*Harvard Journal of Law & Technology*
Volume 38, Number 1 Fall 2024

# ELEMENTS OF STYLE: COPYRIGHT, SIMILARITY, AND GENERATIVE AI

*Benjamin L.W. Sobel**

ABSTRACT

"You can't copyright style" is a shibboleth in today's debate over generative AI. This slogan is, at best, meaningless. More likely, it's wrong. Sometimes, what we call "style" is copyrightable. "Substantial similarity" is the doctrine that assesses when stylistic copying becomes infringement, but it is notoriously erratic, and judges find it especially hard to apply to images. Current law obfuscates artists' rights to control their works and the public's rights to use generative AI trained on these works.

Part II explains how image-generating AI works and debunks the prominent metaphor that it is a "collage machine." The metaphor erroneously posits that it is possible to differentiate "mechanical" reproductions of works of visual art from "intellectual" reproductions, and it erroneously implies that the distinction has legal significance. Generative AI is clearly learning to reproduce *something* from its training data: what matters is what that something is.

Part III defines style as a holistic attribute of a work, or a group of works, that comprises a constellation of expressive choices. These expressive choices might be unprotectable individually, but in combination, they may constitute protectable expression. Protectable style is not necessarily limited to expression in one discrete "work," either: courts regularly find copyrightable expression in aggregations of multiple works. Part III documents courts' struggles to assess the scope of copyright protection in visual art and attributes these struggles to the substantial similarity test's near-irreconcilable demands: courts must simultaneously dissect images into their constituent elements — a task

* Postdoctoral Fellow, Cornell Tech Digital Life Initiative. The author thanks BJ Ard, Saptarishi Bandopadhyay, Robert Brauneis, Bay ByrneSim, Mala Chatterjee, Terry Fisher, Kat Geddes, James Grimmelmann, Kevin Hong, Lily Hu, Diana Liebenau, Grant Meyer, Ruth Okediji, Rachel Orol, Sarah Rosenthal, Joel Sobel, David Stein, Rebecca Tushnet, David Gray Widder, Kathryn Woolard, Jonathan Zittrain, and the participants in the 2024 Tri-State IP Workshop at NYU Law and the 2023 Intellectual Property Scholars Conference at Cardozo Law. Much earlier versions of this research were presented, and much helpful feedback received, at the 2018 Works-in-Progress Intellectual Property (WIPIP) Colloquium at Case Western Reserve University School of Law and at the IP Scholars' Roundtable at the University of New Hampshire School of Law in October 2019. This Article was supported by a grant from the Honorable Nancy F. Atlas Intellectual Property American Inn of Court in Houston, Texas and the University of Houston Law Center's Institute for Intellectual Property & Information Law.

judges claim they are unable to do — while also assessing works' aesthetic appeal holistically and intuitively. Style has always been a challenge for substantial similarity because it is the form of expression least susceptible to analytical dissection and most likely to elicit inarticulate aesthetic intuitions. Generative AI models' replication of style is a hard problem for copyright law because the models are purpose-built to identify and reproduce precisely the forms of similarity that are hardest to analyze legally.

TABLE OF CONTENTS

I. INTRODUCTION .....51

II. GENERATIVE AI TECHNOLOGY: THE "COLLAGE MACHINE?" .....55
*A. The Data that Power AI* .....57
*B. Responses from Artists and Industry* .....59
*C. Diffusion Models: Mapping Interstitial Space in Our Words and Images* .....61
*D. Forget About The "Collage Machine"* .....66

III. WHAT WE TALK ABOUT WHEN WE TALK ABOUT STYLE .....71
*A. What is Style?* .....72
*B. Style, Substantial Similarity, and the Idea-Expression Distinction* .....76
1. Stylistic Similarity is Hard to Adjudicate .....79
2. Stylistic Similarity Can Be Infringement .....82
*C. Where is Style? Scoping the "Work"* .....87
1. Courts Look Past the Bounds of the "Work" .....88
2. The "Work" is Contingent, and Generative AI Changes the Contingencies .....93
3. A Substantially-Similar-To-One-Work Rule Just Duplicates the Idea-Expression Distinction .....98
*D. Summing Up: Similarity You Can Describe Versus Similarity You Can See* .....101

IV. CONCLUSION .....104

## I. INTRODUCTION

Hollie Mengert is a professional illustrator and designer who works in the animation industry.[1] In late 2022, she became unwitting proof of AI's power to emulate artistic style. A user named "MysteryInc152" made a post on the social media site Reddit: "2D Illustration Styles are scarce on Stable Diffusion so i created a dreambooth model inspired by Hollie Mengert's work."[2] MysteryInc152[3] had taken image-generating

1. Hollie Mengert, *About*, THE ART OF HOLLIE MENGERT, https://hollie mengert.com/aboutme [https://perma.cc/B546-FJH4].

2. MysteryInc152, *2D Illustration Styles are Scarce on Stable Diffusion so I Created a Dreambooth Model Inspired by Hollie Mengert's Work*, REDDIT (Oct. 22, 2022), https://www.reddit.com/r/StableDiffusion/comments/yaquby/2d_illustration_styles_are_ scarce_on_stable/ [https://perma.cc/U775-X82H].

3. MysteryInc152 is the alias of Ogbogu Kalu, a mechanical engineering student in Canada. Andy Baio, *Invasive Diffusion: How One Unwilling Illustrator Found Herself Turned into an AI Model*, WAXY.ORG (Nov. 1, 2022), https://waxy.org/2022/11/invasive-diffusion-how-one-unwilling-illustrator-found-herself-turned-into-an-ai-model/ [https://perma.cc/7B3W-B5F9].

AI software called StableDiffusion and "fine-tuned" it to mimic Mengert's style by providing it with thirty-two of Mengert's illustrations.[4] His post provided a link to download the model, as well as some sample images.



Figure 1: Hollie Mengert Illustrations (left) Alongside the Outputs of "a DreamBooth model inspired by Hollie Mengert's work" (right).[5]

The results are striking. Take a look for yourself: Figure 1 shows a side-by-side comparison of four works by Mengert (on the left) and four outputs of MysteryInc152's model (on the right).

Using Mengert's illustrations, MysteryInc152's model learned to produce novel images that have easy-to-identify compositional dissimilarities from the pictured Mengert illustrations but that nevertheless share an overall aesthetic similarity. How should copyright law treat this sort of similarity? Some may view it as clear-cut noninfringement: despite overall aesthetic commonalities, the images evince significant compositional dissimilarities. But this sort of aesthetic similarity is something many others now view as morally and legally objectionable, and an evenhanded application of copyright doctrine indicates that stylistic similarity may indeed be cognizable infringement.[6]

Variants of the hard problem of stylistic similarity are now being litigated. In early 2023, two blockbuster lawsuits were filed against the artificial intelligence company Stability AI. Stability AI produces

4. *Id.*; ogkalu, *Illustration-Diffusion*, HUGGING FACE, https://huggingface.co/ogkalu/Illustration-Diffusion [https://perma.cc/HV6R-37A3].

5. Andy Baio, *Invasive Diffusion: How One Unwilling Illustrator Found Herself Turned into an AI Model*, WAXY.ORG (Nov. 1, 2022), https://waxy.org/2022/11/invasive-diffusion-how-one-unwilling-illustrator-found-herself-turned-into-an-ai-model/ [https://perma.cc/U775-X82H].

6. *See, e.g.*, Robert Brauneis, *Copyright and the Training of Human Authors and Generative Machines*, 47 COLUM. J.L. & ARTS (forthcoming 2025) (manuscript at 30–37), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4909592 [https://perma.cc/9YCN-FSTV].

StableDiffusion, an open-source generative artificial intelligence model that can turn a user's text "prompt" into a striking visual depiction of whatever the prompt describes. StableDiffusion "learned" to perform this feat by consulting billions of images copied from across the Web — including, allegedly, copyrighted works owned by two different sets of plaintiffs, who assert in these lawsuits that Stability AI copied their work without permission and created unauthorized derivative works. One suit, a putative class-action filed by visual artists in the United States District Court for the Northern District of California, alleges copyright infringement and violation of California's right of publicity arising out of Stability AI's copying of the plaintiffs' works of visual art and generation of similar works.[7] Another suit, filed by Getty Images in the District of Delaware, alleges copyright infringement arising out of Stability AI's copying of Getty's photographs to train its model as well as StableDiffusion's "produc[tion of] images that are highly similar to and derivative of the Getty Images proprietary content."[8]

The legal risks of training generative AI on unauthorized copies of copyrighted works have been apparent for many years[9] even if high-profile legal challenges to the practice emerged only recently. But these lawsuits against Stability AI also present a distinct, and under-examined, issue: how should we determine when the *output* of a generative AI model infringes copyright?

Commentators frequently downplay the possibility of copyright infringement by image-generating AI by arguing that these programs copy only the "style" of the artworks they train on and that style cannot be copyrighted.[10] Thus, the reasoning goes, the output of image-generating AI does not infringe the reproduction rights of the visual artists whose works trained the model even if that output is stylistically similar to earlier works.

This framing, however, is misleading. It relies on the slipperiness of a word that has long bedeviled copyright law: "style." Because style has such a broad meaning, it undoubtedly encompasses generic attributes of works that are beyond copyright's monopoly, like "vanitas" or

7. Complaint at 30–32, 36–39, Andersen v. Stability AI Ltd., No. 3:23-cv-00201 (N.D. Cal. Jan. 13, 2023) [hereinafter Andersen Compl.].

8. Complaint at 13, 17–18, 23, Getty Images (US), Inc. v. Stability AI, Inc., No. 1:23-cv-00135 (D. Del. Feb. 3, 2023) [hereinafter Getty Images Compl.].

9. *See, e.g.*, Benjamin L. W. Sobel, *Artificial Intelligence's Fair Use Crisis*, 41 COLUM. J.L. & ARTS 45, 45 (2017).

10. *See infra* note 14; *cf.* Jacob Alhadeff, Cooper Cuene & Max Del Real, *Limits of Algorithmic Fair Use*, 19 WASH J.L., TECH. & ARTS 1, 28–29 (2024) (asserting that "[b]ecause style and similar attributes are akin to ideas, not expression, they don't fall under the protection of copyright" but acknowledging that sufficiently similar output images could infringe).

"anime."[11] But in the visual arts in particular, "style" can also refer to more particularized expression that does fall within the scope of an artist's copyright.

Part II explains how image-generating AI works and documents early extrajudicial responses to the legal and ethical concerns the technology presents. Part II also debunks one of the most prominent metaphors used to describe generative AI: that it is, or is not, a "collage machine." The "collage machine" metaphor is unhelpful for two reasons. First, it erroneously posits that it is possible to differentiate so-called "mechanical" reproductions of works of visual art from so-called "intellectual" reproductions. Second, it implies that the distinction matters, when blackletter copyright doctrine holds that it does not. Generative AI is clearly learning to reproduce *something* from its training data: what matters is what that something is, and how the law ought to treat the copying. This requires discussing what "style" is, which in turn is the focus of Part III.

Part III defines style as a holistic attribute of a work, or a group of works, that comprises a constellation of expressive choices. These expressive choices might be unprotectable individually, but in combination, they may constitute protectable expression. Protectable style may not be limited to single "works," either; in fact, courts routinely aggregate plaintiffs' works to find emergent copyrightable interests. Part III documents courts' struggles to assess similarity in visual art and attributes these struggles to the substantial similarity test's near-irreconcilable demands: courts must simultaneously dissect images into their constituent elements — a task judges claim they are unable to do — while also assessing works' aesthetic appeal holistically and intuitively. Style has always been a challenge for substantial similarity because it is the form of expression least susceptible to analytical dissection and most likely to elicit inarticulate aesthetic intuitions. Generative AI models' replication of style is such a hard problem for copyright law because the models are purpose-built to identify and reproduce precisely the forms of similarity that are hardest to analyze legally.

A word of warning before we begin in earnest: this paper's project is descriptive, not normative. It rebuts the pervasive and uninformative assertion that image-generating AI only reproduces "style," that this thing called "style" is categorically uncopyrightable, and that copyright doctrine strictly forbids courts from finding copyrightable expression

11. *See* Williams v. 3DExport, No. 19-12240, 2020 WL 532418, at *3 (E.D. Mich. Feb. 3, 2020) (applying the idea-expression dichotomy to reject a plaintiff's claim to "own[] a copyright in the anime art style"); *see also* Yankee Candle Co. v. Bridgewater Candle Co., LLC, 259 F.3d 25, 32 n.2, 35 (1st Cir. 2001) ("copyright does not provide protection for the particular style of photography chosen by [plaintiff]," where photographic style in question was the use of a photograph with "no border or visual separation between the photograph and the perimeter of the label").

that transcends the bounds of any single "work." To the contrary: an honest application of copyright law requires us to acknowledge that some of what we call style is copyrightable some of the time and that in some legal contexts, courts regularly protect emergent copyright interests that span multiple works. No descriptive account can deny this: Courts really have reached these conclusions.

However, by describing current copyright doctrine, I am not endorsing the status quo. There are powerful arguments for narrowing the scope of copyright protection and for limiting copyright interests to a single work. But those are normative arguments, just as the slogan "you can't copyright style" is a normative argument. The problem with this normative slogan is that it is currently masquerading as a descriptive statement. This Article merely shows that the slogan is descriptively incorrect; its normative merits are another issue entirely.

## II. GENERATIVE AI TECHNOLOGY: THE "COLLAGE MACHINE?"

This Article focuses on image-generating artificial intelligence, primarily with reference to two of the most prominent image-generating AI services: OpenAI's DALL•E 2 and Stability AI's StableDiffusion. A third service I sometimes discuss, Midjourney, has used a "model" — or underlying technical architecture — based on StableDiffusion.[12] All of these services produce appealing works of visual art in response to text "prompts" entered by users. Under the hood, these image-generating artificial intelligence services employ a technology called a diffusion model.[13] Generative AI models are capable of producing output that bears a striking stylistic resemblance to the oeuvre of particular artists; to discrete, copyrighted works of art; and to recognizable human likenesses. Indeed, perhaps because it is often difficult to articulate specifically what generative AI has learned to reproduce, the technology is described as replicating artists' "styles."[14]

---

12. @EMostaque, X (Aug. 22, 2022, 11:25 PM), https://x.com/EMostaque/status/1561917541743841280 [https://perma.cc/FDG5-LHP6].

13. *See, e.g.*, Alex Nichol, Prafulla Dhariwal, Aditya Ramesh, Pranav Shyam, Pamela Mishkin, Bob McGrew et al., GLIDE: Towards Photorealistic Image Generation and Editing with Text-Guided Diffusion Models 1 (Mar. 8, 2022) (unpublished manuscript) (on file with arXiv), http://arxiv.org/abs/2112.10741 [https://perma.cc/PH4R-VJ8J]; Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser & Björn Ommer, High-Resolution Image Synthesis with Latent Diffusion Models 1 (Apr. 13, 2022) (unpublished manuscript) (on file with arXiv), http://arxiv.org/abs/2112.10752 [https://perma.cc/9K9M-Q6TU].

14. *See, e.g.*, Andres Guadamuz, *A Scanner Darkly: Copyright Infringement in Artificial Intelligence Inputs and Outputs*, 73 GRUR INT'L 111, 125 (2024); James Vincent, *AI Art Tools Stable Diffusion and Midjourney Targeted with Copyright Lawsuit*, THE VERGE (Jan. 16, 2023, 6:28 AM), https://www.theverge.com/2023/1/16/23557098/generative-ai-art-copyright-legal-lawsuit-stable-diffusion-midjourney-deviantart [https://perma.cc/6X3G-ZYP3];

The characteristic that makes StableDiffusion and similar technologies most controversial is that they are powered by massive repositories of images and text, culled from all over the World Wide Web. These media were created and published by unsuspecting Internet users unaffiliated with Stability AI. It is virtually certain that many of the images that trained StableDiffusion are protected by copyright, and that many of the holders of those copyrights did not authorize the use of their works to train generative AI models.[15]

Seven years ago, I described the copyright liabilities that may result simply from training an AI model to generate expressive content using unauthorized copies of copyrighted works, irrespective of similarities between the model's outputs and its training data.[16] I suggested that the fair use doctrine may not permit training of expressive machine learning on unauthorized copies of copyrighted works.[17] The distinct focus of this Article is on the liabilities that may result when a trained model produces output that resembles the data that trained the model. Through statistical analysis of these "training data," generative AI models "learn" to produce images that reflect what users' text prompts describe. For example, by analyzing Hollie Mengert's work, the software can generate images that seem to reproduce Mengert's artistic style. For this reason, critics of generative AI models disparage them as "collage machines" that simply rehash the artistic works that they ingested as training data.

In the following sections in Part II, I briefly inventory the data that generative AI companies use to train their models; describe how these companies, and other actors, have begun to respond to legal and ethical concerns about generative AI; and offer a brief technical explanation of how diffusion models work. Lastly, I build on this technical exposition to explain that arguments over the "collage machine" metaphor are misguided. What the "collage machine" debate is really about is not *whether* generative AI mechanically reproduces the work of human artists, but rather about *what* aspects of those works it mechanically

---

Andres Guadamuz, *Artificial Intelligence and Copyright*, WIPO MAGAZINE (Oct. 2017), https://www.wipo.int/wipo_magazine/en/2017/05/article_0003.html [https://perma.cc/89DA-78HN]; Stephen Wolfson, *The Complex World of Style, Copyright, and Generative AI*, CREATIVE COMMONS (Mar. 23, 2023), https://creativecommons.org/2023/03/23/the-complex-world-of-style-copyright-and-generative-ai/ [https://perma.cc/Y7US-JFSU].

15. Several lawsuits so allege. *See* Second Amended Complaint at 12, 16–17, Andersen v. Stability AI Ltd., No. 3:23-cv-00201 (N.D. Cal. Oct. 31, 2024); Second Amended Complaint at 1, Getty Images (US), Inc. v. Stability AI Inc., No. 1:23-cv-00135 (D. Del. Jul. 8, 2023); *see also* Matthew Sag, *Copyright Safety for Generative AI*, 61 HOUS. L. REV. 295, 303 & n.16 (2023).

16. *See* Sobel, *supra* note 9; Benjamin Sobel, *A Taxonomy of Training Data: Disentangling the Mismatched Rights, Remedies, and Rationales for Restricting Machine Learning, in* ARTIFICIAL INTELLIGENCE AND INTELLECTUAL PROPERTY 221, 222 (Jyh-An Lee, Reto M Hilty & Kung-Chung Liu eds., 2021).

17. *See* Sobel, *supra* note 9, at 66–79.

reproduces. Assessing the legal significance of *that* question requires discussing what "style" is, which in turn is the focus of Part III.

### *A. The Data that Power AI*

The leading image-generating AI tools were trained on incomprehensibly large datasets of text and images, culled from all over the Internet without rightsholders' express permission. OpenAI's DALL•E 2 uses an architecture that was trained on the 400,000,000-image WebImageText ("WIT") dataset, which OpenAI says was "collected form [*sic*] a variety of publicly available sources on the Internet" and assembled by "search[ing] for (image, text) pairs as part of the construction process whose text includes one of a set of 500,000 queries."[18] WIT has been characterized as "a web scrape," which would mean that it was the product of automated queries to websites that downloaded webpages' content.[19]

While WIT is not public, comparable datasets are. Perhaps the most notable open-source datasets are those used by StableDiffusion, which are promulgated by the "Large-scale Artificial Intelligence Open Network" ("LAION") organization.[20] LAION offers for download a dataset containing 5.85 billion image-text pairs, as well as LAION-400M, a dataset of 400 million English image-text pairs.[21] LAION also offers a subset of data filtered for high scores on LAION's "Aesthetics_Predictor," a model trained to predict how favorably people would rate the aesthetic appearance of an image. LAION-400M was compiled by parsing data from the Common Crawl project, which is a nonprofit initiative that "maintains a free, open repository of web crawl data that

18. Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal et al., Learning Transferable Visual Models From Natural Language Supervision 3–4 (Feb. 26, 2021) (unpublished manuscript) (on file with arXiv), http://arxiv.org/abs/2103.00020 [https://perma.cc/6U7R-LBA5].

19. Robert Wolfe & Aylin Caliskan, *Markedness in Visual Semantic AI*, 2022 ACM CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY 1269, 1272 (Jun. 2022), https://dl.acm.org/doi/10.1145/3531146.3533183 [https://perma.cc/RHY9-HQGD]. *See generally* Benjamin L. W. Sobel, *A New Common Law of Web Scraping*, 25 LEWIS & CLARK L. REV. 147 (2021) (discussing the practice of web scraping and the legal liabilities it may engender).

20. Marissa Newman & Aggi Cantrill, *The Future of AI Relies on a High School Teacher's Free Database*, BLOOMBERG (Apr. 23, 2023, 11:00 PM), https://www.bloomberg.com/news/features/2023-04-24/a-high-school-teacher-s-free-image-database-powers-ai-unicorns [https://perma.cc/QFS7-F8HR].

21. Christoph Schuhmann, Richard Vencu, Romain Beaumont, Robert Kaczmarczyk, Clayton Mullis, Aarush Katta et al., LAION-400M: Open Dataset of CLIP-Filtered 400 Million Image-Text Pairs 2 (Nov. 3, 2021) (unpublished manuscript) (on file with arXiv), http://arxiv.org/abs/2111.02114 [https://perma.cc/AU5V-59R4]; Christoph Schuhmann, Romain Beaumont, Richard Vencu, Cade Gordon, Ross Wightman, Mehdi Cherti et al., LAION-5B: An Open Large-Scale Dataset for Training next Generation Image-Text Models 1 (Oct. 16, 2022) (unpublished manuscript) (on file with arXiv), https://arxiv.org/abs/2210.08402v1 [https://perma.cc/HB2U-R449].

can be used by anyone."[22] Common Crawl's data are downloaded automatically by a "crawler," a program that visits publicly accessible websites and downloads the data posted to those webpages in bulk. LAION extracted from Common Crawl data all images that had been captioned in accordance with a particular standard and processed the remaining data for quality and usability.[23] Allegedly, the LAION dataset reproduces copyrighted materials without rightsholders' authorization.[24]

Because Stability AI and LAION disclose their datasets, third parties have begun analyzing those datasets.[25] An analysis of one LAION dataset found that it contained images scraped from commercial stock image sites; user-generated content sites like Pinterest, WordPress, Flickr, and Tumblr; and online retailers that sell art prints and other merchandise.[26] That same analysis produced a list of the artists most commonly featured in the dataset: "Of the top 25 artists in the dataset, only three are still living: Phil Koch, Erin Hanson, and Steve Henderson. The most frequent artist in the dataset? The Painter of Light™ himself, Thomas Kinkade, with 9,268 images."[27]

Many of the works in this LAION dataset are probably copyright-protected. Copyright in works created on or after January 1, 1978 generally subsists for the life of the author plus seventy years, while copyright in works that were already copyright-protected as of that date lasts for 95 years from the date their copyright was originally secured.[28] Thomas Kinkade died in 2012.[29] The 28th-most-frequent artist is Bob Ross, a painter and television personality who died in 1995.[30] The portrait artist Kehinde Wiley is 46th, right behind Andy Warhol, who died in 1987.[31] "Pixar" — the computer animation studio responsible for

22. Schuhmann et al., *LAION-400M*, *supra* note 21; Common Crawl, COMMON CRAWL, https://commoncrawl.org/ [https://perma.cc/MHE6-5QZE].

23. Schuhmann et al., *LAION-400M*, *supra* note 21.

24. Andersen Compl., *supra* note 7, at 23, 24.

25. Andy Baio, *Exploring 12 Million of the 2.3 Billion Images Used to Train Stable Diffusion's Image Generator*, WAXY.ORG (Aug. 30, 2022), https://waxy.org/2022/08/exploring-12-million-of-the-images-used-to-train-stable-diffusions-image-generator/ [https://perma.cc/AN8W-MKS7].

26. *Id.*

27. *Id.*

28. 17 U.S.C. §§ 302, 304.

29. Julia Prodis Sulek, *Thomas Kinkade Autopsy: Alcohol, Valium and Smudges of Green Paint*, THE MERCURY NEWS (May 8, 2012, 4:37 AM), https://www.mercurynews.com/2012/05/08/thomas-kinkade-autopsy-alcohol-valium-and-smudges-of-green-paint/ [https://perma.cc/SJ27-QC5G].

30. Naomi Blumberg, *Bob Ross*, BRITANNICA, https://www.britannica.com/biography/Bob-Ross [https://perma.cc/4W9U-6HD7].

31. Kehinde Wiley, *Kehinde Wiley Studio*, KEHINDE WILEY, https://kehindewiley.com/ [https://perma.cc/R2S5-QT5Q]; *Archives + Research*, THE ANDY WARHOL MUSEUM, https://www.warhol.org/andy-warhols-life/ [https://perma.cc/6MNA-URCM].

over a dozen blockbusters from the past thirty years — is listed as 61st.[32]

### *B. Responses from Artists and Industry*

Unsurprisingly, not every artist is overjoyed that generative AI models are able to create images that share aesthetic similarities with their work. By now, even the mainstream media is attuned to legal, artistic, and ethical quandaries that such technologies present — particularly when those technologies were trained without the consent of artists whose works they ingested. An article in *CNN Business* featured quotes from artists "who were unhappy to learn that pictures of their work were used [to train AI like StableDiffusion] without someone informing them, asking for consent, or paying for their use."[33] One artist remarked, "I don't want to participate at all in the machine that's going to cheapen what I do."[34] An article in *Slate* quoted the illustrator Iris Luckhaus on her discovery that 3,280 of her illustrations had been used to train StableDiffusion, "including early drafts of illustrations that she hadn't realized were findable on the internet":

> "I felt violated and terribly angry," Luckhaus said. Just because she puts artworks online, that's "not an invitation to just do with them as you please" — in this case, to create a tool that she is concerned will replace human artists. She refuses to use A.I. image generators because exploiting other artists makes her feel just "as dirty as buying cheap clothes made with child labor," she said.[35]

By contrast, those who generally oppose requiring copyright owners to authorize the use of their works to train generative AI argue that copyright "has never included a monopoly over the basic building blocks of creativity: ideas, concepts, style, artistic technique, language, or grammar," and that it is these properties that AI extracts from

---

32. John M. Cunningham, *Pixar*, BRITANNICA, https://www.britannica.com/topic/Pixar-Animation-Studios [https://perma.cc/B8ZS-8NEZ].

33. 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/3FBK-QYLT].

34. *Id.*

35. Heather T. Murphy, *A.I. Is Sucking the Entire Internet In. What If You Could Yank Some of It Back Out?*, SLATE MAGAZINE (Mar. 27, 2023, 5:50 AM), https://slate.com/technology/2023/03/how-holly-herndon-and-mathew-dryhurst-brokered-an-a-i-deal-with-stable-diffusion.html [https://perma.cc/8LYZ-6N3M].

training data.[36] In that vein, Mark Lemley and Bryan Casey comment that "a copyright only controls certain uses: copying, distributing, publicly performing, and the like. Notably absent from that list are certain activities fundamental to learning, such as watching, reading, and discussing a work and communicating its unprotectable elements to others" — although they observe elsewhere that "fair use is unlikely to save" AI applications that have the purpose and effect of producing output similar to the work of a particular human artist.[37]

As the above suggests, essential legal questions have yet to be adjudicated, and popular moral intuitions haven't yet coalesced. Industry has taken some precautions that seem designed to limit liability for copyright infringement.[38] Some companies have announced, in vague terms, initiatives that would share revenue with copyright holders whose works were used to train their models.[39] All of the major image-generating services — DALL•E 2, Midjourney, and StableDiffusion — place contractual and technological limits on functionality. Their terms of service typically set certain categories of imagery — such as violence and pornography — off-limits.[40] Both DALL•E and MidJourney employ filtering technology that automatically blocks certain prompts from causing output to be generated.[41] StableDiffusion interposes fewer centralized restrictions on its users, but its developers have modified the software to make it more difficult to generate "nude and pornographic output, photorealistic pictures of celebrities, and images that mimic the artwork of specific artists."[42] Stability AI has also permitted

---

36. Letter from American Library Association et al. to Members of the U.S. House of Representatives and U.S. Senate (Sept. 11, 2023) (on file with the Authors Alliance), https://www.authorsalliance.org/wp-content/uploads/2023/09/AI-Coalition-Letter-9.11.2023-updated.pdf [https://perma.cc/8YCF-RPMJ].

37. Mark A. Lemley & Bryan Casey, *Fair Learning*, 99 TEX. L. REV. 743, 773, 778 (2021).

38. Platforms certainly haven't taken *every* precaution. Members of Microsoft's now-disbanded Ethics and Society team reportedly recommended that Microsoft "block users from inputting the names of living artists as prompts" in its Bing Image Creator service, "or create a marketplace to support artists whose work was surfaced in search." Microsoft adopted neither recommendation. Lauren Leffer, *Microsoft Scraps Entire Ethical AI Team Amid AI Boom*, GIZMODO (2023), https://gizmodo.com/microsoft-ai-ethical-ai-bing-chatgpt-layoffs-1850223358 [https://perma.cc/8D6G-GNL2].

39. Shutterstock, *Shutterstock Partners with OpenAI and Leads the Way to Bring AI-Generated Content to All*, SHUTTERSTOCK (Oct. 25, 2022), https://www.shutterstock.com/press/20435 [https://perma.cc/3Y27-FQ2F].

40. *DALL·E Content Policy: Are There Any Restrictions to How I Can Use DALL·E 2? Is There a Content Policy?*, OPENAI, https://help.openai.com/en/articles/6338764-are-there-any-restrictions-to-how-i-can-use-dall-e-2-is-there-a-content-policy [https://perma.cc/475Y-5P7A]; Terms of Service, MIDJOURNEY, https://docs.midjourney.com/docs/terms-of-service [https://perma.cc/9K6V-FZYV].

41. *See supra* note 40; *DALL-E Content Policy: I Received a Warning while Using DALL·E 2. Will I Be Banned?*, OPENAI, https://help.openai.com/en/articles/6338765-i-received-a-warning-while-using-dall-e-2-will-i-be-banned [https://perma.cc/D43G-LLK3].

42. James Vincent, *Stable Diffusion Made Copying Artists and Generating Porn Harder and Users Are Mad*, THE VERGE (Nov. 24, 2022, 10:17 AM),

artists to opt out from being included in training data for an updated version of their software.[43]

### *C. Diffusion Models: Mapping Interstitial Space in Our Words and Images*

The previous sections explained in broad terms that generative AI ingests vast amounts of images and text in order to "learn" how to generate outputs that resemble those images in response to written prompts. This section describes that technical process in greater detail to explain why the technology presents particularly intricate problems for questions of style and similarity.

Both DALL•E 2 and StableDiffusion employ a neural network[44] component that analyzes naturally-occurring pairings of images and text — collected from all over the Web — in order to predict what sorts of images are most likely to correspond to given textual descriptions.[45] This component technology, when supplied with a mountain of image-caption pairings, uses statistical methods to match images with their corresponding captions.[46] It does so by encoding images and corresponding text captions into formats that facilitate such analysis, then training to maximize a measurement of similarity between an image and its correctly paired caption and to minimize that similarity measurement as between all other image-caption pairs.[47]

An important conceptual detail about models like StableDiffusion and DALL•E 2 is that they, like many applications of machine learning, manipulate their inputs and outputs by first representing them as vectors.[48] That is, inputs and outputs are manipulated not as strings of text

---

https://www.theverge.com/2022/11/24/23476622/ai-image-generator-stable-diffusion-version-2-nsfw-artists-data-changes [https://perma.cc/LF45-X79R].

43. Jonathan Kemper, *Artists Remove 80 million Images from Stable Diffusion 3 Training Data*, THE DECODER (Mar. 8, 2023), https://the-decoder.com/artists-remove-80-million-images-from-stable-diffusion-3-training-data/ [https://perma.cc/T5BN-FK9F].

44. "A neural network is a method in artificial intelligence that teaches computers to process data in a way that is inspired by the human brain. It is a type of machine learning process, called deep learning, that uses interconnected nodes or neurons in a layered structure that resembles the human brain." *What is a Neural Network? — Artificial Neural Network Explained — AWS*, AMAZON WEB SERVICES, INC., https://aws.amazon.com/what-is/neural-network/ [https://perma.cc/Q3U7-EXB9].

45. This component technology is called CLIP, short for "Contrastive Language-Image Pre-training." Radford et al., *supra* note 18, at 2; *CLIP: Connecting Text and Images*, OPENAI (Jan. 5, 2021), https://openai.com/blog/clip/ [https://perma.cc/VZ3H-LDGZ].

46. Radford et al., *supra* note 18, at 4 ("Given a batch of N (image, text) pairs, CLIP is trained to predict which of the N x N possible (image, text) pairings across a batch actually occurred.").

47. *Id.*

48. *See, e.g.*, *id.*

or as pixels, but instead as vectors.[49] The components of those vectors represent "features" in the data they represent[50] which, for present purposes, can be understood as properties of the data that statistical analysis reveals to be salient. As the inputs are processed, they are abstracted in a way that reduces their "dimensionality." High dimensionality can make data computationally taxing to analyze, and it may obscure relationships in the data — computer scientists call this problem the "curse of dimensionality."[51]

One dimensionality-reduction technique used in StableDiffusion and architectures like it is the autoencoder.[52] Autoencoders function by "first encod[ing] [an] image into a lower dimensional latent representation, then decod[ing] the latent representation back to an image."[53] Because autoencoders typically "are restricted in ways that allow them to copy only approximately, and to copy only input that resembles the training data[,] . . . [they] often learn[] useful properties of the data."[54] Importantly, these learned features needn't correspond to attributes that humans believe to be salient about the works in question; indeed, they need not correspond even to attributes that humans can articulate.[55]

In lay terms, the encoding process described in the previous paragraph amounts to asking the model to determine what attributes of the data are most salient and then to represent those attributes efficiently. This analysis yields a "latent space": a multi-dimensional space that represents data according to how similar they are along their many dimensions. This latent space can encode similarities that may not be ascertainable from inspecting raw input or output information itself, as represented in text or in pixels. To generate images, models like

---

49. Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez et al., Attention Is All You Need 5 (June 12, 2017) (unpublished manuscript) (on file with arXiv), https://arxiv.org/abs/1706.03762 [https://perma.cc/8Q5D-CGUE]; 30 ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS, 5 (2017); Matthew Sag, *Copyright Safety for Generative AI*, 61 HOUS. L. REV. 295, 318, 319 (2023).

50. Radford et al., *supra* note 18, at 5.

51. *See* Jacob Murel & Eda Kavlakoglu, *What is dimensionality reduction?*, IBM (Jan. 5 2024), https://www.ibm.com/topics/dimensionality-reduction [https://perma.cc/YR46-LEAF].

52. Rombach et al., *supra* note 13, at 2.

53. *Intro to Autoencoders*, TENSORFLOW, https://www.tensorflow.org/tutorials/generative/autoencoder [https://perma.cc/2MFZ-VU2B].

54. IAN GOODFELLOW, YOSHUA BENGIO & AARON COURVILLE, DEEP LEARNING 499 (2016), https://www.deeplearningbook.org/contents/autoencoders.html [https://perma.cc/93FN-L9XU].

55. *Cf.* Sag, *supra* note 15, at 316–19; A. Feder Cooper & James Grimmelmann, The Files Are in the Computer: On Copyright, Memorization, and Generative AI 30–35, (Apr. 19, 2024) (unpublished manuscript) (on file with arXiv), http://arxiv.org/abs/2404.12590 [https://perma.cc/R9VC-8FK8].

StableDiffusion and DALL•E 2 decode an embedding from the incomprehensible latent space into the visually comprehensible pixel space.[56]

The latent space is space, which means that it can be traversed as one might traverse a physical area. Each point within the space can be decoded into a perceivable image.[57] And any two points in the space can be connected by a path that itself comprises points that are themselves valid, decodable images.[58] This property of the latent space makes it possible to visualize the space between two points. The graphic in Figure 2 shows some of the images that researchers generated by traversing the latent space in StableDiffusion between `watercolor painting of a Golden Retriever at the beach` and `still life DSLR photo of a bowl of fruit`.[59] As is perhaps easier to see than to describe, the images reflect intermediate degrees of similarity between the two text prompts that mark the start and end of the journey through latent space.

56. *See, e.g.*, Aditya Ramesh, Prafulla Dhariwal, Alex Nichol, Casey Chu & Mark Chen, Hierarchical Text-Conditional Image Generation with CLIP Latents 3 (Apr. 13, 2022) (unpublished manuscript) (on file with arXiv), http://arxiv.org/abs/2204.06125 [https://perma.cc/7J67-MHP9].

57. Ian Stenbit, François Chollet & Luke Wood, *A Walk Through Latent Space with Stable Diffusion*, KERAS (Sept. 28, 2022), https://keras.io/examples/generative/random_walks_with_stable_diffusion/ [https://perma.cc/WR4K-JLCR].

58. *Id.*

59. *Id.*

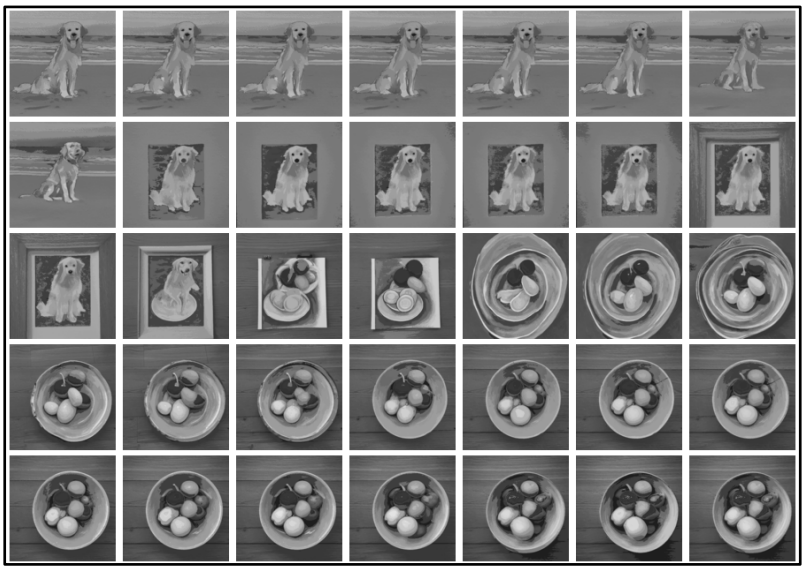


Figure 2: Images Extracted from Ian Stenbit, François Chollet & Luke Wood.[60]

This brings us to a point that's profound but not complex: generative AI is a means of navigating the spaces between and within our linguistic and visual conventions of representation. Because models like DALL•E 2 and StableDiffusion have learned from our own linguistic and visual conventions, as represented by the public Internet, they provide a means of mapping the interstices in our own expressive vocabulary. Even if we have no pre-existing term or concept to describe an image halfway between pointillism and abstract expressionism in appearance, this technology allows us to situate that hypothetical image in a representational space "between" the areas where those pointillist images and abstract expressionist images cluster, and then to generate a compelling visual representation of the chimerical image. This technique also allows generative AI models to distill and reproduce characteristic aspects of artists' styles. Even if the precise features that unify an artist's oeuvre are difficult to articulate, generative AI's statistical techniques can compellingly replicate stylistic attributes. Representing text and images as coordinates in space makes it possible to generate

60. Ian Stenbit, François Chollet & Luke Wood, *A Walk Through Latent Space with Stable Diffusion*, KERAS (Sept. 28, 2022), https://keras.io/examples/generative/random_walks_with_stable_diffusion/ [https://perma.cc/WR4K-JLCR].

images that are, in a new and meaningful sense, "like" other images or categories of images.[61]

Of course, other, more familiar encodings of images also represent them spatially. A digital image file encodes a grid that instructs a computer to display one red pixel here, one green pixel there, and so on. But if one simply "interpolates" two images in pixel-space rather than in latent space, the result is a mess that looks like two images thrown on top of one another (see Figure 3). By contrast, performing the interpolation in latent space — that is, navigating to an intermediate location in latent space and then decoding that image into pixel space — yields an image that looks much more coherent and appealing. If interpolating images in pixel space gets you half an orange glued to half a pomelo, then interpolating images in latent space gets you a tangelo.

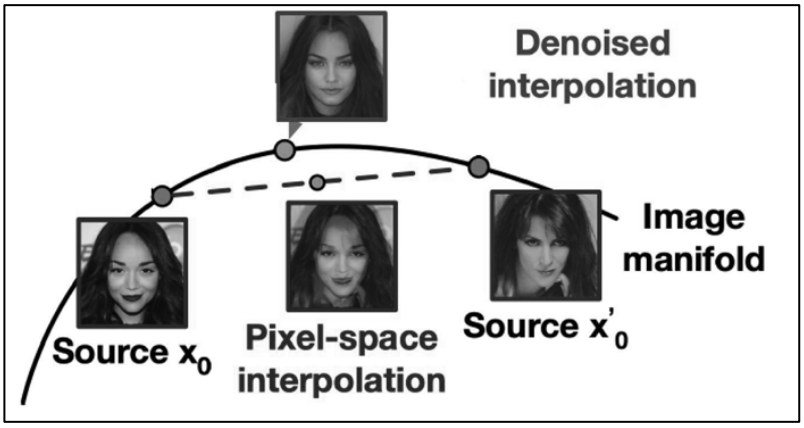



Figure 3: Interpolating in Pixel vs. Latent Space[62]

The process of interpolating images in latent space has proven to be a site of deeply contested conceptual dispute between and among lawyers, scholars, and artists. Is interpolating an image mere copying, like throwing two images together in pixel space, or does the process of interpolation and diffusion turn generative AI into something different in kind from copying and pasting? In my view, there's no meaningful, descriptive answer to this question — and that's fine, because despite the urgency of today's debate, the question has no legal or aesthetic significance. The next section justifies that view.

---

61. *See, e.g.*, Jonathan Ho, Ajay Jain & Pieter Abbeel, Denoising Diffusion Probabilistic Models 8 (Dec. 16, 2020) (unpublished manuscript) (on file with arXiv), https://arxiv.org/abs/2006.11239v2 [https://perma.cc/6W4V-BQR3].

62. *Id.*, reproduced in Andersen Compl., *supra* note 7, at 18 ¶ 81 (original annotations removed for clarity).

### *D. Forget About The "Collage Machine"*

A central metaphor in the *Andersen* complaint is that StableDiffusion is "merely a complex collage tool."[63] The complaint alleges that any given image outputted by StableDiffusion is "necessarily a derivative work, because it is generated exclusively from a combination of the [users' text prompts] and the latent images, all of which are copies of copyrighted images. [StableDiffusion] is, in short, a 21st-century collage tool."[64] The "collage machine" metaphor implies (a) that there is a distinction between "collage machines" and other modes of creative production, and (b) that from this distinction there follows a meaningful difference. Thus far, the commentary I've seen — on either side of the issue — seems to assume that (a) such a distinction can be proven or disproven and (b) it makes a difference. Both assumptions are wrong. Explaining why illuminates long-misunderstood concepts of copying and similarity that are central to the copyright issues that generative AI raises.

The "collage machine" metaphor did not originate in the *Andersen* complaint. Before it appeared in a court filing, the hypothesis emerged as the subject of an online debate about the moral, legal, and aesthetic propriety of image-generating AI and its users' obligations to visual artists. On one side of this debate are commentators who claim that generative AI, by its very nature, outputs images that derive improperly from source materials. This is so, such critics maintain, because while human intellects use source material as "inspiration," generative AI uses source images as "samples."[65] Opponents respond that "AI isn't an automatic collage machine," because "the AI is not storing the image information anywhere to copy and paste bits and pieces of it."[66]

The "collage machine" debate is in fact a debate about two undertheorized questions. The first is whether we should regard image-generating AI as being more like a photocopier or a sketch artist. In other words, do the outputs of generative AI *mechanically reproduce* training data, or is their resemblance to real-world phenomena the result of some process other than "mere" copying and pasting?

The second unexamined question is why we should care whether generative AI is a collage machine. A tacit axiom held by commentators on both sides of the debate seems to be that forms of creativity that involve a bodily appropriation or mechanical reproduction of

63. Andersen Compl., *supra* note 7, at 1.

64. Andersen Compl., *supra* note 7, at 20.

65. @ZedEdge, X (Nov. 19, 2022, 7:59 PM), https://x.com/ZedEdge/status/1594133323710070785 [https://perma.cc/CSN3-XHMY].

66. @ReachArtwork, X (Aug. 31, 2022, 3:30 AM), https://x.com/ai_curio/status/1564878372185989120 [https://perma.cc/JNZ3-8N2V] (capitalization altered); @ReachArtwork, X (Aug. 31, 2022, 3:49 AM), https://x.com/reachartwork/status/1564883115822268416 [https://perma.cc/6MMD-FHH2].

preexisting works of art are in some way inferior to, and/or legally distinct from, modes of expression that do not. In other words, cutting and pasting is suspect, but paraphrasing is more acceptable. As we will see, this proposition is legally questionable.[67]

The problem with the first question — "is StableDiffusion a collage machine?" — is that its framing suggests that a technical answer is possible. The implication is that if we could just disentangle the complex inner workings of StableDiffusion, we could figure out whether it acts like a collage machine. Is it "copy[ing] and past[ing] bits and pieces of" images in there, or not?

It is misleading to suggest that there is a descriptive answer to whether StableDiffusion is a collage machine. For starters, the inquiry presupposes that we can identify what a "collage machine" is. It's not clear that we can. Presumably, to those who use the "collage machine" metaphor to disparage StableDiffusion, the human brain is not a collage machine. Humans can reproduce expression through an *intellectual* process — like an art student sketching a masterwork in a museum — rather than just cutting out a picture and pasting it elsewhere.[68] But explaining precisely what differentiates "intellectual" reproductions from merely mechanical reproductions is tricky. Is the human brain a collage machine when the human it belongs to is making a collage?

If proponents of the "collage machine" metaphor have failed to explain why StableDiffusion *is* a collage machine, opponents of the metaphor have done just as bad a job explaining why StableDiffusion *isn't* a collage machine. For example, Andres Guadamuz argues that the "collage machine" metaphor "couldn't be further from the truth" because "the outputs [of a generative AI model] are similar [to the model's training data], not the same."[69] And? Say I set out to make a collage of images from a glossy, color magazine. Does my creation cease to be a collage if I apply a wash to the canvas after arranging my collage upon it, giving the entire work a blueish hue absent from the original magazine photographs? What if, instead of cutting pictures out of the magazine directly, I use black-and-white photocopies of the images in the magazine? Or hyper-color-saturated photocopies? What if I make the photocopied images larger or smaller to use in my collage, or change their orientation? What if I photocopy onto matte paper, not glossy? I'm still making a collage in these scenarios, notwithstanding

67. It's probably aesthetically questionable, too, although that question is outside this Article's scope.

68. *Cf.* Bridgeport Music, Inc. v. Dimension Films, 410 F.3d 792, 802 (6th Cir. 2005) (describing the sampling of sound recordings as a "physical taking rather than an intellectual one").

69. Guadamuz, *supra* note 14, at 113.

that the images I'm using are just "similar to," and not "the same as" the original images in the magazine.[70]

Just as unsuccessful are attempts to refute the "collage machine" metaphor on seemingly technical grounds. James Vincent writes in *The Verge* that tools like StableDiffusion are not collage machines because "AI art models do not store images at all, but rather mathematical representations of patterns collected from these images. The software does not piece together bits of images in the form of a collage, either, but creates pictures from scratch based on these mathematical representations."[71] The logic that allows Vincent to claim that AI models don't "store images" also dictates that laptop computers don't "store images" either. If you pry open your laptop, you won't find a sheaf of images in there — no matter how many pictures you have saved on your hard drive. A digital photo library is a collection of bytes that instructs a computer to display particular images under particular circumstances. Conventional image file formats are just "mathematical representations of patterns" that appear in the images to which they correspond. An image file might instruct a computer, "Display a 50x50 pixel grid of alternating rows of white and red pixels." That's a mathematical representation of a white-and-red-striped square. It's just a mathematical representation that happens to be more comprehensible to us than latent images in a diffusion model are.

Nor can Vincent's distinction between "piec[ing] together bits of images in the form of a collage" and "creat[ing] pictures from scratch" hold. It has all the problems of Guadamuz's "similar to" versus "same as" analysis, and then some. Once we acknowledge that *all* digital images are mathematical representations of visual phenomena, it becomes impossible to distinguish between "piec[ing] together bits of images" and "creat[ing] pictures from scratch" unless we can explain why we ought to treat some mathematical representations differently from others. Vincent's and Guadamuz's claims are really that the bits and pieces of image-related data that StableDiffusion pieces together aren't the bits and pieces that *they* think are constitutive of the images in its training data. Despite sometimes being cloaked in technical terminology, this is just an unsupported normative assertion.

The reason that both the proponents and opponents of the "collage machine" metaphor are unsuccessful is that both sides are simply trying

---

70. Consider the Second Circuit's account in *Blanch v. Koons* of how the artist Jeff Koons made a series of collages: "To create the . . . paintings, Koons culled images from advertisements or his own photographs, scanned them into a computer, and digitally superimposed the scanned images against backgrounds of pastoral landscapes. He then printed color images of the resulting collages for his assistants to use as templates for applying paint to billboard-sized, 10' x 14' canvasses." Blanch v. Koons, 467 F.3d 244, 247 (2d Cir. 2006). Clearly, in the Second Circuit's view, Jeff Koons was still making a "collage" even though it involved scanning images from magazines and manipulating them digitally.

71. Vincent, *supra* note 14.

to define what forms of reproduction are "mechanical" and what forms are "intellectual." This is a question with no descriptive answer. Consider that today, photographs are commonly regarded as mechanical reproductions of their subject matter — but that wasn't always true. Courts in the 19th century heard arguments about whether photographs constituted direct evidence or hearsay.[72] And even our attitudes about supposedly mechanical modes of reproduction may not be wholly coherent: modern copyright law posits that photographs bear a photographer's intellectual imprint. On multiple occasions, courts have accepted that photographs can be more than mechanical reproductions of the visual world, and thus copyrightable as original works of authorship.[73]

If the history of photography illustrates the difficulty of deciding when a new technology is one of "mechanical" or "intellectual" reproduction, then the present state of photography and similar technologies shows that we haven't yet found the correct classification. For example, so-called "Space Zoom" technology in state-of-the-art Samsung smartphone cameras recently led to accusations of "fake" photographs, because "Space Zoom" deploys extensive computational processing to enhance the detail of blurry photographs of the moon (or, arguably, to "create[]" details not present in the optical data the camera's sensor receives).[74] Is it also "fake" when computational processing stitches together multiple images to produce a single picture with more detail in focus, or more variation in lighting, than could be captured by a lens?[75] We might also be surprised to learn that modern copy machines don't necessarily "copy" all of the visual information on the scanning bed. Instead, they record a compressed image that reuses visually similar information in order to save space.[76] Almost always, we don't notice the difference. But sometimes, when small details do encode salient information — like numerical labels in blueprints — this compression algorithm can cause copiers to produce "copies" we recognize as unfaithful.[77]

The answer, then, to our first question — "is StableDiffusion a collage machine?" — is that there's no meaningful answer. The question,

---

72. Rodney G. S. Carter, *"Ocular Proof": Photographs as Legal Evidence*, 69 ARCHIVARIA 23, 33 (2010).

73. *See, e.g.*, Burrow-Giles Lithographic Co. v. Sarony, 111 U.S. 53, 59–60 (1884); Falk v. Donaldson, 57 F. 32, 33–34 (C.C.S.D.N.Y. 1893).

74. James Vincent & John Porter, *Samsung Caught Faking Zoom Photos of the Moon*, THE VERGE (Mar. 13, 2023 10:16 AM) https://www.theverge.com/2023/3/13/23637401/samsung-fake-moon-photos-ai-galaxy-s21-s23-ultra [https://perma.cc/4MQ6-KWTD].

75. *See* Rebecca Tushnet, *Worth a Thousand Words: The Images of Copyright*, 125 HARV. L. REV. 683, 729 (2012).

76. Ted Chiang, *ChatGPT Is a Blurry JPEG of the Web*, THE NEW YORKER (Feb. 9, 2023), https://www.newyorker.com/tech/annals-of-technology/chatgpt-is-a-blurry-jpeg-of-the-web [https://perma.cc/2FAT-64JK].

77. *Id.* (describing a copy machine erroneously reproducing dimensional labels on blueprints).

although couched as a factual proposition, is just a proxy for a normative question: "should we agree to treat StableDiffusion like a paintbrush, or like a photocopier?" It is clear that when generative AI models produce output, they mechanically reproduce *something* about the works on which they trained. The real debate is over what that something is, and how we should treat it. The shorthand that has emerged to describe some of what is being reproduced is "style"; the "collage machine" debate is in many respects merely a proxy for the weightier question of what style is and how the law ought to treat it.

The insight that the collage machine metaphor is impossible to prove or to refute suggests an answer to our second question, "What does it matter if StableDiffusion is a collage machine or not?" As the following sections argue, it doesn't matter at all. A copyist can infringe copyright in a work of visual art by reproducing it with a paintbrush or with a photocopier. What matters in determining whether infringement has taken place is what the copyist copied from the artwork.[78] So, to

78. In a recent manuscript, Matthew Sag briefly presents an alternative to the "collage machine" metaphor that correctly focuses on the question of what information is being copied, rather than the question of how the copying occurs. "Generative AI," Sag writes, "is more like [p]apier-mâché than collage." Matthew Sag, *Copyright Safety for Generative AI*, 61 HOUS. L. REV. 295, 321 (2023). This is so, Sag explains, because a papier-mâché object created from printed text — like the output of generative AI — may "reveal[] hints of the copyrighted works in its substrate, [but] it has no meaningful similarity to any of them." *Id.* "The critical difference," Sag concludes, "is that for collage, it is plausible that the new work could infringe on the copyright of some underlying image; for papier-mâché it is not." *Id.* at 322.

Sag's analogy is flawed because it implies that AI is indifferent to the expressive content of its training data. Papier-mâché does not require paper with text on it; in fact, professional mixes use recycled paper shreds with no discernible print. *See, e.g.*, The ARTnews Recommends Editors, *The Best Paper-Mache Mixes for Inventive Sculpting*, ARTNEWS.COM (Oct. 3, 2020), https://www.artnews.com/art-news/product-recommendations/best-paper-mache-mixes-1234572358/ [https://perma.cc/TX99-3TCX]. Moreover, the instances in which a papier-mâché project does require paper with text printed on it are precisely the instances in which the resulting work may evince a "meaningful similarity" to the printed works it comprises. Sag, *supra* note 15, at 321. Imagine an artist sculpts a papier-mâché tree out of news articles on climate change and leaves the result unpainted, so that the newsprint is legible. Such a sculpture could be substantially similar to the literary works it displays (although in this hypothetical, fair use and the first-sale doctrine would almost certainly negate any copyright liability).

Unlike papier-mâché, expressive AI depends on the presence of expressive media in its raw materials. Blank paper is at least as good as newsprint for papier-mâché, but AI trained on blank documents or gibberish text would be useless. *See* Sina Alemohammad, Josue Casco-Rodriguez, Lorenzo Luzi, Ahmed Imtiaz Humayun, Hossein Babaei, Daniel LeJeune et al., Self-Consuming Generative Models Go MAD 14–15 (Jul. 4, 2023) (unpublished manuscript) (on file with arXiv), http://arxiv.org/abs/2307.01850 [https://perma.cc/5M45-YCMT] (recommending that "real data [be] present in the training as much as possible"). For AI to excel at emulating expression, it must train on expressive data.

This infelicity in Sag's papier-mâché analogy reveals how dramatically it reframes training data's role in expressive AI. Rather than treating text as expression, the papier-mâché analogy equates it with pulp paper. The papier-mâché analogy implies that text is merely a container for expression, like a book's binding. That implication is incorrect. The expression in a literary work *is* its text. That expression is often humdrum, but that doesn't mean copyright

figure out what it is that generative AI models are mechanically reproducing — and to figure out how the law ought to treat that reproduction — we have to figure out what "style" is. As Part III shows, this is a question that copyright has long struggled to address.

## III. WHAT WE TALK ABOUT WHEN WE TALK ABOUT STYLE

"You can't copyright style" is an old chestnut of copyright law.[79] This rule would offer excellent guidance if anyone knew exactly what it meant. In truth, the legal meaning of style, insofar as it has stable meaning at all, is secondary to its implication: unprotectability. Often, though not always, jurists hurl "style" around not to invoke any positive signification, but rather to justify an outcome. If the style-like material at issue in a case was something a court deemed copyrightable, then it wasn't style to begin with; if it wasn't copyrightable, then it was style all along.[80]

When it comes to potential copyright infringement by image-generating AI, "you can't copyright style" is a common refrain.[81] Generative AI's boosters invoke this slogan to suggest that AI mimicry of visual artistic style does not infringe copyright. Relatedly, supporters of the fair use defense for training generative AI on unauthorized copies of copyrighted works argue that such training only implicates unprotectable qualities of copyrighted works.[82] "You can't copyright style" is on some level necessarily true, and on some level utterly false. My essential contention, therefore, is that the phrase is an unhelpful

---

equates it with paper and glue. Copyright's protection of the humdrum is an intended, and central, feature of the regime. *See* Feist Publ'ns, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340, 345 (1991) ("[T]he requisite level of creativity is extremely low; even a slight amount will suffice. The vast majority of works make the grade quite easily, as they possess some creative spark, 'no matter how crude, humble or obvious' it might be." (quoting Melville B. Nimmer & David Nimmer, COPYRIGHT § 1.08[C][1] (1990)); *cf.* Bleistein v. Donaldson Lithographing Co., 188 U.S. 239, 251 (1903) ("It would be a dangerous undertaking for persons trained only to the law to constitute themselves final judges of the worth of pictorial illustrations, outside of the narrowest and most obvious limits.").

79. *See, e.g.*, Judith Ripka Designs, Ltd. v. Preville, 935 F. Supp. 237, 238 (S.D.N.Y. 1996) ("Copyright laws do not protect styles, but only particular original designs."); McDonald v. West, 138 F. Supp. 3d 448, 455 (S.D.N.Y. 2015), *aff'd*, 669 F. App'x 59 (2d Cir. 2016); Jewelry 10, Inc. v. Elegance Trading Co., No. 88 CIV. 1320, 1991 WL 144151, at *4 (S.D.N.Y. July 20, 1991); Hayuk v. Starbucks Corp., 157 F. Supp. 3d 285, 291 (S.D.N.Y. 2016).

80. *Compare* Williams v. Gaye, 895 F.3d 1106, 1138 (9th Cir. 2018) ("Our decision does not grant license to copyright a musical style or 'groove.'") *with id.* (Nguyen, J., dissenting) ("The majority allows the Gayes to accomplish what no one has before: copyright a musical style.").

81. Guadamuz, *supra* note 13, at 26 ("So, are styles protected by copyright? Roughly speaking no, a style is more of an idea, and copyright does not protect an idea, only the expression of that idea, this is because protecting an idea would potentially lead to monopolies."); Motion to Dismiss at 13, Andersen v. Stability AI Ltd., No. 3:23-cv-00201 (N.D. Cal. Jan. 13, 2023) ("'[A]rtistic styles' are unprotectable under the Copyright Act.").

82. *See, e.g.*, Sag, *supra* note 78, at 307–08.

shibboleth: because it is ambiguous and therefore uninformative, the slogan risks providing AI users with a false sense of security and misinforming copyright holders about the extent of their legal rights.

### *A. What is Style?*

Style is an ambiguous concept, linguistically and legally.[83] The word's pertinent dictionary definitions include "[t]he manner of expression characteristic of a particular writer," "[a] person's characteristic bearing, demeanour, or manner," and, "[t]hose features of literary composition which belong to form and expression rather than to the substance of the thought or matter expressed."[84] This range of possible meanings gives style a precarious position within the cosmology of intellectual property law. Copyright's idea-expression distinction holds that the law protects the original expression of ideas, but it does not protect ideas themselves.[85] Style is difficult to situate in this dichotomy; is it idea, expression, both, neither?[86] Style may not always be expression *per se*, but it is indisputably *part of* expression.[87] To alter style is to alter what a work expresses: "Amazing Grace" played as an uptempo polka might not strike the right tone for a funeral. On the other hand, style can be characterized as merely a process by which some underlying content is expressed; in these terms, style would seem to resemble an unprotectable "idea, procedure, process, [or] system[.]"[88]

---

83. Law is not the only discipline that agonizes over style, and non-legal discourse provides helpful insights. Philosophers, for example, discuss the "problem of style": if we can agree that several distinct depictions can represent the same phenomenon, in art or in science, how do we account for stylistic differences between those representations? In what ways are these stylized representations similar and dissimilar, and why don't their dissimilarities interfere with their representative function? *See* Roman Frigg, *Scientific Representation and the Semantic View of Theories*, 21 THEORIA INT'L J. THEORY, HIST. AND FOUND. SCI. 49, 50 (2006).

84. OED ONLINE *style, n.* (Oxford University Press), http://www.oed.com/view/Entry/192315 [https://perma.cc/2MLH-4L52].

85. 1 MELVILLE B. NIMMER & DAVID NIMMER, NIMMER ON COPYRIGHT § 2.03[D][1] (2018) [hereinafter NIMMER ON COPYRIGHT].

86. *See* Michelle Brownlee, *Safeguarding Style: What Protection Is Afforded to Visual Artists by the Copyright and Trademark Laws*, 93 COLUM. L. REV. 1157, 1159–60 (1993) (explaining the difficulties of defining style for the purposes of copyright in the visual arts, and suggesting that style, explained broadly, is typically treated as an idea, but, defined more narrowly, may qualify as expression).

87. *See* Chicago Rec.-Herald Co. v. Trib. Ass'n, 275 F. 797, 798–99 (7th Cir. 1921) (suggesting that, while a "bare recital of . . . facts" cannot be copyrighted, such a statement is copyrightable "in so far as [it] involves authorship and literary quality and style"); Steinberg v. Columbia Pictures Indus., Inc., 663 F. Supp. 706, 712 (S.D.N.Y. 1987) ("style is one ingredient of 'expression'").

88. 17 U.S.C. § 102(b). *Cf.* Garcia v. Google, Inc., 766 F.3d 929, 942 (9th Cir. 2014) (Smith, J., dissenting) ("An acting performance resembles the 'procedure' or 'process' by which 'an original work' is performed. Therefore, '[i]n no case does copyright protection' extend to an acting performance, 'regardless of the form in which it is described, illustrated,

Style's ambit is similarly hazy. In its broadest conventional sense, the word denotes a genre, like the Romantic style of musical composition or the pointillist style of painting.[89] More narrowly, one could refer to the style of a particular artist or author, the characteristics that distinguish a work as a member of a specific oeuvre.[90] Finally, at its most narrow, style might describe only the attributes of one particular work, or even just a component thereof.[91]

Defining style is a vexing task, not only for jurists, but for philosophers and art historians as well.[92] Some attempt to clarify style by placing it in uncomfortable dichotomies with other attributes of a work. A work's style might be distinguished from its "content" or its "subject," for example, but these concepts aren't easily divisible. A style-content division fails because the way in which something is expressed influences precisely what is expressed.[93] Moreover, a choice about what "content" to depict may itself amount to a stylistic choice; for instance, a painter's signature style might be characterized by a choice to depict women in cafés as the subjects of his paintings.[94] As Susan Sontag puts it, "To speak of style is one way of speaking about the totality of a work of art."[95]

---

or embodied in' the original work."), *on reh'g en banc*, 786 F.3d 733 (9th Cir. 2015); Whitehead v. CBS/Viacom, Inc., 315 F. Supp. 2d 1, 11 (D.D.C. 2004) ("While similar writing styles may contribute to similarity between works' total concept and feel, a particular writing style or method of expression standing alone is not protected by the Copyright Act.").

89. Brownlee, *supra* note 86, at 1159–60.

90. *See* Justin Hughes, *The Personality Interest of Artists and Inventors in Intellectual Property*, 16 CARDOZO ARTS & ENT. L.J. 81, 127–28 (1998) (discussing artists' personal styles); *see also* Brownlee, *supra* note 86, at 1159 ("'[S]tyle' denotes the formal characteristics of the work of an individual artist when the work has 'an inner coherence, or unity, [and] possess[es] a sense of wholeness, of being all of a piece.'").

91. *See* ROBERT PASCALL, *Style,* GROVE MUSIC ONLINE ("Style, a style or styles (or all three) may be seen in any conceptual unit in the realm of music, from the largest to the smallest . . . .").

92. *See, e.g.*, Nelson Goodman, *The Status of Style*, 1 CRITICAL INQUIRY 799, 799–800 (1975); James Elkins, *Style*, GROVE ART (2003), http://www.oxfordartonline.com/view/10.1093/gao/9781884446054.001.0001/oao-9781884446054-e-7000082129 [https://perma.cc/GB4J-PZUA] (calling style "one of the most difficult concepts in the lexicon of art, and one of the chief areas of debate in Aesthetics and Art history.").

93. *See* GRAHAM HOUGH, STYLE AND STYLISTICS 4 (1969) ("[T]he more we reflect on it, the more doubtful it becomes how far we can talk about *different ways of saying*; is not each different way of saying in fact the saying of a different thing?"); *quoted in* Goodman, *supra* note 92, at 800.

94. *See generally* Romm Art Creations Ltd. v. Simcha Int'l, Inc., 786 F. Supp. 1126 (E.D.N.Y. 1992) (granting a preliminary injunction for trade dress infringement against a defendant that sold artworks in the distinctive style of the plaintiff's "Women in Cafés" series of paintings); *contra* Leigh v. Warner Bros., a Div. of Time Warner Ent. Co., L.P., 10 F. Supp. 2d 1371, 1380 (S.D. Ga. 1998).

95. Arjun Gupta, *"I'll Be Your Mirror" — Contemporary Art and the Role of Style in Copyright Infringement Analysis*, 31 U. DAYTON L. REV. 45, 60 (2005) (quoting Susan Sontag, *On Style*, *in* AGAINST INTERPRETATION 15, 17 (Farrar, Straus & Giroux 1966)).

Style, in other words, can at the same time be both a "what" and a "how."[96] To confuse matters further, style is multifarious. Rather than being any one, discrete "thing," style is better understood as the cumulative effect of a collection of many attributes — as Sontag says, a "totality."[97] These attributes might be too numerous to catalog, and some of them might elude verbal description entirely.[98] For instance, that a painting is done in watercolor does not alone constitute its style, but that fact may well be a component of the painting's overall style. Style may be broken down into discrete elements, but it is fundamentally a gestalt. This is why it presents such a challenge in court: dissecting style in an analytical, lawyerly way is apt to leave a pile of unprotectable ideas and techniques, even if the aggregate effect of those techniques is expressive.

In a legal context, the capaciousness of the word "style" can be misleading. Because style can be construed so broadly as to include material that is obviously uncopyrightable, one might conclude that style is categorically unprotectable. This is true if style is understood as a synonym of genre. But the reality of stylistic protection is more complicated. Sometimes, style refers to an idea, and sometimes it can encompass expression.

Style is particularly vexatious in visual art. Judges report that it is especially difficult for them to apply copyright's idea-expression dichotomy to images. Copyright jurisprudence centers on the written word, and strategies of abstraction and textual analysis that are second nature to jurists may not necessarily translate to other media.[99] If the

---

96. Goodman, *supra* note 92, at 806 ("My purpose has not been to impose an elaborate and rigid system of classification upon features of style, but rather to free the theory of style from the warping constraints of prevalent dogma-from the misleading opposition of style and subject, of form and content, of what and how, of intrinsic and extrinsic.").

97. *See, e.g.*, Steinberg v. Columb. Pictures Indus., Inc., 663 F. Supp. 706, 712 (S.D.N.Y. 1987) (discussing a "striking stylistic relationship" between two works in terms of individual stylistic components, such as vantage point, colors and techniques used to represent the horizon, and lettering); Arjun Gupta, *"I'll Be Your Mirror" — Contemporary Art and the Role of Style in Copyright Infringement Analysis*, 31 U. DAYTON L. REV. 45, 66 (2005) ("Style may be understood as the sum of artistic decisions involved in the creation of a work of art."); *cf.* Annie Kim, *Digital Art Says: Protect Our "Groove" Too*, B.C. INTELL. PROP. & TECH. F., March 21, 2022, at 1, 8 ("In music, style can be broken down into and measured by objective elements . . . therefore, one could argue that current copyright laws can protect these elements individually, thereby protecting the overall style."); Michelle Brownlee, *Safeguarding Style: What Protection Is Afforded to Visual Artists by the Copyright and Trademark Laws?*, 93 COLUM. L. REV. 1157, 1159 (1993) (describing style as "some recognizable *combination of characteristics* that allows us to classify [a] work as a creation of [a particular] artist" (emphasis added)).

98. *See* FLINT SCHIER, DEEPER INTO PICTURES: AN ESSAY ON PICTORIAL REPRESENTATION 29 (1986); Joseph G. Cook, *The Fine Arts: What Constitutes Infringement*, 16 ALA. L. REV. 41, 54 (1963); Brownlee, *supra* note 86, at 1162.

99. Tushnet, *supra* note 75, at 710 ("The official story is now one of media neutrality, except where specified otherwise . . . . Nonetheless, the written text remains the prototypical copyrighted work. Perhaps judges, whose output is written, have a particularly easy time seeing the worth and creativity of writing and analogizing other types of creation to words.").

usual tools of legal argumentation are especially ill-suited to describing works of visual art in analytical terms, then those works are more likely to be treated as gestalts in court. And because "style" itself refers to a gestalt expressive effect, it will assume a greater legal role in categories of media that courts treat as gestalts.

In sum, "style" as I use the term refers to the cumulative effect of a number of discrete, expressive choices made by a particular artist. Although it is a gestalt, style is not necessarily *inexpressible*, and one's ability to catalog the discrete, expressive choices that comprise a style will vary with one's analytical expertise in the relevant subject matter. And although analytical expertise is required to *analyze* style, it is not required to *recognize* style. An ordinary observer doesn't have to know the words "sfumato" or "pointillism" in order to recognize that the "Mona Lisa" and "A Sunday Afternoon on the Island of La Grande Jatte" are rendered in distinct styles.

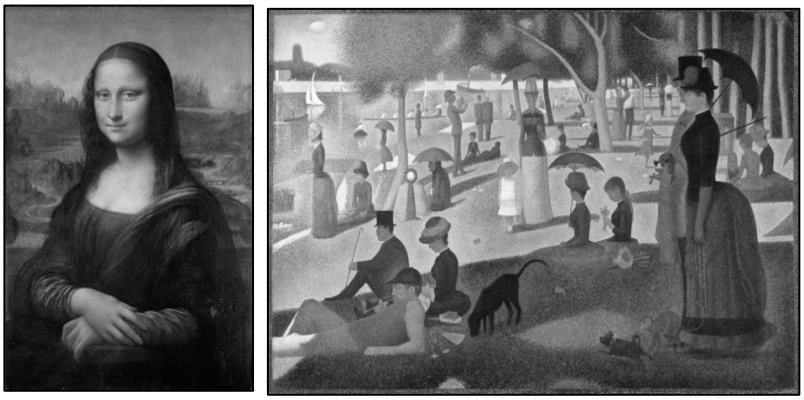


Figure 4: Leonardo, "Mona Lisa,"[100] and Seurat, "A Sunday Afternoon on the Island of La Grande Jatte."[101]

Indeed, a style will have identifiable elements capable of being described in layman's terms; for example, the mature style of the Dutch painter Piet Mondrian consisted of "the simplest combinations of straight lines, right angles, primary colors, and black, white, and gray."[102] Expert vocabulary can explicate a style with greater precision;

100. *File: Mona Lisa, by Leonardo da Vinci, from C2RMF retouched.jpg*, WIKIPEDIA, https://en.wikipedia.org/wiki/File:Mona_Lisa,_by_Leonardo_da_Vinci,_from_C2RMF_retouched.jpg [https://perma.cc/N9V5-QHN8].

101. *A Sunday on La Grande Jatte, Georges Seurat, 1884*, WIKIPEDIA, https://en.wikipedia.org/wiki/A_Sunday_Afternoon_on_the_Island_of_La_Grande_Jatte#/media/File:A_Sunday_on_La_Grande_Jatte,_Georges_Seurat,_1884.jpg [https://perma.cc/ZE2Q-KTTX].

102. Hans L.C. Jaffe, *Piet Mondrian*, BRITANNICA (Aug. 23, 2024), https://www.britannica.com/biography/Piet-Mondrian [https://perma.cc/NM7R-N8EA].

in the words of an art historian, Mondrian's mature style "us[es] color squares and/or rectilinear patterns drawn in black. Because their pictorial elements are seldom attached to the edges of the canvas, and because the background . . . is always dense and hence the forms do not float but are held taut across the surface, these pictures have a staccato rhythm . . . ."[103] With greater analytical expertise comes greater ability to talk about style in precise terms but an observer can perceive style even if she cannot articulate the basis for that perception.

### *B. Style, Substantial Similarity, and the Idea-Expression Distinction*

A copyrighted work is not a clearly delineated piece of property. Real property has definite boundaries. The land one inch inside a property line is just as much the landowner's property as the land at the center of the plot; entering at the periphery or parachuting into the center are both trespass.[104] But the strength of copyright protections and the scope of permissible copying depend on which aspects of the work are being copied, and how.

A plaintiff alleging a violation of her exclusive right of reproduction must show three things: that a defendant's work was created by copying the plaintiff's, that the reproduction is sufficiently fixed and tangible to constitute a copy,[105] and that the copying amounts to "improper appropriation."[106] The first and last inquiries deal with a distinct form of similarity. The similarity between two works that permits a finding that one was copied from the other is "probative similarity," and it is not this Article's focus.[107] Substantial similarity, on the other hand, is similarity between two works that undergirds a finding of

---

103. Charles W. Millard, *Mondrian*, 25 THE HUDSON REVIEW 270, 272 (1972). Mondrian himself cultivated an abstruse vocabulary that may represent the most technically precise verbal description of his style, which he termed "Neo-Plasticism." Mondrian's tract on "General Principles of Neo-Plasticism" includes such dictates as, "The plastic means must be the rectangular plane or prism in primary colors (red, blue, and yellow) and in noncolor (white, black, and gray)," and "Equivalence in the dimension and color of the plastic means is necessary." General Principals of Neo-Plasticism, OBELISK ART HISTORY, https://arthistoryproject.com/artists/piet-mondrian/the-collected-writings-of-piet-mondrian/general-principals-of-neo-plasticism/ [https://perma.cc/R6WQ-27UF].

104. Of course, even the bounds of real property are not always straightforward to ascertain. Sometimes that uncertainty may derive from the recording method used, *see, e.g.*, Maureen E. Brady, *The Forgotten History of Metes and Bounds*, 128 YALE L.J. 872, 893 (2019), and sometimes it may derive from intervening events that alter boundary lines, *see, e.g.*, United States v. Byrne, 291 F.3d 1056, 1059–60 (9th Cir. 2002) (discussing rules for ascertaining boundary lines set by rivers that shift via "accretion" versus via "avulsion").

105. For obvious reasons, this issue is litigated comparatively rarely, but it has been integral to some holdings. *See* Cartoon Network LP, LLLP v. CSC Holdings, Inc., 536 F.3d 121, 127 (2d Cir. 2008).

106. Arnstein v. Porter, 154 F.2d 464, 468 (2d Cir. 1946).

107. 4 NIMMER ON COPYRIGHT, *supra* note 85, § 13.03[A] (2018).

"improper appropriation."[108] Substantial similarity must be similarity of expression; copying that implicates only a work's ideas or factual elements cannot constitute infringement, no matter how extensive the copying or how much of the original work it reproduces.

In the vocabulary of copyright jurists, copying may be "literal," in which one work is "comprehensive[ly]" or "fragmented[ly]" reproduced "verbatim" by another; or copying may be "non-literal," in which one work bears comprehensive similarities to another without sharing word-for-word, note-for-note, pixel-for-pixel, bit-for-bit identity.[109] Cases of comprehensive literal copying — like, say, peer-to-peer file sharing — are most straightforward to adjudicate as *prima facie* infringement, because in these instances there's no question that a copyrighted work has been reproduced in its entirety. Put another way, the exclusive right to completely duplicate a work is a central and uncontested entitlement of its copyright owner.[110]

In contrast, the periphery of a copyright is marked out by an owner's rights to control non-literal copying. Non-literal copies are similar, but not identical, to the copied work. Substantial similarity marks the bounds of non-literal copying; it is a "vague"[111] inquiry that makes for an uncertain boundary. Determining whether particular non-literal copying will constitute infringement is "ad hoc,"[112] and courts have devised a host of formulations for the test, none of them particularly lucid, satisfying, or predictable.[113] Substantial similarity is still governed by the idea-expression distinction, of course; similarity as to ideas alone is not enough to sustain a finding of improper appropriation. Rather, the relevant consideration is whether one work contains expression substantially similar to that of another work.[114]

The two most influential tests for substantial similarity appear in the Second Circuit case *Arnstein v. Porter* and the Ninth Circuit case *Sid & Marty Krofft v. McDonald's*.[115] Both tests divide the inquiry into two stages: an initial stage that admits expert testimony and "analytical

---

108. *Id.* Unfortunately, courts often fail to separate these forms of similarity in their analysis, or conflate the two entirely. *See* Mark A. Lemley, *Our Bizarre System for Proving Copyright Infringement Part I: Brace Lecture*, 57 J. COPYRIGHT SOC'Y U.S.A. [i], 719–21 (2009–2010).

109. 4 NIMMER ON COPYRIGHT, *supra* note 85, § 13.03[A][1]–[2] (2018).

110. This right is of course qualified by copyright's limitations and exceptions, the most notable of which is the fair use doctrine set forth in 17 U.S.C. § 107.

111. Peter Pan Fabrics, Inc. v. Martin Weiner Corp., 274 F.2d 487, 489 (2d Cir. 1960) (L. Hand, J.) ("[T]he test for infringement of a copyright is of necessity vague.").

112. *Id.* ("Obviously, no principle can be stated as to when an imitator has gone beyond the 'idea,' and has borrowed its 'expression.' Decisions must therefore inevitably be ad hoc.").

113. 4 NIMMER ON COPYRIGHT, *supra* note 85, § 13.03[E][1] (2018) ("Although it is clear that the determination of substantial similarity presents an issue of fact, the correct procedure for that determination remains clouded.").

114. *Id.*

115. Lemley, *supra* note 37, at 719.

dissection"[116] of works, and a second stage that eschews analytical dissection and hinges on the reactions of a lay audience.

According to *Arnstein*, courts must first consider the question of actual copying or copying-in-fact; at this stage, the court may avail itself of expert analysis of the works at issue.[117] In the next stage, courts evaluate whether improper appropriation has taken place. There, expert opinion is irrelevant, and the sole criterion is "whether defendant took from plaintiff's works so much of what is pleasing to" a lay audience.[118]

*Krofft* formulates the process slightly differently. Its first step is a test of "extrinsic" similarity, which "depends not on the responses of the trier of fact, but on specific criteria which can be listed and analyzed . . . . Since it is an extrinsic test, analytic dissection and expert testimony are appropriate."[119] The *Krofft* analysis then proceeds to an "intrinsic" test, which "depend[s] on the response of the ordinary reasonable person" and in which "analytic dissection and expert testimony are not appropriate."[120]

The *Arnstein* and *Krofft* tests reveal that substantial similarity jurisprudence is an effort to balance aesthetic holism with lawyerly dissection. On one hand, the doctrine recognizes that the lay audience is the ultimate arbiter of substantial similarity. What matters is the total aesthetic impression — a gestalt that may be difficult or impossible to express analytically — that a work inspires in the ordinary observer.[121] On the other hand, the doctrine acknowledges that copyright infringement cannot depend entirely on lay observers' intuitions about similarity, so it requires that two works show "objective" indicia of similarity.[122] An analytical justification for a finding of similarity serves two major purposes. First, it makes it harder for jurors to predicate infringement on similarities that are not legally cognizable. Second, and more fundamentally, the requirement helps courts discharge

116. "Analytic dissection" (alternatively, "analytical dissection") is a term of art that appears frequently in copyright decisions. *See* Tushnet, *supra* note 75, at 766. This Article uses it to describe the process of describing a unitary work in terms of its component parts, often but not always with the aid of expert testimony.

117. Arnstein v. Porter, 154 F.2d 464, 473 n.19 (2d Cir. 1946).

118. *Id.* at 473.

119. Sid & Marty Krofft Television Prods., Inc. v. McDonald's Corp., 562 F.2d 1157, 1164 (9th Cir. 1977).

120. *Id.*

121. *See* Boisson v. Banian, Ltd, 273 F.3d 262, 272 (2d Cir. 2001); Williams v. Crichton, 84 F.3d 581, 590 (2d Cir. 1996) (formulating substantial similarity inquiry as "whether a lay observer would consider the works as a whole substantially similar to one another.").

122. Shaw v. Lindheim, 919 F.2d 1353, 1357 (9th Cir. 1990), *overruled on other grounds by* Skidmore as Tr. for Randy Craig Wolfe Tr. v. Led Zeppelin, 952 F.3d 1051 (9th Cir. 2020).

their obligation to offer reasoned decisions, rather than up-or-down decisions based on aesthetic intuitions.[123]

Aesthetic intuitions and reasoned analysis aren't in tension in cases of literal reproduction. We intuit that literal copies are identical because we're steeped in reasoned consensus about what those works' constitutive expressive elements are; we understand on both an intuitive and a rational level that if you copy the text of a literary work, you have made an identical copy of the whole work.[124] But nonliteral reproduction is harder to adjudicate, because unlike literal similarity, explaining how and why two works are nonliterally similar does not per se establish that the similarity is aesthetically significant or legally cognizable. Stylistic nonliteral similarity, in turn, is a particularly troublesome form of nonliteral similarity because it places aesthetic intuition and analytical dissection in especially great tension. Stylistic similarity is a holistic similarity that derives from a coincidence of many elements; it can provoke powerful, holistic, aesthetic intuitions that are difficult to explain in terms of discrete, analytic components.

1. Stylistic Similarity is Hard to Adjudicate

Some types of analytical dissection trouble courts more than others. The Ninth Circuit has observed that "[t]he extrinsic test provides an awkward framework to apply to copyrighted works like music or art objects, which lack distinct elements of idea and expression."[125] While "[l]iterary works, such as books, film, and television shows, are more easily broken into a small number of discrete elements to analyze, namely 'plot, themes, dialogue, mood, setting, pace, characters and sequence of events,'"[126] the court noted that "[m]usic, like software programs and art objects, is not capable of ready classification into only five or six constituent elements; music is comprised of a large array of elements, some combination of which is protectable by copyright."[127]

The Ninth Circuit's trouble with images is not unique. Other judges openly admit that they are more comfortable dissecting texts and that they are particularly uncomfortable dissecting images. Judge Jon Newman of the Second Circuit writes that, while literary works are processed sequentially, "graphic or three-dimensional work[s] [are]

123. *See* Lon L. Fuller & Kenneth I. Winston, *The Forms and Limits of Adjudication*, 92 HARV. L. REV. 353, 366–67 (1978). *Cf.* Tushnet, *supra* note 75, at 740 ("it is unlikely that reviewing courts will accept an unanalyzed gestalt judgment without adding further analysis").

124. We know, for example, that a literary work is its text; no plagiarist would try to negate a charge of plagiarism by admitting he copied the text but insisting that he changed the font.

125. Swirsky v. Carey, 376 F.3d 841, 848 (9th Cir. 2004), *as amended on denial of reh'g* (Aug. 24, 2004).

126. *Id.* at 849 n.15 (quoting Metcalf v. Bochco, 294 F.3d 1069, 1073 (9th Cir. 2002)).

127. *Id.* at 849.

created to be perceived as an entirety," and "one cannot divide a visual work into neat layers of abstraction in precisely the same manner one could with a text."[128] In *Mannion v. Coors*, Judge Lewis Kaplan wrote that the idea-expression distinction "breaks down" when considering visual works. In media that can be abstracted and assessed more analytically, such as literary works, the difficulty in delineating idea and expression "is essentially one of line-drawing," while in a case involving photographs, the "difficulty . . . is not simply that it is not always clear where to draw the line; it is that the line itself is meaningless because the conceptual categories it purports to delineate are ill-suited to the subject matter."[129] Judge Kaplan continued, "[I]t is not clear that there is any real distinction between the idea in a work of art and its expression. An artist's idea, among other things, is to depict a particular subject in a particular way."[130] By contrast, he posited, two authors can explain an underlying idea — say, the theory of special relativity — using different expressive vocabulary, and judges and juries have the analytical skills to separate idea from expression.[131]

Putting the judges' observations together teaches us that courts find it easier to perform nonliteral similarity analysis on works that are "easily broken into a small number of discrete elements to analyze."[132] Jurists struggle to break works of visual art into discrete elements, because artworks comprise many elements that rise to the level of copyrighted expression only when perceived as a combined whole.[133] As

128. Jon O. Newman, *New Lyrics for an Old Melody: The Idea/Expression Dichotomy in the Computer Age*, 17 CARDOZO ARTS & ENT. L.J. 691, 698 (1999).

129. Mannion v. Coors Brewing Co., 377 F. Supp. 2d 444, 457–58 (S.D.N.Y. 2005).

130. *Id.* at 458.

131. *Id.*

132. *Swirsky*, 376 F.3d at 849 n.15.

133. *Cf. id.* Of course, by presenting judges' anecdotal testimony, I mean to suggest only that literary works receive more analytical dissection than images do, and not to claim that literary works are *inherently* more susceptible to analytic dissection than are other expressive media. Art historians are trained in "visual analysis," which equips them with a precise vocabulary for dissecting images in terms of attributes like "scale," "composition," "pictorial space," "form," "line," "color," "light," "tone," and so on. *See, e.g.*, Smarthistory, *How to do visual (formal) analysis in art history*, YOUTUBE (Sept. 18, 2017), https://www.youtube.com/watch?v=sM2MOyonDsY [https://perma.cc/UT7D-442A] (describing a painting in terms of "scale," "composition," "pictorial space," "form," "line," "color," "light," and "tone," and identifying such techniques as "atmospheric perspective"). Musicians and music theorists have a refined analytical apparatus for describing and notating the qualities of music and sound. Rather than a reflection of some ground truth, judges' discomfort applying the idea-expression dichotomy to nonverbal media simply may be a consequence of a legal enterprise dominated by *lawyers*, whose training focuses on reading and writing instead of singing or painting. These hyperverbal professionals may be uneasy discussing works in terms of their holistic impressions, rather than in precise, analytical, and perhaps quantifiable terms. *See* Tushnet, *supra* note 75, at 719. Of course, literary works are not necessarily a cakewalk for jurists to dissect, either. To be sure, courts have noted that written works can be broken into analytical components like "plot, themes, dialogue, mood, setting, pace, characters, and sequence of events," Kouf v. Walt Disney Pictures & Television,

a result, courts have a hard time assessing nonliteral similarity in images.

The reasons why judges struggle to analyze substantial similarity in visual art help explain why artistic style in particular is the most evanescent predicate for substantial similarity. Some similarities in visual works are as easy to catalog as similarities in literary plots. For example, it is easy enough to detect prima facie infringement when a defendant's photograph depicts a nude, pregnant woman in the same pose as a plaintiff's photograph does, with the same skin tone, with her hands identically positioned and a ring on the same finger.[134] But similarities of style may not be so easy to discern and to articulate. Style refers to a "totality" created by a combination of many small aesthetic choices that, in tandem, create an expressive effect.[135] Dissecting visual artistic style into its myriad constituents (a) is hard to do, because it requires special perceptual expertise that judges are unlikely to have; and (b) obscures style's overall aesthetic effect, because style is properly perceived as a gestalt, not a pile of individual attributes.

De facto, then, a rigorous juridical analysis of stylistic similarity in visual media may be difficult, if not impossible. But de jure, expressive stylistics are due the same protection as any other copyrightable expression. It is blackletter law that individually unprotectable elements of a work can, in combination, constitute copyrightable expression.[136] While another creator is free to incorporate one or some of these elements into another work, he cannot copy an expressive combination of elements. Style in visual arts presents such a problem because courts struggle to perceive style holistically, while at the same time identifying

---

16 F.3d 1042, 1045 (9th Cir. 1994), and paraphrased in terms of "patterns of increasing generality," Nichols v. Universal Pictures Corp., 45 F.2d 119, 121 (2d Cir. 1930). But the more literary and less expository the text in question, the harder it becomes to winnow idea from expression. The underlying idea of poetry and literary prose depends greatly on the precise expressive form it takes. *See* Tushnet, *supra* note 75, at 716 n.149 (quoting Leo Tolstoy, "If I wanted to express in words all that I meant to express by the novel, then I should have to write the same novel as I have written all over again."); Cleanth Brooks, "The Heresy of Paraphrase," *in* THE WELL-WROUGHT URN 197 (1947) ("[T]he paraphrase is not the real core of meaning which constitutes the essence of the poem.").

134. *See* Leibovitz v. Paramount Pictures Corp., 137 F.3d 109, 111–12 (2d Cir. 1998); *see also* Gross v. Seligman, 212 F. 930, 931 (2d Cir. 1914) (affirming finding of infringement based on "many close identities of pose, light, and shade" between one photograph and another photograph depicting the same model).

135. *See supra*, Section III.A (discussing style as totality).

136. Skidmore as Tr. for Randy Craig Wolfe Tr. v. Led Zeppelin, 952 F.3d 1051, 1074 (9th Cir. 2020) ("[A] copyright plaintiff may argue 'infringement . . . based on original selection and arrangement of unprotected elements.'" (quoting Metcalf v. Bochco, 294 F.3d 1069, 1074 (9th Cir. 2002))); Three Boys Music Corp. v. Bolton, 212 F.3d 477, 485 (9th Cir. 2000), *overruled on other grounds by* Skidmore as Tr. for Randy Craig Wolfe Tr. v. Led Zeppelin, 952 F.3d 1051 (9th Cir. 2020) (upholding finding of infringement based on combination of five unprotectable musical elements). *See also* Swirsky v. Carey, 376 F.3d 841, 848 (9th Cir. 2004), *as amended on denial of reh'g* (Aug. 24, 2004).

analytically which stylistic elements of style have been appropriated and whether the appropriation is too much.[137]

2. Stylistic Similarity Can Be Infringement

Individually stylistic elements of works may be unprotectable ideas. But particular combinations of stylistic elements may constitute protected expression based on the overall aesthetic impression they create. "Style" is used imprecisely enough that it can denote *both* expressive combinations of aesthetic choices *and* broader, unprotectable conventions. Thus, the slogan that "style is not copyrightable" is, at best, a tautology: "when *I* say 'style,' *I* mean 'uncopyrightable style,' and uncopyrightable style is not copyrightable." The hard question remains: when does a combination of individually unprotectable elements become protected expression?

*a. Stylistic Choices Can Be Expressive*

Given that "style" can encompass copyrightable expression, it is not surprising that courts have found infringement based on what a reasonable English speaker could call similarity of style.[138] A leading case on style in visual art is *Steinberg v. Columbia Pictures*,[139] in which an artist sued a film production company for copyright infringement after it used a pastiche of his work as a promotional poster. Steinberg, the plaintiff, drew a *New Yorker* cover depicting a self-centered urbanite's view of Manhattan: drawn from the east, the cityscape looms; the Hudson River separates it from a dull beige rectangle that represents the rest of the continental United States. Japan, Russia, and China appear as blips on the horizon.[140] The defendants' poster presents Manhattan from the west and shows the movie's protagonists in front of landmarks

137. Tushnet, *supra* note 75, at 740 ("We . . . can't have both analytic dissection and gestalt 'feeling,' and we should stop pretending that we can.").

138. *See, e.g.*, Williams v. Bridgeport Music, Inc., No. CV13-06004, 2015 WL 4479500, at *22 (C.D. Cal. July 14, 2015), *aff'd in relevant part, rev'd in part sub nom.* Williams v. Gaye, 895 F.3d 1106 (9th Cir. 2018); *Williams*, 895 F.3d at 1138 (Nguyen, J., dissenting) ("The majority allows the Gayes to accomplish what no one has before: copyright a musical style."); Tim Wu, *Why the "Blurred Lines" Copyright Verdict Should Be Thrown Out*, New Yorker (Mar. 12, 2015), http://www.newyorker.com/culture/culture-desk/why-the-blurred-lines-copyright-verdict-should-be-thrown-out [https://perma.cc/H6ZB-Y25V] (referring to the similarities between the works in question as "the general style of [the plaintiffs' predecessor-in-interest's] songs"). *See also* Joseph P. Fishman, *Music as a Matter of Law,* 131 HARV. L. REV. 1861, 1869 (2017–2018) (suggesting that despite the case's controversial reception, the verdict in *Williams v. Gaye* was in fact "consistent with" caselaw from appellate and district courts). *But see Williams*, 895 F.3d at 1138 ("Our decision does not grant license to copyright a musical style.").

139. 663 F. Supp. 706 (S.D.N.Y. 1987).

140. *Id*. at 710.

not depicted in Steinberg's work.[141] The court granted summary judgment to Steinberg, noting, "Even at first glance, one can see the striking

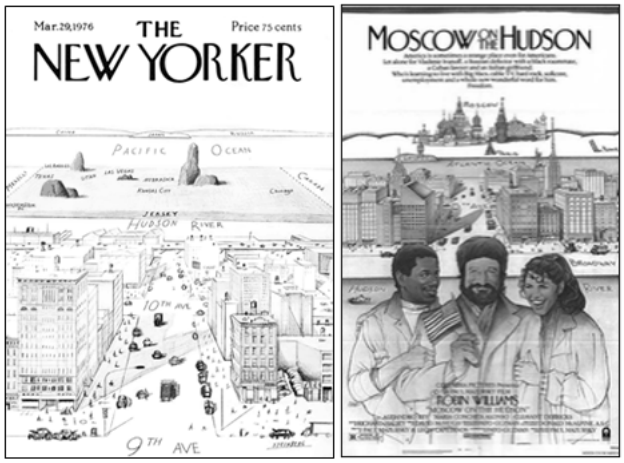



Figure 5: The Images at Issue in *Steinberg*.

stylistic relationship between the posters, and since style is one ingredient of 'expression,' this relationship is significant. Defendants' illustration was executed in the sketchy, whimsical style that has become one of Steinberg's hallmarks."[142] Among the similar stylistic elements the court cataloged are the vantage point; the "minimal[]" detail of faraway objects; the rendering of specific objects such as façades, windows, cars, and signage, many of which "could be mistaken for one another[;]" the handwritten font; and the color and technique used to render the horizon.[143] While the court did also cite similarities in content rather than representational style — e.g., "Both illustrations represent a bird's eye view across the edge of Manhattan and a river bordering New York City to the world beyond" — it was "the striking stylistic relationship between the posters," and its constitutive elements that was the court's focus.[144] *Steinberg* thus reaffirms that an original constellation of individually unprotectable aesthetic elements in a pictorial work can receive copyright protection.

*b. Everything Is Stylized; Not All Style Is Protectable*

Some of the most explicit judicial reasoning about style — and the circumstances under which it may or may not be protectable — appears

141. *Id*.
142. *Id*. at 712.
143. *Id.* at 712–13.
144. *Id.* at 712.

in a Third Circuit case, *Franklin Mint Corp. v. National Wildlife Art Exchange, Inc.*[145] The court affirmed the finding that a wildlife painter did not infringe the copyright in one of his paintings of a cardinal — which he had assigned to a third party — by creating a similar painting of a cardinal.[146] The court reasoned that depicting real-world subject matter with "photograph-like clarity and accuracy" affords only a "weak" copyright in the resulting work.[147] Because the works at issue seemed realistic to the court, their "expression and subject matter converge[d]," and the owner of the rights to a prior painting had less power to enjoin a subsequent, similar work.[148] The genre of ornithological art constrains the forms a painting of a cardinal might take; artists' expressive freedoms are limited, the court reasoned, because they must represent the minute, factual details of the bird's plumage, anatomy, and posture.[149] In other words, the court suggested that the artist's expressive freedom was limited by reality itself.



Figure 6: The Illustrations at Issue in Franklin Mint.[150]

As Tushnet observes, *Franklin Mint* is misguided.[151] It conflated a style that connotes realism with realism itself. To the court, the paintings embodied the "idea" of a cardinal — and that "idea" happened to

145. 575 F.2d 62 (3d Cir. 1978).
146. *Id.* at 65.
147. *Id.*
148. *Id.*
149. *Id.*
150. *The Cardinals Paintings Case*, COPYRIGHT ON! (June 18, 2014), https://paynebritton.wordpress.com/2014/06/18/casenotes-franklin-mint-v-national-wildlife-art-exchange-the-cardinals-paintings-case/ [https://perma.cc/3HAN-4PXM].
151. Tushnet, *supra* note 75, at 728–30.

exist in a particular style.[152] But the paintings' realism comes from adherence to convention, not to some ground truth.[153]

Representational conventions that seem self-evident today are anything but. If you think about it, this is obvious. It would be unsettling to behold a cardinal in the flesh that looked *exactly* like the bird depicted in the *Franklin Mint* paintings: shadowless, two-dimensional, motionless, vividly-colored, positioned atop a lily-white backdrop. This sounds pedantic precisely because we have learned to look past those particular representational conventions to "see" the phenomenon being represented. A more realistic style is one that is, by convention, easier to see "through."[154] The reason it isn't obvious that realism is stylized is that we've been acculturated not to think about it. As Nelson Goodman puts it:

> [W]e must beware of supposing that similarity constitutes any firm, invariant criterion of realism; for similarity is relative, variable, culture-dependent. And even where, within a single culture, judgments of realism and of resemblance tend to coincide, we cannot safely conclude that the judgments of realism follow upon the judgments of resemblance. Just the reverse may be at least equally true: that we judge the resemblance greater where, as a result of our familiarity with the manner of representation, we judge the realism greater.[155]

*Franklin Mint*, for this Article's purposes, is most interesting because of the comments it makes about supposedly non-realistic styles. The court observes:

---

152. *Id.*

153. To be fair, the Third Circuit seemed almost to recognize that realism is dictated by convention. It wrote, "The limitations imposed upon the artist by convention are also factors which must be considered" when assessing the scope of the artist's copyright. But instead of following this statement with the observation that the "realism" of the paintings at issue was dictated by convention, the court seemed to posit in the next sentence that *convention* is dictated by *realism*: "A scientific drawing of a bird must necessarily be more similar to another of the same nature than it would be to an abstract version of the creature in flight." *Franklin Mint*, 575 F.2d at 65. That reasoning is exactly backwards, and the example from which it extrapolates is factually dubious: an MRI and an Audubon-style sketch might both be called "scientific" depictions of a bird, yet they differ no less radically from one another than they might from an "abstract" depiction of the same subject. As Tushnet points out, the *Franklin Mint* paintings use representational conventions developed by the ornithologist John J. Audubon, which, when they debuted, were criticized as deceptive and embellished, and which are criticized as unrealistic in the present, too. Tushnet, *supra* note 75, at 727 n.200.

154. Tushnet, *supra* note 75, at 690 n.17.

155. Nelson Goodman, *Seven Strictures on Similarity*, *in* PROBLEMS AND PROJECTS 438 (1972).

> [I]n the world of fine art, the ease with which a copyright may be delineated may depend on the artist's style. A painter like Monet when dwelling upon impressions created by light on the facade of the Rouen Cathedral is apt to create a work which can make infringement attempts difficult . . . . [I]n the impressionist's work the lay observer will be able to differentiate more readily between the reality of subject matter and subjective effect of the artist's work.[156]

Again, the court's observation here is factually and legally dubious. As a matter of fact, it *isn't* difficult to infringe copyright in a Monet-like painting: one could simply take a photograph of it. And as a matter of law, the difficulty of recreating a work is an erroneous proxy for the strength of copyright in that work. It may be difficult to recreate a Monet *with a paintbrush* because a Monet comprises what the Ninth Circuit would call "a large array of elements, some combination of which is protectable by copyright."[157] While there is a standardized way of reproducing those elements in a photograph, there is no standardized way of reproducing those elements with a paintbrush. The Third Circuit's point about Monet is that the elements of an artwork that the law will deem expressive are the elements that viewers understand to be expressive. This explains why the cardinal paintings' correspondence with realism-connoting conventions made them protected only by thin copyright. And it also explains why stylistic devices that do signify aesthetic expression may, in toto, amount to copyrightable expression.

What is really going on when *Franklin Mint* suggests that "realistic" styles should not be copyrightable, but a style like Monet's might be, is that the court is tacitly opining on the sorts of stylistic details that ordinary observers ought to *overlook*. The Audobon-style drawings are replete with aesthetic choices, but the court discounted those choices as indicia of authorship because it assumed that ordinary viewers, like the court itself, look past those choices and see only a "realistic" representation of a bird. The *Franklin Mint* approach, then, asks what styles are conventionally aligned with realism and then prevents the monopolization of those styles, on the ground that to rule otherwise would permit a monopolization on depictions of reality itself. The reasoning indulges the fiction that what is being copied is just the idea of a real-life bird. But when stylistic choices are not those that denote realism, and are instead details that observers believe themselves to look *at* rather than *through*, one can no longer claim that reproducing these details merely reproduces underlying reality — instead, an audience readily

156. *Franklin Mint*, 575 F.2d at 65.
157. *Cf.* Swirsky v. Carey, 376 F.3d 841, 849 (9th Cir. 2004).

recognizes that the stylistic details themselves are being reproduced. *Franklin Mint* suggests that in these circumstances, similarity as to such stylistic details could sustain a finding of copyright infringement. *Steinberg*'s holding is in accord.

The foregoing doctrinal analysis does not show that every alleged instance of stylistic copying will be a slam-dunk claim for copyright infringement. Far from it! But it does show that the infringement inquiry is much more nuanced and fact-specific, and much less predictable, than the "you can't copyright style" slogan suggests. Applying present-day substantial similarity doctrine to the outputs of generative AI will demand the same messy, fact-specific scrutiny that courts have applied to earlier allegations of improper stylistic similarity. A conclusory rule — "style can't be copyrighted" — obscures the thorny question that actually matters: identifying what style is in the first place.

### *C. Where is Style? Scoping the "Work"*

Another doctrinal rejoinder to claims that generative AI infringes copyright by appropriating artistic style is that copyright protects "works," not oeuvres.[158] Insofar as style inheres in a body of work rather than in a single work, the reasoning goes, stylistic similarity isn't cognizable as infringement. In general, copyright plaintiffs are unable to aggregate multiple works to argue that a defendant has infringed some expressive property present in the aggregate but absent in its constituent parts. The standard approach requires a plaintiff to show substantial similarity between the defendant's work and a *specific* work of the plaintiff's, rather than substantial similarity between the defendant's work and the plaintiff's oeuvre.[159] This rule — call it the anti-aggregation principle — is a serious hurdle for any plaintiff arguing the infringement of an expressive "style" that emanates from multiple

158. Sag, *supra* note 49, at 147 n.154; *Is GenAI allowed to produce outputs "in the style of" a human artist?*, KNOWING MACHINES (Oct. 10, 2023), https://knowingmachines.org/knowing-legal-machines/legal-explainer/questions/is-genai-allowed-to-produce-outputs-in%20the-style-of-a-human artist [https://perma.cc/2VNW-QMJE] ("It is very difficult to protect others from copying the style of human artists, as copyright law covers specific works of art, not general styles, methods, or genres."). *See also* Castle Rock Entm't, Inc. v. Carol Publ'g Grp., Inc., 150 F.3d 132, 138 (noting that "17 U.S.C. § 106 speaks in terms of a singular copyrighted 'work'"); 4 NIMMER ON COPYRIGHT, *supra* note 85, § 13.03[A][3] n.115.22 (2023) ("The rights accorded to copyright owners under 17 U.S.C. § 106 all refer to the 'work' in the singular.").

159. *See infra* note 175.

works.[160] Often, the stated goal of those invoking the anti-aggregation principle is to prevent the protection of ideas.[161]

But the anti-aggregation principle isn't airtight. For one, courts can ignore the boundaries of the "work" and protect expression aggregated from multiple works, sometimes finding infringement even without finding a defendant's work substantially similar to any single work by the plaintiff. For another, the "work" isn't a rigid constraint: courts routinely manipulate its boundaries, and copyright doctrine purposely affords the flexibility to do so. These features of copyright doctrine make the "work" an unprofitable focus.

1. Courts Look Past the Bounds of the "Work"

Even where the boundaries of a "work" are clear, courts have proven willing to find copyright interests that emerge from the aggregation of expression in multiple distinct works, and to find infringement without requiring a plaintiff to show substantial similarity to any particular work. This is most apparent when courts consider claims involving fictional characters.[162] A court in the Central District of California, for example, held that "James Bond is a copyrightable character" because "James Bond has certain character traits that have been *developed over time through the sixteen films in which he appears*."[163] The court enjoined the defendant, Honda, from airing a car commercial featuring a Bond-like character on the ground that "there is substantial similarity between the specific protected elements of the James Bond films" — note the plural "*films*" — "and the [defendant's] commercial."[164] For the most part, the "substantial" similarities the court enumerated were not specific to any particular Bond film: the court cited, for example, that both Bond and the commercial's protagonist were "young, tuxedo-clad, British-looking men with beautiful women in tow and grotesque villains close at hand . . . [who] exude uncanny calm under pressure,

160. Aggregate "style" seems to be at least part of the plaintiffs' theory in *Andersen*. *See* Andersen Compl., *supra* note 7, at ¶ 171.

161. *See, e.g.*, Hayuk v. Starbucks Corp., 157 F. Supp. 3d 285, 289–91 (S.D.N.Y. 2016). Nimmer criticizes *MGM v. Honda* — discussed *infra* — on the ground that "the danger arises that [MGM] could monopolize every scenario involving a suave spy who uses fancy gadgets while fraternizing with beautiful women — with the concomitant claim that every new spy flick that any defendant could develop would be substantially similar to elements of that earlier oeuvre." 4 NIMMER ON COPYRIGHT, *supra* note 85, § 13.03[A][3] n.115.33 (2023). Nimmer is effectively arguing that permitting a plaintiff to aggregate works could afford an unlawful monopoly over ideas (e.g., the idea of "a suave spy who uses fancy gadgets while fraternizing with beautiful women"), rather than expression. *See id.*

162. *See* Sag, *supra* note 78, at 334 ("[P]resenting a case in terms of the infringement of copyrightable characters frees the copyright owner from the burden of pointing to which specific work the defendant's product is infringingly similar.").

163. Metro-Goldwyn-Mayer, Inc. v. Am. Honda Motor Co. (*MGM*), 900 F. Supp. 1287, 1296 (C.D. Cal. 1995) (emphasis added).

164. *Id.* at 1298.

exhibit a dry sense of humor and wit, and are attracted to, and are attractive to, their female companions"; that the films and the commercial involved a "high-speed chase" with a "grotesque villain" and protagonists who "escape with the aid of intelligence and gadgetry;" and that the commercial and the films used dry dialog and exciting horn music.[165] The six similarities to particular Bond films that the court did identify were relatively small allusions to six distinct Bond films, and the court did not suggest that any of these similarities could establish substantial similarity between the commercial and any particular Bond film.[166]

The court's conclusion in *MGM* is unconvincing — a debonair Brit who thwarts villains in car chases and "[is] attracted to, and [is] attractive to, [his] female companion[]" is not a rich tapestry — but the court's willingness to find a copyright that transcends the boundaries of any one "work" is not uncommon, at least as far as fictional characters are concerned.[167] Circuit-court precedent blesses the approach.[168] In *DC Comics v. Towle*,[169] the Ninth Circuit considered whether DC Comics owned a copyright in Batman's vehicle, the Batmobile, which had been depicted with various appearances across decades in comic books, a television series, and a film.[170] The court articulated a three-part test for when a character is copyrightable: (1) the character must have "physical as well as conceptual qualities;" (2) it "must be 'sufficiently delineated' to be recognizable as the same character whenever it appears;" and (3) it "must be 'especially distinctive' and 'contain

---

165. *Id.*

166. Nor does the notion seem plausible: the similarities to specific movies that the court identified almost certainly constitute ideas rather than expression. For example, three of the six similarities to six distinct Bond films were that both the commercial and one Bond movie featured a car with a detachable roof; that both the commercial and another Bond movie depicted a villain with metal hands; and that both the commercial and a third Bond movie showed a villain with metallic teeth who sometimes wears googles. *Id.*

167. *See, e.g.*, Toho Co. v. William Morrow & Co., 33 F. Supp. 2d 1206, 1216 (C.D. Cal. 1998) ("Godzilla has . . . *developed* a constant set of traits that distinguish him/her/it from other fictional characters. While Godzilla may have shifted from evil to good, there remains an underlying set of attributes that remain *in every film* . . . Godzilla is a well-defined character with highly delineated consistent traits." (emphases added)); Anderson v. Stallone, No. 87-0592, 1989 WL 206431, at *7 (C.D. Cal. Apr. 25, 1989) ("The Rocky characters are one of the most highly delineated group of characters in modern American cinema. The physical and emotional characteristics of Rocky Balboa and the other characters were set forth in tremendous detail *in three Rocky movies* . . . . The interrelationships and development of Rocky, Adrian, Apollo Creed, Clubber Lang, and Paulie are central to *all three movies*." (emphases added)). *Toho*, *Anderson*, and *MGM* were all cited by the Ninth Circuit to support the proposition that "[c]haracters that have received copyright protection have displayed *consistent*, widely identifiable traits." Rice v. Fox Broad. Co., 330 F.3d 1170, 1175 (9th Cir. 2003), *overruled on other grounds by* Skidmore as Tr. for Randy Craig Wolfe Tr. v. Led Zeppelin, 952 F.3d 1051 (9th Cir. 2020).

168. *See, e.g.*, *Rice*, 330 F.3d at 1175; DC Comics v. Towle, 802 F.3d 1012, 1015, 1019, 1022 (9th Cir. 2015).

169. 802 F.3d 1012 (9th Cir. 2015).

170. *Id.* at 1015, 1019, 1022.

some unique elements of expression'" rather than be merely a "stock character."[171] Looking across the comic books, the television series, and the film, the Ninth Circuit concluded that the Batmobile satisfied this test.[172] The *DC Comics* test is explicit that a copyrightable character can emanate from multiple works: "[c]onsidering the character *as it has appeared in different productions*, it must display consistent, identifiable character traits and attributes, although the character need not have a consistent appearance."[173] Indeed, a district court even observed that "though perhaps not an explicit requirement, cases finding copyrightable characters have typically seemed to implicitly require that the characters appear several times, in different books, movies, or other presentations."[174]

Admittedly, when not considering fictional characters, courts hesitate to aggregate works to find an emergent copyright interest. A line of district court decisions from the influential Southern District of New York refuses to permit claims predicated on the expressive style of an aggregation of works.[175] An illustrative case, *Hayuk v. Starbucks*[176] dismissed an artist's claim of copyright infringement arising out of a Starbucks advertising campaign that allegedly "appropriate[d] the 'total concept and feel' of [several of the artist's works] in their entirety."[177] Hayuk's works arranged numerous bright colors in triangles,

---

171. *Id.* at 1021 (quoting Walt Disney Prods. v. Air Pirates, 581 F.2d 751, 755; *Rice*, 330 F.3d at 1175; *Halicki*, 547 F.3d at 1224).

172. *Id.*

173. *Id.* (emphasis added); *see also* Metro-Goldwyn-Mayer, Inc. v. Am. Honda Motor Co. (*MGM*), 900 F. Supp. 1287, 1296 (C.D. Cal. 1995) ("James Bond is a copyrightable character" in part because "James Bond has certain character traits *that have been developed over time through the sixteen films* in which he appears." (emphasis added)). *See* Matthew Sag, *Copyright Safety for Generative AI*, 61 HOUS. L. REV. 295, 332–33 (2023).

174. Rice v. Fox Broad. Co., 148 F. Supp. 2d 1029, 1056 (C.D. Cal. 2001) *rev'd on other grounds*, 330 F.3d 1170 (9th Cir. 2003). *But see* Klinger v. Conan Doyle Est., Ltd., 755 F.3d 496, 503 (7th Cir. 2014) (Posner, J.) ("From the outset of the series of Arthur Conan Doyle stories and novels that began in 1887 Holmes and Watson were distinctive characters and therefore copyrightable.").

175. *See* Kroencke v. Gen. Motors Corp., 270 F. Supp. 2d 441, 444 (S.D.N.Y. 2003), *aff'd*, 99 F. App'x 339 (2d Cir. 2004) ("[N]othing in the Copyright Act of 1976 (which refers to the infringed 'work' in the singular) or in the precedents of this Circuit supports the view that a plaintiff's entire oeuvre, or even an aggregated portion of it, may be used as the point of comparison where the works . . . bear little or no relation to one another beyond 'style.'"), *cited with approval in* Dean v. Cameron, 53 F. Supp. 3d 641, 647 (S.D.N.Y. 2014); Camhe v. Dreamworks, LLC, No. CV 07-3741, 2009 WL 10668462, at *5-6, *15 (C.D. Cal. May 14, 2009) (rejecting plaintiffs' argument that they can aggregate a movie screenplay and a television treatment in order to show substantial similarity to the defendant's television series and granting summary judgment to defendant on the ground that "[p]laintiffs' works and [defendant's TV series] contain similarities, but in elements that are not protectable"); Judith Ripka Designs, Ltd. v. Preville, 935 F. Supp. 237, 248 (S.D.N.Y. 1996) ("The copyright laws do not protect styles, but only particular original designs."). *But cf.* Damiano v. Sony Music Ent., Inc., 975 F. Supp. 623, 626, 630 (D.N.J. 1996) (dismissing complaint for lack of substantial similarity but purporting to analyze "as a whole" lyrics from distinct songs by the plaintiff).

176. 157 F. Supp. 3d 285 (S.D.N.Y. 2016).

177. *Id.* at 291.

quadrilaterals, and circles; the Starbucks advertisements contained similar configurations of bright colors.[178] The court rejected Hayuk's argument that she could plausibly show infringement through substantial similarity between the Starbucks campaign and expression present in multiple of her works.[179] It held instead that Hayuk was required to demonstrate that the alleged infringement was substantially similar to a *particular* work of hers, and it suggested that Hayuk's theory amounted to "an assertion that Defendants have copied her style or elements of her ideas, neither of which are protected by copyright law."[180]

178. *Id.* at 287–88. Hayuk also alleged that an advertising agency acting on Starbucks's behalf had approached her about creating artwork for a Starbucks campaign, but she had declined. *Id.* at 288; Complaint at 7–8, Hayuk v. Starbucks, 157 F. Supp. 3d 285 (S.D.N.Y. 2016) (No. 15-cv-04887) [hereinafter Hayuk Compl.].

179. *Hayuk*, 157 F. Supp. 3d at 291–92.

180. *Id.*

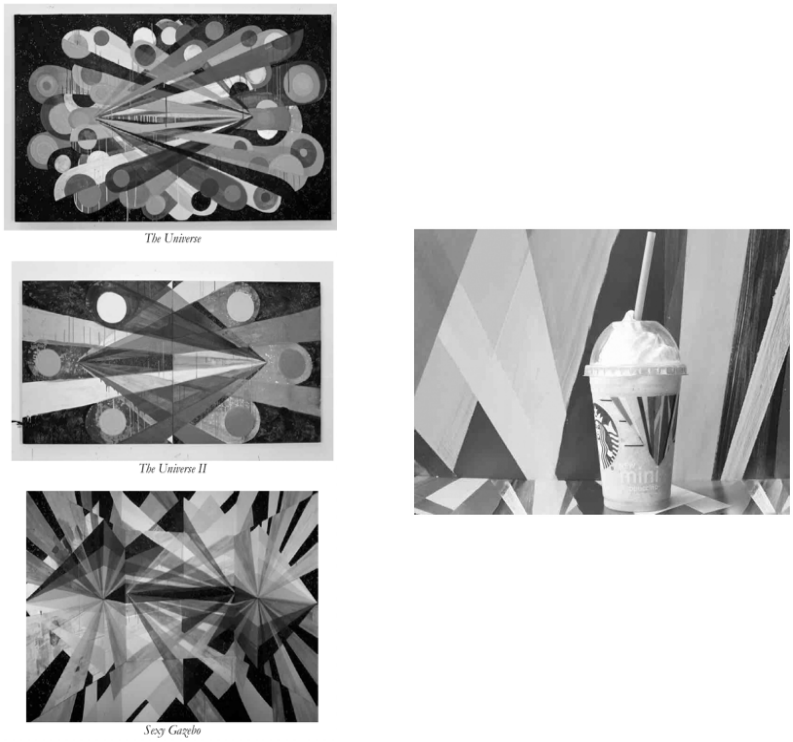



Figure 7: Three of Hayuk's Allegedly Infringed works (left), and an Allegedly Infringing Image from Starbucks's Website (right).[181]

Courts have also refused to aggregate a *defendant's* works, which is a request that plaintiffs and defendants alike have made in different contexts. The Second Circuit rejected a defendant's double-barreled argument that its copying was de minimis because (a) it copied expression from just twenty articles out of the 90,000 the plaintiff, a newspaper publisher, published annually; and (b) the defendant's twenty infringing publications "represent only a tiny fraction of the 17,000" such publications the defendant made that year.[182] A district court in Hawaii also refused a plaintiff's request to aggregate five novels by a defendant — two pairs of which shared common protagonists — for comparison to the plaintiff's novel, reasoning that to adopt the plaintiff's theory would let plaintiffs mix and match "random similarities" between discrete works and thereby create a composite of the

181. Hayuk Compl., *supra* note 178, at 6, 13.

182. Nihon Keizai Shimbun, Inc. v. Comline Bus. Data, Inc., 166 F.3d 65, 71–72 (2d Cir. 1999).

plot in one book by the defendant, the characters in another, the dialog in yet another, and so on.[183]

But none of the aforementioned cases refusing to consider multi-work artistic "style" has been affirmed in a precedential circuit-court opinion.[184] Moreover, the fictional-character cases discussed above belie the argument that courts are categorically barred from recognizing an emergent interest in copyrighted expression that is spread across works. And at the margins, aggregationist impulses creep into decisions involving visual art. For instance, one district court adjudicating a motion to dismiss professed to examine discrete photographs — taken by the same photographer, but published in different venues over a span of twelve years — as an aggregate for the purposes of assessing substantial similarity to an allegedly infringing music video.[185]

### 2. The "Work" is Contingent, and Generative AI Changes the Contingencies

Insofar as the concept of the "work" does constrain courts' determinations of a copyright's scope, that legal concept is an intentionally flexible one. When evaluating media of all kinds, courts frequently make ad hoc judgment calls about the proper scope of the work or works before them. Although courts' decisions about the boundaries of copyrighted works might benefit from clearer guiding principles, judicial manipulation of the shape of the work isn't necessarily chicanery. Rather, the indeterminacy of copyright's concept of the "work" is by design. As the Ninth Circuit explains, "[T]he [copyright] statute

183. Doody v. Penguin Grp. (USA) Inc., 673 F. Supp. 2d 1144, 1151, 1155–56 (D. Haw. 2009).

184. *Kroencke v. General Motors Corp.*, discussed *supra* note 175, was affirmed by the Second Circuit in a nonprecedential summary order. 99 Fed. Appx. 339 (2d Cir. 2004) (summary order); *see* 2d Cir. R. 32.1.1(a) ("Rulings by summary order do not have precedential effect."). The summary order in *Kroencke* did not explicitly discuss the theory of multi-work stylistic infringement that Kroencke advanced below, although Kroencke's appellate briefing renewed the argument. *See* Brief for Plaintiff-Appellant at 33–36, Kroencke v. Gen. Motors Corp., 99 F. App'x 339 (2d Cir. 2004) (No. 03-7839). The summary order acknowledged that "style is . . . one of many considerations in analyzing substantial similarity" but concludes that the "general stylistic sense" in which the parties' works might be similar was not an infringement. *See Kroencke*, 99 F. App'x at 340–41. Finally, the order discussed Kroencke's works in the aggregate for the purpose of rejecting substantial similarity: "although certain details in Kroencke's works are similar to details in defendants' advertisement, these details are not exact copies but instead commonplace depictions." *Id.* at 340–41.

185. LaChapelle v. Fenty, 812 F. Supp. 2d 434, 446 n.77 (S.D.N.Y. 2011). Admittedly, despite the court's statement that its analysis would "treat[] the [plaintiff's] [p]hotographs collectively," its reasoning in fact rested on individualized determinations that three particular photographs were plausibly substantially similar to three particular scenes of the music video. *See id.* at 446–47, 446 n.77. Thus, while it's tempting to speculate that the defendants' motion to dismiss might not have been denied if the photographer had alleged similarity between just one of his photographs and just one scene of the music video — rather than similarity between multiple distinct scenes of the video and multiple distinct photos from his oeuvre — the court's reasoning doesn't support that conjecture.

purposefully left 'works of authorship' undefined to provide for some flexibility,"[186] and scholars argue that "[w]e should generally embrace this flexibility."[187]

A flexible definition of the "work" allows courts to delimit a unit of copyrighted expression to match dynamic expectations in culture and commerce. Precedent demonstrates that courts determine the scope of a work with sensitivity to aesthetics, authorial intent, and business practices although they might not always do so explicitly, let alone uniformly. It is precisely these cultural and commercial conventions that generative AI has upended. Thus, those who oppose copyright liability for image-generating AI should invoke the anti-aggregation principle with some measure of humility. The corners of a canvas or the start- and end-markers of a JPEG file have never defined the boundaries of a work.[188] Moreover, to the extent that "one file, one work" has been a default presumption of the 2010s, our copyright regime was designed to ensure that this view of the work wouldn't necessarily bind us into the 2020s.

Issues of both copyright ownership and substantial similarity can hinge on decisions about the boundaries of a "work" that privilege pragmatism above doctrinal rigidity. In *Garcia v. Google*,[189] an en banc Ninth Circuit considered whether an actress held a copyright in a five-second acting performance in a film. A majority of judges concluded that the claim was not likely to succeed.[190] But the *Garcia* majority didn't hold that the actress made no original contribution to the scene,

---

186. Garcia v. Google, Inc., 786 F.3d 733, 741 (9th Cir. 2015).

187. Margot E. Kaminski & Guy A. Rub, *Copyright's Framing Problem*, 64 UCLA L. REV. 1102, 1168 (2017).

188. For instance, two Manet paintings — *The Bullfight* and the masterpiece *The Dead Toreador* — were originally portions of the same painting, *Incident at a Bullfight*. *Édouard Manet's The Dead Toreador*, WORCESTER ART MUSEUM, https://www.worcesterart.org/exhibitions/past/toreador.html [https://perma.cc/4HR5-44L2]. Manet sliced *Incident at a Bullfight* into the two works after critics "eviscerated" the larger canvas. *Id.*; *Manet's* The Dead Toreador *and* The Bullfight: *Fragments of a Lost Salon Painting Reunited,* THE FRICK COLLECTION, https://www.frick.org/exhibitions/past/1999/manets-dead-toreador-and-bullfight-fragments-lost-salon-painting-reunited [https://perma.cc/9F7M-E3H5]. Other examples, analog and digital, are easy to come by. *See, e.g.*, Blanch v. Koons 467 F.3d 244, 246 (2d Cir. 2006) (describing a collage on canvas by the artist Jeff Koons that reproduced a photographic work by the plaintiff); *The Million Dollar Homepage*, WIKIPEDIA (July 22, 2023), https://en.wikipedia.org/wiki/The_Million_Dollar_Homepage [https://perma.cc/3NKR-AAHM] ("The [Million Dollar Homepage] consists of a million pixels arranged in a 1000 × 1000 pixel grid; the image-based links on it were sold for $1 per pixel in 10 × 10 blocks. The purchasers of these pixel blocks provided tiny images to be displayed on them . . . .").

189. 786 F.3d 733 (9th Cir. 2015).

190. *Id.* at 740. It's worth noting that the Garcia's lawsuit against Google wasn't just some shakedown. The court explained the case's tragic circumstances: "By all accounts, Cindy Lee Garcia was bamboozled when a movie producer transformed her five-second acting performance into part of a blasphemous video proclamation against the Prophet Mohammed." *Id.* at 737. After the producer uploaded the video to YouTube, the actress received death threats, and she filed a copyright lawsuit seeking to enjoin Google to remove the film from its platforms. *Id.* at 738.

and it offered little to justify its outcome-determinative assertion that the actress "played no role in [the] fixation" of her performance.[191] Instead of anchoring its decision in copyright's first principles, the court primarily appealed to the practical concern that allowing an actress to claim authorship in a scene of a film would upend the film industry. Recognizing the actress's claim to authorship, the majority explained, "would result in [a] legal morass . . . splintering a movie into many different 'works,' even in the absence of an independent fixation."[192] The court continued, "Treating every acting performance as an independent work would not only be a logistical and financial nightmare, it would turn cast of thousands into a new mantra: copyright of thousands."[193]

The *Garcia* majority's refusal to disaggregate the actress's creative contribution from the film is arguably in tension with elementary copyright principles. As a dissenting judge observed, the actress's performance was minimally creative and fixed in a tangible form, and copyright doctrine does not ordinarily require a performer to record her performance herself in order to qualify as an author.[194] These details would suggest that the actress was indeed entitled to an independent copyright in the original elements of her performance that were reflected in the film. The dissent suggested that the basis for the majority's ruling was not first principles or doctrinal bedrock, but rather "speculati[on] that a contrary rule might curb filmmaking and burden the internet."[195]

Courts fudge the boundaries of the work for pragmatic reasons at the substantial similarity stage, too. In cases involving serialized media, courts will treat discrete, freestanding episodes as a unitary work for the purposes of assessing similarity, and sometimes even turn around and assess statutory damages on a per-episode basis in the same opinion.[196] In *Castle Rock Entertainment v. Carol Publishing*,[197] the Second

191. *Id.* at 744.

192. *Id.* at 742.

193. *Id.* at 743.

194. *Id.* at 749–50 (Koziniski, J., dissenting). *See also* 1 NIMMER ON COPYRIGHT, *supra* note 85, § 2.03[B][3] (2023) ("[I]t is not necessary for a singer to personally set up the microphone, an actor to operate the camera, or the like. To the extent that a band member voluntarily participates in a recording session, or an actor voluntarily participates in a film shoot, the resulting product may be considered pursuant to the requisite authority.").

195. *Garcia*, 786 F.3d at 753.

196. *See, e.g.*, Twin Peaks Prods., Inc. v. Publ'ns Int'l., Ltd., 996 F.2d 1366, 1372–73, 1381 (2d Cir. 1993) (opining that a book about an episodic television show could be found comprehensively, nonliterally similar to the scripts for the show because "[e]very intricate plot twist and element of character development appear in the Book in the same sequence as in the teleplays" but later holding separately that "ours is the easy case of infringement of eight separate works that warrants eight statutory [damages] awards"). *See also* Castle Rock Ent., Inc. v. Carol Publ'g Gr., Inc., 150 F.3d 132, 138 (2d Cir. 1998) (stating that the *Twin Peaks* case "f[ound] substantial similarity between infringing book and 8 episodes of Twin Peaks weekly television series seen as a whole").

197. 150 F.3d 132 (2d Cir. 1998).

Circuit considered whether the defendant's book of trivia questions about the television series *Seinfeld* infringed the plaintiff's copyright in each episode of the series.[198] The book comprised 643 trivia questions drawn from eighty-four different *Seinfeld* episodes.[199] The court rejected the argument that the copying from any given episode was de minimis and instead "analyze[d] in the aggregate the amount copied from the eighty-four *Seinfeld* episodes."[200] Because the plaintiff's expression appeared in "an entire continuous television series," the court saw "no basis for looking in isolation at the amount copied from each separately copyrighted episode."[201] Had the defendants "copied a few fragments from each of 84 unrelated television programs," the court remarked, ". . . [they] would have a stronger case under the *de minimis* doctrine."[202]

Several treatise-writers criticize *Castle Rock* for aggregating a plaintiff's works.[203] Nimmer urges "caution" before permitting such aggregation, because "the broader the series, the more all-encompassing plaintiff's copyright becomes, thereby squelching new expression in direct defiance to copyright's mandate of stimulating the production of new works."[204] In *Warner Brothers Entertainment v. RDR Books*,[205] a court in the Southern District of New York addressed — and discounted — Nimmer's concern. *RDR* held that an encyclopedia of "persons, places, spells, and creatures from the *Harry Potter* works" was substantially similar to the seven-book *Harry Potter* series as a whole.[206] The court noted explicitly that it chose to "analyze[] the amount of expression copied from the *Harry Potter* series in the aggregate."[207] It dismissed Nimmer's worry about aggregating multiple works: "because the *Harry Potter* novels tell one coherent narrative in a series, rather than tell discrete tales, the danger identified by Nimmer is less likely to exist."[208]

Courts may also aggregate a defendant's works when assessing substantial similarity. In *Warner Brothers v. American Broadcasting*

198. *Id.* at 135.
199. *Id.* at 135–36.
200. *Id.* at 138.
201. *Id.*
202. *Id.*
203. *See* 4 NIMMER ON COPYRIGHT, *supra* note 85, § 13.03[A][3] ("[A] danger exists here — the broader the series, the more all-encompassing plaintiff's copyright becomes, thereby squelching new expression in direct defiance to copyright's mandate . . . . Together with the statutory language . . . that favors a focus on each individual work, great caution is required before adopting any rule that plaintiff may aggregate the works of a series into a single claim."); 3 William F. Patry, PATRY ON COPYRIGHT § 9:66.
204. 4 NIMMER ON COPYRIGHT, *supra* note 85.
205. 575 F. Supp. 2d 513 (S.D.N.Y. 2008).
206. *Id.* at 522, 535, 535 n.14.
207. *Id.* at 535 n.14.
208. *Id.*

*Companies*,[209] the Second Circuit affirmed summary judgment in favor of a defendant on the ground that a bumbling superhero character in its television series was not substantially similar to the plaintiff's character, Superman. The defendant's character had been advertised in over a dozen distinct, short television "promos," and the court held it proper to consider "the visual impact of the series of 'promos'" as a whole.[210] To be sure, the court also held that no single "promo" was substantially similar to the plaintiff's expression, and it reserved in dicta the possibility that "within a series of generally non-infringing 'promos' a single 'promo' could be so substantially similar to a copyrighted character as to establish infringement."[211] But *Warner Brothers v. ABC*'s approach to aggregation would seem to help a defendant who establishes, over a series of works, fundamental dissimilarities between its character and a plaintiff's character that might not be evident from viewing just one of the defendant's works in isolation.

Cases like *Garcia*, *Castle Rock*, and *Warner Brothers v. ABC* teach that the "work" isn't a rigid doctrinal form, and that courts aren't searching for ontological truths when they ask whether some quantum of expression fits that form. Instead, their answers depend on whether it makes sense pragmatically — in light of the prevailing conventions for consuming expression and the business practices that accompany them — to treat that quantum of expression as a discrete "work." Kaminski and Rub document that courts, when determining how to delimit a "work," may consider the copyrightability of the discrete work or its components, the market demand for (and supply of) the discrete work, the author's intent for how the work would be consumed, and whether the work is independently registered, among other things.[212]

Many of the factors that courts look to in delimiting the "work" are socially and technologically contingent, and this means that the "work" itself is socially and technologically contingent. Media technologies influence the format of the expression we consume; the format of the expression we consume influences how we expect media to look; market structures cater to our expectations; and each of these forces modulates in response to the others. Some technologies — say, animated GIFs and digital singles — push towards ever-smaller conceptions of the work.[213] Others, like the rise of the "binge-watching" television format, encourage us to regard discrete, episodic works as a unified aesthetic whole.[214] Markets meet these expectations: iTunes emerges to sell songs a la carte; streaming services release entire series at once;

209. 720 F.2d 231 (2d Cir. 1983).
210. *Id.* at 237, 243 (2d Cir. 1983).
211. *Id.* at 244.
212. Kaminski & Rub, *supra* note 187, at 1151–66.
213. *Id.* at 1116–17.
214. *Id.* at 1117.

companies furnish "curated" text and image datasets in bulk for training AI.[215]

Because the "work" in practice is technologically and socially contingent, it's a very good thing that the law's definition is, too. Generative AI is likely to change the technological, social, and commercial factors that courts — and the artists and businesses before them — look to when deciding where to draw the lines defining the "work." The copyright statute's definition of the "work" is deliberately flexible. Why, then, would we react to generative AI by freezing the definition of the "work" in the late 2010's?

3. A Substantially-Similar-To-One-Work Rule Just Duplicates the Idea-Expression Distinction

Skeptics of copyright in visual artistic style might find it expedient to insist upon enforcing particular boundaries for works of visual art and to insist that those boundaries dictate particular legal consequences. This isn't a good strategy. As we've seen, its doctrinal underpinnings are questionable. Caselaw shows that the "work," by design, lacks rigid boundaries, and that courts define those boundaries in pragmatic, context-specific ways. Caselaw further shows that, at least in the fictional-character context, courts sometimes ignore the work's boundaries entirely and instead recognize emergent copyright interests in expression aggregated from multiple discrete works. Moreover, on a practical level, insisting that substantial similarity requires one-work-to-one-work comparison doesn't even thwart claims of infringement by stylistic similarity. Giving dispositive legal weight to expression's packaging will just encourage copyright owners to package their expression in the manner that affords them the broadest rights.

Some hypotheticals about a painter named Artist illustrate that giving the "work" special doctrinal significance won't prevent the protection of expressive style. Over her mature career, Artist has made 500 oil paintings in an original style. She usually paints portraits, and she has always sold her paintings as individual canvases. Without her consent, digital images of each painting were used to train image-generating AI.

(A) A user of the image-generating AI service prompts the AI to render oil painting portrait in the style of Artist. No reasonable jury could find substantial similarity between the generated image and any single painting by Artist.

---

215. *Id.*; *see also, e.g.*, Sarah Whitten, *Netflix's binge-release model is under new scrutiny as the streaming giant struggles*, CNBC (June 15, 2022, 9:39 PM), https://www.cnbc.com/2022/06/15/will-netflix-stop-binge-releases-experts-weigh-in.html [https://perma.cc/WPA2-T2U6]; Cogito, https://www.cogitotech.com/ [https://perma.cc/3PBF-B2DP].

(B) Same setup, except that Artist has also published a coffee-table book containing a full-page reproduction of each of her 500 paintings. A user of the image-generating AI service prompts the AI to render page number 501 in Artist's coffee-table book. The AI service generates an image indistinguishable from the image generated in (A). A reasonable jury could find substantial similarity between the generated image and the coffee-table book.

(B′) Same facts as (B), except Artist has never before sold or displayed any of her canvases. Instead, she publishes her paintings for the first time in a single-volume book.

The straightforward doctrinal answers are that (A) is not copyright infringement and (B) and (B)′ are. The breadth of Artist's copyright shifts with the definition of the relevant work.

A critical reader might contest the terms of the hypothetical and argue that (A) is irreconcilable with (B); if a reasonable jury couldn't find the generated image substantially similar to any one painting by Artist, then no reasonable jury could find the generated image substantially similar to the book of all of Artist's paintings. One way of interpreting this critique is simply as an insistence that the book of paintings cannot be the relevant "work." But (B)′ belies that premise: to conclude that the book in (B)′ is not a unitary work but instead a compilation of discrete works, a court would have to ignore the tangible boundaries of the publication, the fact that the book is marketed as a single work, and Artist's apparent aesthetic intention. Under the circumstances, it seems unlikely that a court would impose its own contrary aesthetic judgment.[216] At the very least, it would be defensible for a court to treat the entire book as the relevant comparator.

A subtler critic might argue that hypotheticals (A) and (B) are inconsistent because, for every pair of substantially similar works $\alpha$ and $\beta$, there is necessarily some sub-component of $\alpha$ that is substantially similar to $\beta$. But a counterexample refutes this argument: if $\alpha$ is a phrase just long enough to contain copyrightable originality, then any subdivision of $\alpha$ is ineligible for copyright protection and thereby not grounds for substantial similarity.[217] Thus, no sub-component of $\alpha$ can be substantially similar to $\beta$. Copyright doctrine adheres to this logic: the

216. *See supra* Section III.C.2 (discussing factors that courts look to when determining the "work"); *see also* Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith, 143 S. Ct. 1258, 1283 (2023) ("A court should not attempt to evaluate the artistic significance of a particular work.").

217. *See* 37 C.F.R. § 202.1 (2014) (providing that "[w]ords and short phrases" are "not subject to copyright").

"total concept and feel" test posits not similarity between *components* of works, but between their *totalities*.[218]

Finally, the most tempting way to conclude that (A) and (B) are inconsistent is simply to ignore the stipulated facts. The reader who bristles at the hypotheticals is probably imagining an AI-generated image that shares with Artist's book only general, high-level similarities. Indeed, an AI user who inputs a prompt like that in (B) might, more often than not, produce output that is only non-infringingly similar. But this imagined situation changes materially the facts of hypothetical (B). The hypothetical is explicit that a reasonable jury could find substantial similarity; thus, the similarities necessarily go beyond the unprotectable. That such similarity might be comparatively unlikely to result in practice does not change what the relevant unit of comparison is: it's Artist's book, not the individual pages in that book.

To insist that certain visual artists must always prove similarity to a single image is to single them out for unfavorable treatment. Treating hypothetical (B) as infringement is consistent with copyright's treatment of pictorial works like comic strips and graphic novels, as well as other non-pictorial media.[219] An infringing sequel to *Catcher in the Rye* need not be substantially similar to any single paragraph or page of the book authored by J.D. Salinger. Indeed, a court may find infringement even if a defendant's work is not substantially similar to any *single* work of the plaintiff's.[220]

To be clear, I think *MGM v. Honda* is an unfortunate decision, but I base my conclusion on its outcome, not its method.[221] Copyright protects expression from being copied, not from being referred to or evoked. *MGM* effectively permitted a plaintiff to use copyright to prevent a defendant from using ideas from the Bond franchise to evoke James Bond. The court admitted as much in its formulation of the intrinsic similarity test: "it appears likely that the average viewer would immediately *think of James Bond* when viewing the [defendant's] commercial."[222] That a viewer would "think of James Bond" tells us nothing

---

218. *See* Tufenkian Imp./Exp. Ventures, Inc. v. Einstein Moomjy, Inc., 338 F.3d 127, 134 (2d Cir. 2003) ("[T]he total-concept-and-feel locution functions as a reminder that, while the infringement analysis must *begin* by dissecting the copyrighted work into its component parts in order to clarify precisely what is not original, infringement analysis is not *simply* a matter of ascertaining similarity between components viewed in isolation.").

219. *See* Walt Disney Prods. v. Air Pirates, 581 F.2d 751, 754–55 (9th Cir. 1978) (holding Disney cartoon characters copyrightable where Disney's copyrights "in some instances . . . cover[ed] a book and others an entire strip of several cartoon panels").

220. *See* Metro-Goldwyn-Mayer, Inc. v. Am. Honda Motor Co. (*MGM*), 900 F. Supp. 1287, 1298–99 (C.D. Cal. 1995); *see supra* Section III.C.1 (discussing *MGM*).

221. My disapproval of *MGM* may not be unique. *See* Aaron Schwabach, FAN FICTION AND COPYRIGHT: OUTSIDER WORKS AND INTELLECTUAL PROPERTY PROTECTION 38 (2011) ("The [*MGM*] court's reasoning is deeply, deeply disturbing."); Tze Ping Lim, *Beyond Copyright: Applying a Radical Idea-Expression Dichotomy to the Ownership of Fictional Characters*, 21 VAND. J. ENT. & TECH. L. 95, 113 (2018).

222. *MGM*, 900 F. Supp. at 1299 (emphasis added).

about substantial similarity. A viewer would "think of James Bond" if Honda had named its car the "Honda James Bond," but that doesn't make it *copyright infringement* to do so.[223]

But just because *MGM* protected something unprotectable doesn't mean that its analytical method necessarily does. *MGM* still would have been a bad decision if all the challenged elements of the Honda commercial had appeared in one single Bond film. If the objection to copyright in artistic style is that it protects unprotectable ideas, then that objection holds whether those ideas are expressed in a single work or across multiple works. Insisting that the relevant expression be present in a single work is a formalism that overemphasizes the easily manipulated, contextually contingent, and sometimes-ignored legal category of the "work." It is a formalism that does no useful work that the idea-expression distinction doesn't already do.

### *D. Summing Up: Similarity You Can Describe Versus Similarity You Can See*

Style is a holistic attribute of a work, or a group of works, that comprises a constellation of expressive choices. Non-experts may lack the vocabulary to describe style, despite being able to "know it when they see it." Perhaps because visual art "comprise[s] . . . a large array of elements, some combination of which is protectable by copyright," rather than a few discrete constituent elements, courts seem to find images harder to dissect analytically.[224] And because "style" also describes a gestalt effect that unites many discrete expressive choices, that term can encompass a breadth of aesthetically significant information in visual art. Notwithstanding the frequent assertion that copyright doesn't protect style, the broad and imprecise sweep of "style" in visual art means that one can infringe copyright by copying elements of a pictorial work that could be described as "style."

In this summation, a subtle but vital point emerges: the most novel and remarkable power of generative AI is its ability to produce outputs that resemble its training corpus in the ways that are the hardest to express analytically. Sometimes those aesthetic properties might be unrelated to copyrightable expression; it can be hard, for example, to verbalize why an image looks more or less like an authentic human face, but photorealistic depiction of immutable human anatomy is not

223. Trademark infringement is another story. *See, e.g.*, Danjaq LLC v. Sony Corp., No. CV97-8414, 1998 WL 957053, at *4 (C.D. Cal. July 29, 1998), *aff'd*, 165 F.3d 915 (9th Cir. 1998) (discussing the "James Bond service mark").

224. *Cf.* Swirsky v. Carey, 376 F.3d 841, 849 (9th Cir. 2004) (discussing music in the quoted language but stating that music is "like . . . art objects").

copyright's heartland.[225] Other times, however, those similarities might squarely implicate expressive style. Recall the Hollie Mengert-style images reproduced in Figure 1 at the beginning of this Article.[226] At least to this lay observer, the AI-generated images share meaningful aesthetic similarities with Mengert's oeuvre, even if those similarities are hard to articulate precisely.[227] The same is true of AI-generated sound: we now have songs that really do sound like Homer Simpson singing Adele[228] or Johnny Cash covering Aqua.[229] But if asked to explain why the recordings sound like the voices they mimic, or why the Hollie-Mengert-Dreambooth-model outputs resemble her work, most observers — myself included — probably couldn't offer much better than, "they just do."[230]

Generative AI can identify, and replicate, forms of similarity that we can see, but not necessarily describe. As Part II's technical exposition explained, the technology can be understood to function by mapping the interstices in our conceptual vocabulary. It allows us to render images "in between" visual concepts that we have mapped to words. It can reveal aesthetic commonalities between images, and riff on those common features, even if we have not identified the commonalities as such. This architecture allows us to visualize similarities that are rooted in our knowledge, vocabulary, and conventions, but which emerge only as higher-order, systemic properties of that body of knowledge,

---

225. Anecdotal evidence suggests that not-quite-right representations of human likenesses provoke revulsion on an "I know it when I see it" basis. The so-called "uncanny valley" phenomenon hypothesizes that objects that bear close, but imperfect, resemblances to genuine human likeness provoke eerie unease. Emily Kendall, *Uncanny Valley*, BRITANNICA (Sept. 30, 2024), https://www.britannica.com/topic/uncanny-valley [https://perma.cc/AZ8C-VGQL]. But attempts to prove the phenomenon are inconclusive, and a variety of competing explanations have been offered. *See id.* To the extent we experience the "uncanny valley," it seems to trigger an eeriness that we just *feel*, rather than something we can explain. For research on using AI to generate realistic-looking images of faces, *see generally* Tero Karras, Samuli Laine & Timo Aila, A Style-Based Generator Architecture for Generative Adversarial Networks (Mar. 29, 2019) (unpublished manuscript) (on file with arXiv), https://arxiv.org/abs/1812.04948 [https://perma.cc/6XAJ-QU84] (using artificial intelligence to generate realistic-looking images of faces).

226. *See supra* notes 1–4 and accompanying text.

227. Notably, Mengert herself attributes these similarities to "AI . . . kind of mimic[king] brush textures and rendering, and pick[ing] up on some colors and shapes," but she notes, "As far as the characters, I didn't see myself in it. I didn't personally see the AI making decisions [] that I would make, so I did feel distance from the results. . . . [I]t isn't actually mimicking my style." Baio, *supra* note 3.

228. Poqunawff, *Homer Simpson — Somebody That I Used To Know ft. Marge Simpson (AI Cover)*, YOUTUBE, https://www.youtube.com/watch?v=qOTWpKitlPg [https://perma.cc/Q26G-ZZ4T].

229. Merasmus Entertainment, *Johnny Cash — Barbie Girl (Cover by There I Ruined it) Restoration*, YOUTUBE, https://www.youtube.com/watch?v=MAFdzBTe2lg [https://perma.cc/Q35N-NLBN].

230. I appreciate that AI emulation of performers' voices may present different doctrinal issues than AI emulation of visual artists' styles. *See* Midler v. Ford Motor Co., 849 F.2d 460, 462 (9th Cir. 1988) ("A voice is not copyrightable. The sounds are not 'fixed.'").

vocabulary, and convention. It is perhaps no accident, then, that the most remarkable capabilities of image-generating AI are precisely the ones that defy verbal description: generative AI's defining attribute may be its ability to navigate the latent space between our words and concepts, and to conjure apparently meaningful outputs from within that wordless interstitial space.

It is also perhaps no accident, then, that generative AI researchers often measure the success of their image-generating models by the gestalt, inarticulate, and intuitive reactions that humans have to the models' outputs. For example, pioneering work on "style transfer" from 2016 observed candidly, "[W]e consider style transfer to be successful if the generated image 'looks like' the style image but shows the objects and scenery of the content image. We are fully aware though that this evaluation criterion is neither mathematically precise nor universally agreed upon."[231] More recently, researchers from OpenAI have conducted "human evaluation" of their models, which involves asking humans to preference the output of one model against the output of another model for "caption similarity" and "photorealism."[232] And a landmark StableDiffusion paper undertook something similar: to measure a model's ability to improve photo resolution and edit existing images, the authors displayed competing images side-by-side and asked survey respondents, "Which of the two images is a better high quality version of the low resolution image in the middle?" and "Which of the two images contains more realistic inpainted regions of the image in the middle?'"[233] In image generation, inarticulate human intuition about when one picture "looks like" another (or is "better" or "more realistic") remains the ultimate measure of success. It is all "intrinsic" similarity, with no "extrinsic" test.[234]

What makes today's image-generating AI such a striking phenomenon and such an intriguing problem for the law of similarity is that its greatest power is in identifying and replicating precisely the similarities we fail to capture verbally. The results are easy to see, yet difficult to articulate. The law is an analytical, verbal profession; judges struggle to reason juridically about phenomena that defy verbal analysis. When a user of a text-to-image AI service inputs the prompt, `illustration of a princess in an airport, holliemengert artstyle`, that user in a sense performs the very task that copyright jurists claim to be unable to do.[235] The user seems to have accomplished

231. Leon A. Gatys, Alexander S. Ecker & Matthias Bethge, *Image Style Transfer Using Convolutional Neural Networks* 2414, 2241 (2016), http://www.cv-foundation.org/openaccess/content_cvpr_2016/html/Gatys_Image_Style_Transfer_CVPR_2016_paper.html [https://perma.cc/J544-74V3].

232. Nichol et al., *supra* note 13, at 13.

233. Rombach et al., *supra* note 13, at 27.

234. *See supra* Section III.B, (discussing *Krofft* and extrinsic/intrinsic test).

235. This riffs on Baio's example. *See* Baio, *supra* note 3.

what Judge Jon Newman called "divid[ing] a visual work into neat layers of abstraction," or what the Ninth Circuit would call breaking an image down "into only five or six constituent elements."[236]

The temptation in such a case is to focus on how the discrete elements that the prompt articulates, and which the corresponding image will reflect, differ from the easiest-to-identify elements of any extant Mengert illustrations. "How could there be substantial similarity — she's never even drawn an airport!"[237] Indeed, this reasoning follows the roadmap set out in copyright's usual test for substantial similarity: extrinsic similarity may frequently be decided as a matter of law, whereas intrinsic similarity is paradigmatically a jury question.[238] Thus, insufficient extrinsic similarity can block a plaintiff from getting her case before a jury, while — at least in the Ninth Circuit — insufficient intrinsic similarity cannot defeat a case as a matter of law.[239] But focusing on the extrinsic elements that are easiest to identify analytically (the `princess`, the `airport`) ignores the innumerable aesthetic elements that `holliemengert artstyle` adds to an image, and it disregards the judicial admonition that "infringement analysis is not simply a matter of ascertaining similarity between components viewed in isolation."[240] It's surely true that invoking the elements of an artist's style doesn't *always* produce an output image that is substantially similar to copyrighted expression. But to assert that infringement is farfetched or impossible under these circumstances derogates both from the bedrock doctrinal principle that copyright law protects original combinations of unprotectable elements and also from the case-by-case, holistic similarity analysis that this same doctrine mandates.

## IV. CONCLUSION

Substantial similarity in images has long been a vexing question in copyright law. Generative AI is remarkable precisely because it can derive indicia of aesthetic similarity, abstract those indicia of similarity into efficient encodings, and transpose them to novel images. Even if we can't always name these similarities, we know them when we see

236. Newman, *supra* note 128, at 698; Swirsky v. Carey, 376 F.3d 841, 849 (9th Cir. 2004).

237. I don't know for a fact that Hollie Mengert has never drawn a princess in an airport. To facilitate this example, please assume that she hasn't.

238. *See* Shaw v. Lindheim, 919 F.2d 1353, 1358 (9th Cir. 1990), *overruled on other grounds by* Skidmore as Tr. for Randy Craig Wolfe Tr. v. Led Zeppelin, 952 F.3d 1051 (9th Cir. 2020); Copeland v. Bieber, 789 F.3d 484, 490 (4th Cir. 2015) (declining to address "whether a district court may grant an infringement defendant's motion to dismiss, or motion for summary judgment, under the intrinsic prong alone").

239. *Shaw*, 919 F.2d at 1359 ("[W]e hold that it is improper for a court to find, as the district court did, that there is no substantial similarity as a matter of law after a writer has satisfied the extrinsic test.").

240. Tufenkian Imp./Exp. Venture v. Einstein Moomjy, 338 F.3d 127, 134 (emphasis omitted).

them. When our vocabulary comes up short, we've taken to calling the similarities "style." But "style" is an unhelpful word. Its vagueness and breadth come in handy for denoting an ineffable gestalt. But those qualities of "style" make the word an undesirable vehicle for a legal conclusion that depends on a case-by-case analysis of aesthetic appeal. An axiom that style is categorically uncopyrightable only means anything if we know what style is.

If this Article has a moral, it's that technologists and commentators shouldn't be so categorical in proclaiming that image-generating AI only appropriates some uncopyrightable quality called "style." That insight, in turn, suggests a further warning: the nuanced view of artistic style that copyright doctrine requires will also complicate generative AI companies' invocation of the fair use defense. Insofar as AI companies assert fair use on the ground that their models do not generate output substantially similar to their training data — an argument that, for the record, caselaw contradicts[241] — the possibility of substantial similarity through stylistic appropriation weakens that defense.

Defining style has been a hard question since long before image-generating AI. Differentiating between unprotectable stylistics and copyrightable expression has proven particularly difficult in visual art. Unlike literary works, there's little consensus that lets us break images down into components that we treat as constitutive of the work. In images, every visual detail counts; this explains why courts have recognized *copyists'* original expression present in copies of images, when no court would find any originality in copying a book out in handwriting.[242] Because we treat images as gestalts, any detail of which can embody creative expression, and because we struggle to dissect images into their constituent parts, the vague sweep of the word "style" can encompass copyrightable expression in art objects.

The hazy status of images in copyright law also helps explain why the debate over whether generative AI operates as a "collage machine" has been so fraught and so unprofitable. Courts routinely refer to "literal" or "verbatim" copying of nonverbal visual media — "word-for-word" copying[243] — but the metaphor is an awkward byproduct of a text-centric legal regime.[244] Art objects have no literal elements

241. *See* Benjamin L.W. Sobel, *Copyright Accelerationism*, 100 CHI.-KENT L. REV. (forthcoming 2025) (manuscript at 21–23), (available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4658701 [https://perma.cc/F3F6-SVRE]).

242. *See* Alfred Bell & Co. v. Catalda Fine Arts, 191 F.2d 99, 104–05 (2d Cir. 1951).

243. *Verbatim,* ONLINE ETYMOLOGY DICTIONARY (Sept. 28, 2017), https://www.etymonline.com/word/verbatim [https://perma.cc/EK2S-5R7E].

244. *See, e.g.*, Walt Disney Prods. v. Air Pirates, 581 F.2d 751, 757 (9th Cir. 1978) (referring to "a verbatim copy of [a graphical] depiction" of a cartoon character); Northland Fam. Plan. Clinic, Inc. v. Ctr. for Bio-Ethical Reform, 868 F. Supp. 2d 962, 966 (C.D. Cal. 2012) ("Defendants made a series of videos using unaltered segments of the [plaintiff's video]

because no series of standardized symbols defines the identity of a painting as the text does for a work of literature. For this reason, it makes little sense to focus on whether generative AI "copies and pastes" images. Our focus should not be on *how* generative AI distills and reproduces salient information from images, it should be on *what* that information is, and whether it is of a sort that ought to be of consequence for copyright law.

It turns out that in this regard, the task of the copyright jurist and the task of the AI developer (or perhaps the AI model) converge. In a sense, substantial similarity analysis is an exercise in "dimensionality reduction:" it asks tribunals to abstract works, filter out non-cognizable elements, and compare the residue to determine whether one infringes another.[245] Dimensionality reduction is precisely the task that generative AI excels at: it can identify similarities that unify particular categories of images, represent those similarities in an efficient format, and generate remarkable media that interpolate those identified characteristics. But the dimensions in which generative AI encodes salient visual information aren't ones that humans can navigate or even necessarily name.

While technologists don't have to explain why their models capture visual similarities when everyone can see that they do, jurists have to make arguments. The very reason image-generating AI is so remarkable, then, is the reason it short-circuits the juridical mood: it navigates the interstitial space between and among our words to derive meaningful, higher-order patterns that may elude our vocabulary. The substantial similarity inquiry — reducing two non-verbal works of art to a verbal assemblage of similarities and dissimilarities — is an exercise in dimensionality reduction. The present challenge is to figure out whether the reduced dimensions in which generative AI has learned to operate constitute a legally cognizable form of similarity. That challenge is necessarily one that's difficult to talk about.

---

without [plaintiff's] permission. [Defendant] . . . created a 1 minute and 17 second video that uses several verbatim segments of the [plaintiff's video]."); SHL Imaging, Inc. v. Artisan House, Inc., 117 F. Supp. 2d 301, 311 (S.D.N.Y. 2000) ("While plaintiff's photographs meet the minimal originality requirements in *Feist,* they are not entitled to broad copyright protection . . . Practically, the plaintiff's works are only protected from verbatim copying.").

245. *See* Comput. Assocs. Int'l, Inc. v. Altai, Inc., 982 F.2d 693, 706 (2d Cir. 1992).