users As Authors

Another possibility is to attribute authorship to the user who enters text prompts in order to generate specific images. Given the time taken by users to learn which permutations of text prompts will generate specific imagery, many would argue that this effort should be rewarded with exclusive ownership. However, the "sweat of the brow" approach to copyright ownership has long been abandoned by U.S. courts,[64] and the Copyright Office has already indicated that it does *not* regard the manipulation of text prompts as authorship. In correspondence with an artist seeking copyright registration of images generated by Midjourney, the Copyright Office explained that while text prompts can *influence* diffusion outputs, this process "is not controlled by the user because it is not possible to predict what Midjourney will create ahead of time."[65] Because text prompts do not *dictate* specific results, Midjourney users are not "authors" of the generated images.[66] Given the "significant distance" between a user's text prompts and Midjourney's ultimate output, Midjourney users "lack sufficient control over generated images to be treated as the "master mind" behind them."[67]

---

[60] Jane C. Ginsburg and Luke A. Budiardjo, *supra* note 55.
[61] Ibid.
[62] Ibid.
[63] Ibid ("when the upstream creator's creative plan for the work does not limit the downstream user's creative autonomy, and instead relies on the downstream creator to endow the work with additional (and unforeseeable) creative content, the upstream creator cannot claim to be the sole author of the resulting work because she has not crafted a complete creative plan for the work's production").
[64] Daniel J. Gervais, The Machine as Author, 105 Iowa Law Review, 2053, at 2079 (2020).
[65] Letter from US Copyright Office to Kristina Kashtanova, dated February 21, 2023, quoted in Richard Lawler, "The US Copyright Office says you can't copyright Midjourney AI-generated images," The Verge, Feb. 22, 2023.
[66] See, e.g. *Burrow-Giles Lithographic v Sarony,* 111 U.S. 53 (1884).
[67] Letter from US Copyright Office to Kristina Kashtanova, dated February 21, 2023, quoted in Richard Lawler, "The US Copyright Office says you can't copyright Midjourney AI-generated images," The Verge, Feb. 22, 2023.

Ginsburg and Budiardjo agree that the user of a generative model should not be considered the sole author of the model's outputs.[68] Although the user has a specific creative vision ("conception") of the work they wish to create, outsourcing the execution of that conception to the model vitiates the user's claim to ownership over the final output because the user has *no control* over the model's execution of their vision.[69] In general, an artist can retain authorship over a final work, despite outsourcing the execution of their creative vision, only where they exercise such a high degree of control over its execution that the agent carrying it out (human or machine) exercises no creative autonomy.[70] Users of generative models, in contrast, exercise no control over the process by which their creative vision is executed, because they cannot influence the model's inner workings.[71] Ginsburg and Budiardjo compare this to Google Translate, where users may supply the text to be translated, but they have no control over *how* the algorithm converts the text into their chosen language.[72] For this reason, they cannot be considered the "author" of those translations. In general, Ginsburg and Budiardjo argue, when the user of a machine "supplies her creative contribution without influencing how the machine translates that contribution into a final work, then the user does not execute the final work and thus cannot claim authorship."[73]

Pamela Samuelson argues, however, that there are compelling doctrinal and normative reasons to attribute authorship to users of generative models, even when their creative contributions are minimal.[74] The first is that copyright law has traditionally attributed authorship of a work to the individual who *fixes* the work in material form. For example, someone who records a live performance of improvised jazz on their smartphone is the author of that sound recording, despite the fact that their labor was confined to pressing a single button.[75] Similarly, the user of a generative model, by entering text prompts, has *caused* synthetic images to be generated, and is therefore the "instrument of fixation."[76] Indeed, there is a surprisingly robust body of case law concerning psychographic works, or works authored by spirits channeled through a human medium.[77] Despite strong disavowals of human authorship by parties to the case, courts have consistently held that the human medium (who transcribes the spirit's message, and thus fixes it in material form) is the author for copyright purposes.[78]

---

[68] Jane C. Ginsburg and Luke A. Budiardjo, *supra* note 55.
[69] Jane C. Ginsburg and Luke A. Budiardjo, *supra* note 55.
[70] Ibid. See, e.g., Lindsay v. Wrecked and Abandoned Vessel R.M.S. Titanic, 1999 WL 816163 (1999).
[71] Jane C. Ginsburg and Luke A. Budiardjo, *supra* note 55.
[72] Ibid.
[73] Ibid.
[74] Pamela Samuelson, *supra* note 51.
[75] Ibid. The performer has no copyright in the performance unless and until it has been fixed in tangible form. See also Feist Publications, Inc. v. Rural Telephone Service Co (1991) (regarding copyright law's very low standard for originality).
[76] Pamela Samuelson, *supra* note 51.
[77] Annemarie Bridy, Coding Creativity: Copyright and the Artificially Intelligent Author, 5 STAN. TECH. L. Rev. 12 (2012).
[78] Ibid. See, e.g., *Urantia Foundation v. Maaherra* (sufficient human creativity in the selection and arrangement of the revelations); *Penguin Books U.S.A., Inc. v. New Christian Church of Full Endeavor, Ltd* (sufficient creativity in editing of transcription, also, dictation from a non-human source should not be a bar to copyright).

Copyright law has a long history of allocating rights to individuals who did not actually create the expression for which they are credited as authors.[79] Joint authors, for example, are co-owners of the *whole* work, not just the part that they contributed, even if those contributions are "interdependent parts of a unitary whole." This means that a joint author receives ownership over parts of a work that they did not create.[80] Similarly, the "work made for hire" doctrine vests authorship of works created by employees in their employers.[81] These examples illustrate that the concept of authorship is flexible enough to accommodate an instrumental approach to copyright ownership; one that is determined to secure access to valuable works, by rewarding those who *instigate* (but may not be directly responsible for) their creation.[82] Based on this logic, a user who *initiates* or *originates* the creation of computer-generated expression falls "well within the constitutional dimensions" of authorship.[83]

Annemarie Bridy argues that the work made for hire doctrine is the optimal vehicle for accommodating artificially intelligent authors because it vests copyright *as a matter of law* in a party who is known not to be the author-in-fact.[84] The work made for hire doctrine explicitly acknowledges the "disidentity" between the author-in-fact (the employee) and the author-in-law (the employer).[85] Nevertheless, it bypasses the author-in-fact for practical reasons.[86] A similar legal fiction could be created in the context of generative models to vest ownership of artificial images in the user of the model (the author-in-law) rather than the model itself (the author-in-fact). In order to sustain this legal fiction, however, the existing statutory regime would need to be modified, since diffusion outputs do not currently meet the definition of a work made for hire.[87] Additionally, it would be strange to characterize machines (which lack legal personhood) as "employees" for the purposes of the work made for hire doctrine.[88] It may be simpler to vest sole authorship in the user of the generative model (by recognizing text prompts as meeting the low standard for creativity required by *Feist*) without stretching the work made for hire doctrine. Robert Denicola argues that it should be sufficient for copyright authorship that the user *initiated* or *originated* the creation of the computer-generated expression.[89] Several foreign jurisdictions have already taken this position.[90]

---

[79] Robert Denicola, Ex Machina: Copyright Protection for Computer-Generated Works, 69 RUTGERS L. REV. 251 (2016).
[80] Robert Denicola, Ex Machina: Copyright Protection for Computer-Generated Works, 69 RUTGERS L. REV. 251 (2016).
[81] Pamela Samuelson, *supra* note 51.
[82] Robert Denicola, *supra* note 80.
[83] Ibid.
[84] Annemarie Bridy, *supra* note 77. See also Shlomit Yanisky-Ravid, Generating Rembrandt: Artificial Intelligence, Copyright, and Accountability in the 3A Era--The Human-like Authors are Already Here- A New Model, 2017 MICH. ST. L. REV. 659 (2017) (advocating for a work made for hire model for AI-generated works).
[85] Annemarie Bridy, *supra* note 77.
[86] Ibid.
[87] See 17 USC §101.
[88] Robert Denicola, *supra* note 80.
[89] Ibid.
[90] See, e.g. Copyright, Designs and Patents Act, 1988, c. 48, § 9(3), 178 (U.K.); Copyright Act of 1994, § 2, 5 (N.Z.); Copyright and Related Rights Act 2000, Part I, § 2 (Act. No. 28/2000) (Ir.) (the copyright in a computer-generated work vests in "the person by whom the arrangements necessary for the creation of the work are undertaken," and this person is "taken to be" the author for statutory purposes).

Secondly, if the user is not considered the author of the images generated by their text prompts, then they have little incentive to share those images with the world, and are more likely to withhold them.[91] Samuelson argues that society has an interest in such works being made available to the public because only innovations that are revealed and disseminated "promote the progress of science and the useful arts."[92] Accordingly, if someone must be incentivized to bring the work forward, the user is "best situated to respond to [this] motivation."[93] Robert Denicola argues that a denial of copyright protection to computer-generated works would "threaten" the production of works that are "indistinguishable" in value from human-authored works, and consign a "diverse library of works to the public domain."[94]

(iii) Joint Authorship

Alternatively, the user and the programmer could receive joint authorship in the images produced by a generative model.[95] These images would be classified as "inseparable" joint works because the contributions of the user and the programmer cannot be disaggregated.[96] In order to qualify as joint authors, the user and the programmer would need to prove that they *contemporaneously intended* to merge their contributions into a unitary whole.[97] This would be difficult to prove given that their individual contributions occur asynchronously, and given that remote users almost never interact with the programmers of generative models.[98] In other words, it is unlikely that the "kind of collaborative animus that typifies joint authorship" would ever arise between the programmer of a generative model, and a remote user.[99] Unless Congress is willing to revisit the concept of joint authorship, and protect the inseparable combination of uncopyrightable contributions made by "asynchronous unacquainted contributors," then joint authorship is unlikely to vest in AI-generated artworks.[100]

(iv) Machine Authorship

Finally, copyright law could attribute authorship of diffusion outputs to the generative model itself.[101] This would represent an historic admission that machines

---

[91] Pamela Samuelson, *supra* note 51.
[92] Ibid.
[93] Ibid.
[94] Robert Denicola, *supra* note 80.
[95] Pamela Samuelson, *supra* note 51.
[96] Jane C. Ginsburg and Luke A. Budiardjo, *supra* note 55.
[97] Ibid.
[98] Ibid.
[99] Pamela Samuelson, *supra* note 51.
[100] Jane C. Ginsburg and Luke A. Budiardjo, *supra* note 55.
[101] Pamela Samuelson, *supra* note 51 ("[t]he copyright standard of originality is sufficiently low that computer-generated works, even if found to be created solely by a machine, might seem able to qualify for protection").

are no longer "inert tools of creation" but relatively autonomous authors.[102] Attributing legal authorship to a machine would raise issues of legal personhood, and contradict the Copyright Office's current position that human authorship is a prerequisite to copyright protection. However, neither of these hurdles is insurmountable; in 2022, the Canadian Intellectual Property Office (CIPO) registered an AI program as the co-author of an artistic work, alongside a human creator.[103] Ryan Abbott and Elizabeth Rothman argue that AI models should be considered authors for copyright purposes, but simply receive a shorter term of protection than human-authored works.[104] In contrast, Daniel Gervais argues that machines should never be endowed with rights (such as the bundle of exclusive rights attached to copyright) because they cannot be held liable for their creations ("no rights without responsibilities").[105] Ginsburg and Budiardjo assert that machines cannot be "authors" because they "lack the initiative that characterizes human authorship," that is, they cannot form creative visions ("conceptions") that inform their execution of expressive works.[106]

If machine authorship, joint authorship, and sole authorship of AI-generated artworks are equally infeasible under the current copyright regime, there is a chance that these works will simply be authorless, as a matter of law.[107] Reasonable minds will disagree about the likelihood and acceptability of this outcome. Those who believe that AI-generated art *should* receive copyright protection may lobby for changes to the existing statutory regime in order to overcome the authorship hurdles described above. However, changes to general authorship requirements will affect *all* copyright subject matter, not just AI-generated works, and the implications of those changes will draw even more affected parties into the ring. Perhaps our "reluctance to strand computer-enabled outputs on authorless shores" should prompt a deeper discussion about whether copyright law is the optimal vehicle for protection of these works.[108]

Carys Craig and Ian Kerr argue that doctrinal discussions about the *capacity* for copyright law to accommodate AI-generated works overlook the deeper and more fundamental question of what copyright law is *for*.[109] It is perfectly feasible, they argue, that utilitarian copyright jurists will ultimately stretch existing doctrine to accommodate AI authors, in the same way that copyright law has adopted other legal fictions in response to emerging technologies.[110] But such accommodations overlook the very *purpose* of copyright law, which is to promote *human* authorship as a socially

---

[102] *Cf* Report of the National Commission on New Technological Uses of Copyrighted Works (CONTU) (1976) and the 1986 Report of the Office of Technology Assessment (OTA) (interactive computer programs might be considered co-authors of the output they produce).
[103] Oyen Wiggs Green & Mutala LLP, "Canada: CIPO Recognizes An AI As Co-Author In A Copyright Registration," Mondaq, 15 March 2022.
[104] Ryan Abbott, and Elizabeth Rothman, Disrupting Creativity: Copyright Law in the Age of Generative Artificial Intelligence, Florida Law Review (forthcoming) (2022).
[105] Daniel J. Gervais, The Machine as Author, 105 Iowa Law Review, 2053 (2020).
[106] Jane C. Ginsburg and Luke A. Budiardjo, *supra* note 55.
[107] Ibid.
[108] Ibid.
[109] Carys J. Craig & Ian R. Kerr, The Death of the AI Author, 52 OTTAWA L. REV. 31, 42 (2019).
[110] Ibid.

situated, relational, and dialogic practice.[111] To insist upon this purpose is not to resurrect a Romantic notion of authorship, but to recognize that the creation, exchange, and transformation of cultural works between humans is mutually constitutive of meaning, identity, and social values.[112] Accordingly, in protecting and promoting *human* authorship, copyright law lays the discursive foundations for a "robustly participatory culture and democratic civil society."[113]

Generative models cannot contribute to social discourse in the same manner as human authors because they have no intent or desire to communicate.[114] Naturally, some human artists may disagree with this view of the dialogic (in)capacity of generative models; they might argue that their ability to efficiently generate a specific image using Midjourney or Stable Diffusion in fact *enhances* their ability to cultivate their selfhood, in dialogue with other artists. Again, reasonable minds will disagree about the communicative value of AI-generated art, and thus the utility of bringing these works within the existing copyright paradigm. It may be conceptually and doctrinally easier to establish a *sui generis* system for the legal protection of AI-generated works that does not involve fundamentally restructuring the existing copyright regime.[115]

[111] Ibid.
[112] Carys J. Craig & Ian R. Kerr, *supra* note 109.
[113] Ibid.
[114] Ibid.
[115] Ibid.