SMU Science and Technology Law Review

Volume 26 | Number 2 Article 5

2023

Generative AI Art: Copyright Infringement and Fair Use

Michael D. Murray
University of Kentucky

GENERATIVE AI ART: COPYRIGHT INFRINGEMENT AND FAIR USE

Michael D. Murray[*]

ABSTRACT

The discussion of AI copyright infringement or fair use often skips over all the required steps of the infringement analysis in order to focus on the most intriguing question, "Could a visual generative AI generate a work that potentially infringes a preexisting copyrighted work?" and then the discussion skips further ahead to, "Would the AI have a fair use defense, most likely under the transformative test?" These are relevant questions, but without considering the actual steps of the copyright infringement analysis, the discussion is misleading or even irrelevant. This neglecting of topics and stages of the infringement analysis fails to direct our attention to a properly accused party or entity whose actions prompt the question. Making a sudden transition from a question of infringement in the creation of training datasets to the creation of foundation models that draw from the training data to the actual operation of the generative AI system to produce images makes a false equivalency regarding the processes themselves and the persons responsible for them. The questions ought to shift focus from the persons compiling the training dataset used to train the AI system and the designers and creators of the AI system itself to the end users of the AI system who conceive of and cause the creation of images.

The analysis of infringement or fair use in the generative AI context has suffered from widespread misunderstanding concerning the generative AI processes and the control and authorship of the end-user. Claimants, commentators, and regulators have made incorrect assumptions and inaccurate simplifications concerning the process, which I refer to as the Magic File Drawer theory, the Magic Copy Machine theory, and the Magic Box Artist theory. These theories, if they were true, would be much easier to envision and understand than the actual science and technology that goes into the creation

---

https://doi.org/10.25172/smustlr.26.2.4

* The author is the Spears Gilbert Associate Professor of Law and the principal investigator of the Artificial Intelligence and the Law Project at the University of Kentucky, J. David Rosenberg College of Law. The author thanks Robert Brauneis (George Washington) for his thoughtful review and comments on the article, and he thanks the participants in the Texas A&M IP Scholars Roundtable, especially Nikola Datzov (N. Dakota), Danny Friedmann (Peking Transnational), Timothy Hsieh (Oklahoma City), Eric Johnson (Oklahoma), Lee Jyh-An (Chinese Univ. Hong Kong), Barbara Lauriat (Texas Tech), Marshall Leaffer (Indiana), Doris Long (Illinois Chicago), Timothy McFarlin (Cumberland), Alina Ng Boyte (Mississippi College), Emma Perot (West Indies), Guy Rub (Ohio State), Saurabh Vishnubhakat (Cardozo), and Peter Yu (Texas A&M), for their comments and input. He also thanks Dean Mary Davis (Kentucky) and Associate Dean for Faculty Research Michael Healy (Kentucky), and the Rosenberg College of Law faculty research fund for their on-going support of his research.

and operation of a contemporary visual generative AI system. Throughout this Article, I will attempt to clarify and correct the understanding of the science and technology of the generative AI processes and explain the different roles of the training dataset designers, the generative AI system designers, and the end-users in the rendering of visual works by a generative AI system. Part II will discuss the requirements of a claim of copyright infringement including each step from the copyrightability of the claimant's work, the doctrines that limit copyrightability, the requirement of an act of copying, and the infringement elements. Part III will summarize the copyright fair use test paying particular attention to the purpose and character of the use analysis, 17 U.S.C. § 107(1), and the current interpretation of the "transformative" test after *Andy Warhol Foundation v. Goldsmith*, particularly in circumstances relating to technology and the use of copyrighted or copyrightable data sources. Part IV will analyze potential infringement or fair use by the creators of generative AI training datasets. Part V will analyze potential infringement or fair use by the creators of visual generative AI systems. Part VI will analyze potential infringement or fair use by the end-users of visual generative AI systems.

For all their complexity, visual generative AI systems are tools that depend on an end-user who conceives of and designs the image and provides the system with a prompt to set the generative process in motion. The end-users are responsible for crafting the prompt or series of prompts used, for evaluating the outputs of the generative AI, for adjusting and editing the iterations of images offered by the AI system, and ultimately for selecting and adopting one of the images generated by the AI as the final image. The end-users then make further decisions about the actual use and its function and purpose for the images the end-users selected and adopted from the outputs of the AI. While working with the AI tool to try to produce a certain image, an end-user might steer the system to produce a work that could, under an infringement analysis, be regarded as potentially infringing, which would lead us again to the fair use analysis based on the end-user's use of the image.

I. INTRODUCTION: THE ADVANCEMENT OF ARTIFICIAL INTELLIGENCE IN ARTISTIC CREATION

Generative AI is "a branch of artificial intelligence [(AI)] that enables [computerized systems] to quickly and convincingly create original content ranging from images and artwork to poetry, music, text, video, dialog, and even computer code."[1] By all accounts, AI developers have made quantum advances in the generative creation of art and other expressive media, many of which came to fruition with consumer-ready applications in 2022.[2] Programs such as

1. Louis Rosenberg, *Generative AI: The technology of the year for 2022*, Big Think (Dec. 20, 2022), https://bigthink.com/the-present/generative-ai-technology-of-year-2022/ [https://perma.cc/A3ER-MT8B].

2. *Id.*; Jacob Bourne, *Generative AI made its public debut in 2022—it could be an internet earthquake in 2023*, Insider Intelligence (Dec. 22, 2022), https://

DALL-E 2,[3] Stable Diffusion,[4] and Midjourney[5] have enabled artists and enthusiasts to produce innumerable works of great visual interest with simple textual instructions—prompts—which could be as short as "cat playing poker," which produced the image[6] at left below using Stable Diffusion.

An AI also can mimic an actual artist's style, as seen in the work generated by the prompt "oil painting in the style of Renoir of a cat playing poker" (at right).[7]

The first point to observe here is that these images did not already exist; the AI did not simply search the internet or the corpus of image-related data it was trained on for an image that matched the terms in the prompt. The AI followed its generative process to render a new, never-before-seen image that followed the design prompt provided by the human user.[8]

www.insiderintelligence.com/content/generative-ai-made-its-public-debut-2022-it-could-internet-earthquake-2023 [https://perma.cc/ZFR9-GC9R].

3. OpenAI, DALL-E 2, https://openai.com/product/dall-e-2 [https://perma.cc/A3E5-R6ET].

4. Stable Diffusion Online, https://stablediffusionweb.com/ [https://perma.cc/8Y7X-T6MG] (the author used this program to generate the art included in this article).

5. Midjourney, https://www.midjourney.com/home/ [https://perma.cc/6ALB-84SP].

6. Image of Cat Playing Poker (2023), conceived of and prompted by Michael D. Murray, rendered by Stable Diffusion Playground (Dec. 3, 2023), https://stablediffusionweb.com/#demo [https://perma.cc/NH4P-L3W4].

7. Image of Cat Playing Poker in the Style of Renoir (2023), conceived of and prompted by Michael D. Murray, rendered by Stable Diffusion Playground (Dec. 3, 2023), https://stablediffusionweb.com/#demo [https://perma.cc/NH4P-L3W4].

8. In this sentence, I avoided using the term "created" as in "The AI created the new image." "Creation" in the discussion of copyright involves the concept of conceived of in the mind of the creator, *see* Feist Publ'ns, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340, 345, 346, 362–63 (1991); Trade-Mark Cases, 100 U.S. 82, 94 (1879) (copyright extends to works "founded in the creative powers of the mind"); U.S. Copyright Office, Copyrightable Authorship: What Can Be Registered, Compendium of the U.S. Copyright Office Practices §§ 306,

An AI design prompt can be very detailed with lots of adjectives one might desire for a contemplated image, such as the painting of Margaret Thatcher playing poker (below, left).[9]

One can even write a design prompt intended to render a fake news photograph of Queen Elizabeth II and Vladimir Putin playing poker (right).[10]

Have several AIs suddenly become sentient and excessively passionate about flexing their artistic muscles to create paintings, drawings, and photographs? The answer is most certainly "no," none of the AIs currently in use are sentient, they do not decide to paint or compose on their own, and they do not design and create artworks spontaneously:

308, 313.2 (3d ed. 2021), https://www.copyright.gov/comp3/chap300/ch300-copyrightable-authorship.pdf [https://perma.cc/NAV5-X5UE], which does not happen with AI. As noted below, the AI is not imagining and conceiving of these images on its own—it only works in response to a human prompt. *See* Rosenberg, *Generative AI*, *supra* note 1.

9. Image of Margaret Thatcher Playing Poker in the Style of John Singer Sargent (2023), conceived of and prompted by Michael D. Murray, rendered by Stable Diffusion Playground (Dec. 3, 2023), https://stablediffusionweb.com/#demo [https://perma.cc/NH4P-L3W4]. The prompt used was: "a vibrant professional studio portrait painting of Margaret Thatcher playing poker, in the style of John Singer Sargent, attractive, friendly, casual, delightful, intricate, gorgeous, wears gold earrings and gold necklace, curated collection, award winning, breathtaking, groundbreaking, superb, outstanding, dramatic lighting."

10. Image in the style of a news photograph of Queen Elizabeth II and Vladimir Putin Playing Poker (2023), conceived of and prompted by Michael D. Murray, rendered by Stable Diffusion Playground (Dec. 3, 2023), https://stablediffusionweb.com/#demo [https://perma.cc/NH4P-L3W4]. Thank goodness Stable Diffusion placed Putin's bowler hat at an unnatural angle, or the now deceased Her Royal Highness might have had some explaining to do.

> The output [of generative AIs] is so impressive that it is easy to imagine that we've suddenly created sentient machines with a creative spirit, but that is not the case. [AI] systems are master imitators of human creativity. They have been trained on millions upon millions of human artifacts such as documents, articles, drawings, paintings, movies, or whatever else can be stored in databases at scale. These systems have no conceptual understanding of the information they process—to a computer, it's all just patterns of data—and yet, these Generative AI tools can create new pieces of content that are original and awe-inspiring.[11]

The discussion of AI copyright infringement or fair use often skips over all of the required steps of the infringement analysis in order to focus on the most intriguing question, "Could a visual generative AI generate a work that potentially infringes a preexisting copyrighted work?" and then the discussion skips further ahead to, "Would the AI have a fair use defense, most likely under the transformative test?"[12] These are relevant questions, but in isolation from the actual steps of the copyright infringement analysis, the discussion is misleading or even irrelevant. This skipping of topics and stages of the infringement analysis does not train our attention to a properly accused party or entity whose actions prompt the question. The leaping from a question of infringement in the creation of training datasets to the creation of foundation models that draw from the training data to the actual operation of the generative AI system to produce images makes a false equivalency regarding the processes themselves and the persons responsible for them. The questions ought to shift focus from the persons compiling the training dataset used to train the AI system and the designers and creators of the AI system itself to the end users of the AI system who conceive of and cause the creation of images.

The analysis of infringement or fair use in the generative AI context has suffered from widespread misunderstanding concerning the generative AI processes and the control and authorship of the end-user. Claimants,

---

11. Rosenberg, *Generative AI*, *supra* note 1.

12. *E.g., Generative AI Generates Excitement—and Copyright Concerns*, JONESDAY.COM (Apr. 2023), https://www.jonesday.com/en/insights/2023/04/generative-ai-generates-copyright-concerns# [https://perma.cc/S2HP-SCL3]; James Vincent, *The scary truth about AI copyright is nobody knows what will happen next*, THE VERGE (Nov. 15, 2022), https://www.theverge.com/23444685/generative-ai-copyright-infringement-legal-fair-use-training-data [https://perma.cc/TXL4-WZCY]; Jessica L. Gillotte, *Copyright Infringement in Ai-Generated Artworks*, 53 U.C. DAVIS L. REV. 2655, 2658-59, (2020), https://lawreview.law.ucdavis.edu/issues/53/5/notes/files/53-5_Gillotte.pdf [https://perma.cc/FFM3-XBSC] (in general, the discussion in this note jumps from the topic of copying images as training data to the creation of works that might be substantially similar to one or more works in the training data).

commentators, and regulators have made incorrect assumptions and inaccurate simplifications concerning the process:

- The "Magic File Drawer" theory[13]: It is simple to imagine that the designer of a training dataset just downloaded and copied whole images, .jpg, .png, .gif, and other image files, so someone could rifle through them later and select one or two image files for copying or collaging and spit out a "new" work incorporating large copyrightable parts of the selected works it found in the drawer.
- The "Magic Copy Machine" theory: This theory incorporates the magical thinking that if a foundation model for a visual generative AI system was trained on image data from billions of images, and the AI system using the data can produce images that resemble preexisting images whose data was incorporated into the training set and foundation model, then there must have been copying of individual images whose data went into the training data.
- The "Magic Box Artist" theory: This theory engages in magical thinking that the generative AI system itself does all the design work and simply generates the artwork, with no human authorship required.

These theories, if they were true, would be much easier to envision and understand than the actual science and technology that goes into the creation and operation of a contemporary visual generative AI system. Throughout this Article, I will attempt to clarify and correct the understanding of the science and technology of the generative AI processes and explain the different roles of the training dataset designers, the generative AI system designers, and the end-users in the rendering of visual works by a generative AI system.

Part II will discuss the requirements of a claim of copyright infringement including each step from the copyrightability of the claimant's work, the

13. Each of these three theories incorporates the term "magic" by which I mean to suggest the concept of "magical thinking" that is a fallacy in logic and reasoning where obvious disconnects in causation or correlation are ignored in favor of a belief that the causations simply exist as if by magic. It is also referred to as "associative thinking" or the "Post hoc" fallacy (*post hoc, ergo propter hoc*). *See Magical thinking*, Art & Popular Culture (accessed Jun. 13, 2023), http://www.artandpopularculture.com/Magical_thinking [https://perma.cc/Q598-YGK9]; *Post hoc, ergo propter hoc*, Rational Wiki (accessed Jun. 13, 2023), https://rationalwiki.org/wiki/Post_hoc,_ergo_propter_hoc [https://perma.cc/UV39-UL5E]**.** Although it would be apt, I am not specifically referring to what Arthur C. Clarke stated as his third law of prediction of the future: "Any technology that is sufficiently advanced is indistinguishable from magic." Arthur C. Clark, *Hazards of Prophecy: The Failure of Imagination*, *in* Profiles of the Future: An Enquiry into the Limits of the Possible 12, 21 n.1 (New York, Harper & Row eds., 1973).

doctrines that limit copyrightability, the requirement of an act of copying, and the infringement elements.

Part III below will summarize the copyright fair use test paying particular attention to the purpose and character of the use analysis, 17 U.S.C. § 107(1), and the current interpretation of the "transformative" test after *Andy Warhol Foundation v. Goldsmith*,[14] particularly in circumstances relating to technology and the use of copyrighted or copyrightable data sources.

The discussion of infringement and fair use will then proceed in three parts based on the entity or person whose activity relating to image generation is being assessed for infringement or fair use:

Part IV – Potential infringement or fair use by the creators of generative AI training datasets. This part will explore the process of creation of training datasets such as LAION-5B that use image data from copyrighted or copyrightable images from the internet to compile a training model or foundation model for later use by a visual generative AI system.

Part V – Potential infringement or fair use by the creators of visual generative AI systems. Using the specific examples of OpenAI's DALL-E 2 and Stability AI's Stable Diffusion visual generative AI systems, I will examine the requirements of copyright infringement claims in light of how these systems actually operate and evaluate the potential fair use defenses of the developers of these systems.

Part VI – Potential infringement or fair use by the end-users of visual generative AI systems. For all their complexity, visual generative AI systems are tools that depend on an end-user who conceives of and designs the image and provides the system with a prompt or more often a series of prompts to set the generative process in motion. The end users are responsible for crafting the prompt or series of prompts used, for evaluating the outputs of the generative AI, for adjusting and editing the iterations of images offered by the AI system, and ultimately for selecting and adopting one of the images generated by the AI as the final image. The end-users then make further decisions about the actual use and the use's function and purpose for the images the end-users selected and adopted from the outputs of the AI. While working with the AI tool to try to produce a certain image, an end-user might steer the system to produce a work that could, under an infringement analysis, be regarded as potentially infringing, which would lead us again to the fair use analysis based on the end-user's use of the image.[15]

---

14. *Andy Warhol Foundation v. Goldsmith*, 598 U.S. 508 (2023).

15. Just as point of clarification, another extremely popular question raised by visual generative AI is whether the outputs of these AI systems are copyrightable. This question is not the subject of this article. Instead, *see* Michael D. Murray, *Tools Do Not Create: Human Authorship in the Use of Generative Artificial Intelligence*, SSRN (Jul. 5, 2023), https://ssrn.com/abstract=4501543 [https://perma.cc/JDU2-UUDE]; Michael D. Murray, *Generative and AI Authored Artworks and Copyright Law*, 45 Hastings Comm. & Ent. L. J. 27 (2023).

II. THE REQUIREMENTS OF A COPYRIGHT INFRINGEMENT CLAIM

There are five separate requirements for a copyright infringement action[16] that are relevant to this discussion of infringement and fair use:

A. The plaintiff must be the owner of a valid, copyrightable work that is registered with the U.S. Copyright office; the work must be original, created, expression, fixed in a tangible medium, that is not limited or precluded by the idea-expression distinction, the originality doctrines of merger and scènes à faire, or the useful articles doctrine.
B. The defendant must have made an unauthorized copy of original elements of the valid, copyrightable work owned by the plaintiff.
C. The portion copied must be substantial and material and more than de minimis.
D. The defendant's copy must be substantially similar to the original and copyrightable portions of plaintiff's work that were copied.
E. The defendant must not have a fair use.

Each of these requirements will be explained in the subsections below in the context of the plaintiffs and defendants in actions for copyright infringement involving visual generative AI.

A. Plaintiff Must Own and Register A Valid, Copyrightable Work

The starting point of an infringement analysis is not the defendant's actions but the nature and qualifications of the plaintiff's work. There are two conceptual requirements and two formal, physical requirements for copyrightability. The conceptual requirements are that the plaintiff's work be original and that it be a work of authorship,[17] meaning a work conceived of and created by an act of authorship,[18] which together are referred to as the **originality** and **creativity** requirements. Under copyright law, original means the work is not copied from another preexisting work, and that the work contains copyrightable subject matter—e.g., pictorial, graphic, and sculptural expression[19]—as opposed to noncopyrightable subject matter—e.g., an idea, procedure, process, system, method of operation, concept, principle, or discovery.[20]

The two formal requirements are **expression** and **fixation in media**. "Expression" means the work has to have some communicative potential

16. *See* Michael D. Murray, A Short and Happy Guide to Copyright 171-86 (2d ed. 2022); Leonard D. DuBoff & Michael D. Murray, Art Law: Cases and Materials 125–32 (3d ed. 2023).

17. 17 U.S.C. § 102(a).

18. Google LLC v. Oracle Am., Inc., 141 S. Ct. 1183, 1196 (2021).

19. 17 U.S.C. § 102(a)(5).

20. 17 U.S.C. § 102(b).

for one of the senses.[21] Copyright is looking for an author to communicate a concept that can exist as an idea in the mind of the author and be communicated to the mind of someone else through some communicative media.[22] And "fixation in media" means that the expression has to be in some form in which it can be perceived by one of the senses for long enough that we can tell what the creation is and receive its communication. The law defines "fixation" as: "authorship fixed in any tangible medium of expression, now known or later developed, from which [the works] can be perceived, reproduced, or otherwise communicated, either directly or with the aid of a machine or device."[23]

The originality and creativity requirements of the owner's expression have prompted the development of several doctrines that limit copyrightability by focusing on what and how much of an artist's creation truly was original to the artist, not preexisting and not borrowed or adopted from earlier works. The **scènes à faire** doctrine in visual art refers to work that contains stock scenes or stock images and commonplace expressions or elements that are firmly rooted in a style or genre's traditions, that are not original to the artist, and that the artist copied or at least adapted for her own expressions.[24] **Merged** ideas and expression subject to the **merger doctrine** in copyright are also not original to one artist and are not copyrightable because they function as section 102(b) ideas, not section 102(a) expression.[25] A merged idea and expression in the visual arts means that the visual expression of a certain concept, or the depiction of a certain scene or object, or the use of a certain artistic technique, process, or procedure dictates that the outputs of this activity will naturally and predictably resemble each other because the work incorporates the visual features of the merged concept.[26] Artists following a certain genre or style or school of art are almost inevitably incorporating merged, scènes à faire, or uncopyrightable imagery because the artists are following preexisting methods of depiction, and stylistic and genre-specific formulas, processes, or procedures (*i.e.*, ideas) in their works.[27]

---

21. *See Oracle Am., Inc.*, 141 S. Ct. at 1196.

22. *See* Georgia v. Public.Resource.Org, Inc., 140 S. Ct. 1498, 1510-11 (2020).

23. 17 U.S.C. § 101 (definition of "fixed").

24. Gray v. Hudson, 28 F.4th 87, 97-98 (9th Cir. 2022); Design Basics, LLC v. Kerstiens Homes & Designs, Inc., 1 F.4th 502, 503 (7th Cir. 2021).

25. Ets-Hokin v. Skyy Spirits, Inc., 225 F.3d 1068, 1082 (9th Cir. 2000); Morrissey v. Procter & Gamble Co., 379 F.2d 675, 678–79 (1st Cir.1967); 17 U.S.C. § 102(a)-(b).

26. *See* Murray, *supra* note 16, at 59–69; DuBoff & Murray, *supra* note 16, at 79–89.

27. *See* Murray, *supra* note 16, at 59–69; DuBoff & Murray, *supra* note 16, at 79–89.

In addition, there is a practical analysis that will test the functionality and utility of the owner's copyrighted material under the **useful articles** doctrine. If the work has utility, the court will look to see if there are creative, expressive parts that are physically separable, such as the statuette bases for lamps,[28] or decoration or ornamentation that is conceptually separable from the functioning of the work, such as the pictorial design of a piece of flooring[29] or the designs on a cheerleader's uniform.[30] If the form and the function are not separable, the work is not copyrightable.

All of these tests—originality, idea-expression, merger, scènes à faire, and functionality-utility—are a stress on the plaintiff's copyright that might turn a properly thick copyright with a broad scope of protection against duplicates and unauthorized derivative works into a thin copyright.[31] A **thin copyright** prevents little except nearly exact duplicates.[32] The court might find the plaintiff's copyright to be so thin that it cannot possibly preclude the defendant's work, and dismiss the lawsuit.

The relevance of establishing that a plaintiff in a copyright infringement suit must meet the above requirements of a valid copyrightable work before we even consider if a subsequent work might infringe the work is that many artists work within a certain style or genre of art, and many others use techniques, processes, and artistic procedures that result in significant reductions in the copyrightable portions of their works. Many artists seek to depict a certain kind of scene or composition in which preexisting conventions of depiction and actual visual elements of the artists' works are not original to the artists. It is therefore essential in any claim against the creators of a training dataset, a foundation model, an image generating system, or the end-users of such a system, that the claimant reveal their work that they claim has been copied so that the work can be analyzed or parsed for its uncopyrightable elements. After that the various accused parties can defend against the allegation that their activities copied substantial and material portions of the plaintiffs' *copyrightable* elements of their works.

B. An Unauthorized Act of Copying

A successful claim of infringement requires that the court find that the defendant engaged in an unauthorized act of copying: To proceed with a copyright infringement action, "the plaintiff must, as a factual matter, prove that the

28. *See* Mazer v. Stein, 347 U.S. 201, 217–19 (1954).

29. *See* Home Legend, LLC v. Mannington Mills, Inc., 784 F.3d 1404, 1408 (11th Cir. 2015).

30. *See* Varsity Brands, Inc. v. Star Athletica, LLC, 799 F.3d 468, 471 (6th Cir. 2015).

31. Corbello v. Valli, 974 F.3d 965, 973-74 (9th Cir. 2020).

32. *See, e.g.*, Satava v. Lowry, 323 F.3d 805, 812 (9th Cir. 2003).

defendant 'actually used the copyrighted material to create his own work.'"[33] And not just any form of copying will do, as the court must find the defendant copied elements of the plaintiff 's work that are themselves original and copyrightable.[34]

"Copying" itself is not defined in the statute, but "copies" are defined as "material objects . . . in which a work is fixed by any method now known or later developed, and from which the work can be perceived, reproduced, or otherwise communicated, either directly or with the aid of a machine or device."[35] Further, in defining the rights possessed by copyright owners, the owner has the right to preclude others from actions "(1) to reproduce the copyrighted work in copies or phonorecords; (2) to prepare derivative works based upon the copyrighted work; (3) to distribute copies or phonorecords of the copyrighted work to the public . . ."[36] These two definitions together indicate that the act of copying must be one that produces a copy or derivative work that is expressive and communicative and that contains and incorporates copyrightable portions of the claimant's work.[37]

Often there is no direct evidence nor a concession or stipulation regarding actual copying,[38] so the law has substituted a test for the "likelihood of copying" rather than proof of actual copying. The test is (a) proof that the defendant had access to the work, and (b) substantial similarity between the two works.[39] In a generative AI copyright dispute, it is possible to find that data from a specific work was *included* in a dataset that was then used to train the foundation model of a generative AI system by use of a search tool such as "Have I Been Trained" which searches the LAION-5B dataset that was used to train Stable Diffusion.[40] (Whether the image was *copied* or not is discussed immediately below). Assuming that the plaintiff's suspect image shows up in the "Have I Been Trained" results, then that would be evidence that the image was *included* in the billions of materials collected at the initial stage of the formation of LAION-5B which was then used

33. Gen. Universal Sys., Inc. v. Lee, 379 F.3d 131, 141 (5th Cir. 2004).

34. *Feist Publ'ns, Inc.*, 499 U.S. at 361.

35. 17 U.S.C. § 101 (definition of "copies").

36. 17 U.S.C. § 106(1)-(3).

37. *E.g.*, Sega Enters. Ltd. v. Accolade, Inc., 977 F.2d 1510, 1527-28 (9th Cir. 1992) (reproduction of copyrighted works, if incidental to a nonexpressive purpose, was non-infringing fair use).

38. And then there are cases such as *Rogers v. Koons*, 960 F.2d 301 (2d Cir. 1992), and *Cariou v. Prince*, 714 F.3d 694 (2d Cir. 2013), where there was ample direct evidence of copying.

39. *See* Murray, *supra* note 16, at 11–14.

40. Have I Been Trained, https://haveibeentrained.com/ [https://perma.cc/W8SY-BQ46]. An example search would be https://haveibeentrained.com/?search_text=kelly mckernan [https://perma.cc/4ZYY-FWX6].

by Stable Diffusion.[41] But while this fact gives the *possibility* of access, it is not complete proof that the author of the allegedly infringing work actually copied and incorporated copyrightable parts of the first work.

Because a showing of potential *access* at the proof of copying stage is not proof of actual *copying*, the inquiry must turn to the second step of the analysis which introduces the concept of "substantial similarity." A test with this name is part of the overall copyright infringement requirements, and at the proof of an act of copying stage the elements of the analysis are the same.[42] At the proof of an act of copying stage, substantial similarity requires, first, that the court must determine whether the two works are "extrinsically similar because they contain substantially similar ideas that are subject to copyright protection"; and second, the court must ask whether the works are "intrinsically similar" in the sense that they express those ideas in a substantially similar manner from the perspective of the intended audience of the work.[43] This analysis requires that the allegedly infringed work must be presented side-by-side with the allegedly infringing work for extrinsic and intrinsic comparison.[44]

What is relevant about this requirement is that for copyright infringement to be established, the defendant must have created a work that is expressive and fixed that can be compared side-by-side with the allegedly infringed work. And with that observation, we are now so far removed from the actual workings of visual generative AI training and the process of image generation that it strains the imagination to discuss these requirements; but I will endeavor to do just that in sections IV, V and VI below.

III. COPYRIGHT'S FAIR USE TEST AFTER ANDY WARHOL FOUNDATION V. GOLDSMITH

The U.S. Supreme Court in *Andy Warhol Foundation v. Goldsmith*[45] has clarified the copyright fair use transformative test for only the second time

41. As far as my research indicates, there is no such service for searching OpenAI's CLIP and DALL-E 2's dataset or Midjourney's dataset. The founder and CEO of Midjourney admits that it used "a big scrape of the internet." Rob Salkowitz, *Midjourney Founder David Holz on The Impact of AI on Art, Imagination and the Creative Economy*, Forbes (Sep. 16, 2022), https://www.forbes.com/sites/robsalkowitz/2022/09/16/midjourney-founder-david-holz-on-the-impact-of-ai-on-art-imagination-and-the-creative-economy/?sh=465c981a2d2b [https://perma.cc/BGR3-FBEY]; Christian Heidorn, *What We Know About the Midjourney Model*, Tokenized (May 28, 2023), https://tokenizedhq.com/midjourney-model/ [https://perma.cc/Q6A5-K7CJ].

42. *See Gray*, 28 F.4th at 96 (equating the requirements of substantial similarity at the proof of copying stage with substantial similarity at the infringement stage).

43. Lyons P'ship, L.P. v. Morris Costumes, Inc., 243 F.3d 789, 801 (4th Cir. 2001).

44. *Gen. Universal Sys., Inc.*, 379 F.3d at142.

45. Andy Warhol Found. for the Visual Arts, Inc., 598 U.S. 508 (2023) [hereinafter *Warhol*].

since it was adopted in 1994 in *Campbell v. Acuff Rose Music*.[46] *Campbell* provided the foundational discussion of the test and how it should operate, and *Google v. Oracle*[47] explained and applied the test in the context of computer code and the fair use copying of code for a new function and purpose in a new computer application. *Warhol* affirmed both *Google* and *Campbell*, and early commentary on the *Warhol* case appears to agree that the Supreme Court did not significantly reinterpret the transformative test nor did it fundamentally alter the way the test operates in fair use analyses.[48]

A fair use discussion begins, of course, with the text of 17 U.S.C. § 107:

> [T]he fair use of a copyrighted work . . . for purposes such as criticism, comment, news reporting, teaching . . . , scholarship, or research, is not an infringement of copyright. In determining whether the use made of a work in any particular case is a fair use the factors to be considered shall include—
>
> (1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes;
> (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.
>
> The fact that a work is unpublished shall not itself bar a finding of fair use if such finding is made upon consideration of all the above factors.[49]

A. The Andy Warhol Foundation Ruling

The basic statement of the transformative test accepted by *Warhol* and *Google* and promulgated by *Campbell* is "whether the copier's use 'adds something new, with a further purpose or different character, altering' the copyrighted work 'with new expression, meaning or message.'"[50] *Campbell* created and applied the test primarily in reference to the first fair use factor, the purpose and character of the use. Factor one draws on Justice Story's formulation

46. Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569 (1994).

47. *Oracle Am., Inc.*, 141 S. Ct. 1183.

48. *E.g.*, Corynne McSherry et al., *What the Supreme Court's Decision in Warhol Means for Fair Use*, Electronic Freedom Frontier (May 23, 2023), https://www.eff.org/deeplinks/2023/05/what-supreme-courts-decision-warhol-means-fair-use [https://perma.cc/BC2V-KNLG]; Anthony J. Dreyer et al., *Supreme Court Addresses Copyright Fair Use Defense in Goldsmith*, Skadden (May 19, 2023), https://www.skadden.com/insights/publications/2023/05/supreme-court-addresses-copyright-fair-use-defense [https://perma.cc/XK6T-96AJ].

49. 17 U.S.C. § 107.

50. *Accord Warhol*, 143 S. Ct. at 1274-75.

of the fair use analysis in *Folsom v. Marsh*, that when considering "the nature and objects of the selections made" in the new work copying the first work, does the new work merely "supersede the objects of the original creation."[51] To this, *Campbell* added the words of Judge Pierre Leval, asking whether the new work "instead adds something new, with a further purpose or different character, altering the first with new expression, meaning, or message; it asks, in other words, whether and to what extent the new work is 'transformative.'"[52]

Later in the same opinion, *Campbell* broadened the scope of the test to relate to all of the fair use factors so that there would be an equilibrium between the fair use factors with no one factor, such as commerciality, being "dispositive" or "conclusive."[53] The Court referred to the "preamble" (sentence one) of section 107 in defining the transformative test, and connected the test to the public policies favoring free expression and the creation of new, original expression.[54] Transformation is not tied to one factor because a properly transformative use of original work would tip the scales in favor of fair use on all of the factors when considered together.[55] This broader scope helps the transformative test to fulfill the goal of copyright, to promote science and the arts, because this goal is furthered by the creation of transformative works. The Court held that transformative works lie at the heart of the fair use doctrine's guarantee of breathing space within the confines of copyright, and the more transformative the new work, the less will be the significance of other factors, like commercialism, that may weigh against a finding of fair use.[56]

The lower federal courts have, of course, worked to interpret and apply the transformative test in the years after *Campbell*.[57] My study of the application and interpretation of the transformative test in the federal appellate courts between

51. Folsom v. Marsh, 9 F. Cas. 342, 348 (C.C.D. Mass. 1841) (Story, J., sitting as circuit justice); *see* Harper & Row, 471 U.S. at 562 ("supplanting" the original).

52. *Campbell*, 510 U.S. at 579 (quoting Pierre N. Leval, *Commentary, Toward a Fair Use Standard*, 103 Harv. L. Rev. 1105, 1111 (1990)).

53. *Id.* at 578, 584–85, 594 ("Nor may the four statutory factors be treated in isolation, one from another. All are to be explored, and the results weighed together, in light of the purposes of copyright" and that there are "no hard evidentiary presumption[s]. . . . [T]he commercial . . . character of a work is 'not conclusive,' . . . but rather a fact to be 'weighed along with other[s] in fair use decisions.' . . . No such evidentiary presumption is available to address . . . whether a transformative use . . . is a fair one.").

54. *Id.* at 579.

55. *See id.* at 578–79, 594.

56. *See id.* at 579.

57. *Id.*

1994 and 2011[58] indicated two important lessons about the test: first, the courts regarded a change in the function *and* purpose of the new work compared to those of the first work to be highly relevant and in most cases dispositive; in other words, a change in the purpose and function of the two works is more highly indicative that a second work is transformative than a change in the content, meaning, and expression of the first work.[59] Even significant alteration of the form, genre, theme, tone, or even the overall meaning of the works will not be found to be fair use if some of the creative, artistic, and expressive virtues of the original works are not replaced or overwhelmed by the expression in the second work.[60] If the creative, artistic, and expressive virtues of the original works still are discernable in the second work and still add value to the secondary work, the use of the original work will be deemed unfair.[61]

Second, courts are to consider transformation of the content, context, and the predominant purpose of the original work to evaluate whether the alleged fair use changes the content, context, or purpose in a manner that furthers the public policies reflected in the first sentence of section 107.[62] Otherwise, you are making an unauthorized exploitation of the creative expression of the work for exactly the same reasons and purposes that the original author or artist created the work, and you are depriving the original author or artist of the derivative works right guaranteed by copyright.[63]

---

58. Michael D. Murray, *What Is Transformative? An Explanatory Synthesis of the Convergence of Transformation and Predominant Purpose in Copyright Fair Use Law*, 11 Chi.-Kent J. Intell. Prop. 260, 273 (2012).

59. *Id.* at 276–80; R. Anthony Reese, *Transformativeness and the Derivative Work Right*, 31 Colum. J.L. & Arts 467, 484–85 (2008).

60. *E.g.*, Salinger v. Colting, 607 F.3d 68, 82–83 (2d Cir. 2010); Gaylord v. United States, 595 F.3d 1364, 1372–73 (Fed. Cir. 2010); Bridgeport Music v. UMG, 585 F.3d 267, 277–78 (6th Cir. 2009); Castle Rock Ent. v. Carol Publ'g. Grp., 150 F.3d 132, 142 (2d Cir. 1998); Dr. Seuss Enters v. Penguin Books USA, Inc., 109 F.3d 1394, 1400 (9th Cir. 1997).

61. We can add to the list in note 60 the recent case *Dr. Seuss Enters.*, *v. ComicMix LLC*, 983 F.3d 443 (9th Cir. 2020) (significant alteration of the genre, theme, and meaning of the original Seuss work, *Oh the Places You'll Go*, was insufficient to find transformative fair use because the *Boldly Go* work did not change the function and purpose of the material it copied; *Boldly Go* replicated "the exact composition, the particular arrangements of visual components, and the swatches of well-known illustrations" of the famous Seuss work for their same artistic and expressive purposes), and *Warhol* itself (Warhol's artistic style and genre changes still allowed the basic artistic function of the original Goldsmith work, to portray Prince, to shine through in Warhol's work).

62. *ComicMix*, 983 F.3d at 455.

63. Murray, *supra* note 58, at 292.

This observation of eleven years ago was nearly exactly repeated in the Court's decision in *Warhol*.[64] The Court stated that although the addition of new expression to an existing work may be relevant to whether a copying use has a sufficiently distinct purpose or character, it is not, without more, dispositive of the first factor, the purpose and character of the work.[65] Indeed, in the *Warhol* case, the specific function and purpose of Goldsmith's photograph matched that of Warhol's work: both had the function and purpose to be portraits of Prince used to depict Prince in magazine stories about Prince.[66] The Court noted that the use of an original work to achieve a purpose that is the same as or highly similar to that of the original work is more likely to substitute for, "supersede the objects" of, or "supplant" the original work.[67] A use that has a distinct purpose is justified because it furthers the goal of copyright, namely, to promote the progress of science and the arts, without diminishing the incentive to create,[68] but a use that shares the purpose of the original work is more likely to provide the public with a substantial substitute for matter protected by the copyright owner's interests in the original work or derivatives of it, which undermines the goal of copyright.[69] The Court concluded that, "an overbroad concept of transformative use, one that includes any further purpose, or any different character, would narrow the copyright owner's exclusive right to create derivative works, . . . [and] the degree of transformation required to make "transformative" use of an original must go beyond that required to qualify as a derivative."[70]

Adding new expressive content that adds new meaning to an earlier work will not be sufficient for fair use if the function and purpose of the two works remains the same, or if artistic content of the original work still shines through and adds value to the new work. The *Warhol* court accepted that Andy Warhol had altered Goldsmith's work with new expression, meaning, and message, and that Warhol's work had a different aesthetic from Goldsmith's work.[71] But many "derivative works, including musical arrangements, film and stage adaptions, sequels, spinoffs, and others that 'recast, transfor[m] or adap[t]' the original, . . . add new expression, meaning or message, or provide new information, new aesthetics, new insights and understandings."[72] And to allow all such adaptations and alterations of original content to be transformative fair uses would "swallow

64. *Warhol*, 143 S. Ct. at 1273.

65. *Id.*

66. *See id.*

67. *See id.* at 1274.

68. *Id.* at 1276.

69. *See id.*

70. *Warhol*, 143 S. Ct. at 1275.

71. *See id.* at 1282.

72. *See id.* (citing 17 U.S.C. § 101, definition of "derivative work").

Figure 6 from the Warhol opinion: Warhol's orange silkscreen portrait of Prince superimposed on Goldsmith's portrait photograph.

the copyright owner's exclusive right to prepare derivative works."[73] Goldsmith's depiction of Prince in her photograph was easily discernable in Warhol's adaptation of the work[74]: Warhol's work added a new aesthetic by its coloration and posterization, which fit the work within the genre of pop art portraiture that Warhol himself invented and popularized. But in terms of transformation, Goldsmith's exact composition, her particular arrangements of visual components—the pose, the "attitude" of Prince's depiction (the exact angle of the head and forward-facing gaze)—were replicated in the Warhol work. And the *Warhol* court made note of the fact that both images were created and used for the same commercial function and purpose: the function and purpose to be portraits of Prince used to depict Prince in magazine stories about Prince.[75] Thus, the *Warhol* court found that Warhol's use of Goldsmith's work did not constitute fair use.[76]

B. Transformative Fair Use in the Context of Nonexpressive Copying and Copy-Reliant Technologies

The concept of nonexpressive copying in the context of computer operations of programming and "training," the functioning of algorithms, and data analysis, refers to the incidental duplication of data and raw source material to carry out a function unrelated to the creation, consumption, or distribution of the expressive elements of the material.[77] The situations covered

73. *Id.*

74. *Id.* at 1271, fig. 6.

75. *Id.* at 1273.

76. *Warhol*, 143 S. Ct. at 1258.

77. On nonexpressive copying generally, *see* James Grimmelmann, *Copyright for Literate Robots*, 101 Iowa L. Rev. 657, 662–63, 665 (2015); James Grimmelmann,

by this concept involve a form of copying that is so far removed from the normal copying protected under the rights granted to the copyright owner in 17 U.S.C. § 106, it does not count as an "act of copying" or it is excused from infringement by the fair use doctrine.[78] Examples of this are incidental or intermediate processing of data from expressive works that requires the works to be downloaded (which in the digital context means a copy of the digital work necessarily was made) or it is copied in the functional process of analysis.[79] The copies might be temporarily stored for the purpose of the process, but not consumed or distributed.[80] This incidental or collateral copying is necessary for the process to be carried out, leading to the description that the process is reliant on this form of copying, and so named copy-reliant technology.[81] As discussed in Section IV below, the creation of AI foundation models used to train generative AI systems uses a form of copy-reliant nonexpressive copying.

The newness of the technology of the current generation of visual generative AI has not allowed courts to weigh in on the specific application of the transformative fair use test in the context of generative AI. But the courts have considered extremely similar and analogous uses in other contexts involving the incidental copying of copyrighted works for a functional and nonexpressive purpose.

Sega Enterprises Ltd. v. Accolade[82] was one of the earliest examples of nonexpressive copying in order to allow a permissible use of noncopyrightable source code. *Sega* involved a conflict between Sega, the maker of the "Genesis" gaming console, and Accolade, a video game company that wanted to create Genesis-compatible games without Sega's license. Accolade reverse-engineered a Genesis console and some Sega games to copy the functional code that enabled compatibility. Accolade's reverse-engineering produced exact copies of Sega's source code, but only the functional code related to the Genesis interface was used in Accolade's games. The Ninth Circuit found that Accolade's "intermediate copying" of Sega games was fair use, because it was needed to access the "functional requirements for Genesis compatibility"—a functional aspect of Sega's games not protected by copyright.[83] The ruling

---

There's No Such Thing as a Computer-Authored Work - and It's a Good Thing, Too, 39 Colum. J.L & Arts 403, 403 (2016); Matthew Sag, *Orphan Works as Grist for the Data Mill*, 27 Berkeley Tech. L.J. 1503, 1503 (2012); Matthew Sag, *Copyright and Copy-Reliant Technology*, 103 Nw. U. L. Rev. 1607, 1624–25, 1637–38 (2009).

78. *See Oracle Am., Inc.*, 141 S. Ct. at 1205 (citing *Campbell*, 510 U.S. at 588; New Era Publ'ns Int'l, ApS v. Carol Publ'g Grp., 904 F.2d 152, 158 (2d Cir. 1990)).

79. *See, e.g., id.*

80. *See Warhol*, 143 S. Ct. at 1271, fig.6.

81. Sag, *supra* note 77, at 1608–11.

82. Sega Enters. v. Accolade, Inc., 977 F.2d 1510 (9th Cir. 1992).

83. *Id.* at 1518, 1526.

established that copying copyrighted works for a nonexpressive purpose was fair use.[84]

The use of search engines that crawl the web to scrape images—i.e., to indiscriminately make exact copies of copyrighted images—was evaluated in two cases: *Kelly v. Arriba Soft Corp.*[85] and *Perfect 10, Inc. v. Amazon.com, Inc.*[86] These cases both involved tech companies that ran image search engines and were sued by image owners for copying their images, reducing them to thumbnail size, storing them on their servers, and showing them to online users through their search engine services.[87] Each service not only made exact duplicates of the copyrighted images, in full color, in the same medium—digital photography and digital imagery—but the services also stored these duplicate images and made the duplicates available—i.e., distributed them—to internet users for their viewing and consumption.[88] And in both cases, these activities were found to be transformative fair uses.

The image search cases differed from *Sega*'s case because the defendants copied and displayed all the expressive visual elements of the plaintiffs' works, unlike Accolade, who only copied functional elements (source code). In *Kelly*, the defendant, Arriba Soft Corp. obtained images through the operation of a "crawler," a computer program that automatically browsed and indexed web pages. When the crawler encountered an image, it would download (i.e., copy) a full-size copy to Arriba's servers, copy it again to reduce it to a thumbnail size, delete the full-size copy, and display (distribute) the thumbnail copy in its search results.[89] Google's more famous "image search" works the same way by using a web crawler to locate image files on the web, copy them by downloading them, copy them again by converting them to a smaller, thumbnail size, and storing the thumbnail copies for display (distribution) in the course of reporting the results of the image search.[90]

The courts rejected the plaintiffs' claims that Arriba and Google infringed their expression by emphasizing the nonexpressive function of the defendants' image search engines.[91] The defendants' image search engines turned photographs into "tool[s]"—mere machines—not means of conveying

84. *Id.* at 1527–28.

85. Kelly v. Arriba Soft Corp., 336 F.3d 811 (9th Cir. 2003).

86. Perfect 10, Inc. v. Amazon.com, Inc., 508 F.3d 1146 (9th Cir. 2007).

87. *Perfect 10*, 508 F.3d at 1154–56; *Kelly*, 336 F.3d at 815.

88. *Perfect 10*, 508 F.3d at 1154–56; *Kelly*, 336 F.3d at 815. The thumbnails displaying exact duplicates of the original work also came with links to the original images which facilitated consumption not only of the thumbnails but the original images, too.

89. *Kelly*, 336 F.3d at 815.

90. *Perfect 10*, 508 F.3d at 1155.

91. *Id.* at 1165; *Kelly*, 336 F.3d at 818.

expression.[92] In terms of transformative fair use, the search engines had adapted the expressive visual content of the original images for a completely new function and purpose: pointers directing a viewer to a source of information and as an archival reference tool, which is completely different from their original function and purpose as aesthetic objects.

A key factor in *Kelly*'s fair use analysis was Arriba's lack of artistic or expressive purpose in reproducing and redistributing Kelly's images. The original function and purpose of Kelly's photographs was as "artistic works meant to inform and to evoke an aesthetic response from the viewer[,]" while Arriba's thumbnails are only instrumental: they are part of a "tool to help index and enhance access to images[.]"[93] Furthermore, Arriba's use was not artistic expression: "The thumbnails do not inhibit artistic creativity because they are not used for illustrative or artistic purposes and therefore do not replace the need for the originals."[94] Google's Image Search is the same: it converts visual images into "pointer[s] directing a user to a source of information" as part of an "electronic reference tool[,]" rather than copying and using them as aesthetic objects for viewing and consumption.[95]

Authors Guild v. Google Inc.[96] rounds out the discussion of nonexpressive fair uses by applying the doctrine to literature. Google partnered with libraries to scan over twenty million books, some copyrighted, some public domain, and many out of print.[97] Google did not first seek permission or a license from the authors and copyright owners of the copyrighted works it included in its scans; it simply scanned them along with the others.[98] Google used these scans to create a corpus of machine-readable texts for its "Google Books" service. Google Books is a public search engine that lets users search for keywords in the Google Books corpus and shows a list of books with those keywords.[99] The search results further include bibliographic data, the frequency of the terms searched for in the text of the books, and, if available, links to buy the books.[100] One of the more popular features of Google Books, and the aspect that tested copyright infringement and fair use to the highest degree, is that Google Books copied and displayed to users

92. *Perfect 10*, 508 F.3d at 1165; *Kelly*, 336 F.3d at 818.

93. *Kelly*, 336 F.3d at 818.

94. *Id.*

95. *Perfect 10*, 508 F.3d at 1165.

96. Authors Guild v. Google Inc., 804 F.3d 202 (2d Cir. 2015).

97. *Id.* at 208.

98. *Id.*

99. *Id.* at 208–09.

100. *Id.* at 209.

all or part of a book's text.[101] Google limited its service by only showing the full text of public domain books and books authorized by publishers and copyright owners for full-text display.[102] But Google Books also featured the "Snippet View" that showed keywords and phrases in a book and "a few snippets–a few sentences to display [a] search term in context."[103] This last function copied and displayed (distributed) the expressive content of the original works without the permission or license of the copyright owners.

The *Author's Guild* opinion was written by none other than Judge Pierre Leval, the spiritual father of the transformative use test, and the opinion found Google's unauthorized and unlicensed copying of the expressive text of the works for analysis of the text and for display and distribution of portions of the text was a transformative fair use.[104] Even though the "snippet view" copied and displayed the expressive text that surrounds a search term, this copying and distribution of the text still supports Google's transformative function and purpose by showing how a term is used in a book without exposing enough of the original author's expression to "harm the author's copyright interests[.]"[105]

The *Author's Guild* court held that Google's function and purpose matched that of the libraries that facilitated the scanning (copying) of entire books in the earlier *HathiTrust* case.[106] The *Authors Guild v. Google* court recognized that the libraries in *HathiTrust* and Google Books had downloaded and stored complete digital copies of entire books, but it further noted that such copying was essential to permit searchers to identify and locate the books in which words or phrases of interest to them appeared.[107] The new function and purpose for this copying and storage was to serve the interests of education, research, archiving, and historical preservation that are supported in the preamble of the copyright fair use provision.[108] The court concluded "that the creation of a full-text searchable database is a quintessentially transformative use . . . [as] the result of a word search is different in purpose, character, expression, meaning, and message from the page (and the book) from which it is drawn."[109]

---

101. *Id.*

102. *Authors Guild*, 804 F.3d at 209.

103. *Id.*

104. *Id.* at 216–17.

105. *Id.* at 218.

106. Authors Guild, Inc. v. HathiTrust, 755 F.3d 87 (2d Cir. 2014).

107. *Authors Guild*, 804 F.3d at 217 (citing *HathiTrust*, 755 F.3d at 97).

108. *See id.* at 216–18 (citing A.V. ex rel. Vanderhye v. iParadigms, LLC, 562 F.3d 630, 639–40 (4th Cir.2009); *Perfect 10*, 508 F.3d at 1165; *Kelly*, 336 F.3d at 819; *HathiTrust*, 755 F.3d at 97); *see also* 17 U.S.C. § 107.

109. *Authors Guild*, 804 F.3d at 217.

The lessons of these cases are that nonexpressive copying to facilitate a machine function may not be an act of copying that results in an infringing work, and it is very likely that a copyright infringement analysis would fail at this stage because there may not be a copy of a work to compare side-by-side with the original work.[110] Even if there were to be an interim or intermediate copy that could be identified and that does contain expressive content that could potentially have been adapted from the allegedly infringed image, the infringement action would fail if that copy was made to facilitate a completely new function and purpose compared to those of the original work.[111] The reduction of photos and images to numeric data so that this data can train the machine learning of a generative AI foundation model is much the same as changing the audio visual expression of a video game into source code or changing photos and images to data points in a search engine as seen in *Sega, Kelly*, and *Perfect 10*.[112] To the extent that the transformation of images into machine readable numeric data is found to be a form of copying, it is a machine function that allows the users of the machine to express themselves through the generation of new and original images.[113] The machine's function is copy-reliant on learning what images of various kinds look like by training on hundreds of millions or billions of images, but the copying of image data serves a function and purpose of building machine systems that enable new artists to create original artistic expression. The system creator's copying is completely different from the function and purpose of the original works which was an aesthetic purpose in the display and enjoyment of the works while the system creator only wants the image data embedded on vectors in a nonexpressive numeric form to be available for the diffusion process to occur for the generation of new, original artworks, not copies of works whose data was part of the training dataset (*see* section IV below). This new function and purpose of the machine system is completely supported by the Copyright Act's primary function and purpose which is to promote the progress of science and the arts by encouraging the production of original expressive works.[114]

IV. INFRINGEMENT OR FAIR USE COPYING BY THE DEVELOPERS OF TRAINING DATASETS THAT ARE LATER USED BY VISUAL GENERATIVE AI SYSTEMS

Many of the participants in the current debate on visual generative AI systems have latched onto the idea that generative AI systems have

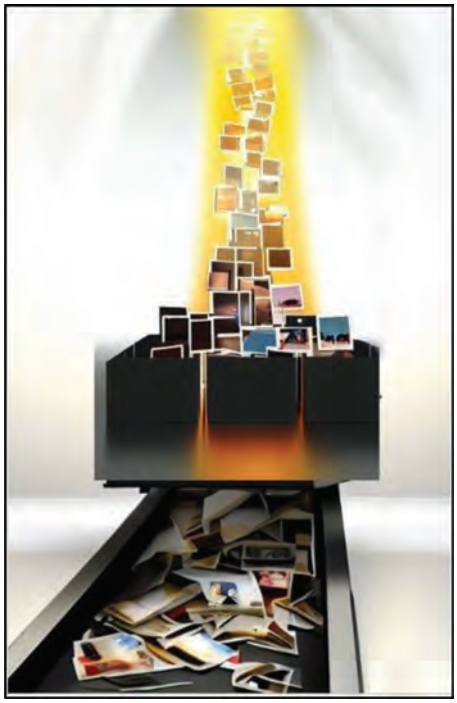
110. *See* Sag, *supra* note 77, at 1617–18.

111. *Id.*

112. *Kelly*, 336 F.3d at 818; *Perfect 10*, 508 F.3d at 1165; *Sega Enters*, 977 F.2d at 1510.

113. *E.g.*, DALL-E 2, supra note 3; Stable Diffusion, supra note 4; Midjourney, *supra* note 5.

114. *Warhol*, 143 S. Ct. at 1276; *Authors Guild*, 804 F.3d at 214; *Campbell*, 510 U.S. at 579.



been trained on datasets and foundation models that contained actual copyrighted image files, .jpgs, .gifs, .png files and the like, scraped from the internet, that *somehow* the dataset or foundation model must have made and stored copies of these works, and *somehow* the generative AI system further selected and copied individual images out of that dataset, and *somehow* the system copied and incorporated significant copyrightable parts of individual images into the final generated images that are offered to the end-user. This is magical thinking.[115] AIs are complex technology, but they are not magic. A visual generative AI is not an impenetrable magic box that takes in whole images and spits out duplicate images (see image at left).[116]

The connected chain of assumptions described above is wrong. The actual steps of the AI image generation process involve building a training dataset, using the dataset to create a foundation model, using the foundation model to supply data to the generative AI service, and an end-user providing design instructions in the form of prompts to start the process of generating images according to the end-user's design.

1. Erroneous assumptions about scraping images in the creation of a dataset

First, using the example of the claims in the complaint filed by three artists as class representatives of all artists whose works were used in the datasets used to train Stable Diffusion, Midjourney, and Deviantart's AI system in *Andersen v. Stability AI*,[117] the artist plaintiffs alleged the following about the defendants' use of their artworks (excerpted below, starting with ¶ 2 of the complaint):

115. On magical thinking generally, *see* sources cited *supra* notes 14–15.

116. Image by Michael D. Murray, using Bing Image Creator (powered by DALL-E), *Magic Box* (2023).

117. Andersen v. Stability AI Ltd., No. 3:23-cv-201 (N.D. filed Cal. Jan. 13, 2023).

2. Stability downloaded or otherwise acquired copies of billions of copyrighted images without permission to create Stable Diffusion, including Plaintiffs'. These images are defined below as "Training Images."
3. By training Stable Diffusion on the Training Images, Stability caused those images to be stored at and incorporated into Stable Diffusion as compressed copies. Stability made them without the consent of the artists and without compensating any of those artists.
4. When used to produce images from prompts by its users, Stable Diffusion uses the Training Images to produce seemingly new images through a mathematical software process. These "new" images are based entirely on the Training Images and are derivative works of the particular images Stable Diffusion draws from when assembling a given output. Ultimately, it is merely a complex collage tool.

* * *

6. All AI Image Products operate in substantially the same way and store and incorporate countless copyrighted images as Training Images.[118]

Each of these allegations in the complaint is fundamentally wrong.

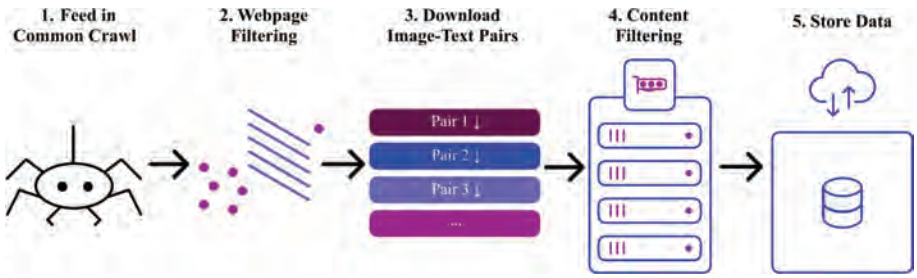
First, Stability AI did not create a training set of images or image data scraped from the internet to create its foundation model, nor did any of the other defendants in the *Andersen* lawsuit. The party that did these actions is LAION-5B, assisted by the CLIP (Contrastive Language–Image Pre-training) technology owned by OpenAI[119] that was used to create the foundation model of the same name, CLIP, that is incorporated in OpenAI's DALL-E 2 system (see section V below). Stable Diffusion, a visual generative AI system owned by Stability AI, trains on the LAION-5B dataset and is a donor to or investor in LAION, but Stability AI did not play a role in gathering image data for the LAION-5B dataset.

Second, LAION did not copy images found on the internet in the manner that the copyright infringement law discussed above contemplates. LAION started the process of creating a training dataset by using image data from Common Crawl, a public web archive.[120] Since 2008, the Common Crawl organization has been crawling the web gathering image

118. Complaint at ¶¶ 2-4, 6, Andersen v. Stability AI Ltd., No. 3:23-cv-201 (N.D. Cal. Jan. 13, 2023).

119. *CLIP: Connecting text and images*, OpenAI (Jan. 5, 2021), https://openai.com/research/clip [https://perma.cc/YEB2-XSC5]; *see* Alec Radford & Jong Wook Kim, *Learning Transferable Visual Models From Natural Language Supervision*, arXiv (Feb. 26, 2021), https://arxiv.org/pdf/2103.00020.pdf [https://perma.cc/9K6T-UG5X].

120. Common Crawl, https://commoncrawl.org/ [https://perma.cc/7ML5-7CH4].



and text information from, in recent years, the approximately 3 billion websites on the World Wide Web that contain images. Common Crawl does not capture or download actual images, it collects "raw web page data, extracted metadata, and text extractions."[121] Common Crawl stores the data in WAT files—Web Archive Transformation files[122]—that use WebAssembly text format to store web page data (including data about images) in an intermediate form that can be reassembled later to binary code so that a web browser or other tools can read it and further process it. [123] The WAT file does not contain actual .jpg or .png files or any other digital format image files. The data Common Crawl stores and publishes in WAT files gives enough information to allow an end-user such as LAION to evaluate the metadata and alt-txt data of images on websites without copying or downloading the actual images.

The LAION Data Assembly Pipeline has the following steps (*see* diagram below):[124]

LAION starts in step 1 with the raw web page data, extracted metadata, and text extractions from Common Crawl.[125] In step 2, LAION filters the raw web page data to select web pages that have images with associated HTML

121. Common Crawl, https://commoncrawl.org/the-data/get-started/ [https://perma.cc/FWS8-KASJ].

122. *Details of extension .wat*, FileDesc.com (Jun. 7, 2023), https://www.filedesc.com/en/file/wat (last visited Sept. 14, 2023) [https://perma.cc/3QGQ-ZX8L]; *WAT file format description*, DataTypes.net (Jun. 7, 2023), https://datatypes.net/open-wat-files [https://perma.cc/5HKF-AD8K].

123. Mozilla, *Converting WebAssembly text format to Wasm*, MDN Web Docs (Jun. 7, 2023), https://developer.mozilla.org/en-US/docs/WebAssembly/Text_format_to_wasm [https://perma.cc/SH2B-AMV8].

124. Christoph Schuhmann et al., *LAION-5B: An open large-scale dataset for training next generation image-text models*, aiXrv, fig. 2 (Oct. 16, 2022), https://arxiv.org/pdf/2210.08402.pdf [https://perma.cc/S576-HK7V].

125. *See id.* at 5.

image (IMG) tags containing the alt-text for the image. If a page does not have images with HTML IMG tags, then its images, if any, will not be included in the downloading. The web page data is analyzed for its language, e.g., English, a different language, or no detectable language.[126] In this step of the process no image files are copied; LAION collects the URLs for the images, 500 million at a time, for processing in step 3.[127]

Step 3 involves downloading WAT files with coded information regarding image and text pairs on the internet using the source information (URLs) from Common Crawl.[128] This step still does not involve "copying" of image files because it is the textual WebAssembly WAT file information that is being downloaded, not digital graphic files from the internet.[129]

Step 4 is the post-processing step where the image URLs from downloaded web page data are filtered. LAION applied several filters including the CLIP (Contrastive Language–Image Pre-training) model to remove data from low-quality image-text pairs, meaning images whose cosine similarity with the text in the image's alt text did not meet CLIP's threshold for semantic similarity between the image and the text.[130] LAION also weeded out data from images on websites whose alt text had fewer than five characters, typically meaning the alt text was a person's name. [131] It eliminated data from duplicate images with the same URLs and applied Not Safe for Work (NSFW) and toxicity detectors to filter out pairs that contained sexualized "adult" images or offensive content, and LAION also purportedly removed pairs that had watermarks, corrupted images, or empty texts. [132] After filtering, LAION had identified data for 5.85 billion image-text pairs on the World Wide Web that could be used for assembling the dataset.[133] Although that number seems staggering, it represents only 10% of the image file data that was included in the Common Crawl WAT files before filtering.[134]

126. LAION uses Google's Compact Language Detector 3 (CLD3) for this language detection process. *Id.; see* Jeroen Ooms, *CLD3: Google's Compact Language Detector 3* (2022), https://rdrr.io/cran/cld3/ [https://perma.cc/8RSC-NW8F].

127. *See* Schuhmann, *supra* note 124, at 5.

128. *Id.*

129. *See id.*; Romain Beaumont, *LAION-5B: A New Era of Open Large-Scale Multi-Modal Datasets*, LAION (Mar. 31, 2022), https://laion.ai/blog/laion-5b/ [https://perma.cc/PS7U-N9S5].

130. *See* Schuhmann, *supra* note 124, at 5; Beaumont, *supra* note 129.

131. *See* Schuhmann, *supra* note 124, at 5; Beaumont, *supra* note 129.

132. *See* Schuhmann, *supra* note 124, at 5; Beaumont, *supra* note 129.

133. *See* Schuhmann, *supra* note 124, at 5; Beaumont, *supra* note 129.

134. *See* Schuhmann, *supra* note 124, at 5; Beaumont, *supra* note 129.

Step 5, the last step, involves storage of the filtered web data containing image URLs. Once again, LAION does not store actual .jpg or .png files or any other digital format image files. It only stores the filtered curated web data taken from the Common Crawl WAT files.[135]

2. Use of the web data in the dataset for image generation

What happens next when a dataset such as LAION-5B is completed is that a client such as Stability AI uses the data in the set as its foundation model for its image generation system. The steps of that process are discussed in section V below.

OpenAI's CLIP (Contrastive Language–Image Pre-training) model can also function as a foundation model when it curates and stores image data from the World Wide Web.[136] As discussed in section V below, OpenAI's current DALL-E 2 model uses CLIP directly as a sorter of image-text pairs with proper cosine semantic similarity that are then used by DALL-E 2's text encoder and image encoder to create the working numeric representations of images in latent space, which enable the DALL-E 2 diffusion process to generate representations that match up with the data points from the text prompt.[137]

3. The developer of a visual generative AI dataset does not copy or store any image files.

Because an act of copying is required for any suit for copyright infringement, it is important to understand that the creators of generative AI datasets—the so-called sources of image files that litigants claim have been copied—have not copied or stored any image files. The narrative allegations that their images were downloaded, copied, and used to make infringing copies or derivative images are simply incorrect. No digital image files are downloaded, none are stored, none are copied, combined, or collaged to make new images.

This observation is not a semantic trick. The technologies involved do not work with actual image files. Even if in the course of parsing the claims of persons who are annoyed by current visual generative AI's ability to generate images that resemble these person's artworks, and the claimants point out that their image was one of the 5.85 billion images whose image URL and metadata were included in the training dataset, the resemblance if any between the two images is not because the training dataset copied and stored and made derivative works of some of the 5.85 billion image files. Instead, as discussed in the next section, the visual generative AI used the data from 5.85 billion images still on the Web to learn what images look like so that the

135. *See* Schuhmann, *supra* note 124, at 5; Beaumont, *supra* note 129.

136. OpenAI, *supra* note 119.

137. *See infra* Section V.

system can generate brand new images in the diffusion process and condition them until they produce final images that are responsive to the requirements communicated by the end-users' text prompts.

IV. INFRINGEMENT OR FAIR USE COPYING BY THE CREATORS OF VISUAL GENERATIVE AI SYSTEMS

The next most popular targets for allegations of copyright infringement are the creators of the visual generative AI systems. The "magic box" thinking here is that copyrighted images were *somehow* acquired by these developers and *somehow* copied in whole or in part to create, compile, or collage derivative works that *somehow* contain substantial and material original, copyrightable, expressive parts of the claimant's images. This magical thinking is incorrect.

A contemporary state-of-the-art AI is a "brain scale"[138] "neural network"[139] learning machine.[140] In terms of what "brain scale" means, the state of the art language processing system in November 2022, OpenAI's GPT-3 (Generative

138. "Brain scale" artificial intelligence is AI run on processor systems capable of training on and processing connections between over 100 trillion parameters of information, *Cerebras Systems Announces World's First Brain-Scale Artificial Intelligence Solution*, BUSINESSWIRE (Aug. 24, 2021, 03:00 PM), https://www.businesswire.com/news/home/20210824005644/en/Cerebras-Systems-Announces-World%E2%80%99s-First-Brain-Scale-Artificial-Intelligence-Solution. [https://perma.cc/LSR5-XLPS]; Matthew Griffin, *Scientists built the largest AI supercomputer yet to create brain scale AI*, FANATICAL FUTURIST (Nov. 27, 2022), https://www.fanaticalfuturist.com/2022/11/scientists-built-the-largest-ai-supercomputer-yet-to-create-brain-scale-ai [https://perma.cc/LVD9-VJLM]. The neuronal processing capacity of the human brain is understood to be capable of processing at least 100 trillion parameters through the brain's synapses. *See Cerebras Systems Announces, supra*; Jeff Dean & Andrew Ng, *Using large-scale brain simulations for machine learning and A.I.*, GOOGLE: THE KEYWORD (Jun. 26, 2012),[https://perma.cc/J8C7-UN2K]; Rebecca Boyle, *Simulated Brain Ramps Up To Include 100 Trillion Synapses*, POPULAR SCIENCE (Nov. 20, 2012, 1:02 AM), https://www.popsci.com/technology/article/2012-11/worlds-fastest-supercomputer-simulates-100-trillion-synapses-many-human-brain/ [https://perma.cc/TN78-9UMG].

139. Contemporary AI models attempt to simulate the learning, processing, recall, and "thinking" processes of the human brain, which is described as a neural network. The human brain is calculated to have synapses capable of processing 100 trillion parameters of information. *See* sources cited *supra* note 138.

140. *See, e.g.*, Sara Brown, *Machine learning, explained*, MIT SLOAN SCH. OF MGMT. (Apr. 21, 2021), https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained [https://perma.cc/K5NT-5Q8C].

Pretrained Transformer-3), processes 175 billion parameters.[141] As of March 2023, OpenAI has released GPT-4 and it is a magnitude larger and more powerful than GPT-3, but the exact number of parameters it processes is not disclosed by its creator, OpenAI. Commentators speculate that the number is over 1 trillion parameters,[142] others speculate that it could be as much as 100 trillion parameters,[143] a claim which OpenAI president Sam Altman denied,[144] but one commentator boldly pronounced that GPT-4 has 170 trillion parameters (without citing a source for this pronouncement).[145] If GPT-4's parameters are indeed over 100 trillion, this number would put GPT-4 in the "brain scale" category. However, the latest rumor in late June 2023 is that GPT-4 is not one huge monolithic large language model but rather a connected "Mixture of Experts" model using eight 220 billion parameter sub-models as the "experts."[146] If this rumor turns out to be correct, GPT-4 technically may not be a singular brain scale neural network, but together it is a model working with 1.76 trillion parameters.

The term "learning machine" is one of the broadest expressions of what artificial intelligence is: a machine that can learn new information and use the

141. "Parameters" refers to connections processed over a neural network, which might also be referred to the number of values that can independently be changed in the process of learning over the neural network. Kizito Nyuytiymbiy, *Parameters and Hyperparameters in Machine Learning and Deep Learning*, Towards Data Science (Dec. 30, 2020), https://towardsdatascience.com/parameters-and-hyperparameters-aa609601a9ac# [https://perma.cc/W26V-HSSA]. OpenAI's GPT-3 is capable of processing 175 billion parameters. Kumar Mehta, *Using The Most Powerful AI Just Got Simpler*, Forbes (Dec. 7, 2022, 1:12pm EST), https://www.forbes.com/sites/kmehta/2022/12/07/using-the-most-powerful-ai-just-got-simpler/?sh=4fce267d6fb9 [https://perma.cc/Z6Q9-43P2].

142. Reed Albergotti, *The secret history of Elon Musk, Sam Altman, and OpenAI*, Semafor (Mar. 24, 2023), https://www.semafor.com/article/03/24/2023/the-secret-history-of-elon-musk-sam-altman-and-openai [https://perma.cc/TN6J-3CNQ.].

143. *GPT-4 Parameters—Is it 100 Trillion?*, ML Learning (Apr. 22, 2023), https://www.mlyearning.org/gpt-4-parameters/ [https://perma.cc/8X53-D8CH]; Ben Lutkevich, *GPT-4*, WhatIs.com (last visited May 9, 2023), https://www.techtarget.com/whatis/definition/GPT-4 [https://perma.cc/GFT6-RBNC].

144. Albergotti, *supra* note 142.

145. Mohammed Lubbad, *The Ultimate Guide to GPT-4 Parameters: Everything You Need to Know about NLP's Game-Changer*, Medium (Mar. 19, 2023), https://medium.com/@mlubbad/the-ultimate-guide-to-gpt-4-parameters-everything-you-need-to-know-about-nlps-game-changer-109b8767855a [https://perma.cc/R84D-NHVQ].

146. *E.g.*, Mandar Karhade, *GPT-4: 8 Models in One; The Secret is Out*, Towards AI (Jun. 24, 2023), https://pub.towardsai.net/gpt-4-8-models-in-one-the-secret-is-out-e3d16fd1eee0 [https://perma.cc/UPJ6-VCNU].

knowledge it learns to perform new tasks.[147] Machine learning is a subfield of artificial intelligence that gives computers (AI systems) the ability to learn without explicitly being programmed.[148] The goal of machine learning is to create AI that can perform complex tasks in a way that is similar to how humans solve problems.[149] Comparatively simple forms of learning algorithms have been around for decades, taking the form of voice-to-text mechanisms on your smartphone and messaging applications that learn to adapt to your speaking input and improve the fidelity of their text output, to GPS navigation systems that learn to take into account events and conditions that slow down travel, from accidents on the side of the road, to speed traps, to construction zones.[150]

Visual generative AI designers and trainers have exposed the brain scale or near brain scale machine learning AI to data from hundreds of millions or billions of images with associated text labels (image-text pairs),[151] and the AI is trained to encode the textual information in the verbal prompt and compare and contrast the encoded information with the encoded image data and text data from the image-text pairs it has been trained on.[152] The process has the following steps[153]:

147. Brown, *supra* note 140.

148. *Id.*

149. *Id.*

150. *See* Mikiko Bazeley, *An Easy Introduction to Speech AI*, NVIDIA Developer (Jun. 23, 2022), https://developer.nvidia.com/blog/an-easy-introduction-to-speech-ai/ [https://perma.cc/54C9-KMSR]; Jeffrey L. Duffany, *Artificial intelligence in GPS navigation systems*, IEEE Explore (Oct. 25, 2010), https://ieeexplore.ieee.org/document/5608862 [https://perma.cc/LYE3-ZZH8].

151. A contemporary AI training set, such as LAION-5B (Large-scale Artificial Intelligence Open Network), has compiled 5.85 billion image-text pairs that can be used to train the "Foundation Model" of an AI to allow the AI to respond to encoded textual prompts and decode the information for the rendering of responsive images. Beaumont, *supra* note 129.

152. The CLIP process - Contrastive Language-Image Pretraining – is used directly by OpenAI (DALL-E 2) and indirectly by Stability AI (Stable Diffusion) because Stable Diffusion is trained on a subset of the CLIP curated LAION-5B dataset. *See CLIP: Connecting Text and Images*, *supra* note 119.; Maximilian Schreiner, *New CLIP model aims to make Stable Diffusion even better*, The Decoder (Sep. 18, 2022), https://the-decoder.com/new-clip-model-aims-to-make-stable-diffusion-even-better/ [https://perma.cc/5B7Z-R5KN].

153. Michael D. Murray, *Diagrams of DALL-E 2 and Stable Diffusion Image Generation Processes* (2023), based on the sources in the corresponding footnotes. I am not discussing Midjourney here because its creators have not been as transparent and forthcoming about the specific operation of this visual generative AI system as have the creators of DALL-E 2 and Stable Diffusion.

Diagram 1 – DALL-E 2 Process

<table>
<tr>
<td>(1) DALL-E 2 is based on a variational autoencoder (VAE) model, which is a type of generative model

(4) Latent Space
Vectors represent overlay of encoded text data and image data.

that learns to encode and decode data. DALL-E 2 uses a large transformer model to encode data from text and image pairs into a shared latent space, and then decodes the data into new images responsive to the text prompt using a diffusion process.[154]</td>
<td>(2) Latent space is a critical concept that differentiates the DALL-E 2-type models from models that attempted to generate and combine images and condition them through diffusion at a pixel by pixel level. Latent space is a purely numeric graphical construct in which image and text data are embedded as points on overlapping vectors. The VAE model can apply the diffusion process to the numeric data on these vectors in latent space without having to render and rerender actual pixel-level images. The image decoder can then use the output of the diffusion process that occurred in latent space to render actual images (with pixels) responsive to the text prompt.[155]</td>
</tr>
</table>

154. Aditya Ramesh et al., *Hierarchical Text-Conditional Image Generation with CLIP Latents*, OpenAI (Apr. 13, 2022), https://cdn.openai.com/papers/dall-e-2.pdf [https://perma.cc/CD4C-TVBG].

155. Robin Rombach et al., *High-Resolution Image Synthesis with Latent Diffusion Models* 3–4, arXiv (Apr. 13, 2022), https://arxiv.org/pdf/2112.10752.pdf [https://perma.cc/Z6Q9-43P2]; Ekin Tiu, *Understanding Latent Space in Machine Learning*, Towards Data Science (Feb. 4, 2020), https://towardsdatascience.com/understanding-latent-space-in-machine-learning-de5a7c687d8d; Joseph Rocca, *Understanding Variational Autoencoders (VAEs)*, Towards Data Science (Sep 23, 2019), https://towardsdatascience.com/understanding-variational-autoencoders-vaes-f70510919f73. [https://perma.cc/CZV9-2B85].

(3) **Text Encoder:** The prompt text – "cat playing poker" – is encoded into numeric data in latent space. Encoded text data is mapped to paired text-image vector representations in latent space.[156]

(4) **Latent Space**
Vectors represent overlay of encoded text data and image data.

157

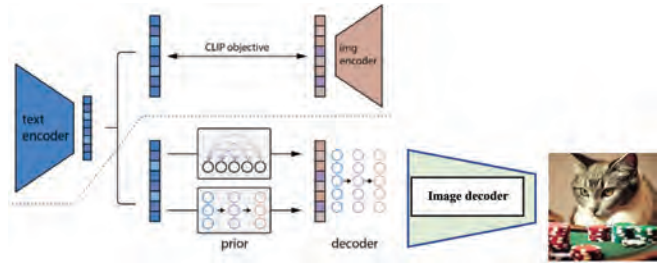
(5) **Image Encoder:** draws from CLIP curated image data (image data sorted by cosine similarity) in the training set and encodes relevant responsive image data into numeric data. Encoded image data is mapped to paired text-image vector representations in latent space.[158]

156. Chitwan Saharia et al., *Photorealistic Text-to-Image Diffusion Models with Deep Language Understanding* 1, 3-4, arXiv (May 23, 2022), https://arxiv.org/pdf/2205.11487.pdf [https://perma.cc/K89T-UJ2G]; *see* sources cited *supra* note 154.

157. *See* sources cited, *supra* note 155; Michael D. Murray, *Diagram of overlay of encoded text data and image data* (2023).

158. Zihao Wang et al., *CLIP-GEN: Language-Free Training of a Text-to-Image Generator with CLIP* 2, 3, arXiv (Mar. 1, 2022), https://arxiv.org/pdf/2203.00386.pdf [https://perma.cc/FY56-9WZY].

(6) The **Text Encoder** combined with the image encoder are used to define the **CLIP objective** (what types of image and text data will be useful in responding to this prompt) which is then used in a **Prior** transformer model of image-text data that can be processed by diffusion.[159]



(7) **Latent Diffusion Process -** A latent diffusion process takes a random image that is obscured by noise and gradually refines it in a series of denoising steps moving toward a target image. The target is defined and conditioned by the image and text data of the prior model. This process occurs in the latent space level to produce responsive image and text numeric data that can be decoded into actual images.[160]

(8) The **Image Decoder** uses the data from the diffusion conditioned by the CLIP objective manifested in the prior model to generate one or more images responsive to the initial text prompt entered into the Text Encoder.[161]

159. Ryan O'Connor, *How DALL-E 2 Actually Works*, AssemblyAI (Apr. 19, 2022), https://www.assemblyai.com/blog/how-dall-e-2-actually-works/ [https://perma.cc/2EAF-YS7N]; *see also* Salim Oyinlola, *DALL-E: Inside the Artificial Intelligence program that creates images from textual descriptions*, Paperspace (Jul. 2022), https://blog.paperspace.com/dall-e-image-generator/ [https://perma.cc/573L-2VNG].

160. Rombach, *supra* note 155; Oyinlola, *supra* note 159; Vaclav Kosar, *OpenAI's DALL-E 2 and DALL-E 1 Explained*, vaclavkosar.com (Apr. 22, 2022), https://vaclavkosar.com/ml/openai-dall-e-2-and-dall-e-1 [https://perma.cc/3BHY-PLCG].

161. *See* sources cited, *supra* note 160.

Diagram 2 –Stable Diffusion Process

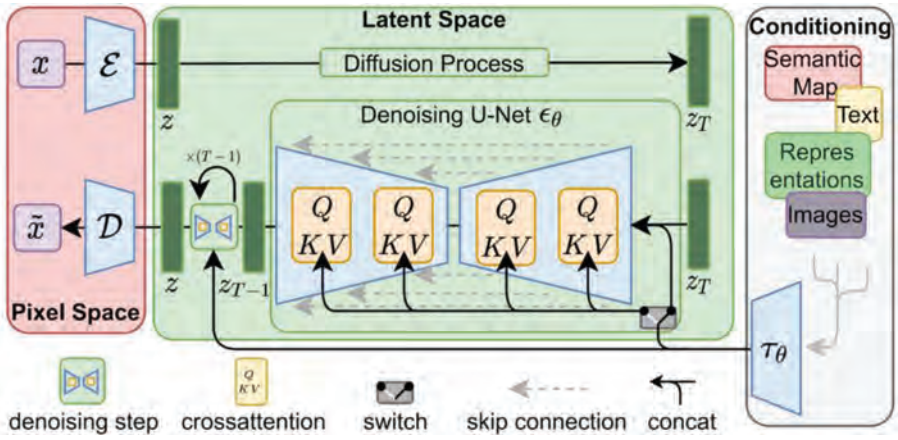
(1) **Latent Diffusion Model:** There are more similarities between Stable Diffusion and DALL-E 2 than differences. Both are text-to-image models that can render images from natural language descriptions. Both use a large-scale latent space diffusion model. Both systems use a variational autoencoder and an image decoder.[162]	(2) **CLIP:** Stable Diffusion and DALL-E 2 both make use of CLIP to guide the image generation process and measure the semantic similarity between text and image. DALL-E 2 uses CLIP directly in its process, while Stable Diffusion uses CLIP indirectly by training on a subset of a large-scale image dataset, LAION-5B, that has itself been curated by the CLIP process to make sure it only contains training data with valid image-text pairings. Stable Diffusion v. 2 also uses CLIP scores of cosine similarity in its diffusion conditioning process.[163]	(3) **Latent Space:** Stable Diffusion trains an autoencoder to map the pixel space into a latent space of vectors where imperceptible details are abstracted away, and the image data is compressed down to a smaller vector, and then performs conditional diffusion in this latent space to account for the text prompt and other conditions.[164]

162. Rombach, *supra* note 155; Louis Bouchard, *How Stable Diffusion works? Latent Diffusion Models Explained*, LOUISBOUCHARD.COM (Aug 27, 2022), https://www.louisbouchard.ai/latent-diffusion-models/[https://perma.cc/BP5Y-NK7A].

163. Beaumont, supra note 151; JIN, *Stable Diffusion: The Mathematical Technique Revolutionizing Photo Generation*, MEDIUM (May 17, 2023), https://jinlow.medium.com/stable-diffusion-the-mathematical-technique-revolutionizing-photo-generation-19199a7ee57d [https://perma.cc/8L3X-VHWA].

164. Marc Päpper, *Everything you need to know about stable diffusion*, PAEPPER.COM (Nov. 16, 2022), https://www.paepper.com/blog/posts/everything-you-need-to-know-about-stable-diffusion/ [https://perma.cc/BZ3Y-AH96]; Sergios Karagiannakos & Nikolas Adaloglou, *How diffusion models work: the math from scratch*, AI SUMMER (Sep. 29, 2022), https://theaisummer.com/diffusion-models/ [https://perma.cc/434S-G68R].

(4)



165

(5) **Conditioning:** Conditioning, also referred to as Guided Diffusion, refers to manipulating the generated samples to incorporate image/text embeddings (vectors) into the diffusion in order to "guide" the generation. The guidance refers to conditioning a prior data distribution with a condition, i.e. the class label or an image/text embedding.[166]	(6) **Cross-attention and Denoising**: Diffusion as a general concept refers to adding Gaussian noise to compressed image data in latent space. This is followed by the denoising process depicted in the center of the diagram where a series of denoising steps in a U-Net configuration are guided by a noise predictor using cross-attention. The output of the text transformer is used multiple times by the noise predictor through the cross-attention mechanism so that the text prompt guides the outcome of the image.[167]

165. This is the primary diagram displayed in Rombach, *supra* note 155, to explain the latent diffusion model later embodied in Stable Diffusion.

166. Karagiannakos & Adaloglou, *supra* note 164.

167. Mario Namtao & Shianti Larcher, *Paper Explained — High-Resolution Image Synthesis with Latent Diffusion Models*, Towards Data Science (Mar. 30, 2023), https://towardsdatascience.com/paper-explained-high-resolution-image-synthesis-with-latent-diffusion-models-f372f7636d42; Andrew, *How does Stable Diffusion work?*, Stable Diffusion Art (Mar. 29, 2023), https://stable-diffusion-art.com/how-stable-diffusion-work [https://perma.cc/NK9G-KR5J].

With regard to the topic of fair use and generative AI, there are two important take-aways from these diagrams and the processes they depict:

(A) Visual generative AI systems do not copy individual images that were used to create the training set.

(B) Visual generative AI systems are designed *not* to copy any images.

A. Visual Generative AI Systems Do Not Copy Individual Images that Were Used To Create the Training Set.

The above diagrams show multi-stage models that rely on a CLIP (Contrastive Language-Image Pretraining) curated set of embedded image data associated with embedded data from text captions, which together exist as points on vectors in the latent space of the model. DALL-E 2 has incorporated the CLIP curation process to assemble numeric information from potentially relevant images (images sorted by cosine similarity) into the main image generation process by using the CLIP method in the creation of the prior model that is then processed by diffusion and denoising.[168] Stable Diffusion makes indirect use of the CLIP process because its LAION-5B training set of images with data from embedded text/image tags is trained, has itself been curated by the CLIP process.[169]

It is true that when the training sets—which are used to train foundation models for visual generative AI systems[170]—were first compiled, the composition of the data in the training dataset drew information from images whose image data, metadata, and alt-text information had been "scraped" from the web by a web crawler,[171] and the visual generative AI systems learned what images look like from the training dataset. But DALL-E 2 or Stable Diffusion do not search the data from the training set for information from an individual image and they do not search the internet for any individual .jpg, .png, or .gif image that matches the text prompt terms. If such an attempt were made, it might properly be called a fool's errand because of the variety and complexity

168. O'Connor, *supra* note 159; Oyinlola, *supra* note 159.

169. Beaumont, *supra* note 163; JIN, *supra* note 163.

170. Rick Merritt, *What Are Foundation Models?*, NVIDIA (Mar. 13, 2023), https://blogs.nvidia.com/blog/2023/03/13/what-are-foundation-models/ [https://perma.cc/4DMQ-U3UA]; Mike Murphy, *What are foundation models?*, IBM (May 9, 2022), https://research.ibm.com/blog/what-are-foundation-models [https://perma.cc/LB7P-KKTX].

171. Jay Alammar, *The Illustrated Stable Diffusion* (Nov. 2022), https://jalammar.github.io/illustrated-stable-diffusion/[https://perma.cc/4HBN-VMVE]; See Also Gulbahar Karatas, Comprehensive Guide to Web Crawling vs Web Scraping in 2023, AI Multiple (Sep. 11, 2023), https://research.aimultiple.com/web-crawling-vs-web-scraping/.

of text prompts that can be used, ranging from the inane[172] to the sublime.[173] An image generating AI does not work like Google's image search which is expressly designed to search for and produce a thumbnail and web address for preexisting images that match the textual search prompt given to Google.[174] Instead, each visual generative AI program draws on its learning of what images need to look like to match up with textual words and phrases that are used to describe images on the internet and in databases—e.g., award winning, studio quality, high resolution photograph, rendered with Octane, etc.[175] And the system proceeds to work on rendering an image that fulfills the prompt requirements.

B. Visual Generative Ai Systems are Designed Not to Copy Any Images.

In DALL-E 2 and Stable Diffusion, a text encoder and image encoder work to form a model of latent space numeric information that can be processed by diffusion and denoising by linking textual and visual semantics to render responsive images.[176] In the case of DALL-E 2, the target of the process is defined and conditioned by the image and text data of the prior model,[177] and in the case of Stable Diffusion, by the guided diffusion process and the cross-attention of the denoising process.[178] But the starting point of the diffusion and denoising process is not the selection of an individual image to diffuse and denoise. Rather, in DALL-E 2, the system starts with a random image embedding (vectors in the latent space) in conjunction with the text input that is used to create the prior from which random noise images are produced

172. Inane prompt example: "orange tabby cat wearing Evel Knievel suit riding a motorcycle." OpenAI, https://labs.openai.com/s/Uh9aHakug5RvyWb0N04h7GGz [https://perma.cc/JJ4A-C5QK].

173. Sublime prompt in Midjourney: "aerial view of a giant fish tank shaped like a tower in the middle of new york city, 8k octane render, photorealistic --ar 9:20." Midjourney, https://prompthero.com/prompt/90abe624731-midjourney-3-aerial-view-of-a-giant-fish-tank-shaped-like-a-tower-in-the-middle-of-new-york-city-8k-octane [https://perma.cc/Q5BM-9VCD].

174. Catherine Dee & Hamish Ogilvy, *Visual search: how does an image finder search engine work?*, Algolia (Feb. 22, 2022), https://www.algolia.com/blog/product/picture-search-how-does-an-image-finder-search-engine-work/ [https://perma.cc/9YUH-3835].

175. *See* Damir Yalalov, *Best 100+ Stable Diffusion Prompts: The Most Beautiful AI Text-to-Image Prompts*, Metaverse Post (Oct. 25, 2022, 6:31 am), https://mpost.io/best-100-stable-diffusion-prompts-the-most-beautiful-ai-text-to-image-prompts/ [https://perma.cc/D9XF-3L3W] (examples of prompts).

176. *See* Oyinlola, *supra* note 159; Kosar, *supra* note 160.

177. *See id.*

178. *See* sources cited *supra* note 160; *see also* sources cited *supra* 162.

for denoising.[179] Stable Diffusion starts by selecting a latent seed from the latent space of numeric image/text data and using the seed to generate random latent image representations that are then conditioned by the text to guide the objective (the final images) of the generation process.[180] This model uses the embedded image and text classifiers from its CLIP-curated training data to generate the random latent image representations which are first multiplied for diversity and then diffused by addition of Gaussian noise and conditioned by guided diffusion and cross-attention in the denoising process.[181]

The process of diffusion always and necessarily renders the subject random latent image representations into unrecognizable messes of Gaussian noise, so it matters very little what any random latent image might have looked like to begin with. The important step comes when these initial random representations are completely obscured (by noise), and then the representations are conditioned in the process of denoising to produce variations of actual images that are not bound to any particular CLIP classified image but still show fidelity to the semantics and style of certain classes and descriptions of embedded images of the training set that are called for by the end-user's text prompt.[182] The Stable Diffusion process is similar to DALL-E 2 in that it learned lessons from the semantics and styles of the CLIP curated LAION-5B dataset, and the decoding process calculates the required number of steps to eliminate "noise" in generated images by varying and deleting the non-essential details of these images to produce one or more final images with minimal loss in visual content values (e.g., photorealism) and text caption similarity.[183] In other words, the entire process for both of these visual generative AI systems is to use the information it has learned about what pictures of certain classes and descriptions look like, and what various combinations of classes and descriptions should look like, so that the systems can render one or more new,

179. Mehul Gupta, *How does DALL-E, the text-to-image generator work?*, Medium (Feb. 2, 2023), https://medium.com/ /how-does-dall-e-the-text-to-image-generator-work-c2d9f4a0f26c [https://perma.cc/SN4C-A8MW] *Generative Ai Art Tools Explained: Dall-E 2 and Stable Diffusion Comparison*, Perzonalization (Jan. 15, 2023), https://www.perzonalization.com/blog/generative-ai-art-tools-explained/ [https://perma.cc/WKR9-DS8B].

180. *See* Suraj Patil et al., *Stable Diffusion with Diffusers*, Hugging Face (Aug. 22, 2022), https://huggingface.co/blog/stable_diffusion [https://perma.cc/CK5L-AEP5].

181. *Id.*

182. The process described here can take the form of unCLIP that inverts the images produced from a CLIP set, *see* Ramesh, *supra* note 154;. *See* Alex Nichol et. al, *GLIDE: Towards Photorealistic Image Generation and Editing with Text-Guided Diffusion Models*, arXiv (Mar. 8, 2022), https://arxiv.org/pdf/2112.10741.pdf [https://perma.cc/4RJQ-3QJ8].

183. *See* O'Connor, *supra* note 159.

visually pleasing images that are faithful to the requirements and descriptions expressed in the text prompt.[184]

As noted above, this process is not a search for a preexisting image that matches the text prompt captions. In contrast, an image generating AI such as DALL-E 2 or Stable Diffusion is expressly designed *not* to use any one preexisting image in the CLIP training set because each image encoded and embedded in the CLIP-curated training must be stochastically decoded (denoised) to generate a new image which maintains the salient features of the targeted image and text data given its embedding.[185] OpenAI has developed an additional process known as GLIDE (Guided Language-to-Image Diffusion for Generation and Editing), in which the DALL-E 2 AI uses newly deconstructed images to guide the composition of new images that will embody all of the text relationships presented by the prompt.[186] Both systems require construction and deconstruction of data from source images to produce a final set of results; neither system selects nor duplicates an image from the training set.

The process of image generation used by DALL-E 2 and other AIs must be noted because copyright infringement vs. fair use claims will depend on a theory that the original work of the complaining artist or creator was copied by the AI and its image generation mechanism. But the very design of the AI is not to copy any preexisting image, but merely to learn from the data of millions of preexisting images what desirable images embodying certain prompt terms should look like so that when it gets a new assignment it can create desirable images that embody the new prompt.

V. INFRINGEMENT OR FAIR USE BY THE END-USERS OF GENERATIVE AI SYSTEMS

The true and proper defendants in a lawsuit involving the creation of an allegedly infringing image rendered by a visual generative AI system are the end-users of the AI system. The end-users are the authors and creators of the images that might arguably be substantially similar to works whose image data was included in the AI system's training data. The end-users design and cause their vision for a work to be produced by the AI system through their prompting; the end-users review the samples produced by the AI system and select an image or further condition the creation process with new prompts and instructions; and the end-users determine the ultimate purpose and function for the works they have designed and adopted as the final work from the

184. OpenAI's DALL-E 2 process is not the only method for drawing on a set of image-text classifications for image generation. *See, e.g.*, Vaclav Kosar, *Multimodal Image-text Classification*, Vaclav Kosar (Nov. 30, 2022), https://vaclavkosar.com/ml/Multimodal-Image-Text-Classification. [https://perma.cc/D6UM-CPRN]

185. O'Connor, *supra* note 159; Ramesh & Dhariwal, *supra* note 154.

186. *See* O'Connor, *supra* note 159; Ramesh & Dhariwal, *supra* note 154.

samples produced by the AI system, which is the critical step in the fair use analysis under *Warhol*. In terms of liability for infringement or exemption for fair use, the end-users are the only realistic and appropriate subjects for the infringement claims or the fair use defenses because they are the only parties with agency making all the decisions relevant to infringement or fair use.

A. Copyright Authorship and AI Generated Works

As noted above in section I, AI systems are not working on their own to outproduce and replace human artists.[187] Instead, AI is best understood as a tool for a human artist or creator to make art in ways that were difficult and time-consuming prior to the advent and widespread adoption of generative AI mechanisms.[188]

At the time I researched my first publication on AI generative artwork and copyright in the summer of 2022,[189] the main concern of copyright and generative AI art was whether an artificial intelligence could be the registered copyright owner of a work of art produced by the AI. At present, the answer is "no,"[190] which begs the questions, who then is the author of such works, and further, what does authorship mean in a world of AI generated art. The answer to the "who is the author" question is the very popular legal answer, "it depends," because factual circumstances can produce an answer ranging from, "There is no author" (meaning the work is a random, accidental creation),[191] to "There is a human author who used the artificial intelligence as a tool to produce artwork that the human author approves and accepts" (and the human author owns the copyright),[192] or to the as yet theoretical situation, "The artificial intelligence itself is the author" (but the AI cannot own the copyright because it is not a human).[193]

Authorship directly impacts the question of infringement or fair use in copyright law because the test examines what the author of the new work appears to have done that is an "act of copying" of the original authors' work that appears to have been used in the new creation. This should not be read to suggest that the law delves into the subjective intentions and motivations

187. Rosenberg, *supra* note 2, at 3.

188. *See* Will Knight, *When AI Makes Art, Humans Supply the Creative Spark*, Wired (Jul. 13, 2022, 7:00 AM), https://www.wired.com/story/when-ai-makes-art/ [https://perma.cc/FJE6-MAXF].

189. Michael D. Murray, *Generative and AI Authored Artworks and Copyright Law*, 45 Hastings Comm. & Ent. L.J. 27 (2023).

190. *Id.* at 38–44.

191. *Id.* at 36–38.

192. *Id.* at 38–41.

193. *Id.* at 41–43.

of the new work's author,[194] but instead focuses on the new work itself and what it shows has been done with and to the earlier work that appears to have been borrowed, copied, or otherwise incorporated in some manner in the new work.[195] If an AI somehow was determined to be the author of an infringing work, then the AI would be the defendant, the real party in interest responsible for the infringing work, which would place the dispute at an impasse because AIs have not been determined to be legal persons.[196] But if the AI is just a sophisticated tool for generating images at the direction of a human artist or creator, then the analysis fits the normal fair use analysis discussed in the next section.

B. End-Users as Artists, The Authors of the Creation

To analyze the infringement complaint against the end-users of visual generative AIs, consider the lessons from the discussion above:

- Generative AI is best understood as a tool for a human artist or creator to use.[197] AI does not make a creative artistic decision about the contents of the art; it only responds to a human prompt and then generates images according to its training and programming. AI should not be personified as the actual author of the generative AI image. The end-user is the author and artist of the image.
- In the process of creation, generative AI does not make creative design decisions, it follows rules and parameters (translated into algorithms) to generate output that the human end-user first directs in the initial prompt, and then evaluates and chooses to accept or reject in each set of samples generated by the AI tool. Human end-users using the AI tool usually are given several image options from which they can choose,[198] or they can rerun the same prompt to generate a new group of images, or they can revise the prompt in multiple iterations and generate a completely new set of images based on each revised prompt

194. Copyright deals with the objective attributes of the examined work, not the motive or creative intentions of its author. Star Athletica, L.L.C. v. Varsity Brands, Inc., 580 U.S. 405, 422–23 (2017).

195. *See generally Campbell*, 510 U.S. at 577-78.

196. *See, e.g.*, David C. Vladeck, *Machines Without Principals: Liability Rules and Artificial Intelligence*, 89 Wash. L. Rev. 117, 121, 124 (2014); Lawrence B. Solum, *Legal Personhood for Artificial Intelligences*, 70 N.C. L. Rev. 1231, 1239 (1992).

197. Rosenberg, *supra* note 2, at 3; Knight, *supra* note 169, at 55.

198. Stable Diffusion Playground produces four images as a default generation set responsive to a creation prompt. *See Stable Diffusion Playground*, https://stablediffusionweb.com/#demo [https://perma.cc/3RBH-67JX].

until the end-user causes the AI to generate the image envisioned and designed by the end-user.

- Human artists and creators control the art generated by the AI by the prompts that they write and revise. Thus, the human artists examine the works produced in the process and either accept the fruits of the process or they keep going with different or revised prompts. This is exactly similar to the process of creating sketches, studies, or drafts (iterations of a creative artistic project) until the artist is happy with the design, composition, framing, perspective, point of view, and the results of the techniques being used.[199]

There are some basic similarities between how an end-user of a contemporary generative AI creates art and how a human artist goes about the same task.[200]

Human Artists	**End-Users of Generative AI**
A human artist conceives of and designs a work.	An end-user conceives of and designs a work.
A human artist envisions what the image should look like drawing from images and other information the artist has been trained on or exposed to, and the human artist may be further guided and inspired by research involving preexisting images and information about schools, genres, and techniques of art, all to determine, "What does this type of image look like."	The end-user's prompt causes the generative AI tool to follow the prompt instructions and conditions and, drawing from the lessons of images and text data it has been trained on, determines, "What does this type of image look like."

199. *See generally* Marion Botella et al, *What Are the Stages of the Creative Process? What Visual Art Students Are Saying*, 9 Frontiers in Psychology 1, 2–6 (Nov. 21, 2018) https://www.frontiersin.org/articles/10.3389/fpsyg.2018.02266/full (discussing stages of research, trials or sketches, techniques, evaluation) [https://perma.cc/JV7C-LZDT]; Matt Fussell, *The Creative Process*, The Virtual Instructor (Jan. 12, 2023) https://thevirtualinstructor.com/blog/the-creative-process (reviewing stages of research, production, critique, rework) [https://perma.cc/2TT5-B4T6].

200. This table follows the steps of the creative process discussed in Botella, Zenasni et al., *supra* note 180, at 57; Fussell, *supra* note 179, at 57–58.

Human Artists	End-Users of Generative AI
A human artist follows her inner vision and creates preliminary sketches, studies, or drafts for evaluation to determine the most desirable elements of the image (composition, style, genre, etc.) and the best techniques to achieve that image – i.e., "What is the best way to render this image."	Following the end-user's prompt instructions and conditions, the AI tool generates an initial diverse set of images with elements (composition, etc.) that match the terms in the prompt – and the end-user evaluates the samples and makes determinations about the most desirable elements (composition, style, genre, etc.) of the image – i.e., "What is the best way to render this image."
A human artist reworks the sketches, studies, or drafts to best meet the requirements of her inner vision – i.e., "What should the final image look like."	The end-user engages in refinement (re-prompting) that causes the AI to delete unnecessary elements of images and retain the best elements based on the directions and conditions of the prompts – i.e., "What should the final image look like."
Ultimately, the human artist creates a work that the artist accepts and adopts as the final iteration of the project – i.e., "This is the image that should be used."	Ultimately, the end-user accepts and adopts a final iteration for the project – i.e., "This is the image that should be used."

It is quite obvious that there is less romance and emotion in the generative AI process,[201] but the steps are most definitely similar if not functionally and substantially the same as the process of creation followed by a human artist.

C. An Act of Copying by the End-User

The claimant's copyright infringement suit will founder if the claimant cannot prove that the end-user copied a substantial, material, original, and copyrightable part of the claimant's work. Direct evidence of copying will not be possible because the end-user of a visual generative AI system never has direct access to any of the image and text data of the training set and foundation model, let alone access to any actual image files whose data was used in the training process. Nor does the generative AI find and copy or collage images using actual image files.

201. Commentary on generative AI ranges from a vision of a "dark forest" of lifeless works of generated content devoid of human emotion, *e.g.*, Maggie Appleton, *The Expanding Dark Forest and Generative AI*, Maggie Appleton-The Garden, https://maggieappleton.com/ai-dark-forest [https://perma.cc/TU97-WP6].

The "Magic File Drawer" theory discussed in section I above is just that—magical thinking. The end-user does not rifle through a drawer of images to find one to copy. An end-user conceives of an image and directs the AI with instructions and conditions about the image she wants to create using the AI tool.

There is a name for the "Magic File Drawer" theory of art creation: it is called a Google image search. It is much simpler and often even faster than the fastest generative AI system for you to write a prompt (i.e., search query) seeking a certain image or a more general description of an image. And through this tool, the end-user is given a selection of actual thumbnail image files and links to full-size image files, which the end-user can right click on, and–magic!–have a brand-new exact copy of an image to work on. From there, a "Magic Paintbrush" can be applied in the form of a highly sophisticated image editing program such as Adobe's PhotoShop®, or the more modest but very user-friendly filters and image-editing tools built right into Microsoft Word and other Office365 applications.

Assuming a court might turn a blind eye to the technical realities of the AI image generation process and the fact that the end-user in the act of using the generative AI tool never had and could not have had access to any individual image whose data was used in the training process and further used in the actual AI generation process, the court might still give the claimant the opportunity to present a case for substantial similarity. Assuming for purposes of this argument that the end-user knew that some of the claimant's artworks might have been scraped and their image data harvested for use in the AI's training set, the end-user could embark on a calculated path to direct the AI to produce an image that is substantially similar to the claimant's image. That argument requires further exploration of generative AI's ability to make works that are substantially similar to works whose data was used in the training dataset and foundation model.

D. End-Users Using a Generative AI Tool Can Match a Human Artist's Ability to Draw on Training and Knowledge to Emulate the Styles and Genres of Other Artists

To say that a generative AI enables an end-user to make works that are substantially similar to works whose data was used in the training dataset and foundation model does not imply that this means the AI generates *infringing* works at the command of the end user. Human end-users using the AI may want to use the AI tool to make a new work that follows the same style and genres of other artists.

It is natural and expected that artists will draw from earlier works of earlier artists.[202] Schools of art are formed from the common approaches used by artists that develop into a similar genre of art.[203] The *Warhol, Google*, and

202. *E.g.*, Campbell, 510 U.S. at 575–76.

203. *E.g., Penguin Books USA, Inc.*, 109 F.3d at 1399.

Campbell courts each acknowledged that all artists draw inspiration from artists' works of the past.[204] The style and genre and the techniques and processes that lead to a certain appearance of works are uncopyrightable elements of prior works, and copying them does not lead to liability for infringement because of the originality doctrine, the idea-expression distinction, and the doctrines of merger and scènes à faire.[205]

Generative AI seems particularly good at helping end-users to create works emulating an artistic style, and can combine it with a recognizable ability to enable end-users to recreate the image and likeness of actual persons:

Cat playing poker in the style of Renoir[206] 	Emma Watson in the style of John Singer Sargent's Portrait of Madame X[207] 
Anne Hathaway as Woman Warrior in the style of Greg Rutkowski[208]	Janelle Monáe in the style of Andy Warhol[209]

204. *Warhol*, 143 S. Ct. at 1286–87; *Oracle Am., Inc.*,141 S. Ct. at 1203–04; *Campbell*, 510 U.S. at 575 (citing Emerson v. Davies, 8 F. Cas. 615, 619 (No. 4,436) (CCD Mass.1845)).

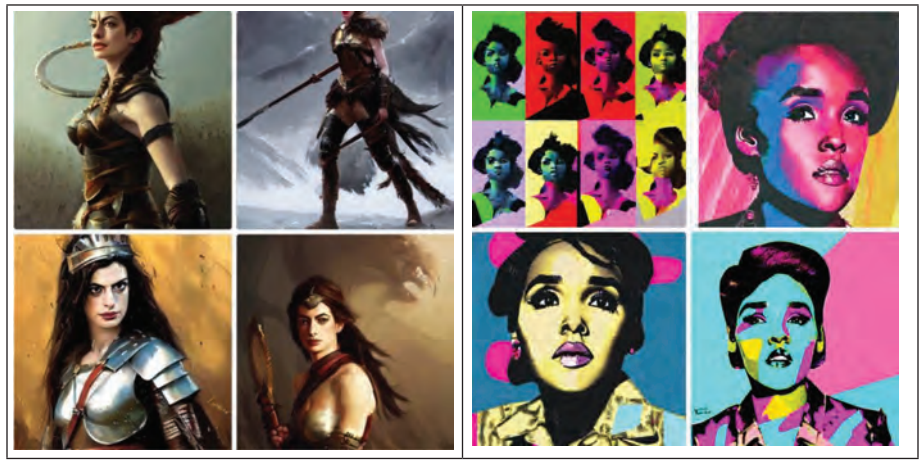
205. *See supra* text accompanying notes 24–27.

206. *Image of Cat Playing Poker in the Style of Renoir* (2023), *supra* note 5, at 4.

207. *Five images of Emma Watson in the Style of John Singer Sargent* (2023), conceived of and prompted by Michael D. Murray, rendered by Stable Diffusion Playground (Jan. 11, 2023), https://stablediffusionweb.com/#demo.

208. *Four images of Anne Hathaway in the Style of Greg Rutkowski* (2023), conceived of and prompted by Michael D. Murray, rendered by Stable Diffusion Playground (Jan. 11, 2023), https://stablediffusionweb.com/#demo.

209. *Four images of Janelle Monáe in the style of Andy Warhol*, (2023), conceived of and prompted by Michael D. Murray, rendered by Stable Diffusion Playground (Jan. 13, 2023), https://stablediffusionweb.com/#demo.



This ability to recreate artistic style and genre is both highly regarded and highly reviled. On the one hand, the generation of a workable image that emulates the style of an artist used to be a project requiring days or even weeks by a talented human artist, but now, a contemporary generative AI can crank out very credible works in 15-30 seconds or fewer.[210] But the same output may be reviled, especially by living artists who do not appreciate having their works used in the training set of images or used later as part of the "in the style of" prompts given to generative AI.[211] In the above illustrations, I included one living artist, Greg Rutkowski, as an "in the style of" artist for a set of images because Greg Rutkowski has complained that his name has been suggested by both Stable Diffusion and Midjourney as a style prompt that will help the AI produce attractive images in the fantasy art genre, and reports are that tens or even hundreds of thousands of artists have used such prompt terms to generate a vast supply of art that resembles Mr. Rutkowski's style of fantasy art.[212] Imitation might be the highest form of flattery, but artists such as Mr. Rutkowski do not want their present and future commissions to dry up simply because an AI can generate "pretty good" and

210. These numbers are based on my own experience painting oil and acrylic paintings that emulate the style of other artists. The designation of "talented" is self-applied. The Stable Diffusion timing is from my own experience using the free demo version of Stable Diffusion Playground. https://stablediffusionweb.com/#demo.

211. Rachel Metz, *These artists found out their work was used to train AI. Now they're furious*, CNN Business (Oct. 21, 2022, 9:06 AM), https://www.cnn.com/2022/10/21/tech/artists-ai-images/index.html [https://perma.cc/7RC9-47B6].

212. *See* Laurie Clarke, *When AI can make art – what does it mean for creativity?*, The Guardian (Nov. 12, 2022, 11:00 AM), https://www.theguardian.com/technology/2022/nov/12/when-ai-can-make-art-what-does-it-mean-for-creativity-dall-e-midjourney [https://perma.cc/7WUC-W4PD].

"close enough" facsimiles of their works.[213] One of the lead plaintiffs in the *Andersen v. Stability AI* lawsuit,[214] Kelly McKernan, reported to the New Yorker magazine that her:

> name was being used with increasing frequency in A.I.-driven image generation. McKernan makes paintings that often feature nymphlike female figures in an acid-colored style that blends Art Nouveau and science fiction. A list published in August . . . suggested "Kelly McKernan" as a term to feed an A.I. generator in order to create "Lord of the Rings"-style art. . . .On the Discord chat that runs an A.I. generator called Midjourney, McKernan discovered that users had included [her] name more than twelve thousand times in public prompts. The resulting images—of owls, cyborgs, gothic funeral scenes, and alien motorcycles—were distinctly reminiscent of McKernan's works. "It just got weird at that point. It was starting to look pretty accurate, a little infringe-y," [she] told me. "I can see my hand in this stuff, see how my work was analyzed and mixed up with some others' to produce these images."[215]

While this narrative undoubtedly reflects an annoying development in the world of visual art and content creation, it is reflective not of a situation of massive copyright infringement but instead a reflection of undesired emulation of works of a certain style, genre, technique, and theme, all of which are distinctly not "infringe-y" under the originality doctrine, the idea-expression distinction, and the doctrines of merger and scènes à faire.

The general attributes and imagery associated with a certain style, genre, school, or theme of art are uncopyrightable and unprotectable ideas on the scale of ideas vs. expressions. The merger and scènes à faire doctrines exist specifically to account for the fact that artists study and learn from the work of other artists in various schools and genres of art, and produce works that are, in Ms. McKernan's words, "distinctly reminiscent of" the style of other artists.[216] Ms. McKernan self-describes herself as following the style and technique of a school of art—art nouveau—combined with themes or scènes à faire elements of science fiction.

The point at which the originality doctrine, the idea-expression distinction, and the doctrines of merger and scènes à faire give way to

213. *Id.*

214. Andersen v. Stability AI, No. 3:23-cv-00201 (N.D. Cal. Jan. 13, 2023).

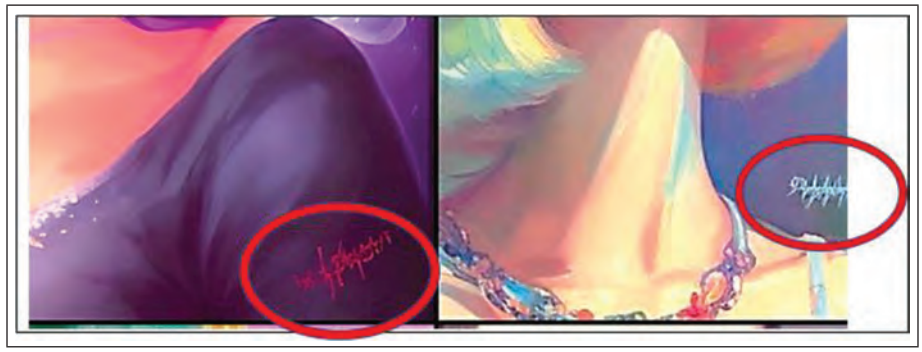
215. Kyle Chayka, *Is A.I. Art Stealing from Artists?*, The New Yorker (Feb. 10, 2023), https://www.newyorker.com/culture/infinite-scroll/is-ai-art-stealing-from-artists [https://perma.cc/LM38-LVZ2].

216. *Id.*

copyright infringement is when a work is substantially similar to another work not just because it shares some common elements relating to genre and theme, but because it copies original, copyrightable parts and captures the "total concept and feel" of the work.[217] The next section addresses this scenario.

E. An AI is Capable of Creating Infringing Works *if* the Human Operator Works Hard Enough on Trying to Achieve this End

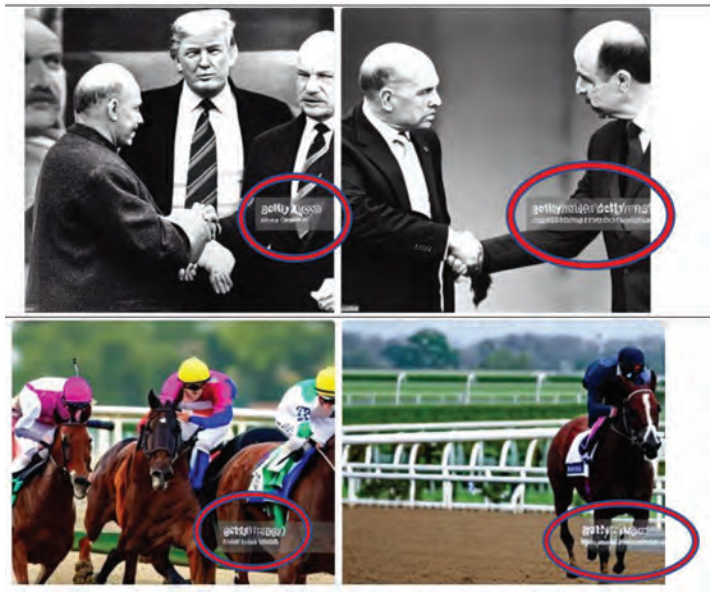
Moving past the question of generating works that emulate the style of an artist, some artists have complained that specific portions of their works appear to have been copied in AI generated works.[218] Artists have noted that their signatures or facsimiles thereof appear in works generated by AI.[219]



217. *E.g.*, Designworks Homes, Inc. v. Thomson Sailors Homes, L.L.C., 9 F.4th 961, 963 (8th Cir. 2021); Abdin v. CBS Broad. Inc., 971 F.3d 57, 66 (2d Cir. 2020).

218. *See* Kyle Wiggers, *Image-generating AI can copy and paste from training data, raising IP concerns*, TechCrunch (Dec. 13, 2022. 7:30 AM), https://techcrunch.com/2022/12/13/image-generating-ai-can-copy-and-paste-from-training-data-raising-ip-concerns/ [https://perma.cc/K298-Y5S6]**.**

219. Shanti Escalante-De Mattei, *Artists Voice Concerns Over The Signatures In Viral LensaAI Portraits*, ARTnews (Dec. 9, 2022, 3:04 pm), https://www.artnews.com/art-news/news/signatures-lensa-ai-portraits-1234649633/ [https://perma.cc/YC6K-S7GM]; Juan Camilo Arroyave Guevara, *Artist shows evidence that AI is possibly stealing others' art*, Level Up (Dec. 16, 2022), https://www.levelup.com/en/news/715571/Artist-shows-evidence-that-AI-is-possibly-stealing-others-art [https://perma.cc/3QJY-V3UH].



220

In my own experience using Stable Diffusion, I, too, found that some generated art produced a facsimile of certain specific portions of works, such as a Getty Images watermark, particularly if the words "news photo" were used in the prompt.

Commentators have explained this occurrence not as an indication that a specific work has been copied but instead as a product of training: if a large amount of image data in an AI's training set reflect that a certain category or classification of images have a signature on them, the generative AI will learn that images in this category or classification of image need a cluster of scrawly lines in the image, and if a significant amount of the data from photos in the dataset bearing an alt text of "news photo" have a Getty Images watermark, then the AI will learn that "news photos" need a facsimile of this watermark to be in the generated images.[221] This is similar to a prompt asking for a cartoon panel, in which the AI attempts to mimic the text it finds in the thousands of cartoon panels it trains on, but the facsimile text is randomly formed and appears as gibberish when generated in the AI process:

220. Michael D. Murray, *Collage of fake news photographs each showing a facsimile of the Getty Images Watermark*, rendered with the assistance of Stable Diffusion Playground (Dec. 2022), https://stablediffusionweb.com/#demo.

221. *See, e.g.*, Jeremy Zaborowski, *The Truth Behind Signatures on AI-Generated Art*, Medium (Jan. 20, 2023), https://node-jz.medium.com/the-truth-behind-signatures-on-ai-generated-art-d40dec8f817b [https://perma.cc/9RD6-ZQWN]; James Vincent, *Getty Images is suing the creators of AI art tool Stable Diffusion for scraping its content*, The Verge (Jan. 17, 2023, 5:30 AM), https://www.theverge.com/2023/1/17/23558516/ai-art-copyright-stable-diffusion-getty-images-lawsuit [https://perma.cc/GYF9-GKS6].

222

I have tried mightily to get Stable Diffusion and DALL-E 2 to reproduce a specific work of art, and every time I have failed. I would attribute this to the fact that the variational text encoders and image encoders directing the diffusion process make it extremely difficult to recreate a specific image, or perhaps guardrails put in place by Open AI or Stability AI try hard to close off this practice. But there will be instances when an end-user manages to produce a work that is substantially similar to a preexisting work whose image data was included in the training data of the AI used by the end-user[223]:

I believe the academic paper from which the images at left were drawn supports my conclusion that it is extremely difficult to get a visual generative AI to replicate a preexisting image because of all the examples shown in the paper, which I assume were cherry-picked to show the most illustrative and indicative examples from the study, only about 20 or so images out of the 9,000 images used in the study actually appear to have copied substantial *copyrightable* elements and the "total

222. Michael D. Murray, *Cartoon Panel Images in the Style of DC Comics Wonder Woman*, rendered with the assistance of Stable Diffusion Playground (2022) https://stablediffusionweb.com/#demo.

223. Gowthami Somepalli & Vasu Singla, *Diffusion Art or Digital Forgery? Investigating Data Replication in Diffusion Models*, arXrv (Dec. 12, 2022), https://arxiv.org/pdf/2212.03860.pdf [https://perma.cc/6LA9-VQT5]. I have shown what I believe to be the two best examples from the study.

concept and feel" of preexisting works beyond the works' style, genre, techniques, and themes, and other merged and scènes à faire elements. So, if 20 out of 9,000 (0.2 %) is the batting average, I would conclude that it is *very* difficult to get a hit in this game.[224]

F. The End-User's Function and Purpose for Any Potentially Infringing Works Will Determine if the End-User Has Authored Any Transformative Artworks

If the planets align, and the plaintiff can get through all the requirements of a copyright infringement action, then the question of fair use will need to be considered. Because, after all is said and done, the end user's copying may be transformative.

As discussed in sections III and III(A) above, the function and purpose of the use for which an allegedly infringing derivative work is created will be the determinative factor in whether the end-user has a fair use. If the function and purpose is comment and criticism, such as a parody of the original work or the artist of the original work, then the end-user will be on fairly firm ground in asserting a fair use claim, assuming that the parodic elements of the new work can be observed in the work itself.[225] However, if the end-user is found to have created a substantially similar visual work created to be used for its aesthetic artistic qualities and it allegedly infringes another visual work created to be used for its aesthetic artistic qualities, and the original work shines through in the new work, then the use will not be found to be transformative and not a fair use.[226]

---

224. Naturally, these figures are based on my evaluation of substantial similarity. A commentator on the study believed that the authors had observed a rate of 1.88% substantial similarity, but that evaluation did not take into account the merged and scènes à faire elements arising from the style, genre, techniques, and themes of the original images which were responsible for much of the similarity in the AI generated images. *See id.*; Wiggers, *supra* note 218, at 67.

225. *Compare Campbell*, 510 U.S. 569 (the rap group 2 Live Crew's song "Pretty Woman" was a parody of Roy Orbison's song "Oh, Pretty Woman" because it targeted the original song for criticism and comment, and used only as much of the original as necessary to conjure up the original), *and* Leibovitz v. Paramount Pictures Corp., 137 F.3d 109 (2d Cir. 1998) (the movie poster for Naked Gun $33^{1}/_{3}$: The Final Insult was a parody of Annie Leibovitz's photograph of a pregnant Demi Moore on the cover of Vanity Fair because it was obvious from the poster that it was making fun of Leibovitz and her artwork), *with Penguin Books USA, Inc.*, 109 F.3d 1394 (the book "The Cat NOT in the Hat!" was not a parody of Dr. Seuss's works because it did not target them for comment or criticism, but merely borrowed their style to tell a story about the O.J. Simpson murder trial).

226. See cases cited, *supra* note 60–61.

G. Contributory or Vicarious Liability for Infringement by the Designer of the Generative AI System

After finding that the end-user is the most appropriate defendant in any copyright infringement suit regarding work generated with an AI tool, the plaintiff in such a dispute will inevitably look around for a defendant with deeper pockets than a typical end-user of an AI system. They will then settle on the designers and creators of the AI system and, presumably, will try to assert a claim for contributory or vicarious liability for the infringement. But the creators of a generative AI system such as OpenAI and Stability AI should not be held liable for contributory infringement or vicarious liability for creating a visual generative AI system such as DALL-E 2, Stable Diffusion, or Midjourney, because the designers are not directly involved in the creation or distribution of infringing content by the end-users of their systems let alone having control over such conduct.

Contributory infringement occurs when a party knowingly induces, causes, or materially contributes to the infringing conduct of another.[227] Vicarious liability occurs when a party has the right and ability to control the infringing conduct of another and derives a direct financial benefit from it.[228] Neither of these theories apply to the creators of generative AI systems for the following reasons.

First, the creators of generative AI systems do not have knowledge of or control over the specific content that end-users of their systems have caused the systems to generate. Generative AI systems are based on complex algorithms and large datasets and foundation models that are designed to produce outputs that are not predictable or predetermined by the AI designers.[229] The design and normal operation of the system are to create new and original artworks according to specifications and requirements provided by the end-user. [230] As discussed in the section above, it is possible for an end user to work at prompt engineering to get the generative AI tool to produce an artwork that could be determined to be infringing, but this outcome defies the normal design for the systems.[231] The AI system creators do not monitor, review, or approve the

227. Sony Corp. of Am. v. Universal City Studios, Inc., 464 U.S. 417, 437 (1984); Leonard v. Stemtech Int'l Inc., 834 F.3d 376, 387 (3d Cir. 2016).

228. Metro-Goldwyn-Mayer Studios Inc. v. Grokster, Ltd., 545 U.S. 913, 937 (2005); *Sony*, 64 U.S. at 437.

229. Damian Brady, *What developers need to know about generative AI*, Github Blog (Apr. 7, 2023), https://github.blog/2023-04-07-what-developers-need-to-know-about-generative-ai/ [https://perma.cc/3A97-GZ2Z],

230. *Id; Generative AI Models Explained*, Altexsoft.com (Oct. 13, 2022), https://www.altexsoft.com/blog/generative-ai/ [https://perma.cc/SC3A-6JN5] (with generative AI, "[t]he main idea is to generate completely original artifacts that would look like the real deal.").

231. *See, e.g.*, Stephen Wolfson, *Fair Use: Training Generative AI*, Creative Commons (Feb. 17, 2023), https://creativecommons.org/2023/02/17/fair-use-training-generative-ai/ [https://perma.cc/E94B-H2SV].

outputs before they are generated or distributed by the end-users. Therefore, the creators do not have the requisite knowledge or control to be liable for contributory infringement or vicarious liability.

Second, the creators of generative AI systems do not induce, cause, or materially contribute to the infringing conduct of their users. The creators do not provide any instructions, guidance, or encouragement to their users on how to use their systems to create *infringing* content. Each service discussed here—DALL-E 2, Stable Diffusion, and Midjourney—provides instructions of a very general nature about writing prompts and editing prompts to guide the AI tool toward the rendering of the work imagined and envisioned by the end user. That is a far cry from giving instructions on how to recreate preexisting works. Other than the generative AI system itself, the creators do not provide any tools or services that facilitate the creation or distribution of infringing content by their users. The creators do not receive any direct financial benefit from the works created by the end-users, whether they are infringing or not. In this manner, the creators do not have the requisite causation or contribution to be liable for contributory infringement or vicarious liability.

A third reason why the creators of generative AI systems, such as OpenAI and Stability AI, should not be held liable for contributory infringement or vicarious liability for creating a visual generative AI system such as DALL-E 2, Stable Diffusion, or Midjourney, is that there are many lawful, non-infringing uses for the technology under Sony.[232] In *Sony*, the Supreme Court held that a technology manufacturer cannot be held liable for its users' copyright infringement if the technology enables substantial non-infringing uses as well.[233] The Court held that Sony was not liable for contributory infringement for selling video tape recorders to consumers who recorded copyrighted television programs because the video tape recorders had substantial non-infringing uses, such as time-shifting and educational purposes.[234]

Similarly, generative AI systems have substantial non-infringing uses, such as artistic expression, education, research, scientific innovation, and entertainment. Generative AI systems primarily are used to create original artworks, music, stories, poems, and games, none of which are copies of preexisting works, and none of which would infringe any copyrighted work. Generative AI systems can be used to conduct research and promote scientific discovery and innovation by generating new structures and forms for analyzing and summarizing data, for generating hypotheses and designing experiments, and for developing solutions to critically important problems for the benefit of the

232. *Sony*, 464 U.S. at 442–47.

233. *Id.*

234. *Id.* at 442–47, 454–56.

world.[235] Therefore, generative AI systems are not merely tools for infringing activities, but rather technologies that enable significant non-infringing activities that benefit society and culture under the *Sony* doctrine.

VI. CONCLUSIONS

As a coda to the above discussion, consider that the tool that the plaintiffs in the *Andersen* class action lawsuit and other aggrieved and disappointed artists are complaining about has been in existence for several decades. It is called the web browser with its image search engine. If one truly wants to create a copy of a preexisting work, nothing is easier than searching for it on the internet and right-clicking on it when you find it. If you want to create an infringing derivative work, then a second step can readily be taken, and the technology for it also has been around for several decades. It called the use of image editing software such as Adobe PhotoShop® or the myriad of other photo editing and painting products. If a collage of derivative works is your caper, you can accomplish this with Microsoft Word and its "Picture Format" tab in conjunction with the snipping tool that shipped with your computer's operating system.

All of this is to say that the plaintiffs in *Andersen* and many other artists are not complaining because now their artworks can be copied and before they couldn't be. They are complaining because technology now enables amateur, untrained, unprofessional, and extremely average persons to compete with professional artists by producing beautiful, complex, painterly artistic works that emulate the genre and style of artists by using a tool that runs on simple textual instructions. Visual generative AI systems have democratized artistic creation to a level never seen in human history. This power has caused the explosion in 2022 and 2023 in the adoption and use of contemporary generative AI devices such as DALL-E 2, Stable Diffusion, and Midjourney.

It is valid for present artists to feel a certain degree of existential pressure which is expressed in the phrase, "the robots are coming for our

235. Scientists believe that AI will be critically important in dealing with the effects of climate change, Megan Mastrola, *How Ai Can Help Combat Climate Change*, Johns Hopkins Univ. HUB (Mar. 7, 2023), https://hub.jhu.edu/2023/03/07/artificial-intelligence-combat-climate-change/ [https://perma.cc/ATQ9-PUW2], in developing new drugs and drug therapies; Maya Yang, *Scientists use AI to discover new antibiotic to treat deadly superbug*, The Guardian, (May 25, 2023), https://www.theguardian.com/technology/2023/may/25/artificial-intelligence-antibiotic-deadly-superbug-hospital [https://perma.cc/SE2F-EPF9]; and in finding cures for diseases, *To accelerate search for an Alzheimer's cure, scientists use artificial intelligence to identify likely drug targets*, Ariz. Coll. of Med. News (May 15, 2023), https://medicine.arizona.edu/news/2023/accelerate-search-alzheimers-cure-scientists-use-artificial-intelligence-identify-likely [https://perma.cc/9SAK-GZTW],

jobs."[236] And in truth, there are some jobs in the visual arts that could be supplemented or in some cases replaced by a robustly trained AI run by a savvy human designer—illustrators, interior designers, advertisers, storyboard artists—or practically any artist or creator whose job it is to produce images in response to a request or directive.[237] The AI is not going to replace all humans involved in the visual and graphic arts, but if there were dozens of persons in a design studio or arts organization whose jobs consisted of interpreting, designing, and rendering in visual media the ideas and directives of others, those person's jobs could theoretically be replaced by one person running a robustly trained generative AI.

Pinning the blame for the elements of copyright infringement on the right parties is essential to an informed and intelligent discussion of whether there should be liability or fair use for AI generated work.

Infringement Element	**Creators of the AI Training Set**	**Creators of the visual generative AI system**	**End-users of the visual generative AI system**
Act of Copying	The creators of the training dataset and foundation model do not copy any of the images on the internet whose data was scraped to create the dataset. No image files are downloaded, copied, or stored.	Generative AI systems do not copy any of the images on the internet whose data was scraped to create the AI's training data and foundation model. The act of generation by the diffusion process does not involve selecting preexisting images from the training set to copy them.	When using the generative AI tool, end-users have no access to the actual image files whose data was used in the AI training data and foundation. But if the law turns a blind eye to this technicality and allows the concept of "access" to be stretched, the end-user still does not create a copy of a preexisting work without going to great lengths in prompt engineering.

236. *E.g.*, Paul Krugman, *Does ChatGPT Mean Robots Are Coming For the Skilled Jobs?*, N.Y. Times, (Dec. 6, 2022), https://www.nytimes.com/2022/12/06/opinion/chatgpt-ai-skilled-jobs-automation.html [https://perma.cc/J3KY-TAY7]; Kevin Roose, *A.I.-Generated Art Is Already Transforming Creative Work*, N.Y. Times, (Oct. 21, 2022) https://www.nytimes.com/2022/10/21/technology/ai-generated-art-jobs-dall-e-2.html [https://perma.cc/AP4U-22MQ].

237. *See* Roose, *supra* note 236, at 76.

Infringement Element	**Creators of the AI Training Set**	**Creators of the visual generative AI system**	**End-users of the visual generative AI system**
Generation of a work that copies a substantial and material portion of copyrightable parts of the first work	The creators of the training dataset do not generate images.	The diffusion process does not generate work that incorporates all or part of preexisting works.	An end-user could engage in significant prompt engineering to cause the AI tool to generate a new work that appears to incorporate all or part of the copyrightable elements of a prior work.
Potential fair use of the copied portions of the first work	To the extent that the activity is (incorrectly) found to be copying, the creation of an AI training dataset or foundation model would most likely be protected by a transformative fair use defense because of the nonexpressive nature of the copying to enable a copy-reliant technology.	The outputs of a generative AI are not preprogrammed or predetermined; the creators of the AI do not see, review, or approve the outputs; they have no control over end-user's actions and do not receive financial benefit from the outputs; the AI systems have substantial noninfringing uses.	End-users who work to produce substantially similar and allegedly infringing works will also determine the function and purpose for the new work; they will determine if the work will be used for education, research, comment or criticism, or another fair use purpose.

The sober truth is that visual generative AI systems are not a Magic File Drawer, they are not a Magic Copy Machine, or a Magic Box Artist. They are not magic at all, but a set of complex algorithms trained on a dataset of information on what images of a certain kind and classification should look like, that work with nonexpressive numeric renderings of image data, rather than copying and editing real digital image files, and that use a complex reiterative diffusion process to work through a sequence of renderings conditioned on the terms and design requirements communicated by the end-user, until the system produces a final image that is accepted and adopted by the end-user. To the extent that the end-user decides to use the generative AI to duplicate a

preexisting work and works through the difficulties of getting the generative AI system to do what it is designed *not* to do—namely duplicate a preexisting image—then the end-user might force the AI system to create an infringing copy or derivative work, after which the end-user's function and purpose for his copy or derivative work will be compared to the function and purpose of the infringed work for evaluation of fair use.