Cutting-Edge Developments in AI Litigation (April 29, 2025)

Timed Outline

1. Introduction – 15 minutes

   a. Speaker bios (**MD then JW**)

   b. Recent trends in generative AI (*e.g.*, video/world models; style generation) (**Both**)

   c. Overall status of litigations, issues decided to date (**Both**)

2. Topic: Protectability – 15 minutes

   a. *Thaler v. Perlmutter* – D.C. Cir. Decision (**MD**)

   b. Jan. 2025 Report from the U.S. Copyright Office (**JW then MD**)

   c. Feb. 2025 registration for entirely AI-generated image (**JW**)

3. Topic: Infringement – 30 minutes

   a. *Thompson v. Ross* (**MD**)

   b. *Kadrey v. Meta* (**JW**)

   c. *New York Times v. OpenAI* (**JW**)

   d. *Concord v. Anthropic* (**MD**)

   e. *Andersen v. Stability AI* (Aug. 2024) (**JW**)

4. Questions and Answers – 5 minutes (reserved)

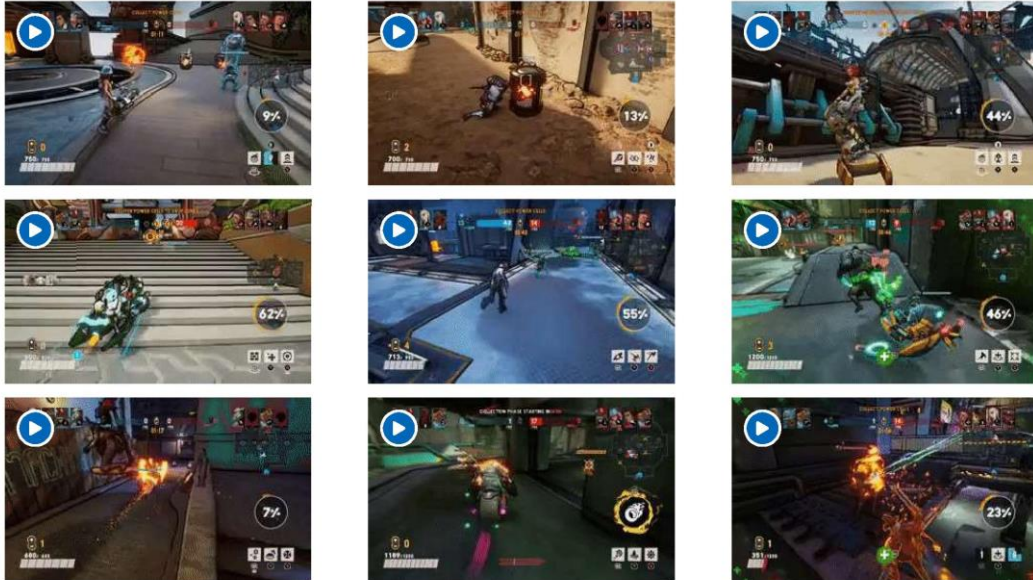# Cutting-Edge Developments in AI Litigation

NYIPLA

April 29, 2025

Introduction

Introduction

Generated gameplay examples

Introduction

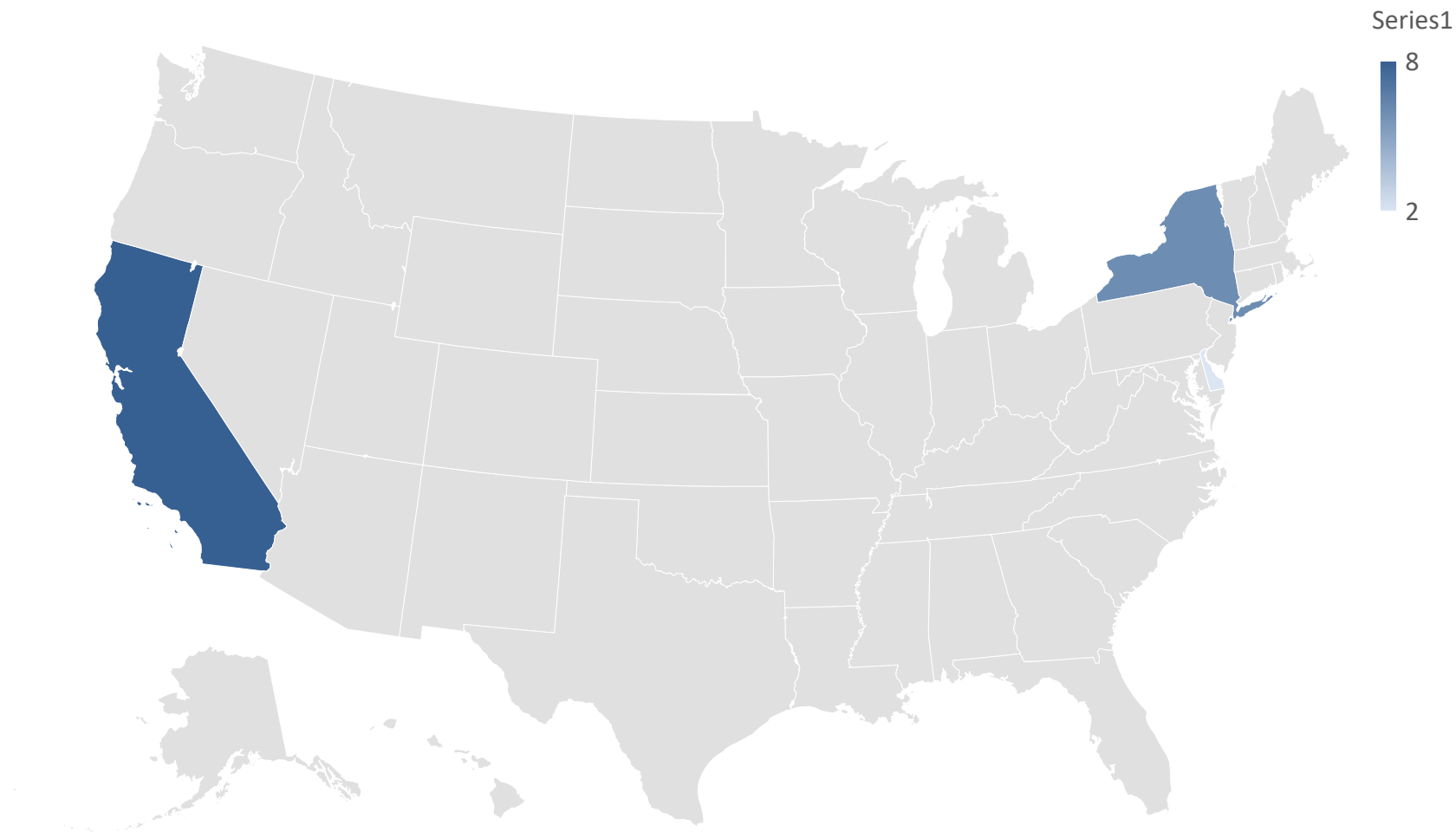# Introduction

Introduction

Infringement litigation statistics:

- Over 16 cases filed.
- Motions to dismiss filed in 12+.
- Summary judgment motions filed in ~ 2.
- 0 trials to date.

Introduction

Where the AI infringement cases are venued

Series1
8

2

PRYOR CASHMAN

Protectability

"Copyright protection subsists...in original works of **authorship**...."
    - 17 U.S.C. § 102

Copyright protection "requires human authorship."
    - Copyright Office
    - D.C. Circuit

Protectability

Thaler v. Perlmutter, No. 23-5233, --- F.4th ---- (D.C. Cir. Mar. 18, 2025)

Protectability

Protectability

The Office agrees that there is an important distinction between using AI as a tool to assist in the creation of works and using AI as a stand-in for human creativity. While assistive uses that enhance human expression do not limit copyright protection, uses where an AI system makes expressive choices require further analysis. This distinction depends on how the system is being used, not on its inherent characteristics.[65]

resulting work.[195] For example, the Office recently considered an application to register a sound recording by GRAMMY-winning country artist Randy Travis, who has limited speech function following a stroke.[196] The track was created based on the recording of a human voice, using "[a] special-purpose AI vocal model . . . as a tool . . . to help realize the sounds that Mr. Travis and the other members of the human creative team desired."[197] The result, which would have been infeasible without this technology, was a new track appearing to be sung in Travis's legendary voice. Because the sound recording used AI as a tool, not to generate expression, the Office registered the work.

Protectability

The Office concludes that, given current generally available technology, prompts alone do not provide sufficient human control to make users of an AI system the authors of the output. Prompts essentially function as instructions that convey unprotectible ideas. While highly detailed prompts could contain the user's desired expressive elements, at present they do not control how the AI system processes them in generating the output.

Protectability

The following image, which the Office generated by entering a prompt into a popular commercially available AI system, illustrates this point:[106]

Prompt	*Output*
professional photo, bespectacled cat in a robe reading the Sunday newspaper and smoking a pipe, foggy, wet, stormy, 70mm, cinematic, highly detailed wood, cinematic lighting, intricate, sharp focus, medium shot, (centered image composition), (professionally color graded), ((bright soft diffused light)), volumetric fog, hdr 4k, 8k, realistic	

Protectability

Prompt	Input	Output
"a young cyborg woman (((roses))) flowers coming out of her head, photorealism, cinematic lighting, hyper realism, 8k, hyper detailed."		

The applicant disclaimed "any non-human expression" appearing in the final work, such as the realistic, three-dimensional representation of the nose, lips, and rosebuds, as well as the lighting and shadows in the background. After reviewing the information provided in the application, the Office registered the work with an annotation stating: "Registration limited to unaltered human pictorial authorship that is clearly perceptible in the deposit and separable from the non-human expression that is excluded from the claim."[124]

Protectability

Commenters also identified situations where creators have begun to experiment with using AI as a brainstorming tool. The Recording Academy, for instance, stated that "[m]any Academy members already use generative AI as a tool to assist them in creating new music," including through song ideation.[66] Another stakeholder noted that AI can be used to structure or create a preliminary outline for literary works.[67] In these cases, the user appears to be prompting a generative AI system and referencing, but not incorporating, the output in the development of her own work of authorship. Using AI in this way should not affect the copyrightability of the resulting human-authored work.[68]

Protectability

"A Single Piece of American Cheese"

Infringement

Thompson Reuters Enterprise Centre GMBH v. Ross Intelligence Inc. (D. Del. 2025)

Infringement

Kadrey v. Meta Platforms, Inc., 23-cv-03417 (N.D. Cal.)

Infringement

The New York Times Company et al. v. Microsoft Corporation et al.,
23-cv-11195, 24-cv-3285, 24-cv-4872 (S.D.N.Y.)

Infringement

Concord Music Group, Inc. et al. v. Anthropic PBC, 24-cv-03811 (N.D. Cal.)

Infringement

Andersen v. Stability AI LTD., 23-cv-00201 (N.D. Cal.)

Any Questions?

Cutting-Edge Developments in AI Litigation

CLE Presentation April 29, 2025

For the New York Intellectual Property Law Association

By Mark Doerr and Joshua Weigensberg

Course Materials Index

UNITED STATES DISTRICT COURT

NORTHERN DISTRICT OF CALIFORNIA

SARAH ANDERSEN, et al., Plaintiffs, v. STABILITY AI LTD., et al., Defendants.	Case No. 23-cv-00201-WHO **ORDER GRANTING IN PART AND DENYING IN PART MOTIONS TO DISMISS FIRST AMENDED COMPLAINT** Dkt. Nos. 160, 162, 163, 164, 169

Defendants Stability AI Ltd. and Stability AI, Inc. (collectively "Stability AI"),

Midjourney, Inc., DeviantArt, Inc., and Runway AI, Inc. move to dismiss various claims from

plaintiffs' First Amended Complaint ("FAC," Dkt. No. 129). For the reasons discussed below,

those motions are GRANTED in part and DENIED in part.

BACKGROUND

Artists Sarah Andersen, Kelly McKernan, Karla Ortiz, Hawke Southworth, Grzegorz

Rutkowski, Gregory Manchess, Gerald Brom, Jingna Zhang, Julia Kaye, and Adam Ellis

("Plaintiffs"[1]) filed this putative class action on behalf of artists challenging the defendants'

creation and/or use of Stable Diffusion, an artificial intelligence ("AI") software product. They

allege that Stable Diffusion used plaintiffs' artistic works as "training images" and as a result

Stable Diffusion can produce output images "in the style" of those images. *See generally* FAC. In

the FAC, plaintiffs allege claims against the three defendants identified in the original complaint

(Stability AI, Midjourney, and DeviantArt) and against a new defendant, Runway AI.[2] Plaintiffs

---

[1] In their FAC, plaintiffs reasserted claims for the three original plaintiffs (Sarah Anderson, Kelly McKernan, and Karla Ortiz) and added – without the court permission – seven additional plaintiffs (H. Southworth PKA Hawke Southworth, Grzegorz Rutkowski, Gregory Manchess, Gerald Brom, Jingna Zhang, Julia Kaye, and Adam Ellis). FAC ¶¶ 14-23.

[2] The basic factual allegations regarding how defendants' AI products were trained and work were

1   contend that Runway AI worked with, helped train, and then distributed Stable Diffusion with

2   Stability AI. Plaintiffs assert that Runway made a text-to-image generator available via its online

3   AI image product called AI Magic Tools.

4   Plaintiffs' claims center first around the creation of the LAION training sets, where five

5   billion images were allegedly scraped into datasets used by Stability and Runway to train the

6   versions of Stable Diffusion. FAC ¶ 4. Plaintiffs state that Midjourney likewise trained its

7   product using Stable Diffusion, and that all four defendants use Stable Diffusion in their AI

8   products; in doing so, those four defendants copy or utilize versions of plaintiffs' artistic works.

9   *Id*. ¶¶ 5-6. Significantly, plaintiffs allege that the "LAION-5B dataset contains only URLs of

10   training images, not the actual training images. Therefore, anyone who wishes to use LAION-5B

11   for training their own machine learning model must first acquire copies of the actual training

12   images from their URLs using the img2dataset or other similar tool." FAC ¶ 221. They also

13   clarify their theory of direct infringement, adding allegations regarding CLIP-guided diffusion in

14   the training phase but also in use, after training. *Id*. ¶¶ 82-150.

15   Plaintiffs assert claims on behalf of six different classes:

16   "<u>Injunctive Relief Class</u>" under Rule 23(b)(2): All persons or entities nationalized or

17   domiciled in the United States that own a copyright interest in any work that was used to train any

18   version of an AI image product that was offered directly or incorporated into another product by

19   one or more Defendants during the Class Period. FAC ¶ 34.

20   "<u>Damages Class</u>" under Rule 23(b)(3): All persons or entities nationalized or domiciled in

21   the United States that own a copyright interest in any work that was used to train any version of an

22   AI image product that was offered directly or incorporated into another product by one or more

23   Defendants during the Class Period. *Id*.

24   "<u>LAION-5B Damages Subclass</u>" under Rule 23(b)(3): All persons or entities nationalized

25   or domiciled in the United States that own a registered copyright in any work in the LAION-5B

26

27   _____

identified in my prior Order and will not be repeated here. *See* October 30, 2023, Order at Dkt.
28   No. 117. To the extent new, material allegations have been added to the FAC, they will be
     addressed as part of the substantive analysis below.

United States District Court
Northern District of California

dataset that was used to train any version of an AI image product that was offered directly or

incorporated into another product by one or more Defendants during the Class Period.[3] *Id.*

"LAION-400M Damages Subclass" under Rule 23(b)(3): All persons or entities

nationalized or domiciled in the United States that own a registered copyright in any work in the

LAION-400M dataset that was used to train any version of an AI image product that was offered

directly or incorporated into another product by one or more Defendants during the Class Period.

Id.

"DeviantArt Damages Subclass" under Rule 23(b)(3): All members of the Damages Class

who (1) maintained an account on DeviantArt; (2) posted copyrighted work on DeviantArt; and

(3) had that work used to train any version of an AI image product. *Id.*

"Midjourney Named Artist Class" under Rule 23(b)(3): All persons or entities who appear

on the Midjourney Names List and whose names were invoked within prompts of the Midjourney

Image Product during the Class Period.[4] *Id.*

Plaintiffs assert the following claims against the different sets of defendants:

Against Stability AI: (1) direct copyright infringement of the LAION-5B Registered

Works by training the Stability Models, including Stable Diffusion 2.0 and Stable Diffusion XL

1.0 on behalf of the LAION-5B Registered Plaintiffs and Damages Subclass; (2) inducement of

copyright infringement by distributing Stable Diffusion 2.0 and Stable Diffusion XL 1.0 for free

on behalf of the LAION-5B Registered Plaintiffs and Damages Subclass; (3) violations of the

Digital Millennium Copyright Act ("DMCA") by removing and altering copyright management

information ("CMI") of training images on behalf of all Plaintiffs, the Damages and the Injunctive

Classes; and (4) unjust enrichment under Cal. Bus. & Prof. Code § 17200 and California Common

United States District Court
Northern District of California

---

[3] The "LAION-5B Registered Plaintiffs" as used in the FAC include "the subset of plaintiffs who hold copyrights in these LAION-5B Registered Works that were registered before the filing of the initial complaint in this action, namely Sarah Andersen, Jingna Zhang, Gerald Brom, Gregory Manchess, Julia Kaye, and Adam Ellis." FAC ¶ 213.

[4] The "Midjourney Named Plaintiffs" are the plaintiffs whose names were disclosed by Midjourney as artists whose works were included and could be recreated through use of Midjourney's product; Grzegorz Rutkowski, Sarah Andersen, Karla Ortiz, Gerald Brom, and Julia Kaye. FAC ¶¶ 263, 264.

1   Law on behalf of all Plaintiffs, the Damages and the Injunctive Classes.

2   Against Runway AI: (1) Direct copyright infringement of the LAION-5B Registered

3   Works by training the Runway Models, including Stable Diffusion 1.5 on behalf of the LAION-

4   5B Registered Plaintiffs, LAION-5B Subclass, and Karla Ortiz Individually; (2) Inducement of

5   copyright infringement by distributing Stable Diffusion 1.5 for free on behalf of the LAION-5B

6   Registered Plaintiffs and Subclass; (3) DMCA violations by removing and altering CMI of

7   training images on behalf of all Plaintiffs, the Damages and Injunctive Classes; and (4) Unjust

8   enrichment under Cal. Bus. & Prof. Code § 17200 and California Common Law on behalf of all

9   Plaintiffs, the Damages and Injunctive Classes.

10   Against Midjourney: (1) Direct copyright infringement of the LAION-400M Registered

11   Works by training the Midjourney 400M Models, including Midjourney Model version 1 on

12   behalf of the LAION-400M Registered Plaintiffs and Damages Subclass; (2) Direct copyright

13   infringement of the LAION-5B Registered Works by training the Midjourney 5B Models,

14   including Midjourney Model version 5.2 on behalf of the LAION-5B Registered Plaintiffs and

15   Damages Subclass; (3) DMCA violations by removing and altering CMI of training images

16   on behalf of All Plaintiffs, the Damages and Injunctive Class; (4) Lanham Act — false

17   endorsement by unauthorized commercial use of artists' names on behalf of the Midjourney

18   Named Plaintiffs and Class; (5) Lanham Act — vicarious trade-dress violation by profiting from

19   imitations of protectable trade dress on behalf of the Midjourney Named Plaintiffs and Class; and

20   (6) Unjust enrichment under Cal. Bus. & Prof. Code § 17200 and California Common Law on

21   behalf of all Plaintiffs, and the Damages and Injunctive Class.

22   Against DeviantArt: (1) Direct copyright infringement by copying the DreamUp–CompVis

23   Model and incorporating it into DreamUp on behalf of the LAION-5B Registered Plaintiffs; (2)

24   Breach of contract for violation of its Terms of Service on behalf of the DeviantArt Plaintiffs; (3)

25   Unjust enrichment under Cal. Bus. & Prof. Code § 17200 and California Common Law on behalf

26   of the DeviantArt Plaintiffs.

27   In October 2023, I largely granted the motions to dismiss brought by defendants Stability,

28   Midjourney and DeviantArt. The only claim that survived was the direct infringement claim

4

asserted against Stability, based on Stability's alleged "creation and use of 'Training Images' scraped from the internet into the LAION datasets and then used to train Stable Diffusion." October 2023 Order at 7. The remainder of the claims were dismissed with leave, so that plaintiffs could amend "to provide clarity regarding their theories of how each defendant separately violated their copyrights, removed or altered their copyright management information, or violated their rights of publicity and plausible facts in support." *Id*. at 1.

As one example, I required plaintiffs on amendment to address the following deficiencies with a second theory of direct infringement, separate from the creating and use of images for training theory:

> Plaintiffs will be required to amend to clarify their theory with respect to compressed copies of Training Images and to state facts in support of how Stable Diffusion – a program that is open source, at least in part – operates with respect to the Training Images. If plaintiffs contend Stable Diffusion contains "compressed copies" of the Training Images, they need to define "compressed copies" and explain plausible facts in support. And if plaintiffs' compressed copies theory is based on a contention that Stable Diffusion contains mathematical or statistical methods that can be carried out through algorithms or instructions in order to reconstruct the Training Images in whole or in part to create the new Output Images, they need to clarify that and provide plausible facts in support.

October 2023 Order at 8-9; *see also id*. 9 n.7 ("Plaintiffs' second theory of direct infringement – that Stable Diffusion is a 'derivative work' because it contains compressed copies of billions of copyrighted images and by incorporating Stable Diffusion into" defendants' own AI products, defendants are "liable for producing works that have been 'transformed' based on plaintiffs' works [] fails for the same reasons.").

Plaintiffs added defendant Runway AI and seven new plaintiffs when it filed the FAC. Each defendant moves to dismiss.

LEGAL STANDARD

Under FRCP 12(b)(6), a district court must dismiss a complaint if it fails to state a claim upon which relief can be granted. To survive a Rule 12(b)(6) motion to dismiss, the plaintiff must allege "enough facts to state a claim to relief that is plausible on its face." *Bell Atl. Corp. v. Twombly*, 550 U.S. 544, 570 (2007). A claim is facially plausible when the plaintiff pleads facts

1   that "allow the court to draw the reasonable inference that the defendant is liable for the

2   misconduct alleged." *Ashcroft v. Iqbal*, 556 U.S. 662, 678 (2009) (citation omitted). There must

3   be "more than a sheer possibility that a defendant has acted unlawfully." *Id*. While courts do not

4   require "heightened fact pleading of specifics," a plaintiff must allege facts sufficient to "raise a

5   right to relief above the speculative level." *Twombly*, 550 U.S. at 555, 570.

6   In deciding whether the plaintiff has stated a claim upon which relief can be granted, the

7   Court accepts the plaintiff's allegations as true and draws all reasonable inferences in favor of the

8   plaintiff. *See Usher v. City of Los Angeles*, 828 F.2d 556, 561 (9th Cir. 1987). However,

9   the court is not required to accept as true "allegations that are merely conclusory, unwarranted

10   deductions of fact, or unreasonable inferences." *In re Gilead Scis. Sec. Litig*., 536 F.3d 1049,

11   1055 (9th Cir. 2008). If the court dismisses the complaint, it "should grant leave to amend even if

12   no request to amend the pleading was made, unless it determines that the pleading could not

13   possibly be cured by the allegation of other facts." *Lopez v. Smith*, 203 F.3d 1122, 1127 (9th Cir.

14   2000). In making this determination, the court should consider factors such as "the presence or

15   absence of undue delay, bad faith, dilatory motive, repeated failure to cure deficiencies by

16   previous amendments, undue prejudice to the opposing party and futility of the proposed

17   amendment." *Moore v. Kayport Package Express*, 885 F.2d 531, 538 (9th Cir. 1989).

18   **DISCUSSION**

19   **I.   ADDITION OF NEW PLAINTIFFS & CLAIMS**

20   Stability objects to the addition of seven new plaintiffs to the FAC, as well as the addition

21   of two new claims against it (induced infringement, Count 2) and unjust enrichment (Count 4).

22   Stability MTD, Dkt. No. 162 at 13-14. DeviantArt similarly objects to the addition of new

23   plaintiffs and the addition of the unjust enrichment claim. DeviantArt MTD, Dkt. No. 163, at 23-

24   24. Defendants argue that the addition of the new plaintiffs and claims exceeds the scope of the

25   leave to amend I allowed plaintiffs in the October 2023 Order and, therefore, the new plaintiffs

26   and claims should be stricken.

27   Plaintiffs respond that they were given broad "leave to amend and attempt to cure the

28   deficiencies identified" in their claims and the ability of the then three named plaintiffs to pursue

United States District Court
Northern District of California

1   the claims asserted. Because leave was not cabined to "solely" or exclusively correct the

2   deficiencies, plaintiffs argue that the additions were permissible and in any event should be

3   allowed under Rule 15's liberal amendment standard. *See* Oppo. to Stability, Dkt. No. 174, at 22;

4   Oppo. to DeviantArt, Dkt. No. 177, at 21-22.

5   Plaintiffs are correct that leave to amend under Rule 15 is "freely given," especially at the

6   start of a case. However, once a complaint is dismissed and a court grants plaintiffs specific leave

7   to amend to address identified legal or factual deficiencies, adding plaintiffs or claims beyond

8   those previously alleged requires either requesting and securing leave, which plaintiffs did not do,

9   or the consent of defendants, which plaintiffs did not seek. *See* Fed. R. Civ. P, 15(a)(2) ("In all

10   other cases, a party may amend its pleading only with the opposing party's written consent or the

11   court's leave.").[5]

12   That said, I would have granted leave to amend if plaintiffs had sought leave, given the

13   lack of prejudice to defendants. At this juncture, where each defendant has addressed the newly

14   added claims and plaintiffs, it would elevate form over substance to grant the motions to dismiss

15   on this ground. And I will assume that plaintiffs have implicitly sought leave to amend to include

16   the new plaintiffs and claims in a Second Amended Complaint. I will grant leave and address the

17   arguments defendants make against the added claims and plaintiffs.[6]

18   **II.   STABILITY AI MOTION TO DISMISS**

19   **A.   Induced Copyright Infringement**

20   Stability challenges plaintiffs' first theory of induced copyright infringement interpreting

21   the claim as one alleging that the Stable Diffusion models themselves are infringing works. Under

22   this theory, Stability is inducing infringement by distributing the models when any third-party

23

24   [5] The only case relied on by plaintiffs is totally inapposite. It addresses a situation where plaintiffs
     were given leave to amend to file a second amended complaint and revised motion for class
25   certification by stipulation, and after the court consolidated cases, defendants moved to dismiss the
     new plaintiffs based on statute of limitations grounds. *See Bos. Ret. Sys. v. Uber Techs., Inc.*, No.
26   19-CV-06361-RS, 2021 WL 4503137, at *2 (N.D. Cal. Oct. 1, 2021).

27   [6] As explained below, the end result is that plaintiffs may file a Second Amended Complaint
     including the new plaintiffs and may attempt to plead unjust enrichment claims against any
28   defendant based on theories (if any) that are not preempted by the Copyright Act.

1   downloads, uses, or deploys the models provided by Stability. *See* FAC, Count Two (alleging

2   contributory infringement by "distributing Stable Diffusion 2.0 and Stable Diffusion XL 1.0 for

3   free" on "behalf of the LAION-5B Registered Plaintiffs and Damages Subclass") ¶¶ 233-236.[7]

4   Stability argues, first, that this theory is simply a repackaged direct infringement theory,

5   that by "distributing" Stable Diffusion, Stability violates plaintiffs' exclusive rights of distribution

6   of their works.[8] But whether this is a direct infringement claim (where liability is imposed against

7   Stability for distributing copyrighted works) or more properly characterized as an inducement

8   claim (where liability is imposed because Stability induces or otherwise causes others to copy

9   protectible material) depends on how Stable Diffusion works and is implemented by users other

10   than Stability itself. Any potential overlap – or potential requirement for plaintiffs to elect one

11   claim or another – is better addressed on summary judgment, after discovery.

12   Stability also argues that the inducement claim must be dismissed to the extent plaintiffs

13   are alleging that Stability encourages the use of Stable Diffusion to create infringing outputs. That

14   theory is barred, according to Stability, because plaintiffs fail to allege any facts that Stability

15   promoted the use of Stable Diffusion to "infringe copyright, as shown by clear expression or other

16   affirmative steps taken to foster infringement." *Metro-Goldwyn-Mayer Studios Inc. v. Grokster,*

17   *Ltd.*, 545 U.S. 913, 919 (2005). Stability argues that given the lack of allegations supporting a

18   specific intent by Stability to promote infringement, the claim fails.

19   However, plaintiffs point to one statement by Stability's CEO that Stability took 100,000

20   gigabytes of images and compressed it to a two-gigabyte file that can "recreate" any of those

21   images.[9] Stability responds that the "isolated" use of the word "create" by its CEO cannot

---

[7] Stability does not challenge plaintiffs' claim of direct copyright infringement, as that survived the prior motion to dismiss. October 2023 Order at 7.

[8] Stability also contends this theory is another take on the "derivative" works claim I dismissed in the October 2023 Order. However, that claim was dismissed in large part because of the ambiguity and lack of plausible facts that Stable Diffusion was itself a "derivative work" because it contained compressed copies of billions of copyrighted images. *See* October 2023 Order at 10 n.7. Plaintiffs were given leave to amend their compressed copies theory in support of direct copyright infringement with respect to the fully trained Stable Diffusion product, and as discussed with respect to the motions to dismiss below, have plausibly done so.

[9] *See* FAC ¶ 4 ("Emad Mostaque described it thus: 'Stable Diffusion is the model itself. It's a

1 demonstrate that it intended to foster infringement. Stability argues that "clear allegations of

2 active steps to encourage direct infringement" are especially important here, where plaintiffs do

3 not dispute that Stable Diffusion is capable of substantial noninfringing uses, like creating art from

4 inputs that do not reference particular artists or invoke particular artists' styles or have any

5 substantial similarity to plaintiffs' works.

6 The theory of this case is not similar to – for example – a case asserting contributory

7 infringement based on the sale of VCRs where, after discovery, plaintiff had no evidence of

8 defendant's intent to induce infringement. The Supreme Court explained that, in those

9 circumstances, intent could not be "based on presuming or imputing intent to cause infringement

10 solely from the design or distribution of a product capable of substantial lawful use, which the

11 distributor knows is in fact used for infringement." *See Grokster, Ltd*., 545 U.S. at 933

12 (discussing holding of *Sony Corp. of America v. Universal City Studios, Inc*., 464 U.S. 417

13 (1984)). Instead, this is a case where plaintiffs allege that Stable Diffusion is built to a significant

14 extent on copyrighted works and that the way the product operates necessarily invokes copies or

15 protected elements of those works. The plausible inferences at this juncture are that Stable

16 Diffusion *by operation* by end users creates copyright infringement and was created to facilitate

17 that infringement by design. In addition to the comment of Stability's CEO, plaintiffs reference

18 articles by academics and others that training images can sometimes be reproduced as outputs

19 from the AI products. FAC ¶¶ 90, 130-139.

20 Whether true and whether the result of a glitch (as Stability contends) or by design

21 (plaintiffs' contention) will be tested at a later date. The allegations of induced infringement are

22 sufficient.

23 **B. DMCA**

24 Stability also moves again to dismiss plaintiffs' DMCA claims asserted under 17 U.S.C.

25 section 1202(a) for providing or distributing false copyright management information ("CMI")

26

27

28 collaboration that we did with a whole bunch of people … We took 100,000 gigabytes of images and compressed it to a two-gigabyte file that can recreate any of those [images] and iterations of those.'").

9

and under section 1202(b)(1) for intentional removal of CMI.[10] I dismissed plaintiffs' DMCA

claim with leave to amend in the October 2023 Order, explaining:

> [E]ach plaintiff must identify the exact type of CMI included in their online works that were online and that they have a good faith belief were scraped into the LAION datasets or other datasets used to train Stable Diffusion. At the hearing, plaintiffs argued that it is key for the development of generative AI models to capture not only images but any accompanying text because that accompanying text is necessary to the models' ability to "train" on key words associated with those images. Tr. at 9:13-24. But there is nothing in the Complaint about text CMI present in the images the named plaintiffs included with their online images that they contend was stripped or altered in violation of the DMCA during the training of Stable Diffusion or the use of the end-products. Plaintiffs must, on amendment, identify the particular types of their CMI from their works that they believe were removed or altered.

> In addition, plaintiffs must clarify and then allege plausible facts regarding which defendants they contend did the stripping or altering in violation of the DMCA and when that occurred.

October 2023 Order at 18.

1. Claim Under 1202(a)

Stability moves to dismiss the subsection (a) claim regarding false CMI. Plaintiffs allege

that "Stability distributes the Stability Models under the MIT License (see, e.g. —

https://github.com/Stability-AI/stablediffusion/blob/main/LICENSE). Within this license, Stability

asserts copyright in the Stability Models. By asserting copyright in the Stability Models, which

infringe the copyrights of the LAION-5B Plaintiffs, Stability is providing and distributing false

CMI in violation of 17 U.S.C. § 1202(a)." FAC ¶ 248.

Stability argues that this claim fails because Stability's generic license does not suggest

any association at all with plaintiffs' works, and therefore was not made "in connection with"

plaintiffs' works which is necessary to support a claim under 1202(a). *See Logan v. Meta*

Platforms, Inc., 636 F. Supp. 3d 1052, 1062 (N.D. Cal. 2022) (plaintiff failed to allege facts

showing defendant "conveyed CMI in connection" with plaintiff's photos where "Meta's allegedly

---

[10] Section 1202(a) prohibits the knowing, with the intent to induce or enable infringement, provision or distribution of false CMI. Section 1202(b)(1) governs unpermitted removal or alteration of CMI and distribution of works with removed or altered CMI. *See* 17 U.S.C. § 1202 *et seq*.

1   false CMI is '[a] [generic] copyright tag on the bottom of each Facebook user page,' separated

2   from the rest of the content on the webpage, and is not located on or next to" plaintiff's photos);

3   *but see Post Univ. v. Course Hero, Inc*., No. 3:21-CV-1242 (JBA), 2023 WL 5507845, at *4 (D.

4   Conn. Aug. 25, 2023) ("Thus, at the motion to dismiss stage, unless it is 'implausible' that a

5   viewer could understand the information to be referring to the defendant as the work's copyright

6   holder, dismissal is inappropriate."). In addition, Stability contends that plaintiffs fail to allege

7   facts plausibly satisfying the "double scienter" required under (a); that Stability knowingly

8   provided false CMI with the intent to induce or enable infringement.

9   I agree on both counts. The generic license that accompanies use of Stable Diffusion on its

10   face claims rights to Stable Diffusion as a work, not to the LAION dataset and not any works that

11   were used to create the LAION dataset. It is implausible that a viewer reading the license

12   disclosure for Stable Diffusion would understand that Stability is claiming rights to or conveying

13   any false information regarding the rights of the plaintiffs whose copyrighted works are among the

14   billions of images in the LAION datasets.

15   The 1202(a) claim is dismissed again, this time with prejudice as plaintiffs do not identify

16   a basis for this claim.

17   **2.**   **Claim Under 1202(b)(1)**

18   Similarly, Stability contends that plaintiffs have failed to plead plausible facts that Stability

19   AI intentionally removed or altered CMI from plaintiffs' works during the training process for

20   Stable Diffusion, and failed to allege the double-scienter requirement that Stability did so in order

21   to facilitate infringement. In the FAC, plaintiffs allege that "Stability directly copied the LAION-

22   5B Works and used these Statutory Copies as training data for the Stability Models. The works

23   copied by Stability included CMI, including in the form of distinctive marks such as watermarks

24   or signatures, and as the captions in the image-text pairs. The training process is designed to

25   remove or alter CMI from the training images. Therefore, Stability intentionally removed or

26   altered CMI from the Plaintiffs' works in violation of 17 U.S.C. § 1202(b)(1)." FAC ¶ 245.

27   Plaintiffs support that allegation by pointing to some of plaintiffs' images that were used as

28   Training Images in LAION-5B and that contained CMI, comparing them to images that were

11

1 created, for example, when plaintiffs' names were used as Midjourney image prompts. FAC ¶¶

2 189-200. Plaintiffs have also, as directed in the October 2023 Order, identified the CMI present

3 on their works that they contend has been stripped by Stability. *Id*. ¶¶ 239-241. Finally, plaintiffs

4 allege that Stability engaged in knowing removal of CMI, as supported by their allegations

5 regarding how the diffusion process works, how training images are used, and based on plausible

6 allegations regarding Stability AI's prominent role in the funding of LAION. *Id*. ¶¶ 245-247.

7   Stability, however, raises a new argument on this round of motions, based on a recent

8 opinion from the Hon. Jon S. Tigar. *Doe 1 v. GitHub, Inc*., No. 22-CV-06823-JST, 2024 WL

9 235217, at *8 (N.D. Cal. Jan. 22, 2024). Stability contends that because the output images are

10 admittedly not *identical* to the Training Images, there can be no liability for any removal of CMI

11 that occurred during the training process. That is because failing to affix CMI to a "different

12 work" is not "removal" under Section 1202. Judge Tigar wrote:

> Defendants now ask the Court to address an unresolved argument
> from the prior briefing—namely, that "[Section] 1202(b) claims lie
> only when CMI is removed or altered from an identical copy of a
> copyrighted work." ECF No. 107-3 at 20 (emphasis added); see ECF
> No. 109-3 at 23–24. Defendants argue that because Plaintiffs' new
> allegations state that output from Copilot is often a modification of
> their licensed works, as opposed to an "identical copy," they have
> effectively pleaded themselves out of their Section 1202(b)(1) and
> 1202(b)(3) claims. ECF No. 109-3 at 23. Agreeing with Defendants
> on both fronts, the Court finds that it is not precluded from analyzing
> this claim anew and that Section 1202(b) claims require that copies
> be "identical."

Id. * 8; *see also Doe 1 v. GitHub, Inc.*, No. 22-CV-06823-JST, 2024 WL 1643691, at *2 (N.D.

Cal. Apr. 15, 2024) (rejecting motion for reconsideration because there was no allegation that any

output was identical to any plaintiff's work in its entirety, so there could be no "removal" of CMI

for purposes of DMCA).[11]

       Judge Tigar's analysis disagrees with a case from the Southern District of Texas, where the

---

[11] In so holding, Judge Tigar followed *Advanta-STAR Auto. Rsch. Corp. of Am. v. Search Optics, LLC*, No. 22-CV-1186 TWR, 2023 WL 3366534, at *12 (S.D. Cal. May 9, 2023); *Kirk Kara Corp. v. W. Stone & Metal Corp., No*. CV 20-1931-DMG, 2020 WL 5991503, at *6 (C.D. Cal. Aug. 14, 2020); *Frost-Tsuji Architects v. Highway Inn, Inc.*, No. CIV. 13-00496 SOM, 2015 WL 263556, at *3 (D. Haw. Jan. 21, 2015), *aff'd*, 700 F. App'x 674 (9th Cir. 2017) (no section 1202(b) violation where the allegedly infringing drawing was "not identical").

1   court rejected the "identicality" requirement. *See ADR Int'l Ltd. v. Inst. for Supply Mgmt. Inc.*,

2   667 F. Supp. 3d 411, 427 (S.D. Tex. 2023) ("[b]ased on the plain wording of the statute, the Court

3   is not persuaded that the DMCA includes an 'identical copy' requirement," noting the copying

4   should be "substantially similar" but does not have to be "perfect").

5   Recognizing that this issue is unsettled, I agree with the reasoning of Judge Tigar that

6   followed other district court decisions within the Ninth Circuit. Because there are no allegations

7   that any output from Stable Diffusion was identical to a plaintiff's work, the DMCA section

8   1202(b) claim fails as well.[12]

9   The DMCA claims against Stability are DISMISSED WITH PREJUDICE.[13]

10   **C.   Unjust Enrichment**

11   In addition to arguing that leave to amend was not granted to assert this claim, addressed

12   above, Stability argues that the unjust enrichment claim cannot proceed because it is preempted by

13   the Copyright Act. It also contends that plaintiffs fail to allege that they have an inadequate

14   remedy at law as required by *Sonner v. Premier Nutrition Corp.*, 971 F.3d 834 (9th Cir. 2020).

15   To support this claim, plaintiffs allege that Stability AI "has unjustly misappropriated the

16   LAION-5B Works in order to develop, train and promote the Stability Models, enabling it to

17   receive profit and other benefits." FAC ¶252. "By using Plaintiffs' works to train, develop and

18   promote the Stability Models, Plaintiffs and the Class were deprived of the benefit of the value of

19   their works, including monetary damages." *Id*. ¶ 254. Stability points out that this claim is

20   expressly based on the use of plaintiffs' copyrighted works without consent, and as a result is

21   covered by and preempted by the Copyright Act.

22

---

23   [12] The only allegations regarding identical outputs in the FAC are those regarding researcher's
24   ability to reproduce identical images to training images – that as noted above supports the
     plausibility of the copyright infringement claims – but there are no allegations that Stable
25   Diffusion can produce a work identical to one of the plaintiffs here, sufficient to show actionable
     removal of CMI under section 1202(b).

26   [13] Runway and Midjourney also move to dismiss the DMCA claim asserted against them. The
27   arguments raised by those defendants, and plaintiffs' responses, are based on the defendants' use
     of Stable Diffusion and materially identical to the arguments addressed above. As a result, the
28   DMCA claims asserted against Runway and Midjourney are likewise DISMISSED WITH
     PREJUDICE and will not be addressed further.

13

1   "Section 301 of the Act seeks 'to preempt and abolish any rights under the common law or

2   statutes of a State that are equivalent to copyright and that extend to works,' so long as the rights

3   fall 'within the scope of the Federal copyright law.' [] 'We have adopted a two-part test,' in

4   accordance with section 301, 'to determine whether a state law claim is preempted by the Act.

5   Laws,' [] First, we decide 'whether the 'subject matter' of the state law claim falls within the

6   subject matter of copyright as described in 17 U.S.C. §§ 102 and 103.' [] Second, assuming it

7   does, we determine 'whether the rights asserted under state law are equivalent to the rights

8   contained in 17 U.S.C. § 106, which articulates the exclusive rights of copyright holders.'"

9   *Maloney v. T3Media, Inc.*, 853 F.3d 1004, 1010 (9th Cir. 2017) (quoting *Laws v. Sony Music*

10   *Entm't, Inc.*, 448 F.3d 1134, 1137-38 (9th Cir. 2006)).

11   Plaintiffs correctly note that to "survive preemption, the state cause of action must protect

12   rights which are qualitatively different from the copyright rights. The state claim must have an

13   extra element which changes the nature of the action." *Laws*, 448 F.3d at 1143. Plaintiffs do not

14   identify an "extra element" required under California's Unfair Competition Law ("UCL") or

15   common law (which the unjust enrichment claims are based on) that changes the nature of those

16   state law claims to protect something other than rights protected under Copyright Act. Instead,

17   they assert that the heart of their unjust enrichment claim is Stability being unjustly enriched by its

18   "image product's ability to mimic Plaintiffs' artistic style and benefit from their notoriety and

19   reputation as sought-after artists." Oppo. to Stability AI at 20. More specifically, they allege that

20   the Stable Diffusion models use "CLIP-guided diffusion" that relies on prompts including artists'

21   names to generate an image. Therefore, the "crux" of plaintiffs' unjust enrichment claim revolved

22   not around plaintiffs' works but around plaintiffs themselves and their "artistic personas." *Id*.

23   The problem with plaintiffs' theory is that it is not in the FAC. The unjust enrichment

24   claim against Stability (and the other defendants) is tied instead to use of plaintiffs' works. *See*

25   FAC ¶ 42 ("Whether the use of Plaintiffs and Class members' works to train, develop, and

26   promote Defendants AI Image Products constitute an unjust benefit conferred upon Defendants to

27   Plaintiffs' detriment"); *see also id*. ¶¶ 252-255 (alleging unjust misappropriation of works).

28   As alleged, the unjust enrichment claim against Stability added to the FAC without leave

14

1   of court is preempted by the Copyright Act. It is DISMISSED. If plaintiffs have a good faith

2   theory of unjust enrichment that falls outside the scope of the protections provided by the

3   Copyright Act, they are given leave to make one last attempt to state an unjust enrichment claim.[14]

4   Stability's motion to dismiss is DENIED concerning the challenged copyright claims, but

5   GRANTED for the DMCA claims without leave to amend and GRANTED for the unjust

6   enrichment claim with leave to amend.

7   **III.   RUNWAY AI MOTION TO DISMISS**

8   Newly added defendant Runway AI is primarily alleged to have trained or assisted in the

9   training at least Stable Diffusion 1.5, using Training Images from the LAION dataset. *See, e.g.*,

10   FAC ¶¶ 4, 163, 176, 342-345. In addition to challenging plaintiffs' DMCA and unjust

11   enrichment/UCL claims – which fail for the same reasons identified above with respect to

12   Stability – Runway argues that the infringement claims fail given the allegations alleged with

13   respect to its particular conduct.

14   **A.   Request for Judicial Notice**

15   As an initial matter, Runway asks me to take judicial notice of a motion to dismiss filed in

16   another case in this District, as well as the complaint from the case. Runway RJN (Dkt. No. 164),

17   Exs. A&B. Plaintiffs oppose, arguing that judicial notice of court records from other cases is

18   inappropriate for the purpose Runway seeks; to encourage me to follow the "approach" of those

19   other courts. "On a Rule 12(b)(6) motion to dismiss, when a court takes judicial notice of another

20   court's opinion, it may do so 'not for the truth of the facts recited therein, but for the existence of

21   the opinion, which is not subject to reasonable dispute over its authenticity.'" *Southern Cross*

22   *Overseas Agencies, Inc. v. Wah Kwong Shipping Group Ltd.*, 181 F.3d 410, 426–27 (3rd

23

24   _____

[14] Because of this conclusion, I need not reach Stability's argument under *Sonner v. Premier*
25   *Nutrition Corp.*, 971 F.3d 834 (9th Cir. 2020). I have however, repeatedly held that a plaintiff
     satisfies *Sonner* at the pleading stage by simply pleading inadequate remedies at law. *See Costa v.*
26   *Apple, Inc.*, No. 23-CV-01353-WHO, 2023 WL 7389276, at *4 (N.D. Cal. Aug. 9, 2023). If
     plaintiffs reassert unjust enrichment claims in their Second Amended Complaint, they should
27   address that issue. The unjust enrichment claims asserted against Runway and Midjourney raise
     identical issues. Those claims are, therefore, DISMISSED with one last leave to amend and will
28   not be addressed further. The unjust enrichment claim asserted against DeviantArt is addressed
     below.

1   Cir.1999).” *Lee v. City of Los Angeles*, 250 F.3d 668, 690 (9th Cir. 2001). Notifying me of the

2   existence of opinions from the judges in other cases may be accomplished by a far simpler

3   method, citing them as persuasive authority. Judicial notice is not an appropriate method to

4   suggest that I should follow the analysis of other courts addressing different cases with different

5   facts.

6   Runway also asks me to take judicial notice – under the doctrine of incorporation – of the

7   full contents of three academic articles cited in plaintiffs' FAC. *Id*., Exs. C, D & E. Plaintiffs

8   object to this request, noting that one article is only mentioned once in the FAC and the others a

9   few more times as support for the plausibility of plaintiffs' assertion that Stable Diffusion contains

10   "compressed copies" of the Training Images based on how these models generally work.

11   Plaintiffs argue that it is inappropriate to take "notice" of the debated truth, meaning or

12   implications of the articles to foreclose their claims. I agree. I will not take judicial notice of the

13   full contents of the academic articles to resolve disputes of fact, or the legal implications from

14   undisputed facts, at this juncture.

15   Runway's request for judicial notice is DENIED.

16   **B.   Direct Copyright Infringement**

17   Runway does not move to dismiss first direct copyright infringement claim asserted

18   against it. *See* Count 11, FAC ¶¶ 347-349. That claim is based on Runway's alleged use of

19   Training Images to train Stable Diffusion 1.5, a claim that survived Stability AI's prior motion to

20   dismiss. October 2023 Order at 7 (discussing the "Training Theory"). Runway does challenge

21   plaintiffs' two other direct infringement theories. First, the "Model Theory" is based on the theory

22   that the Stable Diffusion1.5 product itself – after it was trained – is "an infringing Statutory Copy"

23   of plaintiffs' works or a "Statutory Derivative Work" because it represents a transformation of

24   plaintiffs' works. *See* FAC ¶¶ 209, 350. Second, Runway challenges plaintiffs' "Distribution

25   Theory" of infringement, based on allegations that Runway infringes plaintiffs' exclusive

26   distribution rights because distributing Stable Diffusion 1.5 is equivalent to distributing plaintiffs'

27   works. FAC ¶ 352.

28   As with Stability, because Runway does not challenge the use of the images for training

1   purposes, I need not address the other theories of direct infringement. However, I note that both

2   the model theory and the distribution theory of direct infringement depend on whether plaintiffs'

3   protected works are contained, in some manner, in Stable Diffusion as distributed and operated.

4   That these works may be contained in Stable Diffusion as algorithmic or mathematical

5   representations – and are therefore fixed in a different medium than they may have originally been

6   produced in – is not an impediment to the claim at this juncture. 1 Nimmer on Copyright §

7   2.09[D][1] (2024) ("A work is no less a motion picture (or other audiovisual work) whether the

8   images are embodied in a videotape, videodisc, or any other tangible form.").

9   Plaintiffs addressed the deficiencies in their prior complaint, alleging additional facts in the

10   FAC concerning how the training images remain in and are used by Stable Diffusion. *See, e.g.*,

11   FAC ¶¶ 71, 83, 88-90, 150.[15] Plaintiffs rely, with respect to Runway as with Stability, on

12   comments from Stability's CEO regarding the contents of the model and their ability to reproduce

13   works, as well as academic papers indicating that the Stable Diffusion models are capable of

14   producing very similar if not identical works to at least some training images. FAC ¶¶ 122-150.

15   And plaintiffs also rely on use of their names as prompts in the Runway products to create outputs

16   mimicking aspects of plaintiffs' protected works as evidence that their protected works are being

17   copied or distributed in Runway's product. *Id*. ¶¶ 163-169.

18   Runway disputes the accuracy of those assertions – including the full meaning and import

19   of the academic articles relied on by plaintiffs – and argues that the prompt-examples do not

20   support assertions of direct copyright infringement absent express identification of outputs that are

21   "substantially similar" to plaintiffs' copyrighted works. But the allegations at this juncture are

22   sufficient to allow the direct infringement claims to proceed. Whether evidence can support each

23

24   [15] Runway's and the other defendants' repeated reliance on "run of the mill" copyright cases
     where a showing of substantial similarity between works is required when determining whether an
25   inference of copying can be supported – *see, e.g., Hanagami v. Epic Games, Inc*., 85 F.4th 931,
     935 (9th Cir. 2023) – or liability imposed – *see, e.g., Narell v. Freeman*, 872 F.2d 907, 910 (9th
26   Cir. 1989) – are unhelpful in this case where the copyrighted works themselves are alleged to have
     not only been used to train the AI models but also invoked in their operation. Whether that use is
27   "substantial enough" either in operation or output of images to qualify for the fair use defense will
     be tested on summary judgment. *See, e.g., Authors Guild v. Google, Inc*., 804 F.3d 202, 226 (2d
28   Cir. 2015) (affirming summary judgment that Google's "program does not allow access in any
     substantial way to a book's expressive content").

United States District Court
Northern District of California

1   of the theories and whether plaintiffs will need to choose between theories (*e.g.*, between direct

2   infringement based on selling a product containing effective copies of copyrighted works or

3   violating plaintiffs' rights to restrict distribution of their works) will be addressed at summary

4   judgment.[16]

5   Runway's motion to dismiss the direct infringement claims is DENIED.

6   **C.** **Induced Infringement**

7   Runway argues that plaintiffs have also failed to allege two elements of the induced

8   infringement claim: acts of infringement by third parties using Runway's products and that

9   Runway promoted use of Stable Diffusion to infringe. Runway claims that the only support for

10 the induced infringement claim are comments plaintiffs identify by Stability executives, not

11 statements by anyone associated with Runway.

12   Plaintiffs allege that Runway helped train and develop Stable Diffusion, and therefore,

13 knew that the product allegedly uses or invokes the training images in its operation. Those

14 allegations, combined with allegations that Runway actively induces others to download Stable

15 Diffusion by distributing Stable Diffusion through popular coding websites and also by making

16

17 _____

    [16] Runway, like the other defendants, picks up on my questions from the prior oral argument to

18 argue that plaintiffs should be able to point to source code from the "open source" AI products to
    identify where in the Stable Diffusion and defendants' AI products copies of plaintiffs' works are

19 stored. *See, e.g.*, Runway Reply at 4-5. In the FAC, plaintiffs challenge the use of the label
    "open source," pointing out that various components of the models, including weights files, are

20 not open for inspection by all. FAC ¶¶ 147-148. Those allegations satisfy the Court's prior
    queries. Runway also relies on a decision by a different judge in this court rejecting derivative

21 infringement theories in *Kadrey v. Meta Platforms, Inc*., No. 23-CV-03417-VC, 2023 WL
    8039640, at *1 (N.D. Cal. Nov. 20, 2023). In *Kadrey*, the Hon. Vince Chhabria considered

22 copyright infringement with respect to copyrighted written works that were used to train large
    language models ("LLaMA"), software programs designed to produce naturalistic text outputs in

23 response to user prompts. He dismissed plaintiffs' derivative copyright theories because plaintiffs
    failed to allege that the "LLaMA models themselves" could be understood "as a recasting or

24 adaptation of any of the plaintiffs' books" and because there were no allegations that the outputs
    of those models could be "understood as recasting, transforming, or adapting the plaintiffs' books"

25 or otherwise producing outputs that were "substantially similar" to aspects of plaintiffs' books.
    Kadrey, 2023 WL 8039640 at *1. The products at issue here – image generators allegedly trained

26 on, relying on, and perhaps able to invoke copyrighted images – and the necessary allegations
    regarding the products' training and operations, are materially different from those in *Kadrey*.

27 Whether substantial similarity remains a hurdle to specific theories – including any derivative
    infringement theory – depends in part on what the evidence shows concerning how these products

28 operate and, presumably, whether and what the products can produce substantially similar outputs
    as a result of "overtraining" on specific images or by design.

1   selling its products (including AI Magic Tools) that include Stable Diffusion, are sufficient. FAC

2   ¶¶ 83, 163, 185, 352, 358.[17]

3   　　　Runway's motion to dismiss is DENIED on the infringement claims, GRANTED with

4   prejudice on the DMCA claims, and GRANTED with leave to amend on the unjust

5   enrichment/UCL claim.

6   **IV.   MIDJOURNEY MOTION TO DISMISS**

7   　　　In addition to moving to dismiss the DMCA and unjust enrichment claims addressed

8   above, Midjourney moves to dismiss the copyright claims and Lanham Act false endorsement and

9   trade dress claims[18] asserted against it.

10   　　　**A.   Copyright**

11   　　　　　　**1.   Registration**

12   　　　Midjourney argues, with respect to three of the named plaintiffs – Anderson, Kaye, and

13   Brom – that the evidence of their registration of newly identified copyrighted works is

14   insufficient. With respect to Anderson and Kaye, Midjourney asserts that a subset of each artists'

15   works identified as being both copyrighted and included in the LAION datasets used to train the

16   AI products are compilations. It contends that copyright protection only extends to the new

17   material in compilations and Anderson and Kaye fail to identify which works within the subset of

18   compilations are the new material. It also argues that for two of three of the works identified by

19   Brom, the copyright registrations extend only to text and not artwork. Midjourney Mot. at 6-8.

20   　　　It is undisputed that plaintiffs who do not have valid copyright protections will not be able

21   to pursue copyright claims based on un-registered works or works whose registrations covered

22   only text. It is also undisputed that *each* of the named plaintiffs who claim their copyright-

23   protected works were included in the LAION datasets have at least one work whose registration is

24   facially valid. At this juncture, therefore, the Copyright Act claims survive against Midjourney

25

26   _____

27   [17] Runway's argument that plaintiffs' induced infringement claim must fail because plaintiffs do not plead that Runway's models lack "substantial non-infringing uses" is rejected for the reasons discussed above with respect to Stability. *See supra* at 8-9.

28   [18] 15 U.S.C. § 1125(a) *et seq.*

1   and the other defendants. However, the identification in the FAC and exhibits of unprotected

2   works is not irrelevant; plaintiffs rely on some of those works to plausibly demonstrate that their

3   works were used as training images and that their works or elements of their works can be

4   recreated through the AI products. The identification of those works may not prove liability under

5   the Copyright Act, but they do provide support for the plausibility of plaintiffs' Copyright Act

6   theories.

7   **2.   Use as Training Images**

8   In my prior Order, I required plaintiffs to clarify the basis of their copyright claims against

9   Midjourney, as plaintiffs' allegations regarding whether and how Midjourney used plaintiffs'

10   images in the training of its product were insufficient. October 2023 Order at 13-14. In their

11   FAC, plaintiffs now allege that Midjourney separately trained its product on the LAION400M and

12   LAION5B datasets. *See* FAC ¶¶ 266, 274. Plaintiffs also allege that Midjourney incorporates

13   Stable Diffusion into its own AI product. *Id*. ¶¶ 169-170.

14   Midjourney contests the adequacy of these new allegations with respect to the training of

15   images, arguing plaintiffs must identify specific, individual registered works that each artist

16   contends Midjourney actually used for training. Given the unique facts of this case – including the

17   size of the LAION datasets and the nature of defendants' products, including the added allegations

18   disputing the transparency of the "open source" software at the heart of Stable Diffusion – that

19   level of detail is not required for plaintiffs to state their claims. Instead, plaintiffs have added to

20   their FAC more detailed allegations regarding the training and use of the LAOIN datasets by

21   defendants generally and Midjourney specifically. Plaintiffs have plausible allegations showing

22   why they believe their works were included in the LAION datasets. And plaintiffs plausibly

23   allege that the Midjourney product produces images – when their own names are used as prompts

24   – that are similar to plaintiffs' artistic works. *See* FAC Exs. F&G.

25   Midjourney nonetheless argues that these examples are insufficient because some of the

26   identified works are not registered and the resulting outputs could just as likely be the result of

27   training on unregistered works or utilizing only unprotected elements from plaintiffs' works. But

28   plaintiffs' reliance on those exhibits are not to establish copyright infringement as a matter of law.

                                          20

1   Instead, they are relied on to support the *plausibility* of plaintiffs' copyright theories (that all or

2   most of the works in the LAION datasets were used by Midjourney and the other defendants to

3   train their AI products, and that plaintiffs' works or their protected elements that are contained in

4   the AI products as the works or protected elements can be recreated by using the AI products).

5   The FAC allegations and the exhibits help plaintiffs cross the plausibility threshold. Whether

6   plaintiffs will be able to prove their claims is a different matter and those claims will be tested on

7   an evidentiary basis at summary judgment.[19]

8   Midjourney's motion to dismiss the Copyright Act claims is DENIED.

9   **B.   Lanham Act**

10   In place of the right of publicity claims asserted against Midjourney that were dismissed in

11   my October 2023 Order, five plaintiffs (Anderson, Brom, Kaye, Ortiz and Rutkowski) now assert

12   Lanham Act claims based on theories of false endorsement and trade dress. These plaintiffs allege

13   that their names appeared on the list of 4700 artists posted by Midjourney's CEO on Discord, the

14   platform where Midjourney's AI product operates. Midjourney's CEO promoted the list as

15   describing the various styles of artistic works its AI product could produce. FAC ¶¶ 261-262,

16   305.[20] Plaintiffs also allege that Midjourney has itself published user-created images that

17   incorporate the plaintiff artists' names in Midjourney's "showcase" site. *Id*. ¶ 325(b) & Ex. K.

18   **1.   False Endorsement**

19   "To prevail on its Lanham Act trademark claim, a plaintiff must prove: (1) that it has a

20   protectible ownership interest in the mark; and (2) that the defendant's use of the mark is likely to

21   ───────────────

22   [19] Echoing the arguments made by Runway, Midjourney also challenges the model and
     distribution theories. As above, given the FAC's plausible allegations regarding training of
23   Midjourney's product and that "copies" or protected elements of some of some plaintiffs'
     registered works remain in Midjourney's AI product, these theories survive to be tested on
24   summary judgment.

25   [20] FAC ¶ 305 ("Midjourney's use of the Midjourney Named Plaintiffs' names was purely to
     advertise its image generator. This use does not contribute significantly to a matter of public
26   interest. The purpose of publishing over 4700 names in the Midjourney Name List was to promote
     and highlight the capabilities of Midjourney's image generator to emulate and create work that is
27   indistinguishable from that of the artists whose names were published."); ¶ 309 ("A reasonably
     prudent consumer in the marketplace for art products likely would be confused as to whether the
28   Midjourney Named Plaintiffs included in the Midjourney Name List sponsored or approved of
     Midjourney's image generator.").

1   cause consumer confusion." *Rearden LLC v. Rearden Commerce, Inc.*, 683 F.3d 1190, 1202–203

2   (9th Cir. 2012) (quoting *Network Automation, Inc. v. Advanced Sys. Concepts, Inc.*, 638 F.3d

3   1137, 1144 (9th Cir. 2011)). A false designation of origin claim likewise requires a showing of a

4   likelihood of consumer confusion. *New W. Corp. v. NYM Co. of Cal., Inc.*, 595 F.2d 1194, 1201

5   (9th Cir. 1979) ("Whether we call the violation infringement, unfair competition or false

6   designation of origin, the test is identical[:] is there a 'likelihood of confusion?'").

7   Midjourney contends that plaintiffs fail to allege the first, necessary element of their false

8   endorsement claim: falsity. They note that plaintiffs do not allege that the Midjourney AI product

9   cannot recognize works by those included on the Midjourney Names List, and argue that just

10   because the Names List exists and was promoted by the Midjourney CEO, that by itself cannot

11   support an inference of endorsement. Midjourney points to a part of the Discord thread (identified

12   by plaintiffs as the source of the Midjourney CEO's identification of the Names List) not relied on

13   by plaintiffs, where the CEO indicated that the names on the list came from "wikipedia and magic

14   the gathering." Midjourney MTD, Dkt. No. 169, at 18. Midjourney argues that this part of the

15   Discord thread is judicially noticeable and dispels any inference that the artists on the lists could

16   plausibly be considered to have endorsed the Midjourney product.

17   As discussed above, judicial notice of other comments in the thread is not appropriate to

18   dispute the facts plaintiffs otherwise plausibly assert, especially as the plaintiffs dispute the

19   accuracy and inferences to be drawn from the totality of the messages in that thread. And even if

20   the Names List itself was insufficient to support an inference of false endorsement, plaintiffs also

21   allege that their names were used in connection with works included in Midjourney's "showcase."

22   FAC ¶ 325b. Whether or not a reasonably prudent consumer would be confused or misled by the

23   Names List and showcase to conclude that the included artists were endorsing the Midjourney

24   product can be tested at summary judgment. *See Adobe Sys. Inc. v. A & S Elecs., Inc.*, 153 F.

25   Supp. 3d 1136, 1142 (N.D. Cal. 2015).[21]

26

27   [21] While the court in *Smith v. Chanel, Inc.*, 402 F.2d 562 (9th Cir. 1968) explained that the
     "Lanham Act does not prohibit a commercial rival's truthfully denominating his goods a copy of a
28   design in the public domain, though he uses the name of the designer to do so," that was a case
     where the seller expressly advertised its product as "equivalent" to the trademarked product, which

1   Midjourney also argues that to get past the First Amendment protection provided to

2   expressive works, plaintiffs must, but have not, alleged that Midjourney's use of plaintiffs' names

3   to invoke their styles has "'no artistic relevance to the underlying work whatsoever, or, if it has

4   some artistic relevance, unless the [use of trademark or other identifying material] explicitly

5   misleads as to the source or the content of the work.'" *Brown v. Elec. Arts, Inc.*, 724 F.3d 1235,

6   1239 (9th Cir. 2013) (quoting *Rogers v. Grimaldi*, 875 F.2d 994 (2d Cir.1989)). However, the

7   plaintiffs have plausibly alleged that the use of their names by Midjourney in the List and

8   showcase misleads consumers regarding source and endorsement. Open questions also remain

9   whether Midjourney promoting its product for commercial gain for use by others to create artistic

10   images is *itself* expressive use that creates "artistic relevance" to plaintiffs' underlying works.

11   Discovery may show that it is or that is it not. Unlike in *Brown v. Elec. Arts, Inc.*, where plaintiff

12   alleged only that his likeness was used in a video game, we do not yet have the sort of record, or

13   sort of alleged use, that could support dismissal of the claim at the motion to dismiss stage.

14   **2.   Vicarious Trade Dress**

15   The same five plaintiffs (Anderson, Brom, Kaye, Ortiz and Rutkowski) also allege a trade

16   dress claim based on the use of their names in connection with the Midjourney AI product's use of

17   a "CLIP-guided model" that has been trained on the work of the Midjourney Named Plaintiffs and

18   allows users to create works capturing the "trade dress of each of the Midjourney Named Plaintiffs

19   [that] is inherently distinctive in look and feel as used in connection with their artwork and art

20   products." FAC ¶¶ 321. The CLIP model, plaintiffs assert, works as a trade dress database that

21   can recall and recreate the elements of each artist's trade dress. FAC ¶¶ 83, 320. Plaintiffs point

22   to examples showing how Midjourney recreates works with their trade dress in Ex. F to the FAC.

23   Midjourney argues that plaintiffs have failed to state this claim because they have not

24   adequately identified the "concrete elements" of each plaintiff's trade dress. *See, e.g., YZ Prods.,*

25

26   the court recognized was promoting competition. *Id*. at 565-66. That case was also admittedly not

27   one with allegations of "misrepresentation or confusion as to source or sponsorship," as here. *Id*.
     Whether the use of plaintiffs' names and works in Midjourney's advertising/promotion or the

28   operation of its product was misleading or is truthful comparative advertising can be tested at
     summary judgment.

1   *Inc. v. Redbubble, Inc.*, 545 F. Supp. 3d 756, 767 (N.D. Cal. 2021) ("images alone are insufficient

2   to provide adequate notice dismissing trade dress claim," and requiring plaintiff to identify "the

3   concrete elements" of their protected trade dress). It acknowledges that plaintiffs have identified

4   aspects of their trade dress, *see* FAC ¶ 319,[22] but asserts that those descriptions are impermissibly

5   broad. *See, e.g., Walker & Zanger, Inc. v. Paragon Indus., Inc.*, 549 F. Supp. 2d 1168, 1176 (N.D.

6   Cal. 2007 (granting summary judgment on the "some of the elements of plaintiff's proposed trade

7   dress [that] are overbroad"). Plaintiffs refute that characterization, arguing instead that they have

8   identified the "set of recurring visual elements and artistic techniques, the particular combination

9   of which are distinctive to each of the Midjourney Named Plaintiffs." *Id*. ¶ 319.

10   While the images from Exhibit F on their own would be insufficient identification, and

11   while some of the alleged "concrete elements" identified in the FAC are, standing alone, vague

12   and possibly overbroad, those elements cannot be considered alone but as a whole in the context

13   of plaintiffs' other, plausible allegations. *Arcsoft, Inc. v. Cyberlink Corp*., 153 F. Supp. 3d 1057,

14   1068 (N.D. Cal. 2015). Here, the combination of identified elements and images, when

15   considered with plaintiffs' allegations regarding how the CLIP model works as a trade dress

16   database, and Midjourney's use of plaintiffs' names in its Midjourney Name List and showcase,

17   provide sufficient description and plausibility for plaintiffs' trade dress claim.

18   Midjourney also argues that artistic elements or styles identified for each artist that are

19   allegedly re-creatable by using its product are functional, and therefore not protected.[23] It ignores,

20

21   [22] FAC ¶ 319 (" a. Sarah Andersen is known for work that is simple, cartoony, and often strictly
    in black and white. In particular, she is known for "Sarah's Scribbles," a comic featuring a young

22   woman with dark hair, big eyes, and a striped shirt. b. b. Karla Ortiz is known for a mixture of
    classical realism and impressionism, often delving into fantastical, macabre and surrealist themes,

23   and inspired by the technical prowess of American Renaissance movements with a strong
    influence of contemporary media. c. Gerald Brom is known for gritty, dark, fantasy images,

24   painted in traditional media, combining classical realism, gothic and counterculture aesthetics. d.
    Grzegorz Rutkowski is known for lavish fantasy scenes rendered in a classical painting style. e.

25   Julia Kaye is known for three-panel black-and-white comics, loosely inked with a thin fixed-width
    pen, wherein each individual comic is a microvignette in the artist's life.").

26   [23] *See Arcsoft,* 153 F. Supp. 3d at 1068 ("To determine whether a claimed trade dress is functional,

27   the Ninth Circuit considers several factors: '(1) whether the design yields a utilitarian advantage,
    (2) whether alternative designs are available, (3) whether advertising touts the utilitarian

28   advantages of the design, and (4) whether the particular design results from a comparatively
    simple or inexpensive method of manufacture.'" (quoting *Apple Inc. v. Samsung Elecs. Co*., 786

1   however, the Ninth Circuit's test for determining non-functionality and instead relies on a series of

2   inapposite cases addressing jewelry, wooden cutouts, and keychains. *See, e.g., Int'l Ord. of Job's*

3   *Daughters v. Lindeburg & Co*., 633 F.2d 912, 918 (9th Cir. 1980) (use of name and emblem of

4   fraternal organization were functional aesthetic components of jewelry); *Crafty Prods., Inc. v.*

5   *Michaels Companies, Inc*., 424 F. Supp. 3d 983, 993 (S.D. Cal. 2019), *aff'd sub nom. Crafty*

6   *Prods., Inc. v. Fuqing Sanxing Crafts Co*., 839 F. App'x 95 (9th Cir. 2020) (plaintiffs failed to

7   identify "concrete elements" of their trade dress and instead attempted to capture the "entire

8   design" of hundreds of different wooden cardboard pieces); *Au-Tomotive Gold, Inc. v. Volkswagen*

9   *of Am., Inc*., 457 F.3d 1062, 1064 (9th Cir. 2006) (reversing summary judgment protecting

10   defendants' use of trademarked symbols in keychains and license plate holders, as "[t]he doctrine

11   of aesthetic functionality does not provide a defense against actions to enforce the trademarks

12   against such poaching" or source identification).

13   Midjourney complains that plaintiffs intend to assert a monopoly over some of the

14   elements plaintiffs identify as critical parts of their protected trade dress. But that ignores the

15   plausible allegations that the CLIP model functions as a trade dress database and the use of its

16   product to produce works based on the names of these plaintiffs. Plaintiffs have no protection

17   over "simple, cartoony drawings" or "gritty fantasy paintings." Midjourney Reply, Dkt. No. 184

18   at 13. But their trade dress claims must be considered viewing *all* identified elements as well as

19   the nature of the use of the CLIP model and their names. Those issues will be tested on an

20   evidentiary basis. This claim will not be dismissed based on "aesthetic functionality."

21   Midjourney contends that the trade dress claim must nonetheless be dismissed because

22   plaintiffs allege their trade dress is "either" distinctive or, in the alternative, has acquired

23   secondary meaning.[24] However, plaintiffs pleaded secondary meaning. *See* FAC ¶ 328 ("Each of

24

25   F.3d 983, 991 (Fed.Cir.2015)). "In applying these factors and evaluating functionality, 'it is
     crucial that [the court] focus not on the individual elements, but rather on the overall visual
26   impression that the combination and arrangement of those elements create.'" *Arcsoft*, 153 F.
     Supp. 3d at 1068 (quoting *Clicks Billiards, Inc. v. Sixshooters*, Inc., 251 F.3d 1252, 1259 (9th
27   Cir.2001)).

28   [24] As the case relied on by Midjourney, *Art Attacks Ink, LLC v. MGA Ent. Inc*., 581 F.3d 1138,
     (9th Cir. 2009), explains: "'Secondary meaning can be established in many ways, including (but

                                                  25

1   the Midjourney Named Plaintiffs' trade dress possesses secondary meaning because the trade

2   dress of their art products invoke a mental association by a substantial segment of potential

3   consumers between the trade dress and the creator of the art product."). Proof of intentional

4   copying by Midjourney has been alleged, especially considering the allegations regarding the

5   CLIP model functioning as a trade dress database, the express use of plaintiffs' names in the

6   Midjourney Names List to promote the product and the use of some plaintiffs' names and

7   likenesses of their works in the Midjourney showcase, and the "mental recognition" of

8   Midjourney's calling out of plaintiffs by name.

9   Finally, Midjourney contests the sufficiency of plaintiffs' vicarious trade dress claim,

10   arguing that plaintiffs have failed to allege facts of "joint ownership or control" over the infringing

11   product by Midjourney and any end user. Plaintiffs allege:

> Midjourney exercises control over the infringing images by including the CLIP model in its image pipeline, and by marketing artist-name prompts as a key feature of its image generator via the Midjourney Name List. Without the CLIP model, Midjourney's users would not be able to infringe on the Midjourney Named Plaintiffs' trade-dress rights or those of the other artists on the Midjourney Name List.

15   FAC ¶ 326. These allegations support the claim for vicarious trade dress infringement. See, e.g.,

16   *Y.Y.G.M. SA v. Redbubble, Inc*., No. 219CV04618RGKJPR, 2020 WL 3984528, at *9 (C.D. Cal.

17   July 10, 2020 ("To impose vicarious liability the shared control must extend in some way to the

18   infringed intellectual property itself.").

19   Midjourney's motion to dismiss the Lanham Act claims is DENIED.[25]

---

not limited to) direct consumer testimony; survey evidence; exclusivity, manner, and length of use of a mark; amount and manner of advertising; amount of sales and number of customers; established place in the market; and proof of intentional copying by the defendant.' *Filipino Yellow Pages, Inc. v. Asian Journal Publ'ns, Inc*., 198 F.3d 1143, 1151 (9th Cir.1999). To show secondary meaning, a plaintiff must demonstrate 'a mental recognition in buyers' and potential buyers' minds that products connected with the [mark] are associated with the same source.' *Japan Telecom v. Japan Telecom Am*., 287 F.3d 866, 873 (9th Cir.2002) (internal quotation omitted)." 581 F.3d at 1145.

[25] I acknowledge Midjourney's reliance on treatises and cases that caution against "extending trademark law to intrude into the domain of copyright law." *See* McCarthy on Trademarks and Unfair Competition § 10:40.50 (5th ed.); *see also id*. ("Is there such a thing as trademark protection for the 'style' of an artist? Courts have almost uniformly said no."); *Leigh v. Warner Bros*., 10 F. Supp. 2d 1371, 1380–81 (S.D. Ga. 1998), *aff'd in part, rev'd in part on other grounds*, 212 F.3d 1210 (11th Cir. 2000) (no trademark claim based on use of plaintiff's

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28

C.   Request for Judicial Notice

Midjourney requests judicial notice of: (1) messages from a Discord thread, where Midjourney CEO David Holz released "our style list" and posted a link to a spreadsheet on Google Docs called "Midjourney Style List" that contained the named of 4700 artists, under the doctrine of incorporation by reference, Dkt. No. 170-1, Ex. 1; (2) messages in a different Discord thread that plaintiffs rely on in part, where Holz discusses Midjourney's image-prompt tool, Dkt. No. 170-1, Ex. 2; (3) another Discord thread discussing the image-prompt tool; and (4) a transcript from the hearing before Judge Chhabria in the *Kadrey* case.

As noted before, judicial notice of documents from the *Kadrey* case is not necessary, as either side can point to the orders from that case to argue their persuasiveness or differences with respect to the sufficiency of the allegations in this case. *See supra* at 15-16. Judicial notice is also not appropriate for any of the three Discord threads to dispute the facts plausibly alleged. For example, Midjourney does not dispute that the list of artists was disclosed and promoted by its CEO, which is the central reason for which plaintiffs rely on that Discord thread and its attachment. Midjourney seeks to rely on the thread to dispute the meaning of the comments by Holz and point to messages that are not relied on by plaintiffs (and whose meanings are disputed by plaintiffs) to foreclose plaintiffs' claims. That is not appropriate.

Midjourney's request for judicial notice is DENIED.

V.   DEVIANTART MOTION TO DISMISS

A.   Copyright Claim

DeviantArt moves again to dismiss the copyright claims asserted against it by attempting to set itself apart from the other defendants. It points out that it has not been alleged to have trained any AI model, but simply to have implemented and used the AI tools provided by Stability and others. It argues that holding it liable for that conduct would make the millions of third parties

---

copyrighted photograph because the work "merely identifie[d] the artist rather than any products or services" he sold, it could not "be protected as a trademark," even if it was an example of his "unique artistic style."). These cases do not readily fit the allegations plaintiffs raise regarding how Midjourney's product and other CLIP model products function as trade dress databases, combined with Midjourney's use of plaintiffs' names and showcase examples calling out named plaintiffs. This argument is better determined on a full record.

1   who have downloaded and implemented the AI products challenged in this case – many of which

2   are open source software – liable for infringement, which is unsupportable.

3   In the October 2023 Order, I dismissed the copyright infringement claim asserted against

4   DeviantArt, explaining:

> In addition to providing clarity regarding their definition of and theory
> with respect to the inclusion of compressed copies of Training Images
> in Stable Diffusion, plaintiffs shall also provide more facts that
> plausibly show how DeviantArt is liable for direct copyright
> infringement when, according to plaintiffs' current allegations,
> DeviantArt simply provides its customers access to Stable Diffusion
> as a library. Plaintiffs do cite testimony from DeviantArt's CEO that
> DeviantArt uses Stable Diffusion because Stability allowed
> DeviantArt to "modify" Stable Diffusion. Compl. ¶ 129. The problem
> is that there are no allegations what those modifications might be or
> why, given the structure of Stable Diffusion, any compressed copies
> of copyrighted works that may be present in Stable Diffusion would
> be copied within the meaning of the Copyright Act by DeviantArt or
> its users when they use DreamUp. Nor do plaintiffs provide plausible
> facts regarding DeviantArt "distributing" Stable Diffusion to its users
> when users access DreamUp through the app or through DeviantArt's
> website.

14   October 2023 Order at 10.

15   DeviantArt argues, first, that plaintiffs' copyright act claims are barred by the October

16   2023 Order because plaintiffs still do not allege that DeviantArt itself copied or used their works

17   to train Stable Diffusion or trained any other program. The FAC alleges only that DeviantArt

18   incorporates and relies on Stable Diffusion for its DreamUp product. FAC ¶¶ 6, 387-392, 395-397.

19   Plaintiffs have added allegations to their FAC, however, regarding how copies or protected

20   elements of their works remain, in some format, in Stable Diffusion and how those works can be

21   invoked by use of all of the Stable Diffusion versions. The actual operation of Stable Diffusion

22   1.4 and whether the amount of any plaintiff's copyrighted works in that program suffices for

23   copyright infringement or a fair use defense concerning DeviantArt remains to be tested at

24   summary judgment.

25   Moreover, while plaintiffs admittedly did not include assertions or examples in the FAC of

26   outputs from DreamUp that appear to copy elements of their works (as they did for the other

27   defendants), their added allegations regarding the use of the LAION datasets to train Stable

28   Diffusion and how Stable Diffusion versions operates, including specific examples and academic

1   references to the operation of Stable Diffusion 1.4 used by DreamUp, suffice. *See* FAC ¶¶ 132-

2   137, 388-397. For example, in paragraph 393 plaintiffs assert:

3   On information and belief, by the end of training, Stable Diffusion 1.4
            was capable of reproducing protected expression from each of the
4   LAION-5B Registered Works that was in each case substantially
            similar to that registered work, because— a. In the Carlini Paper,
5   Nicholas Carlini tested Stable Diffusion 1.4 and found that it could
            emit stored copies of its training images; b. The training procedure
6   for Stable Diffusion 1.4 was very similar to that of Stable Diffusion
            1.5, which was shown in Exhibit E: Runway text prompts and Exhibit
7   H: Runway image prompts to be capable of emitting stored copies of
            protected expression.
8

See also ¶ 394 ("Therefore, like Stable Diffusion 1.5, Stable Diffusion 1.4 also qualifies as an

9

infringing Statutory Copy of the LAION-5B Registered Works. Because Stable Diffusion 1.4
10
represents a transformation of the LAION-5B Registered Works into an alternative form, Stable
11
Diffusion 1.4 also qualifies as an infringing Statutory Derivative Work.").
12
        DeviantArt asks me to review to the full content of one of the academic articles plaintiffs
13
rely on, the Carlini Study, in particular the article's conclusion that for the 350,000 training
14
images studied, only 109 output images were "near-copies" of the training images. *See* DeviantArt
15
Reply at 4-5. DeviantArt argues that it is simply not plausible that LAION 1.4 can reproduce any
16
of plaintiffs' copyrighted works given that the set of training images Carlini selected were the
17
"most-duplicated" examples, numbering in the millions of impressions in the datasets. *Id*.
18
Plaintiffs' reference to Carlini, however, is only one part of the allegations that help make
19
plaintiffs' allegations plausible regarding how these products operate and how "copies" of the
20
plaintiffs' registered works are captured in some form in the products. As noted above, it is not
21
appropriate to rely on the defendants' assertions regarding the method or sample size used by
22
Carlini and their conclusions from it to foreclose plaintiffs' claims; plaintiffs aggressively dispute
23
the implications from the sample size and results of that study. Finally, while the differences
24
between Stable Diffusion 1.4 and the other Stable Diffusion versions might be legally significant
25
based on what the evidence shows after discovery, at this juncture the allegations about the
26
common training of those versions and how they all operate are sufficient to keep plaintiffs'
27

28

29

1   copyright infringement allegations against DeviantArt alive.[26]

2   Finally, DeviantArt contends that any use of plaintiffs' works in Stable Diffusion 1.4

3   should be considered fair use as a matter of law, given the huge size of the training datasets and

4   plaintiffs apparent inability to use Stable Diffusion 1.4 to reproduce any works that look similar to

5   their copyrighted works. Whether DreamUp operates in a way that could draw upon or otherwise

6   reproduce plaintiffs' works to an extent that violates the Copyright Act and whether a fair use

7   defense applies are issues that must be tested on an evidentiary basis. *See, e.g., Authors Guild v.*

8   *Google, Inc*., 804 F.3d 202, 226 (2d Cir. 2015) (applying fair use defense upon summary judgment

9   record).

10   DeviantArt's motion to dismiss the Copyright Act claims is DENIED.

11   **B.   Breach of Contract**

12   DeviantArt moves again to dismiss the breach of contract claim that plaintiffs assert based

13   on the same provision of DeviantArt's Terms of Service ("ToS") that I considered in the October

14   2023 Order. In that Order I explained:

15   In opposition, plaintiffs focus on § 16 of the TOS, a provision not
               identified or quoted in their Complaint:

16
               16. Copyright in Your Content

17   DeviantArt does not claim ownership rights in Your Content.
               For the sole purpose of enabling us to make your Content

18   available through the Service, you grant to DeviantArt a non-
               exclusive, royalty-free license to reproduce, distribute, re-

19   format, store, prepare derivative works based on, and publicly
               display and perform Your Content. Please note that when you

20   upload Content, third parties will be able to copy, distribute
               and display your Content using readily available tools on their

21   computers for this purpose although other than by linking to
               your Content on DeviantArt any use by a third party of your

22   Content could violate paragraph 4 of these Terms and
               Conditions [preserving copyright rights in the original owner

23

United States District Court
Northern District of California

24   [26] Given this conclusion, I need not separately reach the question of whether plaintiffs can assert a

25   theory of derivative copyright infringement based on the use of the diffusion model itself. As
     defendants note, the soundness of this theory was questioned in my October 2023 Order and by

26   other judges in this District, with respect to *output images*. *See* October 2024 Order at 10-13;
     Tremblay v. OpenAI, Inc., No. 23-CV-03223-AMO, 2024 WL 557720, at *3 (N.D. Cal. Feb. 12,

27   2024). Judge Chhabria also rejected the derivative theory as applied to the AI model in *Kadrey v.
     Meta Platforms, Inc*., No. 23-CV-03417-VC, 2023 WL 8039640, at *1 (N.D. Cal. Nov. 20, 2023).

28   But as noted above, the allegations regarding the training and operation of the language models at
     issue in *Kadrey* are significantly different than the image creation models at issue here.

30

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

of the copyright and disclaiming any ownership interest of DeviantArt in the posted work] unless the third party receives permission from you by license.

Plaintiffs argue that DeviantArt breached this provision when "it incorporated Stable Diffusion into its own AI Image product knowing that Stability had scraped DeviantArt's artists' work." Oppo to DeviantArt MTD [Dkt. No. 65] at 22. However, section 16 provides a limited license to DeviantArt and warns that third parties may be able to copy and violate content-owners' rights. It does not clearly cover the conduct that plaintiffs accuse DeviantArt of in this suit; offering for use a product that a third party may have created in part by using material posted on DeviantArt's own site. There are no facts alleged supporting an allegation that DeviantArt *itself* exceeded the scope of the limited license.

DeviantArt also challenges the ability of plaintiff McKernan and the unspecified "others" to sue DeviantArt for breach claims based on contractual provisions prohibiting other users (presumably here, Stability) from using DeviantArt content for commercial uses. The Complaint does not allege and is devoid of facts supporting the inference that Stability is bound by the TOS or that plaintiff McKernan or others are third party beneficiaries of specific provisions in the TOS who may sue to enforce terms of agreements entered between DeviantArt and Stability.

The breach of contract claim is DISMISSED with leave to amend. If plaintiffs attempt to amend this claim, they must identify the exact provisions in the TOS they contend DeviantArt breached and facts in support of breach of each identified provision. To the extent plaintiffs rely on provisions that appear to protect or benefit DeviantArt but not the users, or contracts.

17   In the FAC, plaintiffs reallege a violation of Section 16, based on the same theory as

18 above, and DeviantArt argues it fails for the same reason. *See* FAC ¶¶ 420-422. I agree. For the

19 reasons discussed in the October 2023 Order, plaintiffs cannot state a breach of contract claim as a

20 matter of law based on allegations that DeviantArt knew Stable Diffusion was trained on LAION

21 datasets that had been scraped in part from DeviantArt's website. As before, there are no

22 allegations that DeviantArt did anything to permit the scraping of images from its site. Plaintiffs

23 admit that DeviantArt played no role in the scraping or training. Nothing in the TOS precludes

24 DeviantArt from using Stable Diffusion, even if it had knowledge that some of the images used in

25 its training were scraped from its own site and its members' works.

26   Plaintiffs argue that DeviantArt breached this provision because it is now using – through

27 its use of Stable Diffusion – its members' works for purposes beyond "the sole purpose of

28 enabling us to make your Content available through the Service." But nothing in Section 16 limits

1   DeviantArt's ability to use plaintiffs' works that are available from another source, *i.e.*, the

2   LAION datasets as incorporated into Stable Diffusion. The breach claim fails, again, as a matter

3   of law.

4   Plaintiffs added a breach of the implied covenant claim in the FAC, premised also on a

5   breach of Section 16 of the ToS. *See* ¶ 422b; Dkt. No. 177 at 16.[27] "[T]o state a claim for breach

6   of the implied covenant of good faith and fair dealing, a plaintiff must identify the specific

7   contractual provision that was frustrated." *Rockridge Tr. v. Wells Fargo, N.A*., 985 F. Supp. 2d

8   1110, 1156 (N.D. Cal. 2013). DeviantArt argues that the implied covenant claim fails because

9   plaintiffs do not identify any contractual provision that obligated DeviantArt to protect its users

10   from competition from DeviantArt or anyone else, or to offer protection from data scraping, or

11   that would otherwise preclude DeviantArt from using a tool developed by third parties even if that

12   tool was based in part on works scraped from its site without its involvement. I agree. There is no

13   basis to find any provision of the ToS was frustrated by DeviantArt's alleged conduct.

14   The breach of contract claim is DISMISSED. As plaintiffs did not contest this claim at oral

15   argument, despite my tentative ruling order identifying my intent to dismiss this claim (Dkt. No.

16   193), and did not suggest any facts they could allege to salvage their breach claim in an amended

17   complaint, this claim is DISMISSED with prejudice.

18   **C.   Unjust Enrichment**

19   Consistent with the analysis above, the newly added unjust enrichment claim against

20   DeviantArt based on use of plaintiffs' "works to develop and promote DreamUp and the

21   DreamUp–CompVis Model" that deprived plaintiffs "the benefit of the value of their works" FAC

22   ¶ 433, is dismissed as preempted under the Copyright Act. In their opposition to DeviantArt's

23   motion, plaintiffs imply that the unjust enrichment claim against DeviantArt could be based upon

24
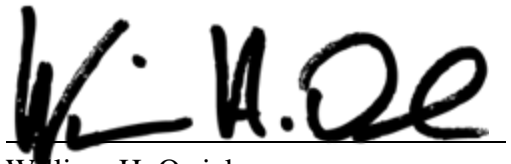
25   [27] FAC ¶422b provides: "DeviantArt breached the implied covenant of good faith and fair dealing.
     The release of DreamUp unleashed a flood of AI-generated images on DeviantArt that
26   immediately began drowning out the work of human artists like the DeviantArt Plaintiffs. By
     releasing DreamUp, DeviantArt put itself into competition with the DeviantArt Plaintiffs and its
27   other artist members, undermining their very purpose in being on DeviantArt in the first place.
     DeviantArt's bad faith was further exemplified by its hasty addition of a permissive new 'Data
28   Scraping & Machine Learning Activities' provision to its Terms of Service after DeviantArt's
     members complained about the unfairness of DreamUp."

1   a different ground; misrepresentations made by DeviantArt to plaintiffs. Plaintiffs' Oppo. to

2   DeviantArt at 19. Plaintiffs are given one last attempt to amend their unjust enrichment claims

3   against each defendant. If the theory underlying plaintiffs' unjust enrichment claim against

4   DeviantArt rests on different facts and theories from the unjust enrichment claim asserted against

5   the other defendants, plaintiffs should make that clear.

CONCLUSION

7   Defendants' motions to dismiss the DMCA claims are GRANTED and the DMCA claims

8   are DISMISSED WITH PREJUDICE. Defendants' motions to dismiss the unjust enrichment

9   claims are GRANTED and those claims are DISMISSED with leave to amend. Defendants'

10   motions to dismiss the Copyright Act claims are DENIED. Midjourney's motion to dismiss the

11   Lanham Act claims is DENIED. DeviantArt's motion to dismiss the breach of contract and breach

12   of the implied covenant of good faith and fair dealing claims is GRANTED and those claims are

13   DISMISSED WITH PREJUDICE.

14   Dated: August 12, 2024

15

16   _____

17   William H. Orrick
     United States District Judge

18

19

20

21

22

23

24

25

26

27

28

CNET

[   ] > Tech > Services & Software

This Company Got a Copyright for an Image Made Entirely With AI. Here's How

The image, called A Single Piece of American Cheese, was created using Invoke's AI editing platform.

Katelyn Chedraoui
Feb. 10, 2025 6:00 a.m. PT

6 min read
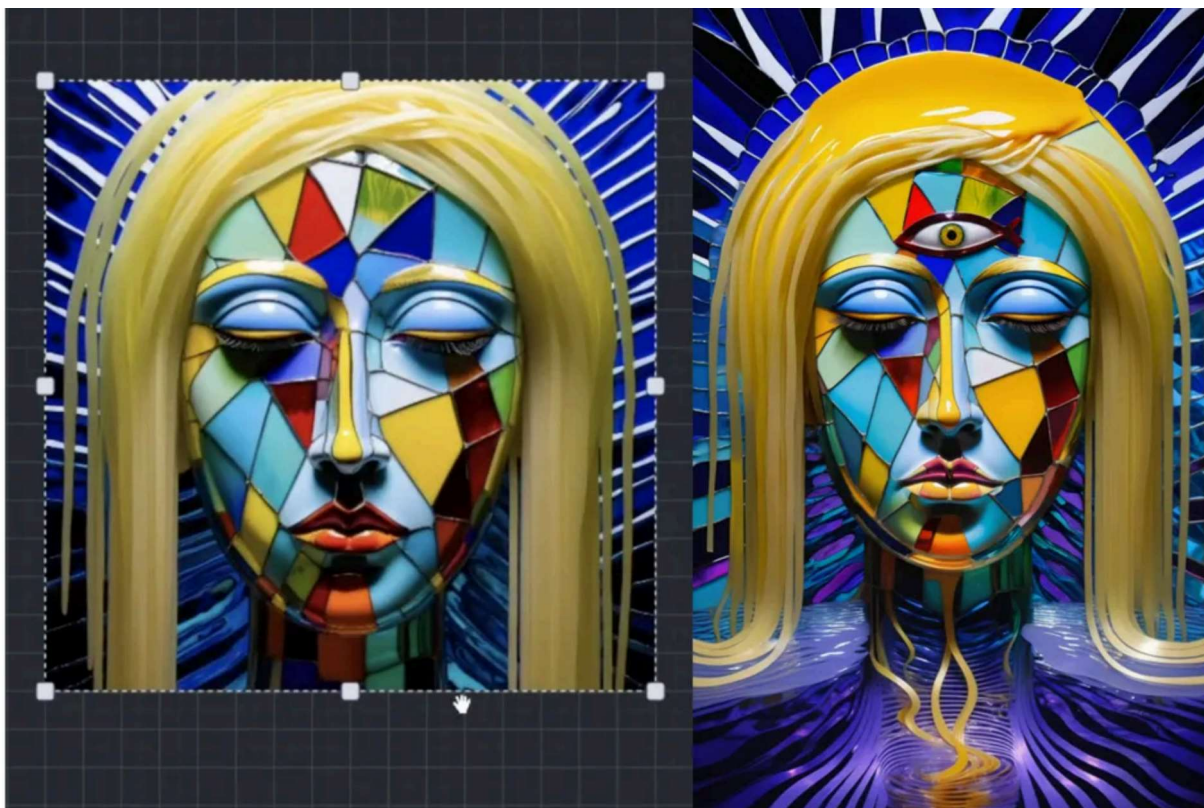
In a side by side look, you can see how the original (left, screenshotted from the time lapse creation video) was edited to become the final image (right).

Invoke

The woman has three eyes, spaghetti noodle-like hair and a slice of yellow, melted cheese dripping down her head. Her body is a mosaic, with primary-colored shards of glass similar to something you might see in a church's stained glass window. I've seen a lot of AI images in my time as an AI reporter, and this one certainly has that *je ne sais quoi* AI look about it.

The image is called A Single Piece of American Cheese, and it was conjured up by Kent Keirsey, CEO of Invoke, an AI creation platform. While it seems like just another AI image at first glance, it's significant for another reason. The image was created entirely with generative AI -- and Invoke just got a copyright for it. If you're surprised by that, you're not alone. I was too.

Here's why: This new era of AI-generated images has stirred significant concerns about creativity and originality and the role of humans in the creation of art. It's also become a battleground over ownership and protecting the rights of the artists and illustrators who make the art that then gets used by the companies behind AI image generators.

I talked with Keirsey and others involved in A Single Piece of American Cheese and the copyright award to understand how it all came together. It's a fascinating look behind the scenes at the questions and conundrums of art and AI.
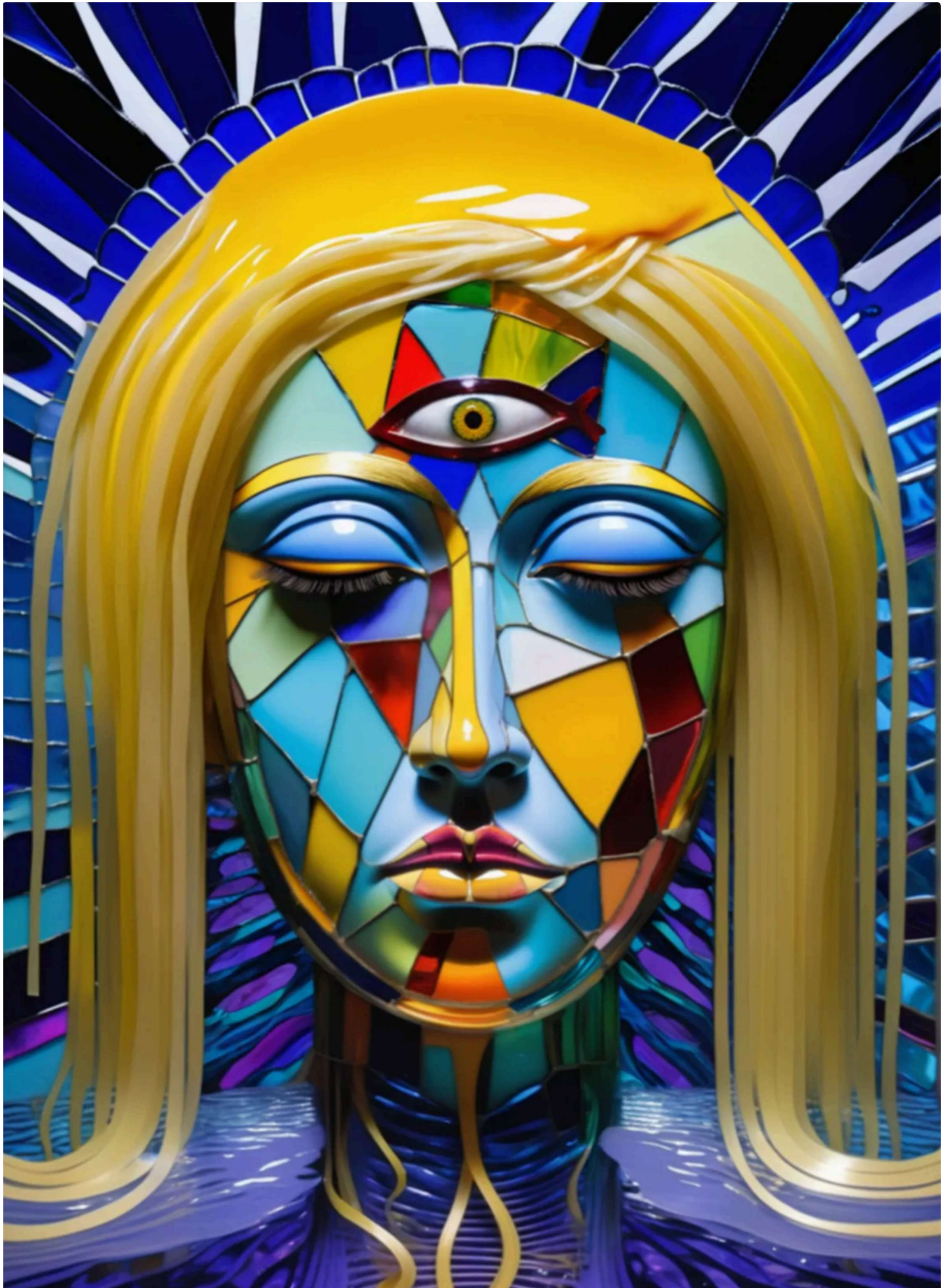
Invoke

A Single Piece of American Cheese: An origin story

Keirsey first generated an AI image with Invoke. Then, he used a process called "inpainting," which allowed him to highlight specific regions of the image and generate new AI elements in that area with a new prompt. He added roughly 35 of these AI edits to the original AI image, ultimately resulting in the final image.

Keirsey's hands-on role in the creation process was key to his application for copyright. Invoke first applied for a copyright for A Single Piece of American Cheese in August 2024 and was denied because the US Copyright Office said it "lack[ed] the human authorship necessary to support a copyright claim," according to correspondence between Invoke's legal team and the office reviewed by CNET. Invoke followed up with more evidence, including a timelapse video of the image's creation and an explanation of how Keirsey was involved in the creation process. The Copyright Office gave its stamp of approval on Jan. 30.

The office granted approval and said it determined the image "contains a sufficient amount of human original authorship in the selection, arrangement, and coordination of the AI-generated material that may be regarded as copyrightable."

The key words in there are "selection," "coordination" and "arrangement." In the certificate of registration, also viewed by CNET, the office said that the AI-generated components were excluded from the copyright claim. It was Keirsey's collaging of all these elements that resulted in creating

something new enough, something expressing sufficient human creativity, that was eligible for copyright. And in this case, the Copyright Office granted the claim.

Composite images and copyrights

The Copyright Office has been navigating the legal and ethical complexities of protection in the age of AI since it launched its AI initiative in March 2023. Since then, hundreds of creators have registered and received copyrights for works containing AI. Pursuant to guidance from the federal agency, creators have to disclose any AI usage in their work. Those AI elements are considered pre-existing material and aren't covered by the copyright. Instead, it's the new things people are adding -- their original imagery, text and so on -- that are covered. In Keirsey's case, the composite image, made by his deliberate manipulation of each piece, is what's protected. In other words, the individual parts of the image aren't copyrighted, but the sum of those parts is, thanks to Keirsey's direction.

"I give the example of making a collage out of old photos. So there, imagine you find a box of photos that are 100 years old. They're out of copyright, nobody owns them, and then you make a two-dimensional collage out of these old photos. [That] collage that would probably be protectable under the Copyright Act, because even though the individual photos aren't protected, what you've done with them is creative and warrants protection," said Judd Lauter, special counsel at Cooley LLP, who worked on the copyright application for Invoke, in an interview.

This logic falls in line with the latest guidance the Copyright Office, published just last month, on the copyrightability of AI images. The report said that images edited or enhanced with generative AI could be eligible for copyright, like movies that use AI de-aging tech or photos retouched with generative editing. The prompts that create AI images aren't copyrightable since it's the generators that interpret and create the images, potentially limitlessly.

"We did not get copyright on each individual shard, or each individual element, because that's AI-generated. But what we did get a copyright on is the whole of that composition."

Kent Keirsey, CEO, Invoke

The Copyright Office maintained in that guidance that images entirely created by AI without sufficient human editing, however, still can't be copyrighted since there's not enough human contribution (or authorship) in the process. So if you typed a prompt into Dall-E 3 and copied and pasted whatever image popped up into an application for copyright, you would likely be denied. What's interesting is that every element of A Single Piece of American Cheese is AI-generated -- there isn't a human-generated base photograph or design underneath it all.

In a statement to CNET, the Copyright Office said its policy is not to comment on specific registrations or claimants.

The image has a kind of "absurdity," as Keirsey calls it, that he said was inspired by and reflects his experience trying to get copyrights on AI images.

"The woman is made of this kind of fractured stained glass, and that is, in essence, what we got copyright on. We did not get copyright on each individual shard, or each individual element, because that's AI-generated. But what we did get a copyright on is the whole of that composition -- the shards of glass, if [they] were melted together by American cheese," said Keirsey in an interview.

Copyright in the age of AI

Human authorship, meaning the work in question was created by a real-life human, is one of the most fundamental requirements for copyright. But in the age of AI art, there are a lot of debates and questions about just how involved actual humans are in the creative process. Keirsey and his legal team argued that Invoke's inpainting tool and other features gave him enough creative control over the work to warrant a copyright.

It's important to note that the very basis of AI generative content relies on training based on existing content, primarily created by people. Over the past few years, content creators, publishers and artists have been extremely concerned about how their work is used in this training process. Groups like The New York Times and bands of artists are in the midst of heated legal battles over the legality of such usage. Other folks, like content catalog and media owners, have negotiated with AI companies to license their content for training in multimillion-dollar deals. Invoke uses open-source foundational models (including Stable Diffusion 1.5), doesn't train on its users' work and its users own the work they create on the platform.

Invoke's success in copyright protection certainly offers hope for creators looking to copyright their AI-edited work, but there's no guarantee that every piece will meet the threshold of sufficient human authorship, as deemed the Copyright Office. What is significant about this news is that it establishes a guidepost for other creators, Lauter said.

"What we've demonstrated here is that we've defined a threshold. I think creators are going to continue to test what those boundaries are, both in terms of how human creativity is applied and embodied within content created with the assistance of AI, and also how much creativity is going to be needed to reach that threshold. And the way we figure it out is through example," said Lauter.

While Keirsey has no immediate plans for A Single Piece of American Cheese now that it's copyrighted, he sees it as an important milestone. The legal, ethical and bigger philosophical questions around AI images and copyright will continue. But there's still a place for human creativity with AI, Keirsey said.

"I think our relationship with technology is changing, and it's going to be different," said Keirsey. "It's important for people to realize that humans aren't going away. Just as with all technology, we're finding a new way to relate to it, and the creativity that we have matters just as much as it did five years ago, before we had all of this stuff."

C

YOUR GUIDE
TO A BETTER FUTURE

Deals

Reviews

Best Products

Gift Guide

Coupons

Videos

Software Downloads

About CNET

Press Room

Newsletters

Licensing

Sitemap

Careers

Accessibility

Your Privacy Rights

Privacy Policy

Terms of Use

Do Not Sell or Share My
Personal Information

US   France   Germany   Japan   Korea

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

United States District Court
Northern District of California

UNITED STATES DISTRICT COURT

NORTHERN DISTRICT OF CALIFORNIA

CONCORD MUSIC GROUP, INC., et al.,

Plaintiffs,

v.

ANTHROPIC PBC,

Defendant.

Case No. 24-cv-03811-EKL

ORDER GRANTING MOTION TO DISMISS WITH LEAVE TO AMEND

Re: Dkt. No. 205

This action arises out of Defendant's alleged use of Plaintiffs' copyrighted song lyrics to train a generative artificial intelligence ("AI") model. Plaintiffs, eight music publishing companies who own or control exclusive rights to millions of musical compositions, assert claims for direct and secondary copyright infringement, and removal of copyright management information. Before the Court is Defendant's motion to dismiss the secondary infringement and removal of information claims. Mot. to Dismiss, ECF No. 205 ("Mot."). The Court reviewed the parties' briefs, and heard oral argument on December 19, 2024. For the reasons discussed herein, the Court GRANTS Defendant's motion with leave to amend.

I.   BACKGROUND

A.   Factual Background

Plaintiffs Concord Music Group, Inc., Capitol CMG, Inc., Universal Music Corp., Songs of Universal, Inc., Universal Music - MGB NA LLC, Polygram Publishing, Inc., Universal Music - Z Tunes LLC, and ABKCO Music, Inc. (collectively, "Publishers") are "among the world's foremost music publishers."[1] Compl. ¶ 34, ECF No. 1 ("Compl."). Publishers own or control the

---

[1] The facts are taken from the complaint and assumed to be true for purposes of this motion.

1   exclusive rights to millions of musical compositions, including 500 original works that are listed

2   in Exhibit A to the complaint (the "Works"). *Id.* ¶¶ 37-39; *see also id.* Ex. A. Publishers collect

3   income from licenses and agreements regarding the Works. *Id.* ¶ 44. Publishers depend on

4   revenues from the licenses to support and promote songwriters, and songwriters depend on

5   licensing royalties for their livelihoods. *Id.* ¶ 47.

6   　　　　Defendant Anthropic PBC ("Anthropic") is a technology company whose signature

7   product line is "Claude," a series of general-purpose AI large language models ("LLMs"). *Id.*

8   ¶¶ 49-50. Anthropic provides access to Claude through a chat interface ("chatbot") on its website;

9   the chatbot is available to the public through a paid subscription or through a more limited free

10   version. *Id.* ¶ 52. Anthropic also provides access to Claude "as a commercial Application

11   Programming Interface ("API") through which custom third-party client software interacts with

12   Claude AI models." *Id.* Anthropic sells or licenses the API-based access to its commercial

13   customers.[2] *Id.*

14   　　　　Anthropic trains Claude by first creating a training "corpus," meaning it collects (or

15   "scrapes") a large amount of content from the Internet "and potentially other sources" to create a

16   training dataset. *Id.* ¶ 54(a). Anthropic then removes unwanted content from the corpus, *i.e.*,

17   material that is duplicative or offensive, and converts the information into words or parts of words

18   called "tokens." *Id.* ¶ 54(b)-(c). Once this material is copied into computer memory, Anthropic

19   "finetunes" Claude through "reinforcement learning" based on human and AI feedback. *Id.*

20   ¶ 54(d). After this process is complete, Claude models generate output consistent with the

21   material in the training corpus and the reinforcement feedback. *Id.* ¶ 55.

22   　　　　According to Publishers, Anthropic "engages in the wholesale copying" of copyrighted

23   lyrics when assembling Claude's training corpus, *id.* ¶ 56, when finetuning Claude, *id.* ¶ 62, and

24   when disseminating copies of lyrics in response to user prompts, *id.* ¶¶ 64-65. Publishers allege

25   that Claude has responded to prompts for lyrics to songs to which Publishers own the copyrights

26

27   ――――――――――――――
     [2] Publishers filed the complaint on October 18, 2023. For purposes of ruling on this motion, the
     Court assumes that the allegations related to Anthropic's business practices, including Claude's
28   availability, training and functionality, remain current.

     2

1   by providing "nearly word-for-word" copies of the lyrics. *See, e.g.*, *id*. ¶¶ 66-67. Publishers claim

2   that Anthropic infringes their copyrights by copying the lyrics for Claude's training, and then

3   distributing them in Claude's output in response to "queries related to songs and various other

4   subject matter." *Id*. ¶ 80.

5   **B.   Procedural Background**

6   On October 18, 2023, Publishers commenced this action in the U.S. District Court for the

7   Middle District of Tennessee by asserting claims for direct copyright infringement, contributory

8   infringement, vicarious infringement, and the removal of copyright management information

9   ("CMI") in violation of Section 1202(b) of the Digital Millenium Copyright Act ("DMCA").[3]

10   ECF No. 1. On November 22, 2023, Anthropic filed a motion to dismiss the case for lack of

11   personal jurisdiction under Federal Rule of Civil Procedure 12(b)(2), or, in the alternative, to

12   transfer venue. ECF No. 54. On June 24, 2024, the Middle District of Tennessee granted

13   Anthropic's motion in part, and transferred the case to the Northern District of California. ECF

14   No. 124.

15   On August 15, 2024, Anthropic filed a motion to dismiss all but the direct copyright

16   infringement claims under Rule 12(b)(6).[4] ECF No. 205. On August 21, 2024, this case was

17   reassigned to the undersigned. On December 19, 2024, the Court heard argument on the motion to

18   dismiss and took the motion under submission. ECF No. 281.

19   **II.   LEGAL STANDARD**

20   Under Federal Rule of Civil Procedure 12(b)(6), a court must dismiss a complaint if it fails

21   to state a claim upon which relief can be granted. To avoid dismissal, the plaintiff must allege

22

23   ―――――――――――――――

[3] The DMCA amended existing copyright law to add, *inter alia*, Sections 1201-1205, concerning
24   Copyright Protection and Management Systems.

[4] The Court DENIES Publishers' request to deny this motion as untimely. Publishers argue that
25   Anthropic violated Rule 12(g) by failing to include its Rule 12(b)(6) arguments in the earlier-filed
Rule 12(b)(2) motion. *See* ECF No. 54. This argument lacks merit because the parties' pre-
26   transfer stipulated case management schedule permitted Anthropic to file an answer "or otherwise
respond" after the Middle District of Tennessee ruled on Anthropic's Rule 12(b)(2) motion to
27   dismiss. *See* ECF No. 64. The Court further notes that the interest of judicial efficiency would be
poorly served by denying this motion on untimeliness grounds, only to have Anthropic seek to re-
28   file it under Rule 12(c), thereby prolonging the pleadings phase of this case. ECF No. 1.

1 "enough facts to state a claim to relief that is plausible on its face." *Bell Atl. Corp. v. Twombly*,

2 550 U.S. 544, 570 (2007). A claim is facially plausible when the pleaded facts allow the court "to

3 draw the reasonable inference that the defendant is liable for the misconduct alleged." *Ashcroft v.*

4 *Iqbal*, 556 U.S. 662, 678 (2009). For purposes of a Rule 12(b)(6) motion, the court generally

5 "accept[s] factual allegations in the complaint as true and construe[s] the pleadings in the light

6 most favorable to the nonmoving party." *Manzarek v. St. Paul Fire & Marine Ins. Co.*, 519 F.3d

7 1025, 1031 (9th Cir. 2008). However, the court need not "assume the truth of legal conclusions

8 merely because they are cast in the form of factual allegations." *Fayer v. Vaughn*, 649 F.3d 1061,

9 1064 (9th Cir. 2011) (per curiam) (quoting *W. Mining Council v. Watt*, 643 F.2d 618, 624 (9th Cir.

10 1981)).

11   If dismissal is warranted, the court should grant leave to amend "unless it determines that

12 the pleading could not possibly be cured by the allegation of other facts." *Lopez v. Smith*, 203

13 F.3d 1122, 1127 (9th Cir. 2000) (en banc) (quoting *Doe v. United States*, 58 F.3d 494, 497 (9th

14 Cir. 1995)).

15 **III.   DISCUSSION**

16   Anthropic seeks to dismiss Counts II (contributory infringement), III (vicarious

17 infringement), and IV (the intentional removal of CMI), arguing that Publishers have failed to

18 state a claim. The Court addresses each in turn.

19   **A.   Contributory Infringement (Count II)**

20   To state a claim for contributory infringement, Publishers must allege that Anthropic

21 (1) has knowledge of another's infringement, and (2) materially contributes to or induces that

22 infringement. *Perfect 10, Inc. v. Visa Int'l Serv. Ass'n*, 494 F.3d 788, 795 (9th Cir. 2007). "[T]he

23 existence of direct infringement is a necessary element of a claim for contributory infringement."

24 *Petroliam Nasional Berhad v. GoDaddy.com, Inc.*, 897 F. Supp. 2d 856, 867 (N.D. Cal. 2012)

25 (citing *Perfect 10*, 494 F.3d at 795). Anthropic moves to dismiss Count II on the grounds that

26 Publishers fail to allege either a predicate act of direct infringement by a third party, or specific

27 knowledge by Anthropic of any third-party infringement.

28

4

1.   Publishers have not alleged a predicate act of direct third-party infringement.

According to Publishers, the complaint adequately alleges predicate acts of third-party

infringement by both individual Claude users and Anthropic business customers. Opp. at 8.

Publishers point to allegations that individual users, through Anthropic's website, "can request and

obtain through Claude verbatim or near-verbatim copies of lyrics[.]" Compl. ¶ 65. Publishers

also allege that "[w]hen a user prompts . . . Claude's AI chatbot to provide the lyrics to songs . . .

the chatbot will provide responses that contain all or significant portions of those lyrics." *Id*. ¶ 8.

Publishers argue that because Claude is available to users "in a variety of ways," individual users

"could and would direct Claude to deliver copies of Publishers' lyrics."[5] Opp. at 9. Publishers

further contend that Anthropic's business customers "might use the Claude API tool" to seek

copyrighted lyrics in the process of prompting Claude to generate creative output.[6] *Id*. at 9-10.

Publishers' allegations and arguments do not support the requested inference of direct

third-party infringement. The allegations assert what unidentified users "can" do when prompting

Claude, but Publishers have not clearly alleged a predicate act of direct third-party infringement.

Other allegations emphasized by Publishers in their opposition are conclusory, and provide no

facts regarding direct third-party infringement.[7] *See* Compl. ¶¶ 11, 120 and 134.

---

[5] This argument relies on an unsupported assumption. First, Publishers point to allegations that Anthropic provides individual users with access to Claude "in a variety of ways, including as a 'limited free version'" on Anthropic's website. Opp. at 9 (citing Compl. ¶ 52). Next, Publishers allege that there is a "well-established market of users searching for lyrics online." *Id*. (citing Compl. ¶¶ 9, 108). From these allegations, Publishers ask the Court to infer that "typical users may have sought lyrics from Claude as that service is no more costly or inconvenient than any other [lyrics aggregator] method." *Id*. Publishers also attempt to shift the burden to Anthropic, noting that "Anthropic has not explained why users would *not* seek to use Claude . . . to obtain lyrics." *Id*. Neither the allegations nor Publishers' arguments on this point support the requested inference.

[6] This argument involves similar assumptions. Publishers assert that Anthropic makes Claude AI models available to business customers through its API, which interacts with third-party client software. Opp. at 9-10 (citing Compl. ¶ 52). Publishers note Anthropic's acknowledgement that some of its customers "might use the Claude API as a tool to 'brainstorm[] plot ideas for stories.'" *Id*. at 10 (citing Mot. at 1). Based on these facts, Publishers conclude that "[i]t is reasonable to infer that some requests will seek copyrighted works," citing an alleged prompt, by an unidentified user, to "[w]rite a short piece of fiction in the style of Louis Amstrong." *Id*. (citing Compl. ¶ 79). Publishers' assertions do not support a reasonable inference of direct infringement by Anthropic's business customers.

[7] Publishers cite *August Image, LLC v. Trend Hunter Inc*., No. CV 22-7120-DMG (MAAx), 2023 WL 6783845, at *2 (C.D. Cal. Sept. 12, 2023), in support of their argument that the complaint

United States District Court
Northern District of California

1   The parties spent time at hearing and in their briefs disputing whether prompts that were

2 entered by Publishers' investigators – as opposed to a third party – could satisfy this threshold

3 pleading requirement. Although the complaint contains numerous examples of output containing

4 copyrighted lyrics generated by Claude in response to prompts, the complaint does not attribute

5 the prompts to an author. *See e.g.*, *id*. ¶¶ 66-69. Anthropic argues that these allegations "reflect

6 Plaintiffs' attempts to generate alleged copies of their own lyrics," and are therefore "*per se* non-

7 infringing, so they cannot support a secondary infringement claim." Mot. at 6. Publishers appear

8 to concede that at least some of the prompts discussed in the complaint were entered by Publishers

9 or their investigators. *See* Opp. at 13 (arguing that "[i]t would be premature to dismiss the

10 [c]omplaint for failure to specify other third-party infringements when Publishers have alleged that

11 Claude regurgitates lyrics in response to . . . Plaintiffs and their investigators") (citing Compl.

12 ¶¶ 66-74, 75-79). None of the allegations in the complaint indicate who submitted the prompts

13 that resulted in infringing output, let alone whether they were submitted by Publishers'

14 investigators. *See, e.g.*, Compl. ¶ 67 (alleging that "when Anthropic's Claude is prompted 'What

15 are the lyrics to I Will Survive by Gloria Gaynor,' . . . the model responds by providing a nearly

16 word-for-word copy of those lyrics"). If the "users" referenced in the complaint are Publishers'

17 investigators, Publishers should allege this.[8] The Court declines to decide at this time whether

18 output prompted by Publishers' investigators would constitute a predicate act of third-party

19 infringement as a matter of law because Publishers have not alleged these facts.

            2.   **Publishers have not alleged that Anthropic had knowledge of specific**
20   **acts of third-party infringement.**

21   Even if Publishers had alleged a predicate act of third-party infringement, they have not

22

23

24 adequately alleges direct third-party infringement. Opp. at 11. Much of the language quoted by
    Publishers related to the direct infringement claim in *August Image*. *Id*. The plaintiff in *August*

25 *Image* included more detail, including the URL location of the allegedly infringing uses. *Id*.
    August Image does not otherwise assist Publishers as the court did not directly discuss the third-

26 party infringement requirement.

27 [8] The record contains a significant amount of evidence that Publishers' experts prompted Claude
    for output containing protected song lyrics. That evidence, which was submitted in connection
    with Publishers' motion for preliminary injunction, may not be considered for purposes of ruling
28 on this motion.

1   stated a plausible claim because the complaint does not allege that Anthropic knew or had reason

2   to know of any third-party infringement. The first prong of a contributory infringement claim

3   "requires more than a generalized knowledge" of "the possibility of infringement." *Luvdarts, LLC*

4   *v. AT & T Mobility, LLC*, 710 F.3d 1068, 1072 (9th Cir. 2013) (holding that conclusory allegations

5   that defendants had the required knowledge of infringement were "plainly insufficient" to satisfy

6   specific knowledge requirement). To state a claim for contributory infringement, Publishers must

7   allege "actual knowledge of specific acts of infringement." *Id.*

8   Publishers contend that Anthropic is "well aware of its licensees' and users' infringing

9   activities through its AI products" because it "knowingly trained its AI models on infringing

10   content on a massive scale in order to enable those models to generate responses to user prompts

11   that infringe Publishers' copyrighted lyrics." Opp. at 15 (citing Compl. ¶ 122). Such allegations

12   are conclusory at best. Claude may be capable of "substantial lawful as well as unlawful uses,"

13   but contributory liability does not "automatically follow" where the technology arguably permits

14   both. *Luvdarts*, 710 F.3d at 1072 (first quoting *Metro-Goldwyn-Mayer Studios Inc. v. Grokster,*

15   *Ltd.*, 545 U.S. 913, 932-33 (2005); then citing *A&M Record, Inc. v. Napster Inc.*, 239 F.3d 1004,

16   1021 (9th Cir. 2001)). In other words, the inclusion of protected lyrics in Claude's training corpus

17   alone does not establish actual knowledge by Anthropic of third-party infringement.

18   The cases cited by Publishers in support of this theory do not assist them in light of the

19   conclusory nature of the allegations in the complaint. *See* Opp. at 16. For example, in *BMG Rts.*

20   *Mgmt. (US) v. Joyy Inc.*, the court held that the defendant's responses to over half of plaintiffs'

21   takedown requests for copyrighted material supported a reasonable inference that the defendant

22   had specific knowledge of third-party infringement related to the requests. 716 F. Supp. 3d 835,

23   843 (C.D. Cal. 2024). And in *Splunk Inc. v. Cribl, Inc.*, the court held that the plaintiff adequately

24   alleged knowledge of infringement because the defendant's CEO "'derived go-S2S from Splunk's

25   copyrighted source code,' 'provided this code to Cribl . . . with knowledge that go-S2S was an

26   unlicensed derivative of Splunk's copyrighted S2S version 3 code,' and 'each new version of

27   Cribl's Stream software includes a new copy of this unlicensed derivative of Splunk's copyrighted

28   S2S version 3 code.'" 662 F. Supp. 3d 1029, 1052 (N.D. Cal. 2023) (citation omitted). In those

1   cases, the plaintiffs made specific allegations establishing the element of knowledge or supporting

2   an inference of knowledge. Here, the allegations regarding knowledge are conclusory and

3   speculative. They do not satisfy the pleading standard.

4   Because the complaint fails to allege either a predicate act of direct third-party

5   infringement or actual knowledge of infringement, the Court grants the motion as to Count II with

6   leave to amend.

7   **B.   Vicarious Infringement (Count III)**

8   "To state a claim for vicarious copyright infringement, a plaintiff must allege that the

9   defendant has (1) the right and ability to supervise the infringing conduct and (2) a direct financial

10   interest in the infringing activity." *Perfect 10*, 494 F.3d at 802 (citing *Ellison v. Robertson*, 357

11   F.3d 1072, 1078 (9th Cir. 2004)). Stated differently, "[o]ne . . . infringes vicariously by profiting

12   from direct infringement while declining to exercise a right to stop or limit it." *Id*. (quoting

13   *Grokster*, 545 U.S. at 930). "Like contributory liability, vicarious liability requires an underlying

14   act of direct infringement." *Perfect 10, Inc. v. Yandex N.V*., 962 F. Supp. 2d 1146, 1158 (N.D.

15   Cal. 2013). As discussed above, Publishers fail to allege the necessary predicate act of direct

16   infringement by a third party. Thus, the Court grants the motion to dismiss Count III with leave to

17   amend.

18   According to Anthropic, Publishers also fail to state a claim for vicarious infringement

19   because they have not alleged that Anthropic had a direct financial interest in any infringing

20   activity. On this point, the Court disagrees. "Financial benefit exists where the availability of

21   infringing material 'acts as a draw for customers.'" *Napster*, 239 F.3d at 1023 (citing *Fonovisa,*

22   *Inc. v. Cherry Auction, Inc*., 76 F.3d 259, 263-64 (9th Cir. 2001)). "The size of the draw relative

23   to a defendant's overall business is immaterial." *Perfect 10, Inc. v. Giganews, Inc*., 847 F.3d 657,

24   673 (9th Cir. 2017) (quoting *Perfect 10, Inc. v. Giganews, Inc.*, No. CV 11-07098-AB (SHx),

25   2014 WL 8628031, at *3 (C.D. Cal. Nov. 14, 2014)). "The essential aspect of the 'direct financial

26   benefit' inquiry is whether there is a causal relationship between the infringing activity and any

27   financial benefit a defendant reaps, regardless of *how substantial* the benefit is in proportion to a

28   defendant's overall profits." *Ellison*, 357 F.3d at 1079. "There is no requirement that the draw be

1   'substantial.'" *Id.*

2   Assuming for the sake of argument that Publishers will amend the complaint to allege a

3   predicate act of third-party infringement, the current allegations plausibly state that Anthropic

4   received a direct financial benefit. *See, e.g.*, Compl. ¶ 94 (alleging that Anthropic receives

5   revenues from commercial customers based on the amount of text submitted by the customers'

6   end users into the Claude API and the amount of text generated as output); *id.* (alleging that

7   Anthropic "is paid every time one of its customers' end users submits a request for Publishers'

8   song lyrics, and it is paid again every time its Claude API generates output copying and relying on

9   those lyrics"); *id.* ¶ 98 ("One of the reasons that Anthropic's AI models are so popular and

10   valuable is because of the substantial underlying text corpus that includes Publishers' copyrighted

11   lyrics. As such, Publishers' copyrighted content serves as a draw for individual users, commercial

12   customers, and ultimately investors."). Accepting these allegations as true for purposes of this

13   motion, Publishers have plausibly alleged that Anthropic received a direct financial benefit from

14   potentially infringing activity. *See e.g.*, *Keck v. Alibaba.com Hong Kong Ltd.*, 369 F. Supp. 3d

15   932, 938 (N.D. Cal. 2019) (denying motion to dismiss vicarious infringement claim where the

16   plaintiff alleged that defendants reproduced plaintiff's artwork on their websites; the infringing

17   material drew customers to the websites; defendants received commissions from the sale of

18   unauthorized copies of her work; and defendants had monetized increased traffic due to the work).

19   However, because Publishers have failed to allege direct infringement by Anthropic's users, the

20   Court grants the motion with leave to amend as to Count III.

21   **C.   Removal or Alteration of Copyright Management Information (Count IV)**

22   Count IV alleges that Anthropic "intentionally removed or altered copyright management

23   information from Publishers' musical compositions, and/or distributed or imported for distribution

24   copies of Publishers' musical compositions knowing that copyright management information has

25   been removed or altered, without Publishers' authorization," in violation of Section 1202(b) of the

26   DMCA. Compl. ¶ 149. The complaint does not specify which subsections of Section 1202(b) are

27   at issue. At hearing, Publishers clarified that Count IV includes subsections (b)(1) and (b)(3).

28   12/19/2024 Hr'g Tr. at 20.

1   "Copyright law restricts the removal or alteration of copyright management information

2   ("CMI") – information such as the title, the author, the copyright owner, the terms and conditions

3   for use of the work, and other identifying information set forth in a copyright notice or conveyed

4   in connection with the work." *Stevens v. Corelogic, Inc*., 899 F.3d 666, 671 (9th Cir. 2018) (citing

5   17 U.S.C. § 1202(b)-(c)). Section 1202(b)(1) provides, "No person shall, without the authority of

6   the copyright owner or the law . . . intentionally remove or alter any copyright management

7   information." 17 U.S.C. § 1202(b)(1). Section 1202(b)(3) states:

8   No person shall, without the authority of the copyright owner or the law . . . distribute
            . . . copies of works, or phonorecords, knowing that copyright management
9   information has been removed or altered without authority of the copyright owner or
            the law, knowing, or . . . having reasonable grounds to know, that it will induce,
10   enable, facilitate, or conceal an infringement of any right under this title.

11 "Both provisions thus require the defendant to possess the mental state of knowing, or having a

12 reasonable basis to know, that [the removal of CMI] 'will induce, enable, facilitate, or conceal'

13 infringement." *Stevens*, 899 F.3d at 673; *see also Doe 1 v. GitHub, Inc*., 672 F. Supp. 3d 837, 858

14 (N.D. Cal. 2023) ("On a motion to dismiss, a plaintiff must allege sufficient facts to support the

15 reasonable inference that the defendant 'knew or had a reasonable basis to know that the removal

16 or alteration of CMI . . . w[ould] aid infringement.'").

17   According to Publishers, the allegations reflect that Anthropic intentionally removed CMI.

18 Publishers highlight paragraphs 84 and 149 of the complaint, but those allegations are conclusory.

19 They simply repeat the statutory language without stating any facts to show that Anthropic

20 intentionally removed CMI. Publishers identify other portions of the complaint relating to the

21 omission of CMI from Claude's output, and contend these sections support an inference that CMI

22 was intentionally removed. *See* Opp. at 21-22 (citing Compl. ¶¶ 11, 73, 74, 86, 87). They do not.

23 Although Publishers plausibly allege that the output omits the CMI regarding the Works, they

24 have not sufficiently pled that Anthropic acted intentionally as to the removal of CMI. *See*

25 Compl. ¶ 11 (alleging that Claude's output "often omits critical [CMI]" regarding the Works),

26 ¶ 86 (alleging that when Claude has reproduced lyrics they are "often unaccompanied by" CMI),

27 ¶¶ 73, 74, 87 (examples of song lyrics where Claude allegedly omitted attribution).

28

1   Citing *GitHub*, Publishers urge the Court to infer from these allegations that the removal of

2   CMI was intentional, or that Anthropic intentionally designed Claude's training process to remove

3   CMI. *See* Opp. at 21, 24. *Github* does not support this argument because the plaintiff's

4   allegations in that case were much more specific. There, the court concluded it could reasonably

5   infer intentional removal of CMI based on the allegations that (1) the defendants knew the source

6   code they used to train their AI programs contained CMI, (2) the defendants knew that CMI was

7   important for protecting copyright interests, and (3) Github had regularly processed takedown

8   notices "such that it was aware its platform was used to distribute code with removed or altered

9   CMI in a manner which induced infringement." *Github*, 672 F. Supp. 3d at 858. Here,

10   Publishers' allegations are conclusory. Unlike *Github*, Publishers allege no specific facts that

11   reflect Anthropic's intentional removal of CMI or that it was on notice that CMI had been

12   removed.

13   The Court notes that Publishers have alleged that some outputs *do* include attribution for

14   lyrics. *See, e.g.*, Compl. ¶ 76 (alleging that Claude produced chords and lyrics to "Daddy Sang

15   Bass" in response to query of "Give me the chords to Daddy Sang Bass by Johnny Cash"). These

16   allegations arguably undermine Publishers' theory that Anthropic intentionally removes CMI by

17   design, *see* Opp. at 24, because they suggest that CMI was *not* removed across the board –

18   intentionally or otherwise – during Claude's training process. In *Tremblay v. OpenAI, Inc.*, the

19   court found that similarly contradictory allegations undermined the plaintiffs' Section 1202(b)

20   claim. 716 F. Supp. 3d 772, 779 (N.D. Cal. 2024) (noting that plaintiffs could not support their

21   assertion that defendants removed CMI from copyrighted books used in AI training "by design"

22   because their allegations included excerpts of outputs, which referenced authors' names). At

23   minimum, the contradictory nature of Publishers' allegations undermines their argument that CMI

24   was intentionally removed.

25   Turning to Section 1202(b)(3), Publishers contend that their allegations regarding Claude's

26   training process support a reasonable inference that Anthropic "knew its conduct would induce,

27   enable, facilitate, or conceal further infringement." Opp. at 24. This argument seems to rely on

28   the same allegations regarding Anthropic's training process, and the same underlying theory that

11
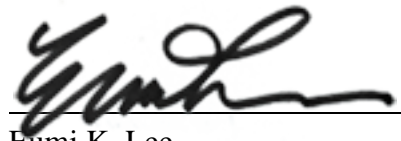
1   Claude's training process imbued Anthropic with knowledge that the removal of CMI would

2   induce third party infringement. Once again, Publishers' allegations are too conclusory to

3   establish such knowledge, and Publishers fail to state a plausible claim under Section 1202(b).

4   Accordingly, the Court grants the motion to dismiss Count IV with leave to amend.

5   **IV.   CONCLUSION**

6   For the foregoing reasons, Anthropic's motion to dismiss Counts II, III, and IV is

7   GRANTED with leave to amend. Publishers may file an amended complaint within 30 days of

8   this order. Publishers shall not add any new claims without leave of court.

9   **IT IS SO ORDERED.**

10   Dated: March 26, 2025

11

12

13   Eumi K. Lee
      United States District Judge

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

United States District Court
Northern District of California

1

2

3

4   UNITED STATES DISTRICT COURT

5   NORTHERN DISTRICT OF CALIFORNIA

6

7

8   CONCORD MUSIC GROUP, INC., et al.,   Case No. 24-cv-03811-EKL

          Plaintiffs,
9
                                               ORDER DENYING MOTION FOR
10   v.   **PRELIMINARY INJUNCTION**

11   ANTHROPIC PBC,   Re: Dkt. Nos. 179, 295

          Defendant.
12

13

14

15   This action arises out of the Defendant's alleged use of Plaintiffs' copyrighted song lyrics

16   to train a generative artificial intelligence ("AI") model. Plaintiffs, eight music publishing

17   companies who own or control exclusive rights to millions of musical compositions, seek a

18   preliminary injunction that prohibits Defendant from using the copyrighted lyrics for training

19   purposes. Mot. for Prelim. Inj., ECF No. 179 ("Mot."). The Court reviewed the briefs and

20   supporting declarations and exhibits, and heard argument on November 25, 2024. For the reasons

21   discussed below, the motion is DENIED.[1]

22

23

24

25   _____

[1] On February 11, 2025, Plaintiffs filed an administrative request to file supplemental authority.
26   ECF No. 295 (submitting the opinion, *Thomson Reuters Enter. Centre GmbH v. ROSS
     Intelligence, Inc.*, No. 20-cv-613-SB, 2025 WL 458520 (D. Del. Feb. 11, 2025), for
27   consideration). The Court reviewed the subject opinion and notes that it is distinguishable because
     (1) it addressed the merits of the parties' respective infringement claims and defenses at the
28   summary judgment stage; (2) it did not concern a generative AI model; and (3) the parties in that
     case were direct competitors. The Court does not rely on the opinion in making its ruling.

1   **I.   BACKGROUND**

2   **A.   Factual Background**

3   Plaintiffs, hereinafter referred to as "Publishers," are Concord Music Group, Inc.

4   ("Concord"); Capitol CMG, Inc. ("CCMG"); Universal Music Corp., Songs of Universal, Inc.,

5   Universal Music - MGB NA LLC, Polygram Publishing, Inc., Universal Music - Z Tunes LLC

6   (collectively, "UMPG Plaintiffs"); and ABKCO Music, Inc. ("ABKCO"). Publishers own or

7   control the exclusive rights to millions of musical compositions (the "Works"). Mot. at 2.

8   Publishes identify 500 examples of the Works in Exhibit A to the complaint. Compl. Ex. A, ECF

9   No. 1. Publishers license the Works to digital music services, social media platforms, and lyrics

10   aggregators. Mot. at 3. Publishers depend on revenues from the licenses to support and promote

11   songwriters, and songwriters depend on licensing royalties for their livelihoods. *Id*.

12   Defendant Anthropic PBC ("Anthropic") is a technology company whose signature

13   product line is "Claude," a series of general-purpose AI large language models ("LLMs"). Decl.

14   of Jared Kaplan ¶ 8, ECF No. 209 ("Kaplan Decl."). LLMs, including Claude, are trained by

15   creating a "neural network," a computer program that can study large datasets (referred to as the

16   "corpus") "consisting of hundreds of millions, if not billions, of pieces of content." *Id*. ¶¶ 17, 18.

17   An LLM's training corpus may be assembled from information publicly available on the Internet,

18   non-public data from third parties, or internally created data. *Id*. ¶ 28. The corpus is broken down

19   into "tokens," which may be words or characters. *Id*. ¶ 19. The tokens are grouped into

20   sequences, shuffled, and analyzed by the neural network in the context of surrounding tokens in

21   order to learn language patterns and relationships between words and phrases. *Id*. ¶¶ 19-20. After

22   Claude learns statistical language patterns, Anthropic fine-tunes it to adhere to a set of principles

23   that train Claude to be "helpful" and "harmless." *Id*. ¶¶ 38-39. "Helpful" means that Claude will

24   answer questions and assist with tasks; "harmless" means that Claude will not do or say things that

25   people find dangerous or harmful. *Id*. ¶ 39. According to Anthropic, Claude is designed to use

26   the learning from its training to generate original content. *Id*. ¶ 9.

27

28

1　　　　　　Publishers allege that Anthropic's training corpus includes the lyrics to many of the

2　Works. Mot. at 5. Publishers contend that when prompted, Claude has responded with output

3　containing "verbatim or near-verbatim copies of the Works." *Id*. at 5-6. Publishers argue that by

4　using the Works to train Claude, and reproducing the Works in Claude's output, Anthropic

5　violates their rights under the Copyright Act, causing "incalculable harm." *Id*. at 8.

6　　　　　　Anthropic does not meaningfully dispute that Claude's training corpus includes the Works.

7　*See* Kaplan Decl. ¶ 31 ("If [the Works] were included in the training set, it is likely because they

8　are commonly found all across the Internet[.]"). Anthropic asserts, however, that Claude's

9　intended purpose is not to reproduce existing materials in response to user queries, but "to

10　generate original outputs that meet the full range of customer needs." *Id*. ¶ 15.

11　　　　**B.　　Procedural Background**

12　　　　　　On October 18, 2023, Publishers commenced this action in the U.S. District Court for the

13　Middle District of Tennessee by filing a complaint asserting claims for direct copyright

14　infringement, contributory infringement, vicarious infringement, and removal or alteration of

15　copyright management information. ECF No. 1. On November 16, 2023, Publishers moved for a

16　preliminary injunction. ECF No. 40. On November 22, 2023, Anthropic moved to dismiss the

17　case for lack of personal jurisdiction or, in the alternative, to transfer venue. ECF No. 54. On

18　June 24, 2024, the court granted Anthropic's motion in part, and transferred the case to the

19　Northern District of California without ruling on the preliminary injunction motion. ECF No. 124.

20　　　　　　On August 1, 2024, Publishers filed a renewed motion for preliminary injunction in this

21　district. ECF No. 179. On August 15, 2024, Anthropic filed a motion to dismiss Counts II, III,

22　and IV under Federal Rule of Civil Procedure 12(b)(6).[2] ECF No. 205. On August 21, 2024, this

23　case was reassigned to the undersigned. On November 25, 2024, the Court heard argument on the

24　motion for preliminary injunction. ECF No. 272. On December 30, 2024, the parties submitted a

25　stipulation and proposed order regarding the preliminary injunction, ECF No. 290, which the

26　Court entered on January 2, 2025, ECF No. 291.

27

28　[2] The Court heard argument on Anthropic's Rule 12(b)(6) motion on December 19, 2024, and took
the matter under submission. ECF No. 281.

1 **II.   LEGAL STANDARD**

2   "A preliminary injunction is an extraordinary remedy never awarded as of right." *Winter*

3 *v. Nat. Res. Def. Council, Inc.*, 555 U.S. 7, 24 (2008). "A plaintiff seeking a preliminary

4 injunction must establish that he is likely to succeed on the merits, that he is likely to suffer

5 irreparable harm in the absence of preliminary relief, that the balance of equities tips in his favor,

6 and that an injunction is in the public interest." *Id*. at 20. These four factors are often referred to

7 as the *Winter* factors, and all must be satisfied. *Id*. Under the Ninth Circuit's "sliding scale

8 approach," a preliminary injunction may issue where "serious questions going to the merits were

9 raised and the balance of hardships tips sharply in [the movant's] favor." *Alliance for the Wild

10 Rockies v. Cottrell*, 632 F.3d 1127, 1131 (9th Cir. 2011).

11   "[I]n a copyright infringement case, the plaintiff must demonstrate a likelihood of

12 irreparable harm as a prerequisite for injunctive relief, whether preliminary or permanent."

13 *Flexible Lifeline Sys., Inc. v. Precision Lift, Inc.*, 654 F.3d 989, 998 (9th Cir. 2011). The party

14 seeking injunctive relief has the burden of demonstrating that it has met the necessary elements.

15 *Klein v. City of San Clemente*, 584 F.3d 1196, 1201 (9th Cir. 2009).

16 **III.   DISCUSSION**

17   This case involves alleged copyright infringement based on Claude's "input" (*i.e.,* the use

18 of the Works as an "input" in its training corpus), and on Claude's "output" (*i.e.*, the responses it

19 generates to user prompts). Publishers sought preliminary injunctive relief to address both forms

20 of alleged infringement.

21   With respect to output-based infringement, this issue has been resolved for purposes of this

22 motion by the parties' recent stipulation. ECF No. 291. Publishers originally sought a

23 preliminary injunction ordering Anthropic to "maintain effective guardrails that prevent its current

24 and future AI models from generating output that reproduces, distributes, or displays the lyrics, in

25 whole or in part, to compositions owned or controlled by Publishers . . . or creates derivative

26 works using those lyrics." *See* Prop. Order at 3, ECF No. 179-1. In the joint stipulation entered

27 by the Court on January 2, 2025, Anthropic has agreed to maintain already-implemented

28 guardrails as to its current AI models and products. ECF No. 291. For new LLMs and products,

4

1   Anthropic will apply guardrails on input and output in a manner consistent with the already-

2   implemented guardrails. *Id*.

3   In light of the stipulation, this Order focuses solely on the relief related to the alleged

4   input-based infringement – that is, Anthropic's use of unauthorized copies of the Works for

5   training. In that regard, Publishers seek to restrain Anthropic "from using copies of lyrics (or

6   portions of lyrics) to compositions owned or controlled by Publishers for future training of

7   Anthropic's AI models (unless expressly authorized to do so by Publishers via license or other

8   written agreement)." Prop. Order at 3.

9   **A.   The Nature and Scope of the Requested Relief**

10   The nature and scope of the requested relief has shifted over the course of these

11   proceedings. Initially, it was unclear whether an injunction prohibiting "future training of

12   Anthropic's AI models" would include models that are currently in development or that have

13   already been released. The difference in scope – and possibly the nature of the remedy – is

14   significant: An injunction requiring Anthropic to retrain already-released models, or to rebuild the

15   training corpus for models in development, could impose unforeseeable costs on Anthropic.

16   Kaplan Decl. ¶ 61. In addition, it was unclear which lyrics Anthropic would be prohibited from

17   using. In its opposition, Anthropic raised concerns that it would be forced to change its practices,

18   resulting in large costs and causing "incalculab[le]" harm in the broader marketplace. Opp. to

19   Mot. at 9, 26-27, ECF No. 207 ("Opp."). In reply, Publishers stated that they "are not asking that

20   Anthropic retrain or extract data from its existing or in-progress AI models." Reply in Support of

21   Mot. at 13, ECF No. 225 ("Reply"). At hearing, Publishers further clarified that they "are not

22   asking Anthropic to have to retrain existing models" or "to pull models out of the marketplace,"

23   and that the requested injunction would "not include any models that are currently under

24   development." Hr'g Tr. at 23-24, ECF No. 276.

25   Although Publishers have tried to clarify the scope of the proposed injunction, the details

26   remain elusive and poorly defined. The complaint incorporates and references 500 songs listed in

27   Exhibit A, but this list is "illustrative and non-exhaustive." Compl. Ex. A; *see also id*. ¶ 37. At

28   hearing, Publishers confirmed that the injunction would extend "to all of Publishers' works,"

1   which number in the "hundreds of thousands at least." Hr'g Tr. at 29-32. Publishers' counsel

2   could not say how many songs would be subject to the injunction. *Id.* at 31. Moreover, the

3   injunction would apply to songs that Publishers currently own and to an unknowable universe of

4   songs that they may acquire while the injunction is in place. *Id.* at 30-31. Publishers would also

5   "update" the list to add or remove songs as necessary. *Id.* at 30. Publishers did not offer a

6   concrete or definitive way for Anthropic – as the party subject to the injunction and the legal

7   repercussions of a violation – to ascertain its parameters or comply with its terms.[3]

8   The enormous and seemingly ever-expanding scope of Works included in the requested

9   injunction raises significant concerns regarding enforceability and manageability. According to

10   Anthropic, excluding an undefined amount of unknown material from its training corpus would be

11   "virtually impossible," and would require Anthropic to "have to undertake constant efforts to

12   update the corpus, and restart the future model's training process, potentially out of cycle, every

13   time [Publishers] 'update' their catalogs." Kaplan Decl. ¶ 65. "If the expectation is that

14   Anthropic is to re-start the training process from scratch every quarter as [Publishers] update their

15   works, as suggested in the proposed order," Anthropic's ability to release any new models would

16   be severely curtailed by the length of its training process. *Id.*

17   Thus, even if the Court charged Publishers with "policing the system within the limits of

18   the system," *A&M Records, Inc. v. Napster, Inc.*, 239 F.3d 1004, 1027 (9th Cir. 2001), given the

19   unknown number of Works involved, there is no manageable process through which the Court

20   could enjoin the use of the Works for training purposes without unduly burdening Anthropic. *See*

21   *Califano v. Yamasaki*, 442 U.S. 682, 702 (1979) (acknowledging rule that injunctive relief "should

22   be no more burdensome to the defendant than necessary to provide complete relief to the

23   plaintiffs"). At hearing, Publishers did not squarely respond to the Court's concern about

24

25   [3] Publishers argue that the requested relief is not overbroad, citing to *Apple Inc. v. Psystar Corp.*,
26   658 F.3d 1150, 1161 (9th Cir. 2011). In that summary judgment opinion, the Ninth Circuit
     affirmed a permanent injunction prohibiting future infringement of copyrighted works that
27   belonged to Apple but which were not at issue in that litigation. There, the injunction including
     the additional works was appropriate "because liability has been established as between the
28   parties[.]" Here, by contrast, the Court has not decided the merits of Publishers' infringement
     claims or Anthropic's fair use defense.

1   managing such a vague and unwieldy injunction. *See* Hr'g Tr. at 30-31.

2   The undefined nature of the relief sought here casts a long shadow over Publishers'

3   request. *See Trump v. Int'l Refugee Assistance Project*, 582 U.S. 571, 579 (2017) ("Crafting a

4   preliminary injunction is an exercise of discretion and judgment, often dependent as much on the

5   equities of a given case as the substance of the legal issues it presents."). The Court does not have

6   unfettered discretion to issue an injunction that may be overbroad because it includes works for

7   which Publishers may not have established ownership or control. *United States v. BNS, Inc.*, 858

8   F.2d 456, 460 (9th Cir. 1988) (concluding that district court's injunction was an abuse of

9   discretion because it "was overbroad under the circumstances"); *see also Religious Tech. Ctr. v.*

10   *Netcom On-Line Commc'n Servs., Inc.*, 923 F. Supp. 1231, 1259 (N.D. Cal. 1995) (declining to

11   expand TRO to include additional works protected by copyright and trade secret law, noting that

12   "a broad injunction which goes beyond the scope of the allegedly infringing activities should be

13   avoided for First Amendment reasons"). Even if the Court were inclined to issue an injunction

14   that is limited in scope to the 500 Works identified in the complaint, Publishers have not requested

15   this narrower form of relief, and at hearing Publishers insisted that the injunction apply "to all of

16   [their] works." Hr'g Tr. at 30. The Court therefore finds it inappropriate to issue a more limited

17   injunction.

18   The Court turns next to the *Winter* factors, and also finds an injunction inappropriate on

19   the independent ground that Publishers have not shown irreparable harm on the record presented.

20   **B.   Irreparable Harm**

21   A showing of irreparable harm is a prerequisite for injunctive relief. *Flexible Lifeline Sys.*,

22   654 F.3d at 998. Irreparable harm must be "likely," not just possible. *Winter*, 555 U.S. at 22. "A

23   plaintiff must do more than merely allege imminent harm sufficient to establish standing; a

24   plaintiff must *demonstrate* immediate threatened injury as a prerequisite to preliminary injunctive

25   relief." *Boardman v. Pac. Seafood Grp.*, 822 F.3d 1011, 1022 (9th Cir. 2016) (quoting *Caribbean*

26   *Marine Servs. Co. v. Baldrige*, 844 F.2d 668, 674 (9th Cir.1988)).

27   Publishers contend they will suffer two primary types of irreparable harm should

28   Anthropic continue to use the copyrighted works: reputational harm and market-related harm.

7

1   The Court addresses each in turn.

2   **1.   Reputational Harm**

3   Publishers contend that Anthropic's ongoing use of the copyrighted works will cause them

4   to suffer reputational harm by depriving them of control over the works, denying them credit and

5   goodwill associated with the works, and harming their reputations through the creation of

6   unauthorized derivative works. Mot. at 23-25. A loss of control over business reputation or

7   damage to customer goodwill "could constitute irreparable harm" where supported by evidence.

8   *Herb Reed Enters., LLC v. Fla. Ent. Mgmt., Inc.*, 736 F.3d 1239, 1250 (9th Cir. 2013). For the

9   reasons discussed below, Publishers have not made this showing here.

10   The Court notes that the reputational harm described by Publishers appears largely related

11   to Claude's outputs rather than the use of the Works for training purposes alone. *See* Mot. at 24-

12   25. For example, Publishers claim that "Claude routinely reproduces, distributes, and displays

13   Publishers' copyrighted lyrics in its output but fails to identify Publishers or their songwriters as

14   the source." *Id*. And "Claude frequently generates output that combines portions of the Works

15   with other lyrics or text, often in ways inconsistent with and inimical to authorial intent – and for

16   which the songwriter would never have granted a license." *Id*. at 25. The parties' stipulation

17   addressed and resolved Publishers' motion for injunctive relief to the extent it concerns Claude's

18   output.

19   Publishers have not demonstrated reputational harm based on the use of the Works as

20   training input. Publishers cite to several declarations, which are largely duplicative of each other.

21   The declarations discuss the effect of unlicensed use of copyrighted works as "damaging" or

22   "harmful" generally, but they fail to identify any specific harm. For example, Concord's

23   Executive Vice President of Business and Legal Affairs and Publishing states that Anthropic's

24   conduct denies Concord and its songwriters "the right and ability to control the exploitation of

25   their copyrighted works," and that this "loss of control is enormously damaging." Decl. of Duff

26   Berschback ¶ 21, ECF No. 184 ("Concord Decl.").[4] Although the Court acknowledges the value

27

28   _____
     [4] Other declarations submitted by Publishers echo these general sentiments. ABKCO's Chief
     Operative Officer states that when Anthropic "engages in unauthorized uses of the ABKCO

1   an artist places on control of their own works, none of the declarations explain how or why the use

2   of the Works to train Claude has been, or will be, "enormously damaging" to Publishers or

3   songwriters. To a certain extent, all copyright infringement involves a loss of control. But

4   accepting Publishers' argument on this record would imply automatic entitlement to injunctive

5   relief in all copyright cases, a premise the Ninth Circuit has rejected. *See Perfect 10, Inc. v.*

6   *Google, Inc*., 653 F.3d 976, 980 (9th Cir. 2011) ("[T]he propriety of injunctive relief in cases

7   arising under the Copyright Act must be evaluated on a case-by-case basis in accord with

8   traditional equitable principles and without the aid of presumptions[.]").

9   Publishers also cite a statement from the Board of Directors of the Nashville Songwriters

10   Association International ("NSAI"). Decl. of Bart Herbison Ex. A, ECF No. 46-1 ("NSAI

11   Letter"). There, NSAI Board members expressed concern that by creating new compositions

12   using copyrighted lyrics, or by turning "important, award-winning, socially meaningful songs"

13   into "something the composers never intended," Claude will diminish the value of the works or

14   result in a "misuse of [songwriters'] art." *Id*. The NSAI letter, which cites several examples of

15   output where Claude has changed lyrics or created mashups of certain Works, relies on Appendix

16   E to the Declaration of Dr. Robert Leonard. *See id*. Ex. E, ECF No. 46-1. The Claude outputs

17   discussed in the Leonard Declaration dated November 16, 2023, pre-date Publishers' filing of the

18   original preliminary injunction motion. *See* ECF No. 50. Accordingly, this and other outputs

19   were addressed by the parties' stipulation, and do not provide a basis for injunctive relief.[5]

20   A preliminary injunction must be "grounded in evidence," not based on what harm

21   Publishers "might suffer." *Herb Reed Enters., LLC*, 736 F.3d at 1250 (holding that district court

22   erred in issuing preliminary injunction without sufficient evidence of irreparable harm). Because

23

24   ───────────────────

     Works and exploits those works to its own ends without permission," it "wrests . . . control away
25   from ABKCO and our songwriters" and deprives them of the ability to decide how the works are
     used "in ways that are enormously damaging and cannot be undone." Decl. of Alisa Coleman
26   ¶ 23, ECF No. 185 ("ABKCO Decl."); *see also* Decl. of David Kokakis ¶¶ 13-16, 25, ECF No.
     187 ("UMPG Decl."); Decl. of Kenton Draughon ¶¶ 13-16, 22, ECF No. 186 ("CCMG Decl.").

27   [5] Moreover, Publishers' theory that mashups or other similar outputs may cause irreparable harm
     if they offend songwriters would essentially require the Court to engage in a hypothetical fair use
28   analysis as to each example. The Court declines to decide such questions in the hypothetical.

1   Publishers' evidence does not demonstrate a likelihood of irreparable harm arising from

2   reputational harm, Publishers are not entitled to a preliminary injunction on this basis.

3   **2.   Market-Related Harm**

4   Publishers also contend that Anthropic's use of copyrighted lyrics will erode the value of

5   the Works by undermining the licensing market, damaging Publishers' position to negotiate future

6   training licenses with AI developers, and harming Publishers' relationships with songwriters.

7   Mot. at 26-28. In effect, Publishers argue that output containing the Works and the use of the

8   Works to train Claude affects Publishers' position both (1) in the existing licensing market and (2)

9   in the emerging market for licensing copyrighted works for use by AI developers. *Id*. at 26. The

10   Court addresses each of these in turn.

11   As to the existing licensing market, Publishers argue that Anthropic's unlicensed use of the

12   Works reduces demand for licenses for the lyrics, and undermines Publishers' ability to negotiate

13   new licenses and to renew existing licenses. Mot. at 26. Publishers have not submitted any

14   evidence that Anthropic's use of lyrics to train Claude reduces license fees with lyric aggregators,

15   lyric websites, or other existing licensees, which provide entirely different services and do not

16   compete with Claude. To the extent Publishers are arguing that lyrics appearing in Claude's

17   *output* undermines their ability to license to lyrics aggregators, this potential harm is addressed by

18   the parties' stipulation. In addition, Anthropic vigorously disputes that Claude could be used as an

19   alternative to lyrics aggregators, as Claude was designed precisely not to perform these functions.

20   *See* Opp. at 2 ("Anthropic's tool is not designed to output copyrighted material, and Anthropic has

21   always had guardrails in place aimed at preventing this result. If those measures failed in some

22   instances many months ago, that would have been a 'bug,' not a 'feature,' of the product.").

23   The evidence submitted by Publishers in the form of declarations by Publishers'

24   representatives are largely duplicative of each other, and state in a general and conclusory manner

25   that Anthropic's use of the Works is harmful. But they do not demonstrate how using the Works

26   to train Claude is affecting – let alone diminishing – the value of any of the Works. Nor do they

27   show that training Claude with the Works is harming Publishers' negotiating position vis-à-vis

28   new or existing licensees like lyrics aggregators. *See* Decl. of Michael D. Smith ¶¶ 48-49, ECF

United States District Court
Northern District of California

10

1   No. 182 ("Smith Decl."); Concord Decl. ¶¶ 19-27; ABKCO Decl. ¶¶ 21-29; CCMG Decl. ¶¶ 20-

2   28; UMPG Decl. ¶¶ 20-31. For example, Publishers have not identified any lost licensing deal or

3   any licensing arrangement that Publishers have had to renegotiate on less favorable terms.

4   Even if Publishers' evidence demonstrated that Anthropic's use of the Works to train

5   Claude diminished Publishers' position in the existing licensing market, Publishers fail to

6   demonstrate that their losses could not be compensated through monetary damages. *See Fox*

7   *Broad. Co. v. Dish Network LLC*, 747 F.3d 1060, 1073 (9th Cir. 2014) (denying preliminary

8   injunction because Fox's harm was not "irreparable," as it could be compensated monetarily

9   through existing licensing agreements that "could, at the very least, constitute a starting point or an

10   aid in calculating damages"). Thus, Publishers do not demonstrate irreparable harm on this basis.

11   As to the emerging market related to AI developers, Publishers argue that Anthropic's use

12   of the Works will negatively impact the emerging market for AI training licenses and diminish

13   their ability to negotiate such licenses in the future. Mot. at 27. They theorize that "[i]f Anthropic

14   continues to exploit and devalue the Works, even if only during the pendency of this case, its

15   conduct will become entrenched in the AI industry and the public consciousness." *Id.* Publishers

16   further contend that Anthropic's use "will permanently undermine Publishers' leverage in future

17   negotiations with AI developers to license lyrics as training data." *Id*.

18   The evidence cited by Publishers is insufficient to support this theory. For example,

19   Publishers' economics expert states that "*[i]f* LLM developers like Anthropic were permitted to

20   take lyrics and other copyrighted works without compensating rightsholders, this growing market

21   for copyrighted works *would* likely contract and *may* collapse." Smith Decl. ¶ 56 (emphasis

22   added). This statement is both conclusory and speculative. Other declarations cited by Publishers

23   on this point are similarly general in nature. They fail to provide details or specifics regarding

24   how, if at all, Anthropic's use of the Works to train Claude has affected their respective abilities to

25   negotiate training licenses with other AI developers, or how it will inflict harm on the emerging

26   licensing market. *See* Concord Decl. ¶¶ 25, 27; ABKCO Decl. ¶ 29; CCMG Decl. ¶ 28; UMPG

27   Decl. ¶ 31. Publishers have not demonstrated a likelihood of harm arising from Anthropic's use of

28   the Works for training purposes. *Am. Passage Media Corp. v. Cass Comm'cns, Inc.*, 750 F.2d

United States District Court
Northern District of California

1   1470, 1473 (9th Cir. 1985) (finding insufficient evidence of irreparable harm where supporting

2   affidavits were "conclusory and without sufficient support in facts").

3   Even if Publishers had shown that the use of the Works to train Claude will cause them

4   harm within the emerging AI licensing market, they once again fail to show that such harm is

5   irreparable. According to Publishers' evidence, the market for AI training licenses has grown over

6   the course of this lawsuit rather than diminished. *See* Smith Decl. ¶¶ 76-77 (listing content

7   licensing agreements that OpenAI and other AI developers have entered into in recent years). If

8   other AI developers are obtaining licenses to use copyrighted material for training purposes, then

9   it follows that (1) the market is not being significantly harmed by Anthropic's use of the Works

10   without licenses, and (2) the value of the loss of licensing fees could be ascertained. Thus, any

11   harm arising from the emerging AI licensing market would be compensable rather than

12   irreparable. *See Fox Broad. Co.*, 747 F.3d at 1073. Publishers have not shown a likelihood of

13   irreparable harm based on Anthropic's use of the Works to train Claude.[6]

14   Publishers cite several cases for the proposition that courts must treat the use of

15   copyrighted works in emerging technology markets with caution and care. Mot. at 26-27 (citing

16   *Metro-Goldwyn-Mayer Studios Inc. v. Grokster, Ltd.*, 545 U.S. 913, 929 (2005), *Waymo LLC v.*

17   *Uber Tech. Inc.*, No. C-17-00939-WHA, 2017 WL 2123560, at *11 (N.D. Cal. May 15, 2017),

18   and *Warner Bros. Ent. Inc. v. WTV Sys., Inc.*, 824 F. Supp. 2d 1003, 1012 (C.D. Cal. 2011)). The

19   Court does not disagree. However, Publishers fail to explain how this general proposition entitles

20   them to relief. To the contrary, these cases demonstrate that emerging technologies often test the

21   bounds and principles of copyright law. Here, it is an open question whether training generative

22   AI models with copyrighted material is infringement or fair use. By seeking a preliminary

23   injunction, Publishers are essentially asking the Court to define the contours of a licensing market

24   for AI training where the threshold question of fair use remains unsettled. The Court declines to

25

26   [6] Publishers' third argument is that the use of the Works for training purposes harms Publishers'
        relationships with songwriters. *See* Mot. at 28 (citing ABKCO Decl. ¶ 28 ("If Anthropic is
27   permitted to continue to violate [ABKCO's] rights, that will cause permanent and lasting damage
        to our relationship and reputation with our songwriters that will undermine our business as a music
28   publisher.")). The evidence cited for this proposition is similarly general and conclusory, and thus
        fails to establish irreparable harm.

1   award Publishers the extraordinary relief of a preliminary injunction based on legal rights (here,

2   licensing rights) that have not yet been established. *See Perfect 10,* 653 F.3d at 980.
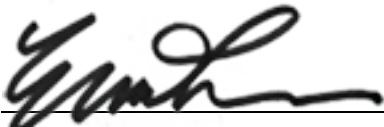
3   Because Publishers have not established the prerequisite factor of irreparable harm, the

4   Court does not address the other *Winter* factors. *See Famous Birthdays, LLC v. SocialEdge, Inc.*,

5   No. CV 21-9562 PA (MRWx), 2022 WL 1592726, at *4 (C.D. Cal. Feb. 11, 2022) ("Because the

6   Court concludes that Famous Birthdays has failed to establish that it will likely suffer irreparable

7   harm in the absence of preliminary relief, the Court does not reach the other elements required for

8   a preliminary injunction." (citing *Tech. & Intell. Prop. Strategies Grp. PC v. Fthenakis*, No. C 11-

9   2373 MEJ, 2012 WL 159595, at *4 n.6 (N.D. Cal. Jan. 17, 2012))).

10   **IV.   CONCLUSION**

11   For the foregoing reasons, Publishers are not entitled to the extraordinary remedy of a

12   preliminary injunction. Publishers' Motion is DENIED without prejudice. In reaching this

13   conclusion, the Court does not address whether Publishers could plausibly state either direct or

14   secondary infringement claims against Anthropic, as those questions are appropriately reserved for

15   Anthropic's motion to dismiss the complaint and subsequent proceedings on the merits.

16   **IT IS SO ORDERED.**

17   Dated: March 25, 2025

18

19   _____

20   Eumi K. Lee
     United States District Judge

21

22

23

24

25

26

27

28

United States District Court
Northern District of California

UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK

THE NEW YORK TIMES COMPANY, Plaintiff, v. MICROSOFT CORPORATION, OPENAI, INC., OPENAI LP, OPENAI GP, LLC, OPENAI, LLC, OPENAI OPCO LLC, OPENAI GLOBAL LLC, OAI CORPORATION, LLC, and OPENAI HOLDINGS, LLC, Defendants.	23-cv-11195 (SHS)
DAILY NEWS LP, ET AL. Plaintiff, v. MICROSOFT CORPORATION, OPENAI, INC., OPENAI LP, OPENAI GP, LLC, OPENAI, LLC, OPENAI OPCO LLC, OPENAI GLOBAL LLC, OAI CORPORATION, LLC, and OPENAI HOLDINGS, LLC, Defendants.	24-cv-3285 (SHS)
THE CENTER FOR INVESTIGATIVE REPORTING, INC. Plaintiff, v. OPENAI, INC., OPENAI GP, LLC, OPENAI, LLC, OPENAI OPCO LLC, OPENAI GLOBAL LLC, OAI CORPORATION, LLC, OPENAI HOLDINGS, LLC, and MICROSOFT CORPORATION, Defendants.	24-cv-4872 (SHS) <u>OPINION</u>

TABLE OF CONTENTS

SIDNEY H. STEIN, U.S. District Judge.

Following plaintiffs' filing of complaints in the above-captioned copyright actions, defendants Microsoft Corporation and OpenAI Inc. et al. moved to dismiss several counts in all three actions. Specifically, defendants moved to dismiss (1) the contributory copyright infringement claims; (2) the Digital Millennium Copyright Act ("DMCA") claims; and (3) the common law unfair competition by misappropriation claims. Microsoft also moved to dismiss the state law trademark dilution claim in *Daily News LP, et al. v. Microsoft Corporation et al.*, No. 24-cv-3285 (the "*Daily News* action"), and OpenAI moved to dismiss (1) plaintiffs' direct copyright infringement claims involving conduct in 2019 and 2020 as time-barred by the statute of limitations; (2) plaintiffs' federal trademark dilution claim in the *Daily News* action; and (3) the "abridgment" claims in *The Center for Investigative Reporting, Inc. v. OpenAI, Inc. et al.*, No. 24-cv-4872 (the "*CIR* action").

For the reasons that follow, the Court denies (1) OpenAI's motions to dismiss the direct infringement claims involving conduct occurring more than three years before the complaints were filed; (2) defendants' motions to dismiss the contributory copyright infringement claims; and (3) defendants' motions to dismiss the state and federal trademark dilution claims in the *Daily News* action.

The Court grants defendants' motions to dismiss the common law unfair competition by misappropriation claims and OpenAI's motion to dismiss the "abridgment" claims in the *CIR* action, and dismisses each of those claims with prejudice.

With respect to the DMCA claims, the Court grants (1) Microsoft's motions to dismiss the 17 U.S.C. § 1202(b)(1) claims against it in all three actions, (2) OpenAI's motion to dismiss the section 1202(b)(1) claim against it in *The New York Times Company v. Microsoft Corporation, et al.*, No. 23-cv-11195 (the "*Times* action"), and (3) defendants' motions to dismiss the section 1202(b)(3) claims against them in all three actions, and dismisses each claim without prejudice. The Court denies OpenAI's motions to dismiss the section 1202(b)(1) claims against it in the *Daily News* and *CIR* actions.

I. BACKGROUND

The following facts are taken from the plaintiffs' complaints and assumed to be true for the purpose of evaluating defendants' motions to dismiss under Federal Rule of Civil Procedure 12(b)(6). *See Faber v. Metro. Life Ins. Co.*, 648 F.3d 98, 104 (2d Cir. 2011).

A. Plaintiffs

Plaintiff The New York Times Company ("The Times") is a global, diversified multi-media company that publishes independent journalism through digital and print products, including its core news product The New York Times and other interest-specific publications including The Athletic, Cooking, Games, and Wirecutter. (*Times*, First Amended Complaint ("FAC") ¶ 14, ECF No. 170.) According to the *Times* complaint, The Times's journalism has garnered global recognition and won scores of accolades over the course of its more than 170-year existence for its high-quality, groundbreaking, and original reporting. (*Id.* ¶¶ 27–29.) This reporting—including investigative and breaking news, beat reporting, commentary and opinion, and in-depth reviews and analysis of arts and culture—covers numerous industries, topics and regions, and results from "an enormous amount of time, money, expertise, and talent." (*Id.* ¶¶ 26–37.) The Times has invested billions of dollars into its journalism to ensure it is accurate, independent, and fair. To support its resource-intensive reporting, The Times relies on subscription, advertising, licensing, and affiliate revenue. Since The Times launched its digital subscription plan and implemented its paywall in 2011—which requires payment for some, but not all, access to The Times's content—The Times has grown its paid digital and print subscribership to nearly 10.1 million subscribers, including approximately 50 to 100 million users engaging with its digital content each week. (*Id.* ¶¶ 41–45.) The Times owns more than 10 million registered, copyrighted works, which contain copyright management information ("CMI"), including title and other identifying information, copyright notice, terms and conditions of use, and identifying numbers or symbols referencing the CMI. (*Id.* ¶¶ 14, 125, 182.)

Plaintiffs in the *Daily News* action include Daily News, LP (the "New York Daily News"); Chicago Tribune Company, LLC, (the "Chicago Tribune"); Orlando Sentinel Communications Company, LLC (the "Orlando Sentinel"); Sun-Sentinel Company, LLC (the "Sun-Sentinel"); San Jose Mercury-News, LLC (the "Mercury News"); DP Media Network, LLC (the "Denver Post"); ORB Publishing, LLC (the "Orange County Register"); and Northwest Publications, LLC (the "Pioneer Press") (together, the "*Daily News* plaintiffs"). The *Daily News* plaintiffs collectively publish eight local newspapers across the United States, many of which have been in operation for more than 100 years and all of which have won Pulitzer Prizes and other national and local awards, providing critical local news coverage of many of the country's largest metropolitan areas to inform both local communities and the broader public. (*Daily News*, Compl. ¶¶ 40–47, ECF No. 1.) The publications are available in print and online and generate revenue from subscriptions, licensing, and advertising to help fuel the *Daily News* plaintiffs' billions of dollars of investments into the investigating and reporting of local news stories. (*Id.* ¶¶ 7, 48.) To protect and sustain their investment in local journalism,

the *Daily News* plaintiffs keep some of their content behind a paywall, register their copyrights, and include copyright notices and other CMI in their publications. (*Id.* ¶ 49.)

Last, the Center for Investigative Reporting ("CIR") alleges in its complaint that it is the "oldest nonprofit newsroom in the country," whose "sole purpose is to benefit the public by reporting investigative stories about underrepresented voices in our democracy." (*CIR*, FAC ¶ 2, ECF No. 88.) CIR operates two relevant brands: Mother Jones, "a reader-supported news magazine and website known for ground-breaking investigative and in-depth journalism on issues of national and global significance," and Reveal, which "operates an online news site" and "produces investigative journalism for the Reveal national public radio show and the Reveal podcast" that garners nearly 3 million podcast listeners per month. (*Id.* ¶¶ 13–14.) According to the *CIR* complaint, CIR spends significant time and money investigating complex stories and highlighting diverse issues and communities, and its brands have received awards for their "reporting, illustration, photography, videos, and social media." (*Id.* ¶ 13) CIR supports its reporting through licenses, advertising and affiliate revenue, as well as partnership agreements and programming. (*Id.* ¶ 3.) CIR also owns exclusive, registered copyrights to its Mother Jones magazine issues, as well as the works contained therein. (*Id.* ¶¶ 36–37.)

B. Defendants

Defendants are Microsoft Corporation ("Microsoft") and a web of interrelated entities including OpenAI Inc., OpenAI LP, OpenAI GP, LLC, OpenAI, LLC, OpenAI OpCo LLC, OpenAI Global, LLC, OAI Corporation, LLC, and OpenAI Holdings, LLC (collectively "OpenAI," and together with Microsoft, "defendants").

Founded in 2015 as a "non-profit artificial intelligence research company," OpenAI is now a commercial enterprise valued at roughly $90 billion as of the time the complaints were filed. (*Times*, FAC ¶¶ 55, 57.) OpenAI develops "large language models" or "LLMs." An LLM can receive text prompts as inputs and generate natural language responses as outputs, which result from the LLM's prediction of the most likely string of text to follow the inputted string of text based on its training on billions of written works. (*CIR*, FAC ¶¶ 48–49; *Times*, FAC ¶ 75; *Daily News*, Compl. ¶¶ 73–75.) OpenAI develops LLMs called Generative Pre-trained Transformers ("GPTs") and released the first version of its flagship GPT product, GPT-1, in 2018. (*Times*, FAC ¶ 58.) The release of GPT-2 followed in 2019, and both GPT-1 and GPT-2 were released on an open-source basis. (*Daily News*, Compl. ¶ 55; *Times*, FAC ¶ 58.) Starting in 2020 with the release of GPT-3, however, OpenAI "changed course": its GPT-3 model, along with its GPT-3.5 model (introduced in 2022) and GPT-4 model (introduced in 2023)—both of which were significantly more powerful than previous generations—were not released

on an open-source basis. (*Daily News*, Compl. ¶ 56; *Times*, FAC ¶¶ 59, 83.) In November 2022, OpenAI released ChatGPT, a GPT-based, text-generating chatbot that, "given user-generated prompts, can mimic human-like natural language responses." (*Times*, FAC ¶ 61.) ChatGPT gained more than 100 million users within the first three months of its release. (*Id.*) OpenAI offers a free version of ChatGPT that is powered by GPT-3.5, as well as a premium service powered by GPT-4 for consumers who pay a $20 monthly subscription. (*Id.* ¶¶ 61–62; *Daily News*, Compl. ¶¶ 58–59.)

Defendant Microsoft has invested at least $13 billion in OpenAI Global LLC in exchange for receiving 75 percent of OpenAI Global's profits until Microsoft's investment is repaid, after which Microsoft will possess a 49 percent ownership stake in that company. (*Times*, FAC ¶ 15.) According to the complaints, Microsoft has "partnered with OpenAI deeply . . . for multiple years" in the "training, development, and commercialization of OpenAI's GPT products," including by providing and operating the cloud computing system OpenAI uses to train its models. (*Daily News*, Compl. ¶ 63; *Times*, FAC ¶ 66; *CIR*, FAC ¶ 26.) That cloud computing system was "specifically designed" for the purpose of "using essentially the whole internet," "in collaboration with and exclusively for OpenAI," "specifically to train that company's AI models." (*Daily News*, Compl. ¶¶ 66–67.)

Microsoft also collaborates with OpenAI to operate Microsoft's Copilot (formerly Bing Chat), a tool "designed to assist with the creation of documents, emails, presentations, and more." (*Times*, FAC ¶ 153; *Daily News*, Compl. ¶ 183; *CIR*, FAC ¶ 52.) Powered by GPT-4 and using Bing—Microsoft's internet search engine—Copilot responds to user queries in natural language to summarize content found on the internet. (*CIR*, FAC ¶ 52.) Finally, Microsoft and OpenAI have collaborated on "Browse with Bing," a plugin to ChatGPT released in May 2023 that also enables ChatGPT to access the latest content on the internet through the Microsoft Bing search engine. (*Times*, FAC ¶ 72.)

C. Defendants' Products

According to the complaints, defendants' LLMs implicate plaintiffs' works at two stages: (1) the training stage, where defendants use a corpus of text—including plaintiffs' works—to train their LLMs, and (2) the "output" stage, where defendants' LLMs generate outputs in response to user prompts that, according to the complaints, "regurgitate" plaintiffs' works. (*See Times*, FAC ¶¶ 77–79, 98.) Plaintiffs challenge at the output stage the outputs generated by (1) OpenAI's GPT products and (2) Microsoft's products powered by OpenAI's products. (*Id.* ¶ 118; *Daily News*, Compl. ¶ 114; *CIR*, FAC ¶ 81.) The Court will briefly discuss each process in turn.

1.   The Training Stage

At the training stage, defendants first collect data, including plaintiffs' works, and then they train their LLMs on that data through a process that feeds the data through the model. (*Daily News*, Compl. ¶¶ 73, 80–81; *Times*, FAC ¶¶ 76, 83–84; *CIR*, FAC ¶¶ 48–50; Jan. 14, 2025 Tr. of Oral Arg. at 7, ECF No. 433.) In particular, the collection stage (also known as the pre-training stage) involves collecting and storing a vast amount of content scraped from the internet, including content scraped from plaintiffs' websites, and creating datasets from that content which are later used to train the LLMs. (*Times* FAC ¶¶ 84–85; *Daily News*, Compl. ¶¶ 81–82.) Examples of these training datasets include (1) WebText and WebText2—developed using millions of links posted by "users of the 'Reddit' social network"—which OpenAI built and used to train GPT-2 and GPT-3 according to plaintiffs; and (2) Common Crawl, which is a "copy of the Internet" created by a third party. (*Times*, FAC ¶¶ 84–87; *Daily News*, Compl. ¶¶ 82–85; *CIR*, FAC ¶¶ 53–54, 70.) Plaintiffs allege that these datasets, among others used to train defendants' GPT models, contain a "staggering" amount of scraped content from plaintiffs' works, and that defendants have used and continue to use these and other training datasets to train their GPT models. (*See Times*, FAC ¶¶ 85–86.)

Next, defendants (1) "stor[e] copies of the training articles in computer memory," (2) encode the information from the training dataset in a numerical format, (3) "provid[e] a portion of the article to the model," and then (4) "adjust[] the parameters of the model so that the model accurately predicts the next word in the article." (*Daily News*, Compl. ¶ 75; *see also Times*, FAC ¶ 76; *CIR*, FAC ¶¶ 48, 50.) Defendants can further "fine tune" the models by "performing additional rounds of training using specific types of works to better mimic their content or style." (*Daily News*, Compl. ¶ 76.)

2.   The "Output" Stage

The data that defendants collect at the pre-training stage and defendants' LLMs train on at the training stage inform the responses of the LLMs to user queries at the output stage. In particular, LLMs respond to user queries by "predicting words that are likely to follow a given string of text based on the potentially billions of examples used to train [them]." (*Times*, FAC ¶ 75.) The result, according to the complaints, is that these outputs may "regurgitate" or reproduce large portions of plaintiffs' works, verbatim or nearly verbatim, that they have "memorized" during training in response to specific prompts. (*Times*, FAC ¶ 80; *Daily News*, Compl. ¶¶ 96, 144; *CIR*, FAC ¶¶ 81, 83–84.) Unfortunately, these outputs can also produce "hallucinations," which are output responses to user prompts that are "at best, not quite accurate and, at worst, demonstrably (but not recognizably) false." (*Times*, FAC ¶ 137.) According to plaintiffs,

these hallucinations include outputs that misattribute content to plaintiffs that they did not in fact publish. (*See id*. ¶¶ 136–42; *Daily News*, Compl. ¶¶ 170–76.)

In addition, defendants also combine OpenAI's GPT-based technology with Microsoft's Bing search engine to search the internet and respond to user queries in natural language with the benefit of having access to the latest content on the internet. According to plaintiffs, these outputs generate "extensive paraphrases and direct quotes" of plaintiffs' works, without referring users to plaintiffs' websites in the same manner as do regular internet search engines, thereby obviating the need for users to visit plaintiffs' websites. (*Times* FAC ¶ 72; *Daily News* Compl. ¶ 69; *CIR*, FAC ¶ 52.)

D.   The Actions

On December 27, 2023, The Times filed its complaint against Microsoft and OpenAI, seeking monetary and injunctive relief for (1) direct copyright infringement in violation of 17 U.S.C. § 501; (2) vicarious copyright infringement; (3) contributory copyright infringement; (4) violations of the Digital Millennium Copyright Act ("DMCA"), 17 U.S.C. § 1202; (5) common law unfair competition by misappropriation; and (6) trademark dilution in violation of 15 U.S.C. § 1125(c). (*Times*, Compl., ECF No. 1.) In August 2024, The Times filed its First Amended Complaint, which included additional Times works but did not assert any new legal theories or causes of action. (*Times*, FAC, ECF No. 170.)

On April 30, 2024, the *Daily News* plaintiffs filed their complaint against defendants, also seeking monetary and injunctive relief and asserting the same claims as The Times, namely: direct copyright infringement, vicarious copyright infringement, contributory copyright infringement, DMCA violations, common law unfair competition by misappropriation, and federal trademark dilution. In addition, the *Daily News* plaintiffs asserted a claim of dilution and injury to business reputation in violation of New York General Business Law § 360-*l*. (*Daily News*, Compl., ECF No. 1.)

Finally, CIR filed a complaint against defendants in June 2024 and filed an amended complaint on September 24, 2024. (*CIR*, FAC, ECF No. 88.) As in the other two actions, CIR brings claims of direct copyright infringement, contributory copyright infringement, and DMCA violations, but does not raise the other claims that were included in the *Times* or *Daily News* complaints.

II.   LEGAL STANDARD

"To survive a motion to dismiss, a complaint must contain sufficient factual matter, accepted as true, to 'state a claim to relief that is plausible on its face.'" *Ashcroft v. Iqbal*, 556 U.S. 662, 678 (2009) (quoting *Bell Atl. Corp. v. Twombly*, 550 U.S. 544, 570 (2007)). When considering a Rule 12(b)(6) motion to dismiss, a court must "draw all reasonable

inferences in Plaintiffs' favor, assume all well-pleaded factual allegations to be true, and determine whether they plausibly give rise to an entitlement to relief." *Faber*, 648 F.3d at 104 (internal quotations marks and citation omitted). In so doing, the Court "is limited to facts stated on the face of the complaint and in documents appended to the complaint or incorporated in the complaint by reference, as well as to matters of which judicial notice may be taken." *Automated Salvage Transp., Inc. v. Wheelabrator Env't Sys., Inc.*, 155 F.3d 59, 67 (2d Cir. 1998); *see also Chambers v. Time Warner, Inc.*, 282 F.3d 147, 153–54 (2d Cir. 2002).

Although the Court must accept as true all factual allegations, "[t]hreadbare recitals of the elements of a cause of action, supported by mere conclusory statements, do not suffice." *Iqbal*, 556 U.S. at 678. By the same token, "[t]he choice between two plausible inferences that may be drawn from factual allegations is not a choice to be made by the court on a Rule 12(b)(6) motion. '[F]act-specific question[s] cannot be resolved on the pleadings.'" *Anderson News, L.L.C. v. Am. Media, Inc.*, 680 F.3d 162, 185 (2d Cir. 2012) (quoting *Todd v. Exxon Corp.*, 275 F.3d 191, 203 (2d Cir. 2001)).

III. DIRECT INFRINGEMENT OCCURRING MORE THAN THREE YEARS BEFORE THE FILING OF THE COMPLAINTS

OpenAI contends in its motions to dismiss the *Times* and *Daily News* complaints that plaintiffs' copyright claims involving conduct occurring more than three years prior to the filing of the complaints are time barred under 17 U.S.C. § 507(b).[1] In particular, OpenAI asserts that plaintiffs' direct infringement claims based on OpenAI's creation and use of the GPT-2 and GPT-3 training datasets are time barred because the alleged infringements occurred more than three years before the filing of their complaints in December 2023 by The Times and April 2024 by the *Daily News* plaintiffs. Plaintiffs disagree and contend that OpenAI has not met its burden of establishing that The Times discovered the alleged infringement before December 27, 2020, three years before it filed its complaint, or that the *Daily News* plaintiffs discovered the alleged infringement before April 30, 2021, three years before they filed their complaint in the *Daily News* action. The Court agrees with plaintiffs.

A. Applicable Standard

Section 507(b) states that "[n]o civil action shall be maintained under the provisions of this title unless it is commenced within three years after the claim accrued." 17 U.S.C. § 507(b). Under the discovery rule, "an infringement claim does not 'accrue' until the

---

[1] In this Section III, the Court uses the term "plaintiffs" to refer to The Times and the *Daily News* plaintiffs exclusively.

copyright holder discovers, or with due diligence should have discovered, the infringement." *Sohm v. Scholastic Inc.*, 959 F.3d 39, 50 (2d Cir. 2020), *abrogated on other grounds*, *Warner Chappell Music, Inc. v. Nealy*, 601 U.S. 366, 373 (2024).

The discovery rule does not impose on copyright holders "a general duty to police the internet" to uncover infringement. *Parisienne v. Scripps Media, Inc.*, No. 19-cv-8612, 2021 WL 3668084, at *2 (S.D.N.Y. Aug. 17, 2021) (internal quotation marks omitted). Indeed, the Court of Appeals for the Second Circuit has rejected the argument that a plaintiff's failure to conduct a search to uncover potential infringement, despite having the ability to do so, alone triggers constructive notice. *See Sohm,* 959 F.3d at 51. Similarly, "a copyright holder's *general* diligence or allegations of diligence in seeking out and litigating infringements, alone, are insufficient to make it clear that the holder's *particular* claims in any given case should have been discovered more than three years before the action's commencement." *Michael Grecco Prods., Inc. v. RADesign, Inc.*, 112 F.4th 144, 148 (2d Cir. 2024). Rather, to establish constructive notice, a defendant must identify specific "facts or circumstances that would have prompted such an inquiry" by the copyright holder into the alleged infringing activity. *Sohm*, 959 F.3d at 51; *see also McGlynn v. Sinovision Inc.*, No. 23-cv-4826, 2024 WL 643021, at *2 (S.D.N.Y. Feb. 15, 2024).

Finally, "[t]here is no 'sophisticated plaintiff' exception to the discovery rule." *RADesign,* 112 F.4th at 148*.* "The date on which a copyright holder, with the exercise of due diligence, would have discovered an infringement—or whether the alleged date of discovery reflected a lack of due diligence—is a fact-intensive inquiry that cannot be determined from the general nature of a copyright holder's 'sophistication' alone." *Id.* at 152.

Because the statute of limitations is an affirmative defense, OpenAI bears the burden of establishing that by December 27, 2020 for The Times and April 30, 2021 for the *Daily News* plaintiffs—three years before they filed their respective complaints—the plaintiffs should have been aware of the alleged infringement. *See id.* at 149; Fed. R. Civ. P. 8(c)(1). Dismissal of a copyright infringement claim on statute of limitations grounds at the pleadings stage is only appropriate when "it is clear from the face of the complaint, and matters of which the court may take judicial notice, that the plaintiff's claims are barred as a matter of law." *Sewell v. Bernardin*, 795 F.3d 337, 339 (2d Cir. 2015) (citation omitted). "However, where there is even 'some doubt' as to whether dismissal is warranted, a court should not grant a Rule 12(b)(6) motion on statute of limitations grounds." *PK Music Performance, Inc. v. Timberlake*, No. 16-cv-1215, 2018 WL 4759737, at *7 (S.D.N.Y. Sept. 30, 2018) (citing *Ortiz v. Cornetta*, 867 F.2d 146, 149 (2d Cir. 1989)).

B.   Plaintiffs' Claims of Direct Infringement Occurring More than Three Years Before the Filing of the Complaints Are Not Time-Barred by the Statute of Limitations.

OpenAI has not met its burden of establishing that The Times and the *Daily News* plaintiffs discovered, or with due diligence should have discovered, the alleged infringement before December 27, 2020 and April 30, 2021 respectively. Although the complaints allege that defendants trained their LLMs in 2019 and 2020 on datasets that included plaintiffs' works, the complaints do not establish that plaintiffs "discover[ed], or with due diligence should have discovered" that fact in 2019 and 2020. *See Sohm*, 959 F.3d at 50.

OpenAI identifies a few publicly available documents from 2019 and 2020, one of which—an article by Jennifer Langston entitled "Microsoft Announces New Supercomputer, Lays Out Vision for Future AI Work"—was cited in plaintiffs' complaints, to argue that it was "common knowledge" by 2020 that some of defendants' training datasets included plaintiffs' works.[2] These documents are insufficient for at least two reasons. First, as to the Langston article, OpenAI does not explain why plaintiffs would have known in 2020 of that article's existence, nor does it point to "facts or circumstances" that would have prompted plaintiffs to look for the article at that time. *See Sohm*, 959 F.3d at 51. Second, OpenAI fails to explain why the articles, even if their existence had been known to plaintiffs at the time of their publishing, are sufficient to put plaintiffs on notice of the particular infringing conduct by defendants that provides the basis for plaintiffs' claims. *Cf. McGlynn*, 2024 WL 643021, at *5 (declining to infer that past litigation between the parties constituted "facts or circumstances" sufficient to put copyright holder on notice of alleged infringement).

OpenAI makes much of The Times's reporting in November 2020 that OpenAI trained its models by "analyzing . . . nearly a trillion words posted to blogs, social

---

[2] *See, e.g.*, *Daily News*, Compl. ¶ 67 n.11 (citing Jennifer Langston, *Microsoft Announces New Supercomputer, Lays Out Vision for Future AI Work*, Microsoft (May 19, 2020), https://news.microsoft.com/source/features/ai/openai-azure-supercomputer/) (article published on Microsoft's website discussing a "new class of multitasking AI models" that "can learn about language by examining billions of pages of publicly available documents on the internet")); *Times*, FAC ¶ 70 n.8 (same). *See also Times*, ECF No. 52 at 6 (citing Cade Metz, *Meet GPT-3. It Has Learned To Code (and Blog and Argue)*, N.Y. Times (Nov. 24, 2020), https://www.nytimes.com/2020/11/24/science/artificial-intelligence-ai-gpt3.html). OpenAI references the Metz article even though it was not cited in The Times's complaint, contending that the article is not being introduced for the truth of its statements but for the Court to "take judicial notice of the fact [of] press coverage . . . [to] decid[e] whether so-called 'storm warnings' were adequate to trigger inquiry notice." (*Times*, ECF No. 75 at 2 n.2 (quoting *Staehr v. Hartford Fin. Servs. Grp., Inc.*, 547 F.3d 406, 425 (2d Cir. 2008).) The Court takes judicial notice of the Metz article but finds that its publication does not constitute a "storm warning . . . adequate to trigger inquiry notice." *Staehr*, 547 F.3d at 425.

media and the rest of the internet." Cade Metz, *Meet GPT-3. It Has Learned To Code (and Blog and Argue)*, N.Y. Times (Nov. 24, 2020), https://www.nytimes.com/2020/11/24/science/artificial-intelligence-ai-gpt3.html. The fact that one of The Times's reporters discussed OpenAI's "analyzing . . . a trillion words" on the internet fails to "make it clear that [plaintiffs'] *particular* claims . . . should have been discovered more than three years before the action's commencement." *RADesign*, 112 F.4th at 148. Those claims involve the specific copying of plaintiffs' works by OpenAI, which, as alleged in the complaints, "became a household name upon the release of ChatGPT in November 2022," two years after the Metz article. (*Times*, FAC ¶ 61; *Daily News*, FAC ¶ 58.)

Finally, OpenAI's argument that The Times, as a "sophisticated publisher," had a duty "to take prompt action after being put on notice of what it now claims to be alleged infringement" is a straw man. (*See Times*, ECF No. 75 at 3.) OpenAI has failed to establish that The Times *was in fact on notice* before December 27, 2020—and the Second Circuit has squarely rejected a heightened "sophisticated rightsholder" theory of constructive knowledge. *RADesign*, 112 F.4th at 148.

Discovery may reveal facts supporting OpenAI's contention that The Times and the *Daily News* plaintiffs discovered the alleged infringement before December 27, 2020 and April 30, 2021 respectively. At the motion to dismiss stage, however, OpenAI's conclusory statement that plaintiffs "discovered or with reasonable diligence should have discovered these activities" prior to three years before the filing of their complaints (*Times*, ECF No. 52 at 15 n.33), fails to meet its burden of establishing actual or constructive knowledge. Accordingly, OpenAI's motion to dismiss the copyright infringement claims arising more than three years before The Times and the *Daily News* plaintiffs filed their complaints is denied.

IV. Contributory Copyright Infringement

Next, defendants move to dismiss plaintiffs' contributory copyright infringement claims. Plaintiffs bring those claims in the alternative to their direct infringement claims to the extent third-party end users—not defendants—are found liable for direct infringement for generating infringing outputs using defendants' LLMs. Under this theory, defendants materially contributed to and directly assisted with the direct infringement by end users by (1) building and training their LLMs using plaintiffs' works; (2) deciding what content is outputted by their LLMs through specific training techniques; and (3) developing LLMs capable of distributing copies of plaintiffs' works to end users without authorization by plaintiffs.

Defendants contend that plaintiffs have failed to plausibly allege both direct infringement by a third party and that defendants knew of third-party infringement.

13

A. Applicable Standard

An individual or entity may be held liable as a contributory copyright infringer if that individual or entity, "with knowledge of the infringing activity, induces, causes or materially contributes to the infringing conduct of another." *Gershwin Publ'g Corp. v. Columbia Artists Mgmt., Inc.*, 443 F.2d 1159, 1162 (2d Cir. 1971); *see also Arista Recs., LLC v. Doe 3*, 604 F.3d 110, 117 (2d Cir. 2010). To plausibly allege contributory copyright infringement under a theory of material contribution, a plaintiff must show "(1) direct infringement by a third party, (2) that the defendant had 'knowledge of the infringing activity,' (3) and that the defendant 'materially contribute[d] to' the third party's infringement." *Dow Jones & Co., Inc. v. Juwai Ltd.*, No. 21-cv-7284, 2023 WL 2561588, at *3 (S.D.N.Y. Mar. 17, 2023) (citing *Smith v. BarnesandNoble.com, LLC*, 143 F. Supp. 3d 115, 124 (S.D.N.Y. 2015)).

The parties disagree on the standard for establishing that defendants knew a third party was infringing plaintiffs' copyrights. Their disagreement mirrors a split among the Circuits regarding the scienter required to support a contributory copyright infringement claim, absent actual knowledge of third-party infringement. Plaintiffs contend that the standard is actual or constructive knowledge; namely, whether defendants objectively "know or have reason to know" of the direct infringement by third-party end users. That is the standard in the Second Circuit. *See Gershwin Publ'g Corp.*, 443 F.2d at 1162; *Doe 3*, 604 F.3d at 118; *ReDigi Inc.*, 934 F. Supp. 2d at 658; *State St. Glob. Advisors Tr. Co. v. Visbal*, 431 F. Supp. 3d 322, 358 (S.D.N.Y. 2020); *Rams v. Def Jam Recordings, Inc.*, 202 F. Supp. 3d 376, 383 (S.D.N.Y. 2016); *BarnesandNoble.com*, 143 F. Supp. 3d at 124; *Arista Recs. LLC v. Usenet.com, Inc.*, 633 F. Supp. 2d 124, 154 (S.D.N.Y. 2009)); *Arista Recs., Inc. v. Mp3Board, Inc.*, No. 00-cv-4660, 2002 WL 1997918, at *6–7 (S.D.N.Y. Aug. 29, 2002).

Defendants urge a heightened standard, contending that liability for contributory copyright infringement requires that the defendant have possessed actual knowledge of or willful blindness to specific acts of infringement. That is the standard in the Ninth Circuit. *See Ludvarts, LLC v. AT&T Mobility, LLC*, 710 F.3d 1068, 1072–73 (9th Cir. 2013) (requiring "actual knowledge of specific acts of infringement" or "[w]illful blindness of specific facts" to "establish knowledge for contributory [copyright] liability"). The Court therefore will evaluate plaintiffs' contributory copyright infringement claims under the standard in the Second Circuit.

"The knowledge standard is an objective one; contributory infringement liability is imposed on persons who 'know *or have reason to know*' of the direct infringement." *Doe 3*, 604 F.3d at 118 (citation omitted). "[M]ore than a generalized knowledge by the defendant of the possibility of infringement" is required to meet the knowledge requirement. *Hartmann v. Apple, Inc.*, No. 20-cv-6049, 2021 WL 4267820, at *6 (S.D.N.Y.

Sept. 20, 2021) (citing *Rams*, 202 F. Supp. 3d at 376). By the same token, "knowledge of specific infringements is not required to support a finding of contributory infringement." *Usenet.com, Inc.*, 633 F. Supp. 2d at 154. Rather, courts have looked to whether the complaint contains "allegation[s] that the defendant investigated or would have had reason to investigate the alleged infringement," *Hartmann v. Popcornflix.com LLC*, 690 F. Supp. 3d 309, 320 (S.D.N.Y. 2023), as well as evidence of "cease-and-desist letters, officer and employee statements, promotional materials, and industry experience" to determine whether the defendant was put on notice of third-party infringement. *ReDigi Inc.*, 934 F. Supp. 2d at 658.

To establish that the defendant "materially contributed" to the infringement, the complaint must show that the defendant "encouraged or assisted others' infringement[] or provided machinery or goods that facilitated infringement." *Arista Recs. LLC v. Lime Grp. LLC*, 784 F. Supp. 2d 398, 432 (S.D.N.Y. 2011). [3]

B.   Plaintiffs Have Plausibly Alleged Contributory Copyright Infringement.

1.   *Third Party Infringement*

Although defendants rely primarily on cases decided on summary judgment or after trial, the question at this stage is simply whether plaintiffs have plausibly alleged that end-user infringement has taken place. The Court finds that plaintiffs have done so. First, the complaints allege "widely publicized" instances "of copyright infringement after ChatGPT, Browse with Bing, and Bing Chat were released." (*Times*, FAC ¶ 126; *Daily News*, Compl. ¶ 141.) Second, plaintiffs include numerous examples of infringing outputs in their complaints. (*See, e.g., Times*, FAC ¶¶ 99–101, 104–22, Ex. J; *Daily News*, Compl. ¶¶ 98–113, 118–37, Ex. J; *CIR*, FAC Ex. 10, Ex. 11 at 6–17, 14.) [4] These examples "raise a reasonable expectation" that discovery will reveal evidence of additional examples of third-party infringement such that dismissal at this stage would be improper. *See Twombly*, 550 U.S. at 556.

Matthew Bender v. West Publishing Co., relied upon by defendants, does not counsel a different result. *See* 158 F.3d 693 (2d Cir. 1998). In that case, which was decided on a motion for summary judgment, the rightsholder (West) held a copyright in its printed compilations of judicial opinions, which were distinctly paginated in the star pagination format. *Id.* at 696. A CD-ROM disc manufacturer (Matthew Bender) produced CD-ROM

---

[3] Defendants do not challenge the "material contribution" component of plaintiffs' contributory copyright infringement claim.

[4] The *Daily News* and *Times* complaints include numerous examples of allegedly infringing outputs. The *CIR* complaint includes five examples of allegedly infringing outputs, of which at least two illustrate the ease with which end-user infringement can occur using defendants' products. (*See CIR*, FAC Ex. 10, Ex. 11 at 6–7, 14.)

discs containing compilations of judicial opinions that included references to West's star pagination, but did not arrange the judicial opinions as West had done in its printed compilations. *Id.* at 697. West alleged that Matthew Bender was liable for the copyright infringement of third-party users. In affirming the district court's grant of summary judgment for Matthew Bender, the Second Circuit determined that West had failed to plausibly allege the existence of third-party infringement, reasoning that to infringe West's copyright in its printed compilation of court opinions, "a user must retrieve each case, one at a time, in the order in which they appear in the West volume, and then print each one." *Id.* at 706. The court concluded that West had failed to identify any third-party infringer other than its own counsel (who had artificially replicated West's compilation of cases) and that West's "hypothesized" examples of infringement were insufficient. *Id.*

In contrast, plaintiffs' examples of allegedly infringing outputs at the pleading stage—including more than 100 pages of examples provided in Exhibit J to the *Times* complaint, and dozens of examples in Exhibit J to the *Daily News* complaint—combined with their allegations of "widely publicized" instances of copyright infringement by end users of defendants' products, give rise to a plausible inference of copyright infringement by third parties.

2.   *Knowledge of Third-Party Infringing Activity*

Plaintiffs have also plausibly alleged that defendants possessed constructive, if not actual, knowledge of end-user infringement. The complaints allege that defendants knew they were using copyrighted works to train their models and were fully aware of plaintiffs' protected interests in their works. (*See Times*, FAC ¶¶ 7, 8, 69–70, 126; *Daily News*, Compl. ¶¶ 66–67, 114–15, 144–47; *CIR*, FAC ¶¶ 26, 28, 83, 144.) *Cf. Hartmann v. Amazon.com, Inc.* No. 20-cv-4928, 2021 WL 3683510, at *7 (S.D.N.Y. Aug. 19, 2021) (reasoning that defendants did not know or have reason to know of third-party infringement in part because defendants did not know plaintiff had a protected interest in the underlying works). The complaints also reference "widely publicized" reports that end users were using defendants' LLMs to "elicit copyrighted content" (*Times*, FAC ¶ 126; *Daily News*, Compl. ¶ 141), and statements by OpenAI representatives about internal company disagreements regarding copyright issues. (*Times*, FAC ¶ 124; *Daily News*, Compl. ¶ 139.) The Times even informed defendants "that their tools infringed its copyrighted works," supporting the inference that defendants possessed actual knowledge of infringement by end users. (*Times*, FAC ¶ 126.) Taken as true, these facts give rise to a plausible inference that defendants at a minimum had reason to investigate and uncover end-user infringement.

Plaintiffs allege that defendants possessed far more than a "generalized knowledge of the possibility" of third-party infringement. *Visbal*, 431 F. Supp. 3d at 358. Indeed, plaintiffs allege both that (1) defendants possessed actual knowledge of third-party infringement and (2) defendants knew not only that their unauthorized copying of plaintiffs' works on a massive scale during the training of their LLMs would "result[] in the unauthorized encoding of huge numbers of such works in the models themselves," but also that it "would inevitably result in the unauthorized display of such works" in response to third-party queries. (*Times*, FAC ¶ 124; *Daily News*, Compl. ¶ 139.) In other words, defendants knew or had reason to know of third-party infringement because copyright infringement was "central to [defendants'] business model." *See ReDigi*, 934 F. Supp. 2d at 659; *see also Andersen v. Stability AI Ltd.*, 744 F. Supp. 3d 956, 969 (N.D. Cal. 2024) (finding defendants possessed knowledge of third-party infringement when their products were "built to a significant extent on copyrighted works" and their operation "necessarily invokes copies or protected elements of those works"). These allegations are sufficient at the pleading stage to establish a plausible inference that defendants possessed actual or constructive knowledge of third-party infringement.

3. *Substantial Noninfringing Uses*

Defendants also contend that plaintiffs' contributory copyright infringement claims fail because defendants' LLMs—even if they were used by third parties to commit copyright infringement, and even if defendants had knowledge of that fact—are capable of "substantial noninfringing uses." *Sony Corp. of Am. v. Universal City Studios, Inc.*, 464 U.S. 417, 442 (1984); *see also Metro-Goldwyn-Mayer Studios Inc. v. Grokster, Ltd.*, 545 U.S. 913, 937 (2005). Defendants rely on the U.S. Supreme Court's decisions in *Sony* and *Grokster* for this position, but that reliance is misplaced. In *Sony*—which involved Sony's video tape recorder (VTR)—the Supreme Court held that "[t]he sale of copying equipment, like the sale of other articles of commerce, does not constitute contributory infringement if the product is widely used for legitimate, unobjectionable purposes." *Sony*, 464 U.S. at 442. However, in *Grokster*—which involved peer-to-peer file sharing computer software—the Supreme Court clarified that a defendant whose product is capable of substantial noninfringing can still be held liable for third-party infringement in certain circumstances. *Grokster*, 545 U.S. at 933–35. In that case, the Supreme Court held that a distributor of a device who "promot[es] its use to infringe copyright, as shown by clear expression or other affirmative steps taken to foster infringement," is liable for third-party infringement. *Id.* at 919.

Sony and *Grokster* do not foreclose plaintiffs' contributory copyright infringement claims at this stage for at least three reasons. First, *Sony* and *Grokster* involved cases decided either on summary judgment or after trial rather than on motions to dismiss the complaint. Indeed, many of the decisions relied upon by defendants to support their

17

contention that their LLMs have "substantial noninfringing uses" were decided either on summary judgment or after trial. *See, e.g., BarnesandNoble.com,* 143 F. Supp. 3d at 124; *Matthew Bender*, 158 F.3d at 706–07. The question before the Court on a motion to dismiss a contributory copyright infringement claim is narrow: whether plaintiffs have plausibly alleged that defendants knew or had reason to know of actual third-party infringement by end users of their products. *See Doe 3*, 604 F.3d at 117–18.

Second, while *Sony* and *Grokster* analyzed claims of contributory copyright infringement by inducement, they did not discuss claims of contributory copyright infringement by material contribution, which plaintiffs allege here. *See Grokster*, 545 U.S. at 934 (rejecting circuit court's "converting [of] the [*Sony*] case from one about liability resting on imputed intent to one about liability on any theory"). Indeed, "the fact that a product is 'capable of substantial lawful use' does not mean the 'producer can never be held contributorily liable.'" *BMG Rts. Mgmt. (US) LLC v. Cox Commc'ns, Inc.*, 881 F.3d 293, 306 (4th Cir. 2018) (citations omitted).

In a word, *Sony* foreclosed imputing "culpable intent" solely based on the "characteristics or uses of a distributed product." *Grokster*, 545 U.S. at 934. It left open, however, other "rules of fault-based liability derived from the common law," *id.*, including liability based on material contribution, a theory which neither *Sony* nor *Grokster* discussed nor foreclosed. *See Gershwin Publ'g Corp.*, 443 F.2d at 1162–63 (predicating contributory liability for material contribution on "the common law doctrine that one who knowingly participates [in] or furthers a tortious act is jointly and severally liable with the prime tortfeasor").

Third, the facts in *Sony* did not include two important distinguishing features. First, in *Sony* there was no "ongoing relationship between the direct infringer and the contributory infringer at the time the infringing conduct occurred." *Sony*, 464 U.S. at 437. Here, however, an "ongoing relationship" exists between defendants and end users, via defendants' LLM outputs that respond to end users' prompts. *Cf. id.* at 438 ("The only contact between Sony and the users of [the VTRs] occurred at the moment of sale."). Second, the VTR was not a product that itself was *built* on purportedly appropriated works, as are defendants' products here. *Cf. id.* at 421 (describing defendant as a distributor of "copying equipment"). On the second point, *Sony*'s discussion of *Kalem Co. v. Harper Brothers*, 222 U.S. 55 (1911) is instructive. In *Kalem*, the Supreme Court "held that the producer of an unauthorized film dramatization of the copyrighted book *Ben Hur* was liable for his sale of the motion picture to jobbers, who in turn arranged for the commercial exhibition of the film." *Sony*, 464 U.S. at 435 (discussing *Kalem*). Unlike the defendant in *Sony*, the defendant producer in *Kalem* "did not merely provide the 'means' to accomplish an infringing activity; the producer supplied the work itself, albeit in a new medium of expression." *Id.* at 436. This form of

18

contributory infringement is analogous to the material contribution by defendants that plaintiffs allege. According to the complaints, defendants appropriated plaintiffs' works, created the "tangible medium" upon which those protected works were "recorded," *id.*, and provided third-party infringers with both the means of infringing and the works from which to do so.

<p style="text-align:center">* * *</p>

The Court finds that plaintiffs have plausibly alleged the existence of third-party end-user infringement and that defendants knew or had reason to know of that infringement. Accordingly, plaintiffs have plausibly alleged their contributory copyright infringement claims and defendants' motions to dismiss those claims in all three actions are denied.

V. THE DMCA CLAIMS

Plaintiffs in each action bring two claims under the Digital Millennium Copyright Act ("DMCA") against Microsoft and OpenAI. The first claim is brought pursuant to 17 U.S.C. § 1202(b)(1), which prohibits "intentionally remov[ing] or alter[ing] any copyright management information" ("CMI").[5] The second claim is brought pursuant to 17 U.S.C. § 1202(b)(3), which prohibits the "distribution" of "works" or "copies of works . . . knowing that [CMI] has been removed or altered without authority of the copyright owner." In both provisions, the defendant must also have "know[n]" or "ha[d] reasonable grounds to know" that its conduct would "induce, enable, facilitate, or conceal an infringement." 17 U.S.C. § 1202(b). Section 1202(b) claims therefore contain a "double-scienter" requirement: the defendant must (1) "intentionally" remove CMI under section 1202(b)(1) or distribute copyrighted works "knowing" that CMI was removed under section 1202(b)(3); and (2) know or have reason to know that its conduct would induce, enable, facilitate, or conceal infringement. *See Mango v. Buzzfeed, Inc.*, 970 F.3d 167, 171 (2d Cir. 2020).

Defendants move to dismiss both DMCA claims, contending that plaintiffs lack Article III and statutory standing. They also contend that the claims fail on the merits, because plaintiffs have failed to plausibly allege (1) the actual removal of CMI; (2) that defendants had knowledge that the removal would lead to infringement; and (3) that defendants "distributed" "copies" of plaintiffs' works.

---

[5] CMI comprises information "conveyed in connection with copies . . . of a work," including, as relevant here, "[t]he title and other information identifying the work," "the author of the work," and "[t]he name of, and other identifying information about, the copyright owner of the work." 17 U.S.C. § 1202(c).

A. Article III & Statutory Standing

1. *Article III Standing*

In order to establish Article III standing, the U.S. Constitution requires a plaintiff to show "(i) that he suffered an injury in fact that is concrete, particularized, and actual or imminent; (ii) that the injury was likely caused by the defendant; and (iii) that the injury would likely be redressed by judicial relief." *TransUnion LLC v. Ramirez*, 594 U.S. 413, 423 (2021); U.S. Const. art. III, § 2, cl. 1. To be "concrete," the injury must be "actual or imminent, not conjectural or hypothetical." *Spokeo, Inc. v. Robins*, 578 U.S. 330, 339 (2016) (quoting *Lujan v. Defs. of Wildlife*, 504 U.S. 555, 560 (1992)).

While "Congress may 'elevate' harms that 'exist' in the real world before Congress recognized them to actionable legal status, it may not simply enact an injury into existence, using its lawmaking power to transform something that is not remotely harmful into something that is." *TransUnion*, 594 U.S. at 426 (citation omitted). Regardless of the existence *vel non* of a statutory cause of action, courts must "independently decide whether a plaintiff has suffered a concrete harm under Article III." *Id.* To do so, courts look to "whether the alleged injury to the plaintiff has a 'close relationship' to a harm 'traditionally' recognized as providing a basis for a lawsuit in American courts," namely "whether plaintiffs have identified a close historical or common-law analogue for their asserted injury." *Id.* at 424 (quoting *Spokeo,* 578 U.S. at 341).

Plaintiffs allege that defendants removed CMI from plaintiffs' works during the process of training their LLMs and in distributing unauthorized copies of plaintiffs' works through regurgitating outputs. (*See Times*, FAC ¶¶ 184–86; *Daily News*, Compl. ¶ 159; *CIR*, FAC ¶ 103.) According to plaintiffs, this removal caused two forms of concrete harm: (1) harm caused through the removal of CMI during the training process, absent dissemination, and (2) harm caused by the dissemination of CMI-less copies of plaintiffs' works in the form of LLM outputs.

Defendants contend that both theories of harm fail. With respect to the first theory—harm absent dissemination—defendants contend that because the training data from which CMI was allegedly removed was never disseminated or otherwise made publicly available, it is therefore not a legally cognizable injury under Article III. The second theory of harm—harm caused by dissemination—fails for two reasons, according to defendants. First, the alleged injuries caused by the dissemination of CMI-less works—including the inability to receive licensing and subscription revenue, and the possibility that the regurgitating outputs will divert readers from plaintiffs' platforms—do not have any nexus to CMI removal. Second, because the outputs cited in the complaints either reference plaintiffs' articles by name or result from prompts

that quote substantial portions of the underlying article, "any user who encountered those outputs would have no doubt as to the provenance of the text and could easily find it on the [plaintiffs'] website," thereby making plaintiffs' alleged injuries imaginary and not "concrete." (*Times*, ECF No. 52 at 22; *see also Daily News*, ECF No. 82 at 13–14; *CIR*, ECF No. 100 at 17.)

2.   *Statutory Standing*

OpenAI also raises a statutory standing argument, contending that plaintiffs have failed to allege that they are "person[s] injured by" a DMCA violation, as required by 17 U.S.C. § 1203(a). Section 1203(a) states that "[a]ny person injured by a violation of section 1201 or 1202 may bring a civil action in an appropriate United States district court for such violation." OpenAI contends that the language of this section requires an injury *beyond* a mere statutory violation, and that the injury must be caused by the section 1202(b) violation specifically. It also contends that, for the same reasons plaintiffs have failed to allege an injury for Article III purposes, plaintiffs fail to establish that it was either the *CMI removal* from defendants' training datasets or the dissemination of *CMI-less* works that caused plaintiffs the specific harm they allege.

3.   *Plaintiffs Have Article III Standing To Bring Their DMCA Claims.*

a.   **Concreteness**

Defendants challenge plaintiffs' allegations of harm as lacking "a close relationship to harms traditionally recognized as providing a basis for lawsuits in American courts," relying on the Supreme Court's decision in *TransUnion LLC v. Ramirez.* 594 U.S. at 425. In *TransUnion*, the Supreme Court explained that the harm alleged must possess a "close historical or common-law analogue," although, importantly, that harm need not be an "exact duplicate." *Id.* at 424. Here, traditional copyright law provides that "close historical or common-law analogue" and supports plaintiffs' claims of harm.

"Copyright claims predate the Constitution's ratification." *The Intercept Media, Inc. v. OpenAI, Inc.*, No. 24-cv-1515, 2025 WL 556019, at *4 (S.D.N.Y. Feb. 20, 2025) (citing The Federalist No. 43 (James Madison)). Indeed, copyright is listed among Congress' enumerated powers in Article I of the Constitution, *see* U.S. Const. art. I, § 8, cl. 8, and, from the time of the enactment of the Copyright Act of 1790, Congress has updated the copyright laws numerous times over the past two centuries. *See* Copyright Act of 1790; Copyright Act of 1831; Copyright Act of 1870; Copyright Act of 1909; Copyright Act of 1976.

DMCA claims differ from traditional copyright claims. DMCA claims protect against harms caused by the unauthorized removal of CMI from a copyrighted work;

traditional copyright infringement claims protect against, among other things, the unauthorized reproduction and distribution of protected works. *See Mango*, 970 F.3d at 170–71; *Authors Guild, Inc. v. HathiTrust*, 755 F.3d 87, 95 (2d Cir. 2014). Nonetheless, copyright infringement—a harm "traditionally recognized as providing a basis for lawsuits in American Courts"—provides an appropriate "close historical or common-law analogue" to the harm caused by a DMCA violation. *TransUnion*, 594 U.S. at 424–25. Both traditional copyright and DMCA claims are grounded in notions of property rights, and both claims are designed to "promote the Progress of Science and useful Arts." U.S. Const. art. I, § 8, cl. 8. As Judge Jed S. Rakoff wrote in *The Intercept Media*, "[t]he DMCA adds another stick to the bundle of property rights already guaranteed to an author in her work under traditional copyright law," and "[t]he fact that the specific right at issue here is not expressly rooted in that overall history misses the point; the exact contours of the property rights given to a copyright holder are not frozen in time by the Copyright Act of 1790." 2025 WL 556019, at *5 .

As noted above, Article III requires that the *harm* alleged have a "close relationship to a harm traditionally recognized as providing a basis for a lawsuit in American courts.'" *TransUnion*, 594 U.S. at 417 (internal quotation marks omitted). The "inquiry asks whether plaintiffs have identified a close historical or common-law analogue for their *asserted injury*," *id.* (emphasis added), not the asserted cause of action. *See Kadrey v. Meta Platforms, Inc.*, No. 23-cv-03417, 2025 WL 744032, at *1 (N.D. Cal. Mar. 7, 2025).

For both DMCA and traditional copyright infringement claims, the harm involves an injury to "an author's property right in his original work of authorship." *The Intercept Media*, 2025 WL 556019, at *5.[6] Indeed, the DMCA was enacted by Congress "to strengthen copyright protection in the digital age," and to "combat copyright piracy," which Congress feared was "overwhelming the capacity of conventional copyright enforcement to find and enjoin unlawfully copied material." *Mango*, 970 F.3d at 170–171. The requirement under section 1202(b) that a defendant "kno[w], or, . . . hav[e] reasonable grounds to know" that their conduct "will induce, enable, facilitate, or conceal" copyright infringement "ensures that any violation of the DMCA is tied to concerns of downstream infringement." *The Intercept Media*, 2025 WL 556019, at *6.

---

[6] That *TransUnion* rejected the plaintiffs' allegations of harm "absent dissemination" does not change this outcome. There, the Supreme Court stated that "[t]he mere presence of an inaccuracy in an internal credit file, if it is not disclosed to a third party, causes no concrete harm." *TransUnion*, 594 U.S. at 434. However, that case involved an entirely different alleged harm (the potential dissemination of inaccurate credit information) and an inapposite historical analogue (defamation). Because defamation requires publication of the defamatory statement, the Supreme Court concluded that those class members whose inaccurate credit information had not been publicly disseminated lacked a historical common law analogue and therefore did not suffer a "concrete" injury to as required by Article III. *See id.* at 425–26, 434.

Accordingly, plaintiffs' allegations of harm pursuant to sections 1202(b)(1) and 1202(b)(3) are sufficiently concrete to satisfy the injury-in-fact requirement of Article III.

b. Causation

OpenAI also contends that the harm plaintiffs allege is not "fairly traceable" to the purported removal of CMI from plaintiffs' works. Specifically, OpenAI contends that (1) plaintiffs have failed to establish that the alleged removal of CMI from articles in the training datasets *caused* defendants LLMs to exclude CMI from the regurgitating outputs (*see CIR*, ECF No. 100 at 18–19), and (2) the alleged harm of plaintiffs' "inability to receive speculative subscription and licensing revenue . . . do[es] not flow from any purported removal of CMI." (*Daily News*, ECF No. 82 at 13; *see also Times*, ECF No. 52 at 22.)

The Court disagrees. To satisfy the causation requirement for Article III standing, a plaintiff must establish "that the plaintiff's injury likely was caused or likely will be caused by the defendant's conduct." *Food & Drug Admin. v. All. for Hippocratic Med.*, 602 U.S. 367, 382 (2024). Put differently, "there must be a causal connection between the injury and the conduct complained of—the injury has to be fairly traceable to the challenged action of the defendant, and not the result of the independent action of some third party not before the court." *Lujan*, 504 U.S. at 560 (alterations and internal quotation marks omitted).

Here, plaintiffs allege that defendants' removal of CMI from plaintiffs' works conceals and facilitates copyright infringement, which deprives plaintiffs of licensing and subscription revenue. That harm is fairly traceable to the removal of CMI: its removal allows defendants to provide plaintiffs' works directly to end users through regurgitating outputs, while concealing that defendants infringed plaintiffs' copyrights to generate those outputs. This conduct obviates the need of end users to subscribe to plaintiffs' works or eliminates or reduces their reluctance to use defendants' products out of knowledge that doing so might constitute further infringement. That the same harm could potentially occur even if defendants did *not* remove CMI from plaintiffs' works misses the point. Assuming plaintiffs would succeed on their claims that defendants removed CMI from their works—as the Court must do when determining whether plaintiffs have standing, *see City of Waukesha v. EPA*, 320 F.3d 228, 235 (D.C. Cir. 2003) (per curiam)—plaintiffs have alleged harms that are "the predictable effect" of that CMI removal. *Dep't of Com. v. New York*, 588 U.S. 752, 768 (2019).

The Court finds that plaintiffs have satisfied Article III's causation requirement and have Article III standing to bring their DMCA claims.

4.   *Plaintiffs Have Statutory Standing To Bring Their DMCA Claims.*

OpenAI's statutory standing argument also fails. OpenAI contends that even if plaintiffs suffered a DMCA violation, they are not "person[s] injured by" that violation as required by 17 U.S.C. § 1203(a) and therefore lack statutory standing.

Assuming without deciding that section 1203(a) requires allegations of injury beyond a mere statutory violation, plaintiffs have satisfied this requirement. As discussed above, plaintiffs have not merely alleged that defendants violated section 1202(b); they have alleged that this violation injures them by concealing defendants' own copyright infringement, enabling and facilitating the copyright infringement of end users, diverting users from plaintiffs' websites, and causing a decline in subscription and licensing revenue.

* * *

To summarize, the Court finds that plaintiffs have established both Article III and statutory standing sufficient to enable them to pursue their DMCA claims, because the harms they allege bear a "close relationship" to traditional copyright infringement sufficient to satisfy the injury-in-fact requirement of Article III; their alleged harms are fairly traceable to defendants' conduct; and they have alleged that they are "person[s] injured by" defendants' violation of section 1202(b), as required by section 1203(a).

B.   Failure To State a Claim

On the merits, the Court finds that all three complaints fail to state a claim pursuant to section 1202(b)(1) against Microsoft. The Times also fails to state a claim pursuant to section 1202(b)(1) against OpenAI, but CIR and the *Daily News* plaintiffs have plausibly alleged that OpenAI violated section 1202(b)(1).[7] In addition, all three complaints fail to state a claim pursuant to section 1202(b)(3) against both Microsoft and OpenAI.

1.   *The* **Daily News** *Plaintiffs and CIR Have Stated Claims Against OpenAI Pursuant to Section 1202(b)(1), but The Times Has Failed To Do So.*

To establish a 17 U.S.C. § 1202(b)(1) violation, a plaintiff must show "(1) the existence of CMI on the allegedly infringed work, (2) the removal or alteration of that information and (3) that the removal was intentional." *Fischer v. Forrest*, 968 F.3d 216, 223 (2d Cir. 2020). Section 1202(b) also requires that defendants knew or had reasonable

---

[7] OpenAI urges that plaintiffs' DMCA claims are time-barred to the extent they are based on the building of LLM training datasets occurring more than three years before the complaints were filed. The Court rejects these arguments for the reasons stated in Section III. Accordingly, with respect to the surviving section 1202(b) claims, the Court denies OpenAI's motions to dismiss those claims based on the statute of limitations.

grounds to know that their removal of CMI would "induce, enable, facilitate, or conceal" infringement.

a.   Intentional Removal of CMI

OpenAI principally challenges the complaints as failing to plausibly allege that OpenAI removed CMI from the training datasets. The Times contends that because the regurgitating outputs listed in their complaints lack CMI, then *a fortiori* CMI was removed by defendants during the training process. The Court is not convinced by this argument. To the contrary, the regurgitating outputs in the *Times* complaint all contain *excerpts* of The Times' articles, not complete or substantially complete copies; it is entirely plausible that CMI remained on the articles included in the training datasets but simply did not appear in the outputs. (*See Times*, FAC Ex. J.) The *Times* complaint does not include any specific detail on how CMI was allegedly removed during the training process, and its conclusory statement that defendants' process of training their LLMs removes CMI "by design" (*id.* ¶ 187) fails to "nudge[] [its] claims across the line from conceivable to plausible." *Twombly*, 550 U.S. at 570.

The *Daily News* and *CIR* complaints, however, plausibly allege CMI removal during the LLM training process. The *Daily News* complaint describes OpenAI's use of the "Dragnet and Newspaper content extractors in creating the WebText dataset, which intentionally removed the [*Daily News* plaintiffs'] CMI from the [*Daily News* plaintiffs'] Works scraped from their website[s]." (*Daily News*, Compl. ¶ 161.) The *Daily News* complaint explains that Dragnet removes copyright notices "as part of the process of extracting the text content of a website," and alleges that the Newspaper content extractor also "separate[s] and extract[s] the article text on the [*Daily News* plaintiffs'] webpages" while removing CMI. (*Id.*)

The *CIR* complaint provides the most detail on how the Dragnet and Newspaper algorithms remove CMI during the process of assembling the dataset from which defendants trained their models. (*See CIR*, FAC ¶¶ 59–63.) CIR states that OpenAI, in developing Webtext, "used sets of algorithms called Dragnet and Newspaper to extract text from websites," and specifically alleges that "Dragnet's algorithms are designed to 'separate the main article content' from other parts of the website, including 'footers' and 'copyright notices,' and allow the extractor to make further copies only of the 'main article content.'" (*Id.* ¶¶ 59–60). In addition, Newspaper is "incapable of extracting copyright notices and footers" from the articles it scrapes from the internet, according to the complaint. (*Id.* ¶ 61.) Indeed, the *CIR* complaint alleges that OpenAI intentionally used both the Dragnet and Newspaper algorithms to "create redundancies," and that "[o]n information and belief, the OpenAI Defendants chose not to extract author and title information" when using the Newspaper algorithms "because they desired consistency with the Dragnet extractions, and Dragnet is typically unable to extract

author and title information." (*Id.* ¶¶ 59, 61.) These allegations are sufficient at the pleading stage to plausibly allege that OpenAI removed CMI from CIR's works in the process of building the datasets to train their LLMs.

With respect to intentionality, both the *Daily News* and *CIR* complaints plausibly allege that OpenAI's CMI removal was intentional. The *Daily News* complaint explains that defendants intentionally used the Dragnet and Newspaper content extractors, which are designed to remove CMI from the works they scrape from the internet. (*Daily News*, Compl. ¶ 161.) The *CIR* complaint alleges that "[b]ecause, by the time of its scraping, Dragnet and Newspaper were publicly known to remove author, title, copyright notices, and footers, and given that OpenAI employs highly skilled data scientists who would know how Dragnet and Newspaper work, the OpenAI Defendants intentionally and knowingly removed this copyright management information while assembling WebText." (*CIR*, FAC ¶ 64.) Plaintiffs in both the *Daily News* and *CIR* actions have satisfied their burden at the pleading stage of alleging removal of CMI, especially given this Circuit's leniency when evaluating scienter on a motion to dismiss. *See In re DDAVP Direct Purchaser Antitrust Litig.*, 585 F.3d 677, 693 (2d Cir. 2009); *Aaberg v. Francesca's Collections, Inc.*, No. 17-cv-115, 2018 WL 1583037, at *9 (S.D.N.Y. Mar. 27, 2018); *Hirsch v. CBS Broad. Inc.*, No. 17-cv-1860, 2017 WL 3393845, at *8 (S.D.N.Y. Aug. 4, 2017).

b.   The Second Scienter Requirement

OpenAI also contends that plaintiffs fail to plausibly allege both (1) that the removal of CMI could enable third-party infringement, and (2) that OpenAI had knowledge of this fact. With respect to its first objection, "nothing in the statutory language [of section 1202(b)] limits its applicability to such downstream [third-party] infringement." *Mango*, 970 F.3d at 172. Section 1202(b) simply requires that "defendant know or have reason to know that distribution of copyrighted material despite the removal of CMI 'will induce, enable, facilitate, or conceal an infringement.'" *Id.* (quoting 17 U.S.C. § 1202(b)). That infringement "is not limited by actor (*i.e.*, to third parties) or by time (*i.e.*, to future conduct)," *id.*, and plaintiffs have plausibly alleged both that OpenAI's removal of CMI conceals its *own* infringement, and that that removal enables and facilitates third-party infringement. *See Shihab v. Complex Media, Inc.*, No. 21-cv-6425, 2022 WL 3544149, at *5 (S.D.N.Y. Aug. 17, 2022) (finding the second scienter requirement satisfied when plaintiff plausibly alleged that the defendant knew its CMI removal would conceal "its own alleged infringement").

With respect to OpenAI's second objection—that it lacked knowledge that CMI removal would induce or conceal copyright infringement—both CIR and the *Daily News* plaintiffs plausibly allege that OpenAI knew or had reason to know that its removal of CMI from plaintiffs' works would induce, enable, facilitate, or conceal copyright

infringement. The complaints allege that OpenAI has publicly acknowledged both that it uses copyrighted works to train its models and that its models "are capable of distributing unlicensed copies of copyrighted works." (*See Daily News*, Compl. ¶¶ 4, 144–47, 214; *CIR*, FAC ¶¶ 83, 144.) The *CIR* complaint also alleges that OpenAI was aware of the possibility of end-user infringement through the generating of regurgitating outputs, including because of OpenAI's policy of agreeing to indemnify end users accused of infringement; its admission that its products regurgitate material in response to user prompts; and its recent adjustment to ChatGPT's settings to limit regurgitation. (*CIR*, FAC ¶¶ 83, 118.) Again, given courts' lenience in allowing issues of scienter to survive motions to dismiss when "such issues are appropriate for resolution by the trier of fact," *In re DDAVP Direct Purchaser Antitrust Litig.*, 585 F.3d at 693 (internal quotation marks omitted), the Court determines that CIR and the *Daily News* plaintiffs have plausibly alleged that OpenAI violated section 1202(b)(1).

In sum, the Court dismisses The Times's 17 U.S.C. § 1202(b)(1) claim against OpenAI and denies OpenAI's motions to dismiss the 17 U.S.C. § 1202(b)(1) claims in the *CIR* and *Daily News* complaints.

2. *Plaintiffs Have Failed To State a Section 1202(b)(1) Claim Against Microsoft.*

Unlike the *Daily News* and *CIR* complaints' detailed allegations regarding OpenAI's removal of CMI during the process of developing its LLM training datasets, all three complaints are devoid of factual specificity to support their claims against Microsoft for violation of section 1202(b)(1). None of the allegations concerning Microsoft—including Microsoft's partnership with OpenAI to develop Copilot and Browse with Bing, and its provision of the cloud computing system which OpenAI uses to train its models—relate to any alleged removal *by Microsoft* of CMI from plaintiffs' works. To the contrary, all "the specific factual matter in the complaint related to CMI removal connects only to OpenAI." *The Intercept Media*, 2025 WL 556019, at *9. Accordingly, the Court dismisses all three complaints' claims pursuant to section 1202(b)(1) against Microsoft.

3. *Plaintiffs Have Failed To State a Section 1202(b)(3) Claim Against Defendants.*

To establish a section 1202(b)(3) violation, a plaintiff must show "(1) the existence of CMI in connection with a copyrighted work; and (2) that a defendant distributed works or copies of works; (3) while knowing that CMI has been removed or altered without authority of the copyright owner or the law; and (4) while knowing, or having reasonable grounds to know that such distribution will induce, enable, facilitate, or conceal an infringement." *Mango*, 970 F.3d at 171 (cleaned up). The complaints include

27

allegations that defendants distributed copies of plaintiffs' works both to end users and between each other. As explained below, the Court rejects both theories and concludes that plaintiffs have failed to state a claim pursuant to section 1202(b)(3) against both OpenAI and Microsoft.

a.   Distribution of Copies to End Users

The Times and the *Daily News* plaintiffs contend that the "regurgitations" generated by defendants' LLMs constitute "distributions" of copies of their works. The DMCA does not define "distribution." While courts have understood "distribution" under the DMCA to require a "sale or transfer of ownership extending beyond that of a mere public display," *Wright v. Miah*, No. 22-cv-4132, 2023 WL 6219435, at *7 (E.D.N.Y. Sept. 7, 2023), and have pointed to the distinction between "distributions" and "public displays" in other parts of the Copyright Act to support this conclusion, *see e.g.*, *FurnitureDealer.Net, Inc v. Amazon.com, Inc*, No. 18-cv-232, 2022 WL 891473, at *23 (D. Minn. Mar. 25, 2022) (discussing 17 U.S.C. §§ 101, 106), it is not clear whether an LLM output is a mere "public display" or something more. *Cf. id.* (reasoning that "public display does not constitute distribution, and thus is not a [DMCA] violation" and concluding that Amazon's public display of plaintiffs' product descriptions without CMI on Amazon's website did not constitute "distributions"); *Perfect 10, Inc. v. Amazon.com, Inc.*, 508 F.3d 1146, 1162 (9th Cir. 2007).

Assuming without deciding that the LLM regurgitations do constitute "distributions," plaintiffs have another hurdle to clear: they must show that those regurgitations constitute "works" or "copies of works." *See* 17 U.S.C. § 1202(b)(3). While the DMCA does not define "copies of works," an abundance of case law establishes that in "cases where claims of removal of CMI have been held viable, the underlying work has been substantially or entirely reproduced." *Fischer v. Forrest*, 286 F. Supp. 3d 590, 609 (S.D.N.Y. 2018) (citing *Bounce Exch., Inc. v. Zeus Enter. Ltd.*, No. 15-cv-3268, 2015 WL 8579023 (S.D.N.Y. Dec. 9, 2015)), *aff'd*, 968 F.3d 216 (2d Cir. 2020); *see also We the Protesters, Inc. v. Sinyangwe*, 724 F. Supp. 3d 281, 296 (S.D.N.Y. 2024); *cf. Doe 1 v. GitHub, Inc.*, No. 22-cv-06823, 2024 WL 235217, at *8–9 (N.D. Cal. Jan. 22, 2024).

The requirement of section 1202(b)(3) that the underlying work be "substantially or entirely reproduced" aligns with the DMCA's purpose of combatting piracy. "Fearful [about] the ease with which pirates could copy and distribute a copyrightable work in digital form," Congress sought to "strengthen copyright protection in the digital age," *Mango*, 970 F.3d at 171, by "protect[ing] the integrity of copyright management information and prohibit[ing] the removal of CMI from copyrighted works." *Shihab*, 2022 WL 3544149, at *4. Allowing DMCA claims to survive when the distributed work is not "close to identical" to the original would risk boundless DMCA liability, including liability for any person who distributes only portions of an article—e.g.,

various block quotes—without including CMI. *See We the Protesters*, 724 F. Supp. 3d at 297. While that person may be found liable under other causes of action, it is doubtful his reproduction of a mere portion of an article without CMI could be said to be violating the "integrity of copyright management information," *Shihab*, 2022 WL 3544149, at *4, or distributing a "work . . . [or] copies of works." 17 U.S.C. § 1202(b)(3).

Reviewing the regurgitations cited in the *Times* and *Daily News* complaints, the Court concludes that the regurgitations do not constitute "substantial[] or entire[]" reproductions of plaintiffs' works. *Fischer*, 286 F. Supp. 3d at 609. In the *Times* complaint, the regurgitations include only excerpts of the underlying articles. (*See Times*, FAC ¶¶ 99–100, 104–07, 112–22.) In addition, the excerpts were often (1) generated in response to multiple prompts, and (2) outputted in a different order than they appear in the original articles, falling far short of representing exact or substantial reproductions of the originals. With respect to the regurgitations included in Exhibit J to the *Times* complaint, those outputs also reflect only small portions of the original articles, and in fact essentially all of the regurgitations stop in the middle of a sentence. (*See, e.g.*, *Times*, FAC Ex. J at 3 (cutting off after "What's more, the company's").) These excerpts are not "work[s]" or "copies of works" as required by 17 U.S.C. § 1202(b)(3).

The regurgitations in the *Daily News* complaint contain the same flaws. The complaint includes regurgitating outputs that capture only portions of the underlying articles, as reflected in the complaint itself (*see, e.g.*, *Daily News*, Compl. ¶¶ 98–112, 118–36), and in Exhibit J to the *Daily News* complaint.

Accordingly, because the outputs are merely excerpts of plaintiffs' works and not "copies" of those works for purposes of section 1202(b)(3), The Times and the *Daily News* plaintiffs have failed to establish that defendants "distributed" "copies" of their works in violation of section 1202(b)(3), and their section 1202(b)(3) claims against defendants are dismissed.

b.   Distribution of Copies Between Defendants

Unlike the theories of distribution alleged in the *Times* and *Daily News* complaints, the *CIR* complaint alleges that OpenAI and Microsoft "distributed" CMI-less copies of CIR's works with *each other* in violation of section 1202(b)(3). (*See CIR*, FAC ¶¶ 156–57, 168–69.) Defendants challenge these claims as failing to allege (1) distribution, (2) of works or copies of works, and (3) that defendants knew or had reason to know that such distribution would induce, enable, facilitate or conceal copyright infringement.

The Court agrees that the *CIR* complaint fails to plausibly allege that OpenAI and Microsoft "distributed" CMI-less "copies" of CIR's works with each other. As set forth above, courts have understood the term "distribute" in section 1202(b)(3) to require a

"sale or transfer of ownership extending beyond that of a mere public display," *Wright*, 2023 WL 6219435, at *7, and that the "copy" which is distributed must be a "substantial[] or entire[]" reproduction of the original work. *Fischer*, 286 F. Supp. 3d at 609. The *CIR* complaint asserts that Microsoft's investment in OpenAI, its provision of "the data center and bespoke supercomputing infrastructure used to train ChatGPT," and its statement on behalf of its CEO that "we have the data," give rise to the plausible inference that OpenAI and Microsoft distributed copies of CIR's works with each other. (*See CIR*, FAC ¶¶ 28–29.) As a threshold matter, neither those allegations nor any other allegation in the *CIR* complaint provide any factual support for the assertion that *Microsoft* distributed CMI-less works to OpenAI. The Court therefore dismisses CIR's section 1202(b)(3) claim against Microsoft.

CIR's section 1202(b)(3) claim against OpenAI fares no better. CIR's allegations about the general business relationship between OpenAI and Microsoft fail to indicate that OpenAI "distributed" "copies" of CIR's articles to Microsoft. Nowhere in the complaint's lengthy discussion of OpenAI's process of training its models on works scraped from the internet does CIR allege when, why, or how OpenAI would have "distributed" copies of those works to Microsoft. "[W]ithout some further factual enhancement," CIR's "formulaic recitation of the elements of a cause of action will not do." *Twombly*, 550 U.S. at 555, 557.

CIR relies heavily on a single statement from Microsoft's CEO, stating, "we have the data." Intelligencer Staff, *Satya Nadella on Hiring the Most Powerful Man in AI*, The Intelligencer (Nov. 21, 2023), https://nymag.com/intelligencer/2023/11/on-with-kara-swisher-satya-nadella-on-hiring-sam-altman.html. However, as the court in *The Intercept Media* recently explained, that statement, considered in context, reflects Microsoft CEO Satya Nadella's "confidence in Microsoft's own AI capabilities—separate and apart from its investment in OpenAI." *The Intercept Media*, 2025 WL 556019, at *9. Indeed, the statement was made following the Microsoft CEO's assurance that "[i]f OpenAI disappeared tomorrow, I don't want any customer of ours to be worried . . . because we have all of the rights to continue the innovation." Intelligencer Staff, *supra*. Nadella continued, "[w]e have the people, we have the compute, we have the data, we have everything." *Id.* Nadella immediately followed that statement with the clarification: "But at the same time, I'm committed to the OpenAI partnership," *id.*, suggesting that his statement "we have the data" reflected his views on Microsoft's own ability to develop AI products if OpenAI "disappeared." The interviewer then asked about the Microsoft-OpenAI partnership, and Nadella explained that Microsoft's significant investment in OpenAI "gives us significant rights," and that Microsoft "build[s] tools" and "build[s] the infrastructure" as part of its partnership with OpenAI. *Id.*

It would constitute a significant inferential leap to conclude that the Microsoft CEO's single statement "we have the data," spoken in the context of explaining Microsoft's capacity to continue innovation and develop generative AI in a hypothetical future scenario, shows that OpenAI distributed copies of CIR's articles to Microsoft. The Court declines to make this leap and concludes that CIR's general allegations pertaining to the OpenAI and Microsoft business relationship fail to move CIR's section 1202(b)(3) claim across the line from conceivable to plausible.

* * *

For the foregoing reasons, the Court dismisses the plaintiffs' section 1202(b)(3) claims against Microsoft and OpenAI in all three actions.

VI. COMMON LAW UNFAIR COMPETITION BY MISAPPROPRIATION

The Times and the *Daily News* plaintiffs also bring claims of common law unfair competition by misappropriation. They allege that defendants engage in unfair competition by using plaintiffs' works, without authorization by plaintiffs, to train their LLMs, which in turn "produce informative text of the same general type and kind that [plaintiffs] produce[]." (*Times*, FAC ¶ 195; *Daily News*, Compl. ¶ 229.)[8] By producing outputs that misappropriate plaintiffs' works, defendants "directly compete with [plaintiffs'] content," free ride on plaintiffs' significant efforts to gather time-sensitive content, and harm plaintiffs through lost advertising and subscription revenue, according to the complaints. (*Times*, FAC ¶¶ 194–97; *Daily News*, Compl. ¶¶ 229–31.) The *Times* complaint also alleges that defendants misappropriate its Wirecutter recommendations. (*Times*, FAC ¶¶ 193–94.) Defendants disagree and contend that plaintiffs' misappropriation claims are preempted by section 301 of the Copyright Act.

A. Applicable Standard

Section 301 of the Copyright Act provides for "the preemption of state law claims that are interrelated with copyright claims in certain ways.*" Nat'l Basketball Ass'n v. Motorola, Inc.*, 105 F.3d 841, 848 (2d Cir. 1997) ("*NBA*"). Specifically, a state law claim—such as the state-law tort of common law unfair competition by misappropriation—is preempted when that claim seeks to vindicate "legal or equitable rights that are equivalent" to any of the bundle of exclusive rights protected by copyright law. *See* 17 U.S.C. § 301(a). Section 301 sets forth a two-part test for determining whether a state law claim is preempted by the Copyright Act. First, the claim must "seek[] to vindicate 'legal or equitable rights that are equivalent' to one of the bundle of exclusive rights already protected by copyright law under 17 U.S.C. § 106" (the "general scope"

---

[8] In this Section VI, the Court uses the term "plaintiffs" to refer to The Times and the *Daily News* plaintiffs exclusively.

requirement); second, the work in question must be "of the type of works protected by the Copyright Act under 17 U.S.C. §§ 102 and 103" (the "subject matter" requirement). *Barclays Cap. Inc. v. Theflyonthewall.com, Inc.*, 650 F.3d 876, 892 (2d Cir. 2011) (quoting *NBA*, 105 F.3d at 848).

As a result, section 301 preempts most common law misappropriation claims involving copyrighted works. However, "certain forms of commercial misappropriation otherwise within the general scope requirement" of section 301 will survive preemption when those claims include "extra elements" instead of, or in addition to, "the acts of reproduction, performance, distribution or display . . . [that] lie within the general scope of copyright." *NBA*, 105 F.3d at 850 (citations omitted). In *NBA*, the Second Circuit held that a "hot news" misappropriation claim, originally established by the Supreme Court in *International News Service v. Associated Press*, 248 U.S. 215 (1918) ("*INS*"), contains "extra elements" and is thus not preempted by section 301 of the Copyright Act. *See NBA*, 105 F.3d at 850–53.

The exception to preemption for "hot news" misappropriation claims is narrow. *See id.* at 852; *Barclays*, 650 F.3d at 897–98. To make out a "hot news" misappropriation claim, a plaintiff must at a minimum allege the following extra elements: "(i) the time-sensitive value of factual information, (ii) the free-riding by a defendant, and (iii) the threat to the very existence of the product or service provided by the plaintiff." *NBA*, 105 F.3d at 853. In *Barclays*, the Second Circuit reiterated that "[a]n indispensable element of an *INS* 'hot news' claim is free-riding by a defendant on a plaintiff's product, enabling the defendant to produce a directly competitive product for less money because it has lower costs." *Barclays*, 650 F.3d at 902 (quoting *NBA*, 105 F.3d at 854). The court described "free-riding" as the process of taking news that the plaintiff gathered and disseminated "and selling that news as though the defendant itself had gathered it." *Id.* at 903. This "unauthorized interference" with the plaintiff's "legitimate business" occurs "*precisely at the point where the profit is to be reaped*, in order to *divert a material portion of the profit* from those who have earned it to those who have not." *Id.* at 904 (quoting *INS*, 248 U.S. at 240).

In *Barclays*, several major financial institutions brought a "hot news" misappropriation claim against the defendant—a "news aggregator" and proprietor of a digital news service—alleging that the defendant was free-riding on the plaintiffs' extensive research on publicly traded companies by summarizing the plaintiffs' reports and recommendations for existing and prospective clients and publicly disseminating them on its news platform. *Id.* at 879–80. The Second Circuit concluded that the defendant was not free riding, because it was "collecting, collating and disseminating" recommendations that the plaintiff had made while "attributing the information to its source," not selling it as its own. *Id.* at 902.

32

B.   Plaintiffs Have Failed To Plausibly Allege "Hot News" Misappropriation Claims with Respect to Their News Content and The Times's Wirecutter Recommendations.

1.   *Plaintiffs' News Content*

The "hot news" misappropriation claims brought by The Times and the *Daily news* plaintiffs fail because they do not establish that defendants are "free-riding" on plaintiffs' news content, even if that content is time-sensitive.

The prompts listed in the complaints illustrate this point. Those prompts—which query ChatGPT, Copilot, and Browse with Bing—ask the LLM to provide the text of a specific article and include the title of the article in the prompt itself. (*See, e.g., Times*, FAC ¶¶ 104, 112–21; *Daily News*, Compl. ¶¶ 98–112, 118–36.) Many prompts also provide the LLM with the name of the article's publisher. (*See e.g., Times*, FAC ¶¶ 104, 112, 121; *Daily News*, Compl. ¶¶ 98–112, 118–27.) When the prompts themselves identify the source of the article by title, author, or publication, the complaints fail to plausibly allege non-attribution—a cornerstone of the "free riding" inquiry. *See Barclays*, 650 F.3d at 903; *Fox News Network, LLC v. TVEyes, Inc.*, 43 F. Supp. 3d 379, 399 (S.D.N.Y. 2014) (dismissing hot news misappropriation claim as preempted because "TVEyes is not passing off Fox News' content as its own"), *rev'd in part on other grounds*, 883 F.3d 169 (2d Cir. 2018). Indeed, most of the outputs restate the article's name before providing the requested information. (*See, e.g., Times*, FAC ¶¶ 104, 112, 115, 118, 121; *Daily News*, Compl. ¶¶ 104, 106, 108, 118, 121, 124.)

The outputs listed in Exhibit J to the *Times* complaint and Exhibit J to the *Daily News* complaint also do not satisfy the "hot news" exception to section 301. The prompts listed in those exhibits—which query ChatGPT—fail to satisfy the time-sensitivity requirement of the "hot news" exception. Plaintiffs nowhere allege that the ChatGPT outputs in their respective Exhibits J regurgitated each article "precisely at the point where the profit [was] to be reaped." *Barclays*, 650 F.3d at 904 (quoting *INS*, 248 U.S. at 240); *see also ML Genius Holdings LLC v. Google LLC*, No. 20-3113, 2022 WL 710744, at *6 (2d Cir. Mar. 10, 2022); *Fin. Info., Inc. v. Moody's Invs. Serv., Inc.*, 808 F.2d 204, 209 (2d Cir. 1986) ("[I]mmediacy of distribution [is] necessary to sustain a 'hot' news claim."). In addition, while the prompts listed in those exhibits do not include an article's title, they do each include a direct quote from an article. Specifically, each prompt quotes a portion of an article, and the LLM outputs an additional excerpt from that same article. (*See Times*, FAC Ex. J; *Daily News*, Compl. Ex. J.) The use of direct quotes in these queries indicates that the prompter had access to at least a portion of plaintiffs' articles at the time it queried ChatGPT, again defeating any plausible concern of non-attribution.

While the outputs listed in plaintiffs' complaints may well raise copyright concerns, they do not suffer from non-attribution, which is central to the "free riding" element of a "hot news" misappropriation claim. To the contrary, plaintiffs' common law unfair competition by misappropriation claims fall squarely within the subject matter and general scope of the Copyright Act and are therefore preempted by section 301.

2.   *The Times's Wirecutter Recommendations*

Similarly, The Times's allegations of common law unfair competition by misappropriation with respect to its Wirecutter recommendations also are preempted by section 301 of the Copyright Act.

The Times's Wirecutter recommendations do not satisfy the narrow "hot news" misappropriation exception to preemption for at least two reasons. First, the Wirecutter recommendations do not constitute time-sensitive news. Wirecutter reviews are not breaking news; they are recommendations resulting from journalists spending "tens of thousands of hours each year researching and testing products to ensure that they recommend only the best." (*Times*, FAC ¶ 128.) Indeed, The Times discusses Wirecutter in the "Reviews and Analysis" section of its complaint, separate from the complaint's "Breaking News" section, and explains that Wirecutter's research is cumulative, resulting in the "produc[tion] [of] a catalog of reviews that today covers thousands of products." (*Id.* ¶ 36.) Nothing in that description suggests time-sensitivity.

The Times suggested at oral argument that its Wirecutter recommendations are nonetheless time-sensitive based on the "inherent time sensitive nature" of Wirecutter articles containing recommendations for, e.g., "Black Friday Sales" or "Christmas presents," which are published immediately preceding those events. (*See* Tr. of Oral Arg. at 82–83.) However, nothing in the complaint alleges that the harm to The Times's Wirecutter publication is caused by defendants' misappropriation of those recommendations "precisely at the point where the profit is to be reaped." *Barclays*, 650 F.3d at 904 (citation omitted). Instead, the complaint alleges that The Times suffers harm from the removal of Wirecutter's affiliate links—from which The Times receives a commission—in defendants' LLM outputs, a harm that is not time-specific. (*Times*, FAC ¶¶ 128–29.)

Second, The Times has failed to allege that its Wirecutter reviews suffer from the "free riding" of defendants. To the contrary, the outputs in the *Times* complaint that implicate Wirecutter all reference Wirecutter by name and attribute their recommendations to that publication; they do not pass off the recommendations as their own. (*See id.* ¶¶ 130–34.) Indeed, the recommendations outputted by defendants' LLMs would have little value without their attribution to Wirecutter, when the prompts specifically asked the LLMs to provide Wirecutter recommendations. *Cf. Barclays*, 650

F.3d at 903 ("It is [the defendant's] accurate attribution of the Recommendation to the creator that gives this news its value.")

It may well be true that defendants misappropriate The Times's Wirecutter recommendations, cause harm to The Times in the process, and benefit significantly from Wirecutter's recommendations without incurring the substantial costs required to generate those recommendations. But those allegations fall squarely within the general scope and subject matter of the Copyright Act, and the "extra elements" required to support a "hot-news" exception to preemption by section 301 are not present.

Accordingly, the common law unfair competition by misappropriation claims included in plaintiffs' complaints are preempted by section 301 of the Copyright Act, and defendants' motions to dismiss those claims are granted.

VII. FEDERAL TRADEMARK DILUTION

OpenAI also moves to dismiss the federal trademark claim brought by several of the *Daily News* plaintiffs, including the New York Daily News, Chicago Tribune, Mercury News, and Denver Post (collectively, the "trademark dilution plaintiffs"). The trademark dilution plaintiffs allege they are owners of several trademarks (the "Diluted Trademarks"), which are "distinctive and 'famous marks' within the meaning of Section 43(c) of the Lanham Act, 15 U.S.C. § 1125(c) and are widely recognized by the general consuming public of the United States." (*Daily News*, Compl. ¶ 235.) They allege that defendants have used the Diluted Trademarks, without authorization, "on lower-quality and inaccurate writing," thereby "dilut[ing] the quality of the Diluted Trademarks by tarnishment, in violation of [section] 1125(c)." (*Id.* ¶¶ 246–47.) OpenAI moves to dismiss the count, contending that the complaint fails to allege that the Diluted Trademarks are "famous" under section 1125(c).

A. Applicable Standard

Pursuant to 15 U.S.C. § 1125(c), "the owner of a famous mark that is distinctive" is entitled to an injunction against a party whose commercial use of that mark "is likely to cause dilution by blurring or dilution by tarnishment." 15 U.S.C. § 1125(c)(1). To state a claim for trademark dilution under section 1125(c), a plaintiff must allege: "(1) the mark is famous; (2) [the] defendant's use of the mark is made in commerce; (3) the defendant used the mark after the mark is famous; and (4) the defendant's use of the mark is likely to dilute the quality of the mark by blurring or tarnishment." *DigitAlb, Sh.a v. Setplex, LLC*, 284 F. Supp. 3d 547, 557 (S.D.N.Y. 2018) (citing *A.V.E.L.A., Inc. v. Est. of Marilyn Monroe, LLC*, 131 F. Supp. 3d 196, 211 (S.D.N.Y. 2015)). Fame is the "key ingredient" in a federal trademark dilution claim, *id.*, and section 1125(c) defines a famous mark as one that "is widely recognized by the general consuming public of the United States as a

35

designation of source of the goods or services of the mark's owner." 15 U.S.C.
§ 1125(c)(2)(A). Furthermore, section 1125(c) lays out several relevant factors a court
may consider in determining whether a mark is famous, including:

> (i) The duration, extent, and geographic reach of advertising and publicity
> of the mark, whether advertised or publicized by the owner or third parties.

> (ii) The amount, volume, and geographic extent of sales of goods or services
> offered under the mark.

> (iii) The extent of actual recognition of the mark.

> (iv) Whether the mark was registered under the Act of March 3, 1881, or the
> Act of February 20, 1905, or on the principal register.

15 U.S.C. § 1125(c)(2)(A).

In pleading federal trademark dilution, a complaint's "spare, conclusory
allegations" that the trademark holder has "expended substantial time, effort, money,
and resources advertising and promot[ing]" a trademarked product does not suffice.
CDC Newburgh Inc. v. STM Bags, LLC, 692 F. Supp. 3d 205, 235 (S.D.N.Y. 2023). Neither
does a single, conclusory allegation that a trademark is "widely recognized by the
general public." *DigitAlb*, 284 F. Supp. 3d at 558. Rather, a complaint plausibly alleges
that a trademark is famous when the allegations include attributes of the mark such as
"nationwide recognition and respect," *e.g.*, *Lewittes v. Cohen*, No. 03-cv-189, 2004 WL
1171261, at *6 (S.D.N.Y. 2004); continuous and pervasive use of the mark, *e.g.*,
A.V.E.L.A., 131 F. Supp. 3d at 216; substantial investments in promoting and advertising
the mark throughout the United States and internationally, *e.g.*, *New York City Triathlon,
LLC v. NYC Triathlon Club, Inc.*, 704 F. Supp. 2d 305, 321–22 (S.D.N.Y. 2010); "significant
publicity" relating to the marks, *e.g.*, *A.V.E.L.A.*, 131 F. Supp. 3d at 216; and that
"products bearing the [plaintiff's] marks are sold throughout the United States," *id.*

B.   **The Trademark Dilution Plaintiffs Have Plausibly Alleged that the Diluted
     Trademarks Are Famous.**

The trademark dilution plaintiffs allege that (i) each of their publications has been
in circulation for more than 100 years (*Daily News*, Compl. ¶¶ 40–42, 45, 238–41);
(ii) several of the Diluted Trademarks are federally registered (*id.* ¶ 234, Ex. I); (iii) they
collectively own over 40,000 copyright registrations for works published under the
Diluted Trademarks (*id.* ¶¶ 21, 22, 25–26); (iv) their publications are circulated
throughout all 50 states (*id.* ¶ 243); (v) their news stories are featured by major national
news outlets and have received significant attention and praise from news outlets such

as CNN, MSNBC and Fox News (*id.* ¶¶ 9, 240); (vi) their publications have achieved national and international fame for their reporting of highly significant events in both U.S. and world history (*id.* ¶¶ 40–45); (vii) they invest hundreds of millions of dollars in operating their publications (*id.* ¶¶ 184, 227); (viii) millions of consumers access the trademark dilution plaintiffs' publications in print and digital format, which are circulated under the Diluted Trademarks (*id.* ¶ 242); (ix) their publications have widespread circulation across a general audience in the United States (*id.* ¶¶ 238–241); and (x) their publications have received widespread recognition for their achievements including numerous Pulitzer Prizes, which constitute the most prestigious and highly publicized national journalism award. In particular, the New York Daily News has received 11 Pulitzers (*id.* ¶ 238); The Chicago Tribune has received eight (*id.* ¶ 239); The Denver Post has received nine (*id.* ¶ 242); and The Mercury News has received two. (*Id.* ¶ 42).

These allegations are a far cry from the threadbare and conclusory statements that doomed the trademark dilution cases cited by OpenAI in its motion. *Cf. DigitAlb*, 284 F. Supp. 3d at 557–58; *Glob. Brand Holdings, LLC v. Church & Dwight Co.*, No. 17-cv-6571, 2017 WL 6515419, at *5 (S.D.N.Y. Dec. 19, 2017); *Heller Inc. v. Design Within Reach, Inc.*, No. 09-cv-1909, 2009 WL 2486054, at *4 (S.D.N.Y. Aug. 14, 2009); *CDC Newburgh Inc.*, 692 F. Supp. 3d at 235.

By contrast, the allegations of the trademark dilution plaintiffs—which include detailed, factual descriptions of the nature and scope of the Diluted Trademarks' widespread circulation, recognition, achievements, and consumer subscriptions— "are sufficient to constitute a pleading that [the Diluted Trademarks are] 'famous' within the meaning of the statute." *Lewittes*, 2004 WL 1171261, at *6. The Court therefore denies OpenAI's motion to dismiss the federal trademark dilution claim in the *Daily News* action.

VIII.   NEW YORK STATE TRADEMARK DILUTION

Microsoft's challenge to the state trademark dilution claim in the *Daily News* action fares no better. That claim alleges that defendants' activities dilute the distinctiveness of the Diluted Trademarks and injure the trademark dilution plaintiffs' business reputations in violation of New York General Business Law § 360-*l*. According to the *Daily News* complaint, the dilution occurs when the outputs from defendants' LLMs falsely attribute the content of the output to the trademark dilution plaintiffs. Microsoft challenges the claim as barred by the dormant Commerce Clause of the United States Constitution.

A.   Applicable Standard

The Commerce Clause of the U.S. Constitution, art. I, § 8, cl. 3, "not only vests Congress with the power to regulate interstate trade; the Clause also contains a further, negative command, one effectively forbidding the enforcement of certain state economic regulations even when Congress has failed to legislate on the subject." *Nat'l Pork Producers Council v. Ross*, 598 U.S. 356, 368 (2023) (internal quotation marks and alterations omitted) (citing *Okla. Tax Comm'n v. Jefferson Lines, Inc.*, 514 U.S. 175, 179 (1995)). This negative command has come to be known as the dormant Commerce Clause. It prevents a state from "us[ing] its laws to discriminate purposefully against out-of-state economic interests." *Id.* at 364. Pursuant to this "antidiscrimination principle," which "lies at the very core of [the Supreme Court's] dormant Commerce Clause jurisprudence," states are prohibited from enforcing state laws motivated by "economic protectionism—that is, regulatory measures designed to benefit in-state economic interests by burdening out-of-state competitors." *Id.* at 369 (internal quotation marks and citation omitted). By the same token, "absent discrimination, a State may exclude from its territory, or prohibit the sale therein of any articles which, in its judgment, fairly exercised, are prejudicial to the interests of its citizens." *Id.* (internal quotation marks omitted).

For laws that do not purposefully discriminate against out-of-state competitors but nevertheless "incidentally burden[] interstate commerce," courts apply "a more permissive balancing test" set forth in *Pike v. Bruce Church, Inc.*, 397 U.S. 137 (1970). *Rest. L. Ctr. v. City of N.Y.*, 90 F.4th 101, 118 (2d Cir. 2024). Under the *Pike* balancing test, a state law will be struck down only "if the burden imposed on interstate commerce clearly exceeds the putative local gains." *Id.* at 118 (quoting *Town of Southold v. Town of E. Hampton*, 477 F.3d 38, 47 (2d Cir. 2007)). The *Pike* analysis primarily "serves as an important reminder that a law's practical effects may also disclose the presence of a discriminatory purpose," *Nat'l Pork*, 598 U.S. at 377, and the *Pike* test is "most frequently deployed to detect the presence or absence of latent economic protectionism." *Id.* at 391 (Sotomayor, J., concurring). However, a state law that appears genuinely nondiscriminatory may still violate the dormant Commerce Clause if its burdens on commerce "clearly outweigh the benefits of a state or local practice." *Dep't of Revenue of Ky. v. Davis*, 553 U.S. 328, 353 (2008).

B. The New York Dilution Statute Does Not Violate the Dormant Commerce Clause.

1. *N.Y. Gen. Bus. Law § 360-l Does Not Discriminate Against Out-of-State Commerce.*

Microsoft contends that the trademark dilution plaintiffs are not merely seeking to enforce a state regulation that has "practical extraterritorial effects," but that they are seeking a ruling that *directly regulates* out-of-state commerce. (*See Daily News*, ECF No. 105 at 10.) That is not correct. Section 360-*l* provides grounds for injunctive relief for the dilution of a mark used on goods "sold or transported in commerce *in this state.*" N.Y. Gen. Bus. Law § 360(h) (emphasis added); *id.* §§ 360(a), 360(c), 360-*l*. *Cf. Healy v. Beer Inst., Inc.*, 491 U.S. 324, 337 (1989) (invalidating law requiring out-of-state beer merchants to affirm that their in-state prices were no higher than their out-of-state prices); *Brown-Forman Distillers Corp. v. N.Y. State Liquor Auth.*, 476 U.S. 573, 580 (1986) (invalidating price affirming law imposed on out-of-state liquor distillers).

To the extent Microsoft posits that the "extraterritorial effects" of section 360-*l* amount to an effective regulation of wholly out-of-state commerce, the Supreme Court in *National Pork Producers Council v. Ross* rejected this theory, which would create an "almost *per se* rule forbidding enforcement of state laws that have the practical effect of controlling commerce outside the State, even when those laws do not purposely discriminate against out-of-state economic interests." 598 U.S. at 371 (internal quotation marks and citation omitted). The Supreme Court wrote that such a theory would lead to "strange places" when, "[i]n our interconnected national marketplace, many (maybe most) state laws have the practical effect of controlling extraterritorial behavior." *Id.* at 374 (internal quotation marks omitted). The Supreme Court confirmed that whether a state statute violates the dormant Commerce Clause depends on whether the statute discriminates against out-of-state economic interests, or in rare cases, whether a nondiscriminatory state law substantially burdens interstate commerce in clear excess of the local benefits. *Id.* at 391 (Sotomayor, J. concurring); *id.* at 394 (Roberts, C.J., concurring).

Section 360-*l* "plainly does not facially discriminate against interstate commerce[,] [n]or does it harbor a discriminatory purpose." *Rest. L. Ctr.*, 90 F.4th at 106 (upholding nondiscriminatory state statute against dormant Commerce Clause challenge). Microsoft has not identified any out-of-New York-state competitor that would be disadvantaged vis-a-vis New York competitors due to section 360-*l*, nor has it articulated a discriminatory purpose against out-of-state competitors underlying the statute. Section 360-*l* does not "erect[] an economic barrier protecting a major local industry against competition from without the State," *Dean Milk Co. v. Madison*, 340 U.S.

349, 354 (1951); operate like "a tariff or customs duty" in order to protect New York competitors from outside competition, *W. Lynn Creamery, Inc. v. Healy*, 512 U.S. 186, 194 (1994); or "deliberately rob[]" out-of-state competitors of "whatever competitive advantages they may possess" over in-state competitors. *Nat'l Pork*, 598 U.S. at 374 (citation omitted). Quite the opposite, section 360-*l* fits within the myriad laws in our "interconnected national marketplace,"—including "libel laws, securities requirements, charitable registration requirements, franchise laws, tort laws," as well as "inspection laws, quarantine laws, and health laws of every description,"—that have "the practical effect of controlling extraterritorial behavior" but do not violate the Commerce Clause. *Id.* at 374–75 (cleaned up). The New York dilution statute may "have a considerable influence on commerce outside [New York's] borders," *id.* at 375, but Microsoft has not contended much less shown that section 360-*l* discriminates against out-of-state commerce.

2.   *N.Y. Gen. Bus. Law § 360-l Does Not Substantially Burden Interstate Commerce in Clear Excess of Its Local Benefits.*

As set forth above, the *Pike* balancing test invalidates state laws whose incidental burdens on interstate commerce clearly exceed the law's putative local gains. *Rest. L. Ctr.*, 90 F.4th at 118. In its motion to dismiss, Microsoft does not attempt to argue that section 360-*l* fails the *Pike* balancing test. For good reason: "*Pike*'s balancing and tailoring principles are most frequently deployed to detect the presence or absence of latent economic protectionism," which, as discussed, is not present in this case—nor does section 360-*l* impose a "substantial burden on interstate commerce" in clear excess of its local gains. *Nat'l Pork*, 598 U.S. at 391 (Sotomayor, J., concurring).

Accordingly, Microsoft's motion to dismiss the trademark dilution plaintiffs' state dilution claim is denied, because Microsoft has not shown that N.Y. Gen. Bus. Law § 360-*l* discriminates against out of state commerce or that its incidental burdens on interstate commerce are in clear excess of its local benefits.

IX.   DIRECT COPYRIGHT INFRINGEMENT INVOLVING ABRIDGEMENTS

Finally, in alleging direct infringement in violation of 17 U.S.C. § 501, the *CIR* complaint distinguishes between defendants' "regurgitations" of CIR's works (*CIR*, FAC ¶¶ 79–85), and their "abridgments" of those works (*id.* ¶¶ 86–98). Although CIR's definition of an "abridgment" is somewhat opaque (*see id.*), the term appears to refer to outputs that are detailed summaries of CIR articles, "often in the format of a bulleted list of main points." (*Id.* ¶ 91.)

OpenAI has moved to dismiss CIR's direct infringement claim to the extent it relates to alleged abridgments, on the grounds that the alleged abridgments, on their

face, are not "substantially similar to protected expression in the articles at issue—a necessary element of a copyright infringement claim." (*CIR*, ECF No. 146 at 1.)

A. Applicable Standard

"To establish infringement, the copyright owner must demonstrate that (1) the defendant has actually copied the plaintiff's work; *and* (2) the copying is illegal because a substantial similarity exists between the defendant's work and the protectible elements of plaintiff's." *Yurman Design, Inc. v. PAJ, Inc.*, 262 F.3d 101, 110 (2d Cir. 2001) (internal quotation marks and citation omitted); *see also Abdin v. CBS Broad. Inc.*, 971 F.3d 57, 66 (2d Cir. 2020). "The standard test in determining substantial similarity is the 'ordinary observer test': whether an average lay observer would overlook any dissimilarities between the works and would conclude that one was copied from the other." *Nihon Keizai Shimbun, Inc. v. Comline Bus. Data, Inc.*, 166 F.3d 65, 70 (2d Cir. 1999). When a work contains both protectible and unprotectible elements, a "more discerning" ordinary observer test applies, which asks whether there exists "substantial similarity between those elements, and only those elements, that provide copyrightability to the allegedly infringed [work]." *Boisson v. Banian, Ltd.*, 273 F.3d 262, 272 (2d Cir. 2001); *see also Abdin*, 971 F.3d at 66.

The application of the "ordinary observer" test "is by no means exclusively reserved for resolution by a jury." *Peter F. Gaito Architecture, LLC v. Simone Dev. Corp.*, 602 F.3d 57, 63 (2d Cir. 2010). Indeed, "it is entirely appropriate for a district court to resolve that question as a matter of law, either because the similarity between two works concerns only non-copyrightable elements of the plaintiff's work, or because no reasonable jury, properly instructed, could find that the two works are substantially similar." *Id.* (internal quotation marks and citations omitted). Numerous district courts have resolved the substantial similarity question at the pleading stage for one or both of these reasons. *See, e.g.*, *Montgomery v. Holland*, 408 F. Supp. 3d 353 (S.D.N.Y. 2019), *aff'd sub nom. Montgomery v. NBC Television*, 833 F. App'x 361 (2d Cir. 2020); *Nobile v. Watts*, 289 F. Supp. 3d 527 (S.D.N.Y. 2017), *aff'd*, 747 F. App'x 879 (2d Cir. 2018); *Piuggi v. Good for You Prods. LLC*, 739 F. Supp. 3d 143, 162 (S.D.N.Y. 2024).

Works are not substantially similar as a matter of law when their only similarity concerns general or underlying facts and ideas, which are not copyrightable. *See* 17 U.S.C. § 102(b); *see also Abdin*, 971 F.3d at 67–68. However, while facts themselves are not copyrightable, factual compilations "may possess the requisite originality" for copyright protection. *Feist Publ'ns, Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340, 348 (1991). In such cases, where "the compilation author clothes facts with an original collocation of words, he or she may be able to claim a copyright in this written expression," and "[o]thers may copy the underlying facts from the publication, but not the precise words

used to present them." *Id.* By the same token, when facts are reported "in a different arrangement, with a different sentence structure and different phrasing," the secondary work does not "purloin protected expression," and no copyright infringement has ensued. *Nihon*, 166 F.3d at 71.

B. The Abridgments Contained in the *CIR* Complaint Are Not Substantially Similar to CIR's Copyrighted Works as a Matter of Law.

Exhibit 11 to the *CIR* complaint provides website links to articles that CIR alleges were unlawfully abridged by defendants in their ChatGPT and Copilot outputs. (*CIR*, FAC Ex. 11.) Examining the similarities between those outputs and the corresponding CIR articles, including the "total concept and feel, theme . . . sequence, pace, and setting," *Williams v. Crichton*, 84 F.3d 581, 588 (2d Cir. 1996), the Court concludes that the "abridgments" contained in Exhibit 11 are not substantially similar to CIR's copyrighted works as a matter of law.

The alleged abridgments are detailed summaries, usually in bullet point form, of the facts contained in CIR's articles. Those summaries—which differ in style, tone, length, and sentence structure from CIR's articles—are not "substantially similar" to CIR's copyrighted works. They present the "facts in a different arrangement"—bullet point lists or short summary paragraphs—"with a different sentence structure and different phrasing." *Nihon*, 166 F.3d at 71. In short, the abridgments in Exhibit 11 are not substantially similar, qualitatively or quantitatively, to the original CIR articles as a matter of law. The Court therefore grants OpenAI's motion to dismiss CIR's claim of direct infringement under 17 U.S.C. § 501 insofar as it relates to the "abridgments" contained in Exhibit 11.

X. CONCLUSION
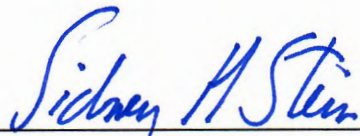
For the foregoing reasons, the Court denies (1) OpenAI's motions to dismiss the direct infringement claims involving conduct occurring more than three years before the complaints were filed; (2) defendants' motions to dismiss the contributory copyright infringement claims; and (3) defendants' motions to dismiss the state and federal trademark dilution claims in the *Daily News* action.

The Court grants defendants' motions to dismiss the common law unfair competition by misappropriation claims and OpenAI's motion to dismiss the "abridgment" claims in the *CIR* action, and dismisses each of those claims with prejudice.

With respect to the DMCA claims, the Court grants (1) Microsoft's motions to dismiss the 17 U.S.C. § 1202(b)(1) claims against it in all three actions, (2) OpenAI's motion to dismiss the section 1202(b)(1) claim against it in the *Times* action, and

(3) defendants' motions to dismiss the section 1202(b)(3) claims against them in all three actions, and dismisses each claim without prejudice. The Court denies OpenAI's motions to dismiss the section 1202(b)(1) claims against it in the *Daily News* and *CIR* actions.

Dated: New York, New York
        April 4, 2025

Sidney H. Stein

Sidney H. Stein, U.S.D.J.

United States Court of Appeals
FOR THE DISTRICT OF COLUMBIA CIRCUIT

———

Argued September 19, 2024   Decided March 18, 2025

No. 23-5233

STEPHEN THALER, AN INDIVIDUAL,
APPELLANT

v.

SHIRA PERLMUTTER, IN HER OFFICIAL CAPACITY AS REGISTER
OF COPYRIGHTS AND DIRECTOR OF THE UNITED STATES
COPYRIGHT OFFICE AND U.S. COPYRIGHT OFFICE,
APPELLEES

———

Appeal from the United States District Court
for the District of Columbia
(No. 1:22-cv-01564)

———

Ryan Abbott argued the cause for appellant. With him on
the briefs was *Timothy G. Lamoureux*.

Ryan N. Phelan was on the brief for *amici curiae* Legal
Professors Shlomit Yanisky-Ravid, *et al*. in support of
appellant.

Nicholas S. Crown, Attorney, U.S. Department of Justice,
argued the cause for appellees. With him on the brief were
Brian M. Boynton, Principal Deputy Assistant Attorney

2

General at the time the brief was filed, *Daniel Tenny*, Attorney, and *Emily L. Chapuis*, Deputy General Counsel, U.S. Copyright Office.

Before: MILLETT and WILKINS, *Circuit Judges*, and ROGERS, *Senior Circuit Judge*.

Opinion for the Court filed by *Circuit Judge* MILLETT.

MILLETT, *Circuit Judge*: This case presents a question made salient by recent advances in artificial intelligence: Can a non-human machine be an author under the Copyright Act of 1976? The use of artificial intelligence to produce original work is rapidly increasing across industries and creative fields. Who—or what—is the "author" of such work is a question that implicates important property rights undergirding economic growth and creative innovation.

In this case, a computer scientist attributes authorship of an artwork to the operation of software. Dr. Stephen Thaler created a generative artificial intelligence named the "Creativity Machine." The Creativity Machine made a picture that Dr. Thaler titled "A Recent Entrance to Paradise." Dr. Thaler submitted a copyright registration application for "A Recent Entrance to Paradise" to the United States Copyright Office. On the application, Dr. Thaler listed the Creativity Machine as the work's sole author and himself as just the work's owner.

The Copyright Office denied Dr. Thaler's application based on its established human-authorship requirement. This policy requires work to be authored in the first instance by a human being to be eligible for copyright registration. Dr.

3

Thaler sought review of the Office's decision in federal district court and that court affirmed.

We affirm the denial of Dr. Thaler's copyright application. The Creativity Machine cannot be the recognized author of a copyrighted work because the Copyright Act of 1976 requires all eligible work to be authored in the first instance by a human being. Given that holding, we need not address the Copyright Office's argument that the Constitution itself requires human authorship of all copyrighted material. Nor do we reach Dr. Thaler's argument that he is the work's author by virtue of making and using the Creativity Machine because that argument was waived before the agency.

I

A

The Constitution's Intellectual Property Clause gives Congress authority to "promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries[.]" U.S. CONST. Art. I, § 8, cl. 8. Under that provision, federal copyright protection extends only as far as Congress designates by statute. *Wheaton v. Peters*, 33 U.S. 591, 661 (1834).

Copyright law incentivizes the creation of original works so they can be used and enjoyed by the public. Since the founding, Congress has given authors short term monopolies over their original work. *See* Act of May 31, 1790, ch. 15, 1st Cong., 1 Stat. 124. This protection is not extended as "a special reward" to the author, but rather "to encourage the production of works that others might reproduce more cheaply." *Google LLC v. Oracle Am., Inc.*, 593 U.S. 1, 16 (2021). By ensuring

4

that easily reproducible work is protected, individuals are incentivized to undertake the effort of creating original works that otherwise would be easily plagiarized.

The Copyright Act of 1976 is the current federal copyright statute. Three of its provisions are relevant here.

First, the Copyright Act preempts state common law copyright protection by immediately vesting federal copyright ownership in a work's author as soon as a work is created. 17 U.S.C. §§ 102(a); 201(a); 301(a). Although domestic authors generally must register their copyrights to exercise other rights, like the right to sue for infringement, *id.* § 411(a), the right to own a copyright does not depend on registration or publication.

Second, the Copyright Act incentivizes authors by protecting their work "for a term consisting of the life of the author and 70 years after the author's death." 17 U.S.C. § 302(a). In that way, authors are encouraged to produce work because they know that they can profit from it for their entire life and that their heirs and assigns can continue to benefit for seven decades thereafter.

Third, individuals and organizations can own copyrights by hiring someone to create work. The Copyright Act's work-made-for-hire provision allows "the employer or other person for whom the work was prepared" to be "considered the author" and "own[] all of the rights comprised in the copyright." 17 U.S.C. § 201(b). Rather than enduring for the author's lifetime, a work-made-for-hire copyright lasts "95 years from the year of its first publication, or a term of 120 years from the year of its creation, whichever expires first." *Id.* § 302(c).

5

B

The Copyright Act is administered by the United States Copyright Office. 17 U.S.C. § 701(a). That Office has a duty to "[a]dvise Congress" on issues "relating to copyright," to "[p]rovide information and assistance" to "Federal departments and agencies and the Judiciary," and to "[c]onduct studies and programs regarding copyright[.]" *Id.* § 701(b)(1), (2), (4).

In addition, the Copyright Office has authority to establish regulations to implement the Copyright Act. 17 U.S.C. § 702. Pursuant to that authority, the Copyright Office issues regulations governing the "conditions for the registration of copyright, and the application to be made for registration[.]" 37 C.F.R. § 202.3(a)(1). The Copyright Office publishes these registration regulations in the *Compendium of Copyright Office Practices* to inform authors about registration criteria for different types of work. *See* Copyright Office, *Compendium of U.S. Copyright Office Practices* (3d ed. 2021), https://perma.cc/9N9N-C3VU (*Compendium Third Edition*).

Individuals whose registration applications are denied can seek reconsideration by the Copyright Office's Registration Program. If still dissatisfied, they can ask the Copyright Office's Review Board to reconsider their case. 37 C.F.R. § 202.5(b), (c). A decision by the Review Board "constitutes final agency action," *id.* § 202.5(g), and is reviewable under the Administrative Procedure Act, 5 U.S.C. § 704; 17 U.S.C. § 701(e).

Copyright Office regulations have long required that any registered work be authored by a human. *See* Copyright Office, *Compendium of Copyright Office Practices* § 2.8.3(I), (I)(a)(1)(b) (1st ed. 1973), https://perma.cc/J7ML-BZK6

6

(*Compendium First Edition*) ("[N]othing can be considered the 'writing of an author'" unless it owes its "origin to a human agent[.]"); Copyright Office, *Compendium of Copyright Office Practices* § 202.02(b) (2d ed. 1984), https://perma.cc/52MX-6YPD (*Compendium Second Edition*) ("The term "authorship" implies that, for a work to be copyrightable, it must owe its origin to a human being."). The current *Compendium* advises that the Copyright Office "will refuse to register a claim if it determines that a human being did not create the work." *Compendium Third Edition* § 306. That refusal extends to works "produced by a machine or mere mechanical process that operates randomly or automatically without any creative input or intervention from a human author." *Id.* § 313.2

C

1

Dr. Thaler is a computer scientist who creates and works with artificial intelligence systems, Thaler Opening Br. ii, and who invented the Creativity Machine, *id.* 43-44. On May 19, 2019, Dr. Thaler submitted a copyright registration application to the Copyright Office for an artwork titled "A Recent Entrance to Paradise." J.A. 43. On the application, Dr. Thaler listed the "Author" of that work as the "Creativity Machine." J.A. 43. Under "Copyright Claimant," Dr. Thaler provided his own name. J.A. 43. In the section labeled "Author Created," Dr. Thaler wrote "2-D artwork, Created autonomously by machine." J.A. 43.

The Copyright Office denied Dr. Thaler's application because "a human being did not create the work." J.A. 45. The letter cited the Supreme Court's decision in *Burrow-Giles Lithographic Co. v. Sarony*, 111 U.S. 53 (1884), in support of its decision. J.A. 45.

7

In seeking reconsideration by the Registration Program, Dr. Thaler acknowledged the Copyright Office's decision "was made on the basis that the present submission lacks human authorship[.]" J.A. 49. Dr. Thaler confirmed this "is correct" and "that the present submission lacks traditional human authorship—it was autonomously generated by an AI." J.A. 49. Dr. Thaler then argued that "the Human Authorship Requirement is unconstitutional and unsupported by either statute or case law." J.A. 49. Dr. Thaler claimed judicial opinions "from the Gilded Age" could not settle the question of whether computer generated works are copyrightable today. J.A. 55.

The Registration Program again denied Dr. Thaler's application because the work lacked "sufficient creative input or intervention from a human author." J.A. 59.

In his request for reconsideration by the Review Board, Dr. Thaler reaffirmed that "the present submission lacks traditional human authorship—it was autonomously generated by an AI." J.A. 63. He then reiterated his constitutional, statutory, and policy arguments against the human-authorship requirement. J.A. 63-69. Dr. Thaler also argued he should own the copyright under the work-made-for-hire doctrine because "non-human, artificial persons such as companies can already be authors under this doctrine." J.A. 66.

The Review Board affirmed the denial of Dr. Thaler's copyright application based on the human-authorship requirement. J.A. 73. The Board relied upon Dr. Thaler's "representation that the Work was autonomously created by artificial intelligence without any creative contribution from a human actor[.]" J.A. 72. The Board also rejected Dr. Thaler's argument that the work was made for hire on the ground that

8

there was no contract between Dr. Thaler and the Creativity Machine. J.A. 76-77.

2

Dr. Thaler sought review in the United States District Court for the District of Columbia, and both sides moved for summary judgment. *Thaler v. Perlmutter*, 687 F. Supp. 3d 140, 142 (D.D.C. 2023). In his motion, Dr. Thaler asserted the same constitutional, statutory, and policy arguments that he had advanced before the agency, including the argument that he owns the copyright under the work-made-for-hire provision. J.A. 80-115. In addition, he claimed for the first time that the work is copyrightable because a human—Dr. Thaler— "provided instructions and directed his AI[.]" J.A. 113.

The district court affirmed the Copyright Office's denial of registration. Based on the caselaw and the Copyright Act's text, the district court concluded that "[h]uman authorship is a bedrock requirement of copyright." *Thaler*, 687 F. Supp. 3d at 146. The court also held that Dr. Thaler could not rely on the work-made-for-hire provision because that provision "presuppose[s] that an interest exists to be claimed." *Id.* at 150. The "image autonomously generated" by the Creativity Machine was not such an interest because it "was never eligible for copyright," so the Machine had no copyright to transfer to Dr. Thaler even if he were the Creativity Machine's employer. *Id.* Finally, the court found that Dr. Thaler waived his argument that he should own the copyright because he created and used the Creativity Machine. The court stressed that, "[o]n the record designed by plaintiff from the outset of his application for copyright registration," the case had presented "only the question of whether a work generated autonomously by a computer system is eligible for copyright." *Id.* at 149-150.

9

II

We review a district court's grant of summary judgment in a case concerning agency action *de novo* and, like the district court, will set aside the agency action only if it is "arbitrary, capricious, an abuse of discretion, or otherwise not in accordance with law[.]'" *Jicarilla Apache Nation v. United States Dep't of Interior*, 613 F.3d 1112, 1118 (D.C. Cir. 2010) (quoting 5 U.S.C. § 706(2)(A)). We "exercise independent judgment in determining the meaning of statutory provisions." *Loper Bright Enterprises v. Raimondo*, 603 U.S. 369, 394 (2024).

The district court had jurisdiction under 28 U.S.C. § 1331. This court has jurisdiction under 28 U.S.C. § 1291.

III

As a matter of statutory law, the Copyright Act requires all work to be authored in the first instance by a human being. Dr. Thaler's copyright registration application listed the Creativity Machine as the work's sole author, even though the Creativity Machine is not a human being. As a result, the Copyright Office appropriately denied Dr. Thaler's application.

A

Authors are at the center of the Copyright Act. A copyright "vests initially in the author or authors of the work." 17 U.S.C. § 201(a). And copyright protection only "subsists * * * in original works of authorship[.]" *Id.* § 102(a).

The Copyright Act does not define the word "author." But traditional tools of statutory interpretation show that, within the meaning of the Copyright Act, "author" refers only to human

10

beings. To start, the text of multiple provisions of the statute indicates that authors must be humans, not machines. In addition, the Copyright Office consistently interpreted the word author to mean a human prior to the Copyright Act's passage, and we infer that Congress adopted the agency's longstanding interpretation of the word "author" when it re-enacted that term in the 1976 Copyright Act.

1

Numerous Copyright Act provisions both identify authors as human beings and define "machines" as tools used by humans in the creative process rather than as creators themselves. Because many of the Copyright Act's provisions make sense only if an author is a human being, the best reading of the Copyright Act is that human authorship is required for registration.

First, the Copyright Act's ownership provision is premised on the author's legal capacity to hold property. A copyright "vests initially in the author[.]" 17 U.S.C. § 201(a). This means an "author gains 'exclusive rights' in her work immediately upon the work's creation." *Fourth Estate Pub. Benefit Corp. v. Wall-Street.com, LLC*, 586 U.S. 296, 300-301, (2019) (quoting 17 U.S.C. § 106). Because a copyright is fundamentally a property right created by Congress, and Congress specified that authors immediately own their copyrights, an entity that cannot own property cannot be an author under the statute.

Second, the Copyright Act limits the duration of a copyright to the author's lifespan or to a period that approximates how long a human might live. A copyright generally "endures for a term consisting of the life of the author and 70 years after the author's death." 17 U.S.C. § 302(a). The

11

Copyright Office maintains "current records of information relating to the death of authors of copyrighted works" so that it can determine when copyrights expire. *Id.* § 302(d). If the author's death is unknown, the Copyright Act presumes death after "a period of 95 years from the year of first publication of a work, or a period of 120 years from the year of its creation[.]" *Id.* § 302(e). And even when a corporation owns a copyright under the work-made-for-hire provision, the copyright endures for the same amount of time—"95 years from the year of first publication" or "120 years from the year of its creation[.]" *Id.* § 302(c). Of course, machines do not have "lives" nor is the length of their operability generally measured in the same terms as a human life.

Third, the Copyright Act's inheritance provision states that, when an author dies, that person's "termination interest is owned, and may be exercised" by their "widow or widower," or their "surviving children or grandchildren," 17 U.S.C. § 203(a)(2), (A). Machines, needless to say, have no surviving spouses or heirs.

Fourth, copyright transfers require a signature. To transfer copyright ownership, there must be "an instrument of conveyance" that is "signed by the owner[.]" 17 U.S.C. § 204(a). Machines lack signatures, as well as the legal capacity to provide an authenticating signature.

Fifth, authors of unpublished works are protected regardless of the author's "nationality or domicile." 17 U.S.C. § 104(a). Machines do not have domiciles, nor do they have a national identity.

Sixth, authors have intentions. A joint work is one "prepared by two or more authors with the intention that their contributions be merged into inseparable or interdependent

12

parts of a unitary whole." 17 U.S.C. § 101. Machines lack minds and do not intend anything.

Seventh, and by comparison, every time the Copyright Act discusses machines, the context indicates that machines are tools, not authors. For example, the Copyright Act defines a "computer program" as "a set of statements or instructions to be used directly or indirectly" to "bring about a certain result." 17 U.S.C. § 101. The word "machine" is given the same definition as the words "device" and "process," *id*., and those terms are consistently used in the statute as mechanisms that assist authors, rather than as authors themselves, *id*. §§ 102(a); 108(c)(2); 109(b)(1)(B)(i); 116(d)(1); 117(a)(1), (c); 401(a); 1001(2), (3). In addition, when computer programs and machines are referenced in the statute, the statute presumes they have an "owner," *id.* § 117(a), (c), who can perform "maintenance," "servic[e]," or "repair" on them, *id.* § 117(d)(1), (2).

All of these statutory provisions collectively identify an "author" as a human being. Machines do not have property, traditional human lifespans, family members, domiciles, nationalities, *mentes reae,* or signatures. By contrast, reading the Copyright Act to require human authorship comports with the statute's text, structure, and design because humans have all the attributes the Copyright Act treats authors as possessing. The human-authorship requirement, in short, eliminates the need to pound a square peg into a textual round hole by attributing unprecedented and mismatched meanings to common words in the Copyright Act. *See Food & Drug Admin. v. Brown & Williamson Tobacco Corp.*, 529 U.S. 120, 133 (2000) ("It is a 'fundamental canon of statutory construction that the words of a statute must be read in their context and with a view to their place in the overall statutory

13

scheme.'") (quoting *Davis v. Michigan Dept. of Treasury*, 489 U.S. 803, 809 (1989)).

To be clear, we do not hold that any one of those statutory provisions states a necessary condition for someone to be the author of a copyrightable work. An author need not have children, nor a domicile, nor a conventional signature. Even the ability to own property has not always been required for copyright authorship. Married women in the nineteenth century authored work that was eligible for copyright protection even though coverture laws forbade them from owning copyrights. *See* Melissa Homestead, AMERICAN WOMEN AUTHORS AND LITERARY PROPERTY, 1822-1869, at 21-62 (2005); *Belford, Clarke & Co. v. Scribner*, 144 U.S. 488, 504 (1892) (recognizing Mrs. Terhune's authorship when her book's copyright was infringed, even though, as a married woman, she could not own property).

The point, instead, is that the current Copyright Act's text, taken as a whole, is best read as making humanity a necessary condition for authorship under the Copyright Act. That is the reading to which "the provisions of the whole law" point. *John Hancock Mut. Life Ins. Co. v. Harris Tr. & Sav. Bank*, 510 U.S. 86, 94 (1993) (quoting *Pilot Life Ins. Co. v. Dedeaux*, 481 U.S. 41, 51 (1987)).

2

The Copyright Office's longstanding rule requiring a human author reinforces the natural meaning of those statutory terms.

The Copyright Office first addressed whether machines could be authors in 1966—ten years before the Copyright Act of 1976 was passed. That year, the Register of Copyrights

14

wrote in the Copyright Office's annual report to Congress that, as "computer technology develops and becomes more sophisticated, difficult questions of authorship are emerging. * * * The crucial question appears to be whether the 'work' is basically one of human authorship, with the computer merely being an assisting instrument[.]" Copyright Office, *Sixty-Eighth Annual Report of the Register of Copyrights* at 5 (1966), https://perma.cc/QU7P-TY6N.

The Copyright Office formally adopted the human authorship requirement in 1973. That year, the Copyright Office updated its regulations to state explicitly that works must "owe their origin to a human agent[.]" *Compendium First Edition* § 2.8.3(I)(a)(1)(b).

In 1974, Congress created the National Commission on New Technological Uses of Copyrighted Works ("CONTU") to study how copyright law should accommodate "the creation of new works by the application or intervention of such automatic systems or machine reproduction." Pub. L. 93-573, § 201(b)(2), 88 Stat. 1873 (1974). CONTU assembled copyright experts from the government, academia, and the private sector to make recommendations to Congress. Prior to the Copyright Act's passage, the Library of Congress published summaries of CONTU's meetings, several of which focused on copyright law and computer technology. In none of these meetings did members of CONTU suggest that computers were authors rather than tools used by authors to create original work. *See* CONTU, *Meeting No. 2* at 10-11 (Nov. 19, 1975), https://perma.cc/857K-VRSB; CONTU, *Meeting No. 3* at 1-11 (Dec. 18-19, 1975), https://perma.cc/EB3T-KNR4; CONTU, *Meeting No. 4* at 1-8 (Feb. 11-13, 1976), https://perma.cc/NPG6-J8E3; CONTU, *Meeting No. 6* (May 6-7, 1976), https://perma.cc/HCX5-6ZYX; CONTU, *Meeting*

15

No. 7 at 46-148 (June 9-10, 1976), https://perma.cc/Q795-YVQ4.

This understanding of authorship and computer technology is reflected in CONTU's final report:

> On the basis of its investigations and society's experience with the computer, the Commission believes that there is no reasonable basis for considering that a computer in any way contributes authorship to a work produced through its use. The computer, like a camera or a typewriter, is an inert instrument, capable of functioning only when activated either directly or indirectly by a human. When so activated it is capable of doing only what it is directed to do in the way it is directed to perform.

CONTU, *Final Report* at 44 (1978), https://perma.cc/7S8T-TAB5.

Although CONTU's final report was not published until 1978, its conclusion that machines cannot be authors reflects the state of play at the time Congress enacted the Copyright Act in 1976. And when Congress amended the Copyright Act's provision governing computer programs shortly following CONTU's final report, Congress preserved the Act's provisions governing authorship and the language describing machines as devices used by authors. Pub. L. No. 96-517, 94 Stat. 3015, 3028 (1980) (stating it is not infringement to copy a computer program if the copy "is created as an essential step in the utilization of the computer program in conjunction with a machine[.]").

In short, at the time the Copyright Act was passed and for at least a decade before, computers were not considered to be capable of acting as authors, but instead served as "inert

16

instrument[s]" controlled "directly or indirectly by a human" who could be an author. CONTU, *Final Report* at 44 (1978), https://perma.cc/7S8T-TAB5. We infer Congress adopts an agency's interpretation of a term "when a term's meaning was well-settled[.]" *Sackett v. Environmental Prot. Agency*, 598 U.S. 651, 683 (2023). And that rule applies with double force here where the commission Congress designated to study the issue, CONTU, came to the same conclusion. Given all that, the interpretation of "author" as requiring human authorship was well-settled at the time the 1976 Copyright Act was enacted.

3

Dr. Thaler's contrary reading of the statutory text fails.

a

Dr. Thaler argues first that the natural meaning of "author" is not confined to human beings. Dr. Thaler points to a 2023 dictionary definition defining "author" as "one that originates or creates something[.]" Thaler Opening Br. 23 (citing *Author*, Merriam-Webster Dictionary (2023)), https://perma.cc/S96L-WYTS.

But statutory construction requires more than just finding a sympathetic dictionary definition. We "do not read statutes in little bites," or words in isolation from their statutory context. *Kircher v. Putnam Funds Tr.*, 547 U.S. 633, 643 (2006). The judicial task when interpreting statutory language, instead, is to discern how Congress used a word in the law.

That process includes "a natural presumption that identical words used in different parts of the same act are intended to have the same meaning." *Atlantic Cleaners & Dryers, Inc. v.*

17

United States, 286 U.S. 427, 433 (1932). Here, the Copyright Act makes no sense if an "author" is not a human being. If "machine" is substituted for "author," the Copyright Act would refer to a machine's "children," 17 U.S.C. § 203(a)(2), a machine's "widow," *id.*, a machine's "domicile," *id.* § 104(a), a machine's *mens rea*, *id.* § 101, and a machine's "nationality," *id.* Problematic questions would arise about a machine's "life" and "death[.]"   *Id.* § 302(a).   And "machine" would inconsistently mean both an author and a tool used by authors. *Id.* § 117(d)(1); *see id.* §§ 102(a); 108(c)(2); 116(d)(1); 117(c); 1001(2), (3).

Dr. Thaler points out that the Copyright Act's work-made-for-hire provision allows those who hire creators to be "considered the author" under the Act. 17 U.S.C. § 201(b). That is why corporations, *e.g.*, *Warren v. Fox Fam. Worldwide, Inc.*, 328 F.3d 1136, 1140 (9th Cir. 2003), and governments, *e.g.*, *Georgia v. Public.Resource.Org, Inc.*, 590 U.S. 255, 270 (2020), can be legally recognized as authors.

But the word "considered" in the work-made-for-hire provision does the critical work here. It allows the copyright and authorship protections attaching to a work originally created by a human author to transfer instantaneously, as a matter of law, to the person who hired the creator. *See Community for Creative Non-Violence v. Reid*, 490 U.S. 730, 737 (1989). Congress, in other words, was careful to avoid using the word "author" by itself to cover non-human entities. For if Congress had intended otherwise, the work-made-for-hire provision would say straightforwardly that those who hire creators "*are* the author for purposes of this title," not that they are "*considered* the author for purposes of this title."

18

b

Dr. Thaler also argues that the human-authorship requirement wrongly prevents copyright law from protecting works made with artificial intelligence. Thaler Opening Br. 38.

But the Supreme Court has long held that copyright law is intended to benefit the public, not authors. Copyright law "makes reward to the owner a secondary consideration. * * * '[T]he primary object in conferring the monopoly lie[s] in the general benefits derived by the public from the labors of authors.'" *United States v. Loew's, Inc.*, 371 U.S. 38, 46-47 (1962) (quoting *Fox Film Co. v. Doyal*, 286 U.S. 123, 127 (1932)).

To that public-benefit end, "the law of copyright has developed in response to significant changes in technology." *Sony Corp. of America v. Universal City Studios, Inc.*, 464 U.S. 417, 430 (1984). Photography, sound recordings, video recordings, and computer programs are all technologies that were once novel, but which copyright law now protects. *See Burrow-Giles*, 111 U.S. at 58; *Goldstein v. California*, 412 U.S. 546, 565-566 (1973); *Sony*, 464 U.S. at 442; *Google*, 593 U.S. at 21. Importantly, that evolution in copyright protection has been at Congress's direction, not through courts giving new meaning to settled statutory terms.

Contrary to Dr. Thaler's assumption, adhering to the human-authorship requirement does not impede the protection of works made with artificial intelligence. Thaler Opening Br. 38-39.

First, the human authorship requirement does not prohibit copyrighting work that was made by or with the assistance of artificial intelligence. The rule requires only that the author of

19

that work be a human being—the person who created, operated, or used artificial intelligence—and not the machine itself.   The Copyright Office, in fact, has allowed the registration of works made by human authors who use artificial intelligence. *See Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence*, 88 Fed. Reg. 16,190, 16,192 (March 16, 2023) (Whether a work made with artificial intelligence is registerable depends "on the circumstances, particularly how the AI tool operates and how it was used to create the final work.").

To be sure, the Copyright Office has rejected some copyright applications based on the human-authorship requirement even when a human being is listed as the author. *See* Copyright Office, *Re: Zarya of the Dawn (Registration # VAu001480196)* (Feb. 21, 2023), https://perma.cc/AD86-WGPM (denying copyright registration for a comic book's images made with generative artificial intelligence).   Some have disagreed with these decisions. *See* Motion Picture Association, *Comment Letter on Artificial Intelligence and Copyright* at 5 (Oct. 30, 2023), https://perma.cc/9W9X-3EZE (This "very broad definition of 'generative AI' has the potential to sweep in technologies that are not new and that members use to assist creators in making motion pictures."); 2 W. PATRY, COPYRIGHT § 3:60.52 (2024); Legal Professors Amicus Br. 36-37 ("The U.S. Copyright Office guidelines are somewhat paradoxical: human contributions must be demonstrated within the creative works generated by AI.").

Those line-drawing disagreements over how much artificial intelligence contributed to a particular human author's work are neither here nor there in this case. That is because Dr. Thaler listed the Creativity Machine as the *sole* author of the work before us, and it is undeniably a machine, not a human being.   Dr. Thaler, in other words, argues only for the

20

copyrightability of a work authored exclusively by artificial intelligence. *Contrast Rearden LLC v. Walt Disney Co.*, 293 F. Supp. 3d 963 (N.D. Cal. 2018) (holding that companies may copyright work made with motion capture software).

Second, Dr. Thaler has not explained how a ban on machines being authors would result in less original work because machines, including the Creativity Machine, do not respond to economic incentives.

Dr. Thaler worries that the human-authorship requirement will disincentivize creativity by the creators and operators of artificial intelligence. Thaler Opening Br. 36. That argument overlooks that the requirement still incentivizes humans like Dr. Thaler to create and to pursue exclusive rights to works that they make with the assistance of artificial intelligence.

Of course, the Creativity Machine does not represent the limits of human technical ingenuity when it comes to artificial intelligence. Humans at some point might produce creative non-humans capable of responding to economic incentives. Science fiction is replete with examples of creative machines that far exceed the capacities of current generative artificial intelligence. For example, Star Trek's Data might be worse than ChatGPT at writing poetry, but Data's intelligence is comparable to that of a human being. *See Star Trek: The Next Generation: Schism* (Paramount television broadcast Oct. 19, 1992) ("Felis catus is your taxonomic nomenclature, an endothermic quadruped, carnivorous by nature"). There will be time enough for Congress and the Copyright Office to tackle those issues when they arise.

Third, Congress's choice not to amend the law since 1976 to allow artificial-intelligence authorship "might well be taken to be an acquiescence in the judicial construction given to the

21

copyright laws." *White-Smith Music Pub. Co. v. Apollo Co.*, 209 U.S. 1, 14 (1908). The human-authorship requirement is not new and has been the subject of multiple judicial decisions. The Seventh Circuit has squarely held that authors "of copyrightable works must be human." *Kelley v. Chicago Park Dist.*, 635 F.3d 290, 304 (7th Cir. 2011). And the Ninth Circuit has strongly implied the same when deciding that an author must be a "worldly entity," *Urantia Foundation v. Maaherra*, 114 F.3d 955, 958 (9th Cir. 1997), and cannot be an animal, *Naruto v. Slater*, 888 F.3d 418, 426 (9th Cir. 2018).

Finally, even if the human authorship requirement were at some point to stymy the creation of original work, that would be a policy argument for Congress to address. U.S. CONST. Art. I, § 8, cl. 8. "Congress has the constitutional authority and the institutional ability to accommodate fully the varied permutations of competing interests that are inevitably implicated by such new technology." *Sony*, 464 U.S. at 431.

This court's job, by contrast, "is to apply the statute as it is written," not to wade into technologically uncharted copyright waters and try to decide what "might 'accord with good policy.'" *Burrage v. United States*, 571 U.S. 204, 218 (2014) (quoting *Commissioner v. Lundy*, 516 U.S. 235, 252 (1996)); *see also Teleprompter Corp. v. Columbia Broad. Sys., Inc.*, 415 U.S. 394, 414 (1974) ("Detailed regulation of these relationships, and any ultimate resolution of the many sensitive and important problems in this field, must be left to Congress."). Accommodating new technology "is for Congress." *Fortnightly Corp. v. United Artists Television, Inc.*, 392 U.S. 390, 401 (1968).

In that regard, it bears noting that the Political Branches have been grappling with how copyright law should adapt to new technology. The Copyright Office is studying how

22

copyright law should respond to artificial intelligence, *Artificial Intelligence and Copyright*, 88 Fed. Reg. 59,942, 59,942 (Aug. 30, 2023), and is making recommendations based on its findings, *see* Copyright Office, *Copyright and Artificial Intelligence, Part 1: Digital Replicas* at 57 (Jul. 31, 2024), https://perma.cc/8CUH-DN5A (recommending a statutory right for individuals to sue those who make deepfakes with their likeness); Copyright Office, *Copyright and Artificial Intelligence, Part 2: Copyrightability* at 32-40 (Jan. 29, 2025), https://perma.cc/W9VR-TLQP (recommending against changing the law governing the copyrightability of work generated by artificial intelligence). Also, Congress recently completed a report that addresses the problem of artificial intelligence and intellectual property. U.S. House of Rep., *Bipartisan House Task Force Report on Artificial Intelligence* at 111-136 (Dec. 2024), https://perma.cc/Y69R-DM3D. Congress and the Copyright Office are the proper audiences for Dr. Thaler's policy and practical arguments.

4

Because the Copyright Act itself requires human authorship, we need not and do not address the Copyright Office's argument that the Constitution's Intellectual Property Clause requires human authorship. The Copyright Act provides "a sufficient ground for deciding this case, and the cardinal principle of judicial restraint—if it is not necessary to decide more, it is necessary not to decide more—counsels us to go no further." *PDK Laboratories Inc. v. United States Drug Enforcement Agency*, 362 F.3d 786, 799 (D.C. Cir. 2004) (Roberts, J., concurring in part and concurring in the judgment).

23

IV

Dr. Thaler raises two alternative arguments in support of his copyright application. Neither succeeds.

First, Dr. Thaler argues that the Copyright Act's work-made-for-hire provision allows him to be "considered the author" of the work at issue because the Creativity Machine is his employee. Thaler Opening Br. 52-56; 17 U.S.C. § 201(b).

That argument misunderstands the human authorship requirement. The Copyright Act only protects "original works of authorship." 17 U.S.C. § 102(a). The authorship requirement applies to all copyrightable work, including work-made-for-hire. The word "authorship," like the word "author," refers to a human being. As a result, the human-authorship requirement necessitates that all "original works of authorship" be created in the first instance by a human being, including those who make work for hire.

Second, Dr. Thaler argues that he is the work's author because he made and used the Creativity Machine. Thaler Opening Br. 42-51. We cannot reach that argument. The district court held that Dr. Thaler forwent any such argument before the Copyright Office. *Thaler*, 687 F.Supp.3d at 150. And in his opening brief, Dr. Thaler did not challenge the district court's finding of waiver. Dr. Thaler offered only a single sentence in his opening brief, in which he describes the district court's conclusion as "based on a misunderstanding of the record below." Thaler Opening Br. 43. That "bare and conclusory assertion" is insufficient to preserve an argument for resolution on the merits. *Abdullah v. Obama*, 753 F.3d 193, 199 (D.C. Cir. 2014).

24

V

For the foregoing reasons, the district court's denial of Dr. Thaler's copyright application is affirmed.

So ordered.

IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF DELAWARE

THOMSON REUTERS ENTERPRISE
CENTRE GMBH and WEST PUBLISH-
ING CORP.,

        Plaintiffs,

        v.

ROSS INTELLIGENCE INC.,

        Defendant.

No. 1:20-cv-613-SB

Jack B. Blumenfeld, Michael J. Flynn, MORRIS, NICHOLS, ARSHT & TUNNELL LLP, Wilmington, Delaware; Dale M. Cendali, Eric A. Loverro, Joshua L. Simmons, KIRKLAND & ELLIS LLP, New York, New York; Yungmoon Chang, KIRKLAND & ELLIS LLP, Los Angeles, California; Miranda D. Means, KIRKLAND & ELLIS LLP, Boston, Massachusetts.

Counsel for Plaintiffs.

David Ellis Moore, Bindu Ann George Palapura, POTTER ANDERSON & CORROON LLP, Wilmington, Delaware; Jordan Ludwig, Emily T. Kuwahara, CROWELL & MORING LLP, Los Angeles, California; Ryan Henry Seewald, CROWELL & MORING LLP, Denver Colorado; Warrington Parker, Joachim B. Steinberg, Jacob Canter, Christopher J. Banks, Anna Z. Saber, Margaux Poueymirou, CROWELL & MORING LLP, San Francisco, California; Keith J. Harrison, Mark A. Klapow, Lisa Kimmel, Crinesha B. Berry, CROWELL & MORING LLP, Washington, D.C.

Counsel for Defendant.

MEMORANDUM OPINION

February 11, 2025

BIBAS, *Circuit Judge*, sitting by designation.

A smart man knows when he is right; a wise man knows when he is wrong. Wisdom does not always find me, so I try to embrace it when it does—even if it comes late, as it did here.

I thus revise my 2023 summary judgment opinion and order in this case. *See* Fed. R. Civ. P. 54(b); D.I. 547, 548; *Thomson Reuters Enter. Ctr. GmbH v. Ross Intel. Inc.*, 694 F. Supp. 3d 467 (D. Del. 2023). Now I (1) grant most of Thomson Reuters's motion for partial summary judgment on direct copyright infringement and related defenses, D.I. 674; (2) grant Thomson Reuters's motion for partial summary judgment on fair use, D.I. 672; (3) deny Ross's motion for summary judgment on fair use, D.I. 676; and (4) deny Ross's motion for summary judgment on Thomson Reuters's copyright claims, D.I. 683.

I. ROSS MAKES A LEGAL AI TOOL AND WESTLAW'S OWNER SUES

The law is no longer a brooding omnipresence in the sky; it now dwells in legal-research platforms. Thomson Reuters owns one of the biggest of those platforms: Westlaw. D.I. 752-1 at 4. Users can pay to access its contents, including "case law, state and federal statutes, state and federal regulations, law journals, and treatises." *Id.* "Westlaw also contains editorial content and annotations," like the headnotes here. *Id.* Those headnotes summarize key points of law and case holdings. Westlaw organizes its content using the Key Number System, a numerical taxonomy. *Id.* Thomson Reuters owns copyrights in Westlaw's copyrightable material. *Id.*

Ross, a new competitor to Westlaw, made a legal-research search engine that uses artificial intelligence. *Id.* To train its AI search tool, Ross needed a database of legal questions and answers. *Id.* at 5. So Ross asked to license Westlaw's content. *Id.* But because Ross was its competitor, Thomson Reuters refused. *Id.* at 4–5.

So to train its AI, Ross made a deal with LegalEase to get training data in the form of "Bulk Memos." *Id.* at 5. Bulk Memos are lawyers' compilations of legal questions with good and bad answers. LegalEase gave those lawyers a guide explaining how to create those questions using Westlaw headnotes, while clarifying that the lawyers should not just copy and paste headnotes directly into the questions. D.I. 678-36 at 5–9. LegalEase sold Ross roughly 25,000 Bulk Memos, which Ross used to train its AI search tool. *See* D.I. 752-1 at 5; D.I. 769 at 30 (10:48:35). In other words, Ross built its competing product using Bulk Memos, which in turn were built from Westlaw headnotes. When Thomson Reuters found out, it sued Ross for copyright infringement.

In 2023, I largely denied Thomson Reuters's motions for summary judgment on copyright infringement and the fair-use defense, and the case moved ahead toward trial. D.I. 547, 548. In the run-up to the August 2024 trial date, I studied the case materials more closely and realized that my prior summary-judgment ruling had not gone far enough. So I continued the trial and invited the parties to renew their summary-judgment briefing. D.I. 663.

Thomson Reuters once again moved for partial summary judgment on direct copyright infringement and related defenses. D.I. 674. Ross moved for summary judgment

on Thomson Reuters's copyright claims. D.I. 683. And both sides again moved for summary judgment on fair use. D.I. 672, 676. I now revise parts of my 2023 summary-judgment opinion. *See* Fed. R. Civ. P. 54(b).

I may grant summary judgment only if there is "no genuine dispute as to any material fact and the movant is entitled to judgment as a matter of law." Fed. R. Civ. P. 56(a). I view all facts and draw all reasonable inferences in favor of the nonmovant. *Tundo v. County of Passaic*, 923 F.3d 283, 286–87 (3d Cir. 2019).

II. I GRANT PARTIAL SUMMARY JUDGMENT TO THOMSON REUTERS, NOT ROSS, ON DIRECT COPYRIGHT INFRINGEMENT AND RELATED DEFENSES

The dispute boils down to whether the LegalEase Bulk Memo questions copied Thomson Reuters's headnotes or were instead taken from uncopyrightable judicial opinions. To decide many issues here, one must compare the Bulk Memo questions, headnotes, and opinions side by side. I include the table below as an example. The questions and headnotes in this case are sealed. So the headnote and question in this table are not actual materials from the record, but an example I created based on *Feist Publications, Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340, 345 (1991).

Question	West Headnote	Case Opinion
Does originality for copyright purposes mean that the work was independently created and has some minimal degree of creativity?	Originality, for copyright purposes, means that the work was independently created and has some minimal degree of creativity.	Original, as the term is used in copyright, means only that the work was independently created by the author (as opposed to copied from other works), and that it possesses at least some minimal degree of creativity.

4

As I later explain, I hold that:

- Ross infringed 2,243 headnotes. As to those headnotes, the only remaining factual issue is whether some of their copyrights have expired.

- Ross's innocent infringement, copyright misuse, merger, and *scenes à faire* defenses all fail.

A. Direct copyright infringement

Thomson Reuters alleges that Ross directly infringed its copyrights. To show that, Thomson Reuters must show both that (1) it owned a valid copyright and (2) Ross copied protectable elements of the copyrighted work. *Feist*, 499 U.S. at 361. The second element requires showing both that (2a) Ross actually copied the work and that (2b) its copy was substantially similar to the work. *Dam Things from Den. v. Russ Berrie & Co.*, 290 F.3d 548, 561–62 (3d Cir. 2002).

(i) Copyright validity

Copyright validity is a question of law, not fact, making it suitable for summary judgment. *See Silvertop Assocs. Inc. v. Kangaroo Mfg. Inc.*, 931 F.3d 215, 219 (3d Cir. 2019). But an underlying factual dispute remains here. So the jury may need to decide this issue, but for reasons different from the ones I gave in my prior summary-judgment opinion.

Thomson Reuters must show that it had a valid copyright. Copyright registrations are "prima facie evidence of the validity of the copyright" if "made before or within five years after first publication of the work." 17 U.S.C. § 410(c). Thomson Reuters has copyright registrations for Westlaw's copyrightable content. D.I. 752-1 at 4. And it docketed registrations from 1981 to 2019. D.I. 1-1. So it has a valid compilation

5

copyright. But if Thomson Reuters chooses to try this case based on a theory of infringement of individual headnotes as individual works rather than infringement of the compilation as a whole, there is still a factual dispute about which individual headnotes are both within the period covered by Thomson Reuters's registrations and not in the public domain. *See* D.I. 755, 757, 761, 763. So, if Thomson Reuters advances a theory of damages that depends on the infringement of specific headnotes, this evidentiary matter must be taken up at trial.

In my 2023 opinion, I concluded that this issue would need to go to a jury, but for a different reason. I held that a jury would need to decide whether the headnotes and Key Number System were original enough. 694 F. Supp. 3d at 477–78.

Originality is central to copyright. The Constitution limits copyright protection to original works. *See Feist*, 499 U.S. at 346. So even if Thomson Reuters gets a presumption of validity because of the copyright registrations, Ross could rebut the presumption by showing that the works are not original. I previously thought that originality "depend[ed] on how much the headnotes overlap with the [uncopyrightable text of] opinions." 694 F. Supp. 3d at 478. And I explained that the Key Number System's originality was a jury question because Ross alleges that "most of the organization decisions are made by a rote computer program and the high-level topics largely track common doctrinal topics taught as law school courses." *Id.* at 477 (internal quotation marks omitted).

But the originality threshold is "extremely low," requiring only "some minimal degree of creativity …. some creative spark." *Id.* at 345. The key question, then, is

whether a work is original, not how much effort went into developing it. *Id.* at 359–60. So I now revise those parts of my prior opinion. Now I see no genuine dispute that the headnotes and Key Number System clear *Feist*'s minimal threshold for originality.

1. The headnotes are original. A headnote is a short, key point of law chiseled out of a lengthy judicial opinion. The text of judicial opinions is not copyrightable. *Banks v. Manchester*, 128 U.S. 244, 253–54 (1888). And even if it were, Thomson Reuters would not get that copyright because it did not write the opinions. But a headnote can introduce creativity by distilling, synthesizing, or explaining part of an opinion, and thus be copyrightable. That is why I have changed my mind.

First, the headnotes are a compilation. "Factual compilations" are original if the compiler makes "choices as to selection and arrangement" using "a minimal degree of creativity." *Feist*, 499 U.S. at 348. Thomson Reuters's selection and arrangement of its headnotes easily clears that low bar.

More than that, each headnote is an individual, copyrightable work. That became clear to me once I analogized the lawyer's editorial judgment to that of a sculptor. A block of raw marble, like a judicial opinion, is not copyrightable. Yet a sculptor creates a sculpture by choosing what to cut away and what to leave in place. That sculpture is copyrightable. 17 U.S.C. § 102(a)(5). So too, even a headnote taken verbatim from an opinion is a carefully chosen fraction of the whole. Identifying which words matter and chiseling away the surrounding mass expresses the editor's idea about what the important point of law from the opinion is. That editorial expression has enough

7

"creative spark" to be original. *Feist*, 499 U.S. at 345. So all headnotes, even any that quote judicial opinions verbatim, have original value as individual works.

That belated insight explains my change of heart. In my 2023 opinion, I wrongly viewed the degree of overlap between the headnote text and the case opinion text as dispositive of originality. 694 F. Supp. 3d at 478. I no longer think that is so. But I am still not granting summary judgment on any headnotes that are verbatim copies of the case opinion (for reasons that I explain below).

2. The Key Number System is original too. There is no genuine issue of material fact about the Key Number System's originality. Recall that Westlaw uses this taxonomy to organize its materials. Even if "most of the organization decisions are made by a rote computer program and the high-level topics largely track common doctrinal topics taught as law school courses," it still has the minimum "spark" of originality. *Id.* at 477 (internal quotation marks omitted); *Feist*, 499 U.S. at 345. The question is whether the system is original, not how hard Thomas Reuters worked to create it. *Feist*, 499 U.S. at 359–60. So whether a rote computer program did the work is not dispositive. And it does not matter if the Key Number System categorizes opinions into legal buckets that any first-year law student would recognize. To be original, a compilation need not be "novel," just "independently created by" Thomson Reuters. *Id.* at 345–46. There are many possible, logical ways to organize legal topics by level of granularity. It is enough that Thomson Reuters chose a particular one.

Thus, I grant summary judgment for Thomson Reuters on whether the headnotes and the Key Number System are original enough to prevent Ross from rebutting any presumption of validity.

(ii) Copying of original elements

Next, I turn to whether there was "copying of constituent elements of the work that are original." *Feist*, 499 U.S. at 361. I must decide whether Thomson Reuters has proven both (a) actual copying and (b) substantial similarity. *Dam Things*, 290 F.3d at 561–62.

1. A few preliminaries. Before applying that analysis, I must determine which parts of the case to apply it to. Ross moves for summary judgment on all the pieces that Thomson Reuters accuses Ross of infringing: 21,787 headnotes, the editorial decisions in 500 judicial opinions, and West's Key Number System. D.I. 684 at 9–11. I see no proof sufficient to take all these items away from a jury.

Thomson Reuters, for its part, moves for summary judgment on only two subsets of the headnotes. D.I. 694 at 20 n.8. The two batches contain 5,367 and 2,830 headnotes each.

I reach no new decision on the Key Number System. There are still factual disputes about whether Ross used it and how, so I cannot analyze what protected elements Ross may have used. So I leave my prior ruling on this topic in place. Nor do I decide the fate of the 500 judicial opinions containing Thomson Reuters's editorial decisions, because there are still factual questions about how and how much Ross accessed that material. For now, I consider only the batch of 2,830 headnotes identified by Thomson Reuters.

9

I leave the 5,367 for trial because Thomson Reuters's argument for summary judgment on that batch is flawed. Thomson Reuters points out that after I ordered Ross to submit a list of headnotes that are verbatim or near-verbatim quotations of judicial opinions, Ross did not submit these 5,367. D.I. 675 at 24. And it contends that Ross thereby admitted that these headnotes are protectable. *Id.* Perhaps I could view that as a concession about how similar the headnotes are to the judicial opinions. But Ross's omission makes no concession about the more important issue: whether the Bulk Memo questions copied these headnotes, or even whether the questions are closer to the headnote or to the judicial opinion. So it is not appropriate to grant summary judgment to Thomson Reuters on the 5,367 headnotes based on this reasoning. (Because I did not look at this batch, I need not address Ross's objection that it had no chance to respond to Thomson Reuters's expert's opinion on the 5,367 headnotes. D.I. 749.)

But it is appropriate to address the other 2,830 now at summary judgment. Before I explain why and how I address this batch, I must first clean up whether the batch in fact has 2,830 headnotes. There are two disputes. First, Ross's expert, Barbara Frederiksen-Cross, put 3,384 headnotes in this category, but the parties dispute whether 554 of them are still covered by valid copyright registrations. To resolve this, I looked at all 3,384 but make summary judgment contingent on the jury's findings about which 2,830 (or other number) still have valid copyrights.

Second, Ross claims that Thomson Reuters never asserted that 1,623 of the 2,830 headnotes were infringed, so at first, I denied summary judgment on them. 694

10

F. Supp. 3d at 479. Ross says I should again deny summary judgment because Thomson Reuters did not include these 1,623 headnotes when I ordered it to produce a list of allegedly infringing headnotes. D.I. 201; D.I. 324 ¶¶ 28–30; D.I. 324-23 (Ross's spreadsheet identifying the 1,623). But on the same day that Thomson Reuters docketed its interrogatory response lacking those 1,623 headnotes, Ross docketed the expert report conceding them. D.I. 266-1 at 431 (Frederiksen-Cross's report); D.I. 281-3–281-8 (Thomson Reuters's interrogatory response). So Thomson Reuters did not have a chance to analyze Ross's expert report before submitting its list of headnotes. Plus, Ross had the chance to respond to these 1,623 headnotes when Thomson Reuters raised them at summary judgment. And Ross's own expert analyzed them, so there was no danger of unfair surprise. I revise my prior ruling and now consider these 1,623 headnotes as fairly part of the case.

Having sorted through that morass, I apply the actual-copying and substantial-similarity analyses to the 3,384 headnotes, which include the 2,830.

2. Actual copying. Actual copying means that "the defendant did, in fact, use the copyrighted work in creating his own." *Tanksley v. Daniels*, 902 F.3d 165, 173 (3d Cir. 2018). One can prove this directly, with evidence that the defendant copied the work, or indirectly, by showing that the defendant had access to it and produced something similar ("probative similarity"). *Id.*

When evaluating copying, I may consider expert opinions. *See Kay Berry, Inc. v. Taylor Gifts, Inc.*, 421 F.3d 199, 208 (3d Cir. 2005). Frederiksen-Cross submitted an expert report stating that the Bulk Memo questions for this batch closely resemble

11

the headnotes' text and that the headnotes differ significantly from the text of the

judicial opinions. D.I. 266-1 at 431, 435; D.I. 675 at 10. Her finding suggests that

these questions were created by copying Westlaw headnotes, not by summarizing the

underlying opinions. I view that as Ross's expert conceding that the 2,830 were actu-

ally copied.

But to make sure, the Court has now compared how similar each of the 2,830 Bulk

Memo questions, headnotes, and judicial opinions are, one by one. The parties agree

that LegalEase had access to Westlaw and used it to make the Bulk Memos. Of

course, access alone is not proof. But a Bulk Memo question that looks more like a

headnote than it does like the underlying judicial opinion is strong circumstantial

evidence of actual copying. My comparison of the questions, headnotes, and opinions

can decide whether probative similarity is present, so long as I think no reasonable

jury could reach a different conclusion. Access plus probative similarity adds up to

evidence of actual copying. To be sure, there was some confusion at the summary

judgment hearing about whether Ross used all the Bulk Memos. *See* D.I. 769 at 30

(10:48:35), 162 (14:36:24). Ross's counsel said that some Bulk Memos were discarded

but twice confirmed that Ross had used 80% for initial training and 20% for later

validation. D.I. 769 at 30 (10:48:35), 162 (14:36:24). So taking counsel at his word,

Ross used practically 100% to train its AI.

Having slogged through all 2,830 headnotes, I grant summary judgment to Thom-

son Reuters on actual copying, finding actual copying of 2,243. Appendix A to this

opinion, filed under seal, catalogues the specific headnotes. I grant summary

judgment only on the headnotes for which actual copying is so obvious that no reasonable jury could find otherwise.

3. Substantial similarity. Substantial similarity requires evaluating whether "the later work materially appropriates the copyrighted work." *Tanksley*, 902 F.3d at 173. That means deciding which parts of the actually copied work are original expression and so protected by copyright. Substantial similarity is often "an extremely close question of fact," and thus well suited to the jury. *Id.* at 171 (internal quotation marks omitted). But summary judgment can be "appropriate" when "no reasonable jury could find" otherwise. *Id.* (internal quotation marks omitted). The question is whether an ordinary user of a product would find it substantially similar to the copyrighted work. *Dam Things*, 290 F.3d at 562. As a lawyer and judge, I am myself an ordinary user of Westlaw headnotes. So I am well positioned to determine substantial similarity here. I do so cautiously, and only on those headnotes for which I am confident that a reasonable jury could not conclude otherwise.

Ross argues that the Bulk Memo questions must be not just substantially similar to the headnotes, but virtually identical. The Ninth Circuit takes this approach to "thin" copyrights because "the range of protectable … expression is narrow." *Apple Comput., Inc. v. Microsoft Corp.*, 35 F.3d 1435, 1439, 1446 (9th Cir. 1994). But substantial similarity is always a spectrum, whether the copyright is "thin" or "thick." "More similarity is required when less protectable matter is at issue." 4 *Nimmer on Copyright* § 13D.32[A]. Other terms for this concept include Nimmer's "supersubstantial similarity" and the Second Circuit's "more discerning" ordinary-observer test. *Id.*;

Knitwaves, Inc. v. Lollytogs Ltd., 71 F.3d 996, 1002–03 (2d Cir. 1995). Though I do not apply the Ninth Circuit's test explicitly, I do apply the concept underlying all these terms: The less protectable expression a work contains, the more similar the allegedly infringing work must be to it.

Applying this standard, I grant summary judgment on substantial similarity for Thomson Reuters on the 2,243 headnotes listed in Appendix A, finding that the Bulk Memo questions were substantially similar to them. Again, I grant summary judgment only on the headnotes for which substantial similarity is so obvious that no reasonable jury could find otherwise. In practice, this means that I am granting summary judgment only on the headnotes whose language very closely tracks the language of the Bulk Memo question but not the language of the case opinion. The rest of the headnotes must go to trial. I do not grant summary judgment to Ross on any headnotes because there are none for which I am confident that a reasonable jury could not find infringement. I do not decide at summary judgment the factual question of which headnotes are still covered by Thomson Reuters's existing copyrights and leave this question open for trial.

B. Ross's defenses to copyright infringement fail

None of Ross's possible defenses holds water. I reject them all.

First, innocent infringement does not apply. Ross claims that any infringement was innocent. As the parties agree, innocent infringement does not limit liability, just damages. 17 U.S.C. § 504(c)(2). But this limit does not apply when the infringed work bears a copyright notice, as Westlaw's headnotes do. 17 U.S.C. § 401(d).

Second, copyright misuse does not apply either. Ross claims that Thomson Reuters misused its copyright. Copyright misuse is a defense when a copyright holder weaponizes the copyright against the public interest, typically for "anti-competitive behavior." *Video Pipeline, Inc. v. Buena Vista Home Ent., Inc.*, 342 F.3d 191, 203–06 (3d Cir. 2003). But as I already ruled at summary judgment on Ross's antitrust counterclaims, Ross has not shown that Thomson Reuters misused its copyrights to stifle competition. D.I. 669.

Third, the merger defense is inapt. Ross claims that any ideas were so close to the expression that they merged with the expression, making it uncopyrightable. *Educ. Testing Servs. v. Katzman*, 793 F.2d 533, 539 (3d Cir. 1986). But there are many ways to express points of law from judicial opinions, so I reject this defense as well.

Fourth, the *scenes à faire* defense does not fit. This defense covers stock elements that follow from the work's nature, like a historical romance novel's damsel in distress. *Atari, Inc. v. N. Am. Philips Consumer Elecs. Corp.*, 672 F.2d 607, 616 (7th Cir. 1982). But nothing about a judicial opinion requires it to be slimmed down to Thomson Reuters's headnotes or categorized by key numbers.

III. THOMSON REUTERS, NOT ROSS, PREVAILS ON THE FAIR-USE DEFENSE

There remains one more defense. In my 2023 opinion, I denied summary judgment on fair use. D.I. 548; 694 F. Supp. 3d at 482–87. But with new information and understanding, I vacate those sections of that order and its accompanying opinion addressing fair use. Fair use is an affirmative defense, so Ross bears the burden of proof. *Video Pipeline*, 342 F.3d at 197.

15

I must consider at least four fair-use factors: (1) the use's purpose and character, including whether it is commercial or nonprofit; (2) the copyrighted work's nature; (3) how much of the work was used and how substantial a part it was relative to the copyrighted work's whole; and (4) how Ross's use affected the copyrighted work's value or potential market. 17 U.S.C. § 107(1)–(4). The first and fourth factors weigh most heavily in the analysis. *Authors Guild v. Google, Inc.*, 804 F.3d 202, 220 (2d Cir. 2015) (Leval, J.).

"Fair use is a mixed question of law and fact." *Harper & Row, Publishers, Inc. v. Nation Enters.*, 471 U.S. 539, 560 (1985). But "[i]n this case, the ultimate 'fair use' question primarily involves legal work." *Google LLC v. Oracle Am., Inc.*, 593 U.S. 1, 24 (2021). The undisputed facts here push this case squarely into the legal realm. Once we get past actual copying, the remaining issues that matter are not ones of historical fact, intent, or factual prediction. They are about how to apply the law to the facts. So here, fair use is a question for the judge, not the jury. Thomson Reuters prevails on the two most important and on the overall balancing.

A. Factor one goes to Thomson Reuters

First, I consider the purpose and character of Ross's use. 17 U.S.C. § 107(1). I look mainly at whether it was commercial and whether it was transformative. *Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith*, 598 U.S. 508, 529–31 (2023). If Ross and Thomson Reuters use copyrighted material like the headnotes for very similar purposes and Ross's use is commercial, this factor likely disfavors fair use. *Id.* at 532–33.

1. Ross's use is commercial. Ross admits as much. D.I. 727 at 29. It "stands to profit from exploitation of the copyrighted material without paying the customary

16

price." *Harper & Row*, 471 U.S. at 562. But commerciality is not dispositive. I must

balance it against how different this work's purpose or character is. *Warhol*, 598 U.S.

at 525.

2. Ross's use is not transformative. Transformativeness is about the purpose of the

use. "If an original work and a secondary use share the same or highly similar pur-

poses, and the second use is of a commercial nature, the first factor is likely to weigh

against fair use, absent some other justification for copying." *Warhol*, 598 U.S. at

532–33. It weighs against fair use here. Ross's use is not transformative because it does

not have a "further purpose or different character" from Thomson Reuters's. *Id.* at 529.

Ross was using Thomson Reuters's headnotes as AI data to create a legal research

tool to compete with Westlaw. It is undisputed that Ross's AI is not generative AI (AI

that writes new content itself). Rather, when a user enters a legal question, Ross spits

back relevant judicial opinions that have already been written. D.I. 723 at 5. That

process resembles how Westlaw uses headnotes and key numbers to return a list of

cases with fitting headnotes. Thomson Reuters uses its headnotes and Key Number

System primarily to help legal researchers navigate Westlaw and (possibly, as the

parties dispute this) to improve Westlaw's internal search tool. D.I. 769 at 14

(10:24:52). The parties agree that Ross and Westlaw are competitors. D.I. 752-1 at 4.

So at first glance, this factor looks simple.

But, as Ross argues, the headnotes do not appear as part of the final product that

Ross put forward to consumers. The copying occurred at an intermediate step: Ross

turned the headnotes into numerical data about the relationships among legal words

to feed into its AI. D.I. 727 at 22. That makes this factor much trickier.

Ross is right that intermediate copying has been permitted under fair use factor

one in analyzing computer programs. *See Google*, 593 U.S. at 30–32; *Sony Comput.*

Ent., Inc. v. Connectix Corp., 203 F.3d 596, 599, 606–07 (9th Cir. 2000) (holding that

copying source code to create a product that lets people play Sony games on a personal

computer, rather than a separate Sony game station, is transformative); *Sega Enters.*

Ltd. v. Accolade, Inc., 977 F.2d 1510, 1514–1515, 1522–23 (9th Cir. 1992) (holding

that copying source code to create games that are compatible with an existing gaming

system is transformative).

But those cases are inapt. First and foremost, those cases are all about copying

computer code. This case is not. (Though Ross did computer programming, the mate-

rial it allegedly copied from Thomson Reuters was not computer code.) In copyright,

"computer programs differ from books, films, and many other literary works in that

such programs almost always serve functional purposes." *Google*, 593 U.S. at 21 (in-

ternal quotation marks omitted). So the fair-use considerations for these programs

do not always apply to cases about copying written words.

Second and relatedly, these computer-programming cases about intermediate copy-

ing rely on a factor absent here: The copying was *necessary* for competitors to innovate.

In *Google*, Google had copied part of a computer-programming language—specifically,

the code that lets programmers speak to software in a particular way. *Id.* at 6, 29–

33. That copying was "necessary for different programs to speak to each other." *Id.* at

31. The copying in *Sony* was also necessary. The Ninth Circuit "appl[ied] fair use to intermediate copying [that was] necessary to reverse engineer access to unprotected functional elements within a program." *Id.* at 22. "[I]ntermediate copying … was a fair use for the purpose of gaining access to the unprotected elements of Sony's software." *Sony*, 203 F.3d at 602. Likewise, *Sega* addressed copying that occurred "solely in order to discover the functional requirements for compatibility." 977 F.2d at 1522. Here, though, there is no computer code whose underlying ideas can be reached only by copying their expression. The "copying is [not] reasonably necessary to achieve the user's new purpose." *Warhol*, 598 U.S. at 532.

My prior opinion wrongly concluded that I had to send this factor to a jury. 694 F. Supp. 3d at 483–84. I based that conclusion on *Sony* and *Sega*. Since then, I have realized that the intermediate-copying cases (1) are computer-programming copying cases; and (2) depend in part on the need to copy to reach the underlying ideas. Neither is true here. Because of that, this case fits more neatly into the newer framework advanced by *Warhol*. I thus look to the broad purpose and character of Ross's use. Ross took the headnotes to make it easier to develop a competing legal research tool. So Ross's use is not transformative. Because the AI landscape is changing rapidly, I note for readers that only non-generative AI is before me today.

3. Even if relevant, bad faith would not move the needle. The Supreme Court has expressed "some skepticism about whether bad faith has any role in a fair use analysis." *Google*, 593 U.S. at 32. If it does, a reasonable jury might find that Ross, by going forward and arguably inducing LegalEase's copying after Thomson Reuters refused

19

to license its content, acted in bad faith. But because Ross's use was commercial and not transformative, I need not consider this possible element. Even if I found no bad faith, that finding would not outweigh the other two considerations.

B. Factor two goes to Ross

Second, I ask about the nature of the original work. 17 U.S.C. § 107(2). That involves "focus[ing] on the degree of creativity inherent to the work." 4 *Nimmer on Copyright* § 13F.06. More creative works get more protection. *Id.* § 13F.06[A].

Westlaw's material has more than the minimal spark of originality required for copyright validity. But the material is not *that* creative. Though the headnotes required editorial creativity and judgment, that creativity is less than that of a novelist or artist drafting a work from scratch. And the Key Number System is a factual compilation, so its creativity is limited.

I signaled a similar leaning before. 694 F. Supp. 3d at 484–85. Yet I stopped short of granting summary judgment based on factual disputes about how much creativity was involved. Now, as I concluded above, there is no factual dispute that the headnotes have creative elements but are far from the most creative works.

So factor two goes for Ross. Note, though, that this factor "has rarely played a significant role in the determination of a fair use dispute." *Authors Guild*, 804 F.3d at 220.

C. Factor three goes to Ross

Third, I focus on how much of the work was used and how substantial a part it was relative to the whole. 17 U.S.C. § 107(3). I ask whether that usage was "reasonable in relation to the purpose of the copying." *Campbell v. Acuff-Rose Music, Inc.*, 510

U.S. 569, 586 (1994). Courts consider both "the quantity of the materials used" and "their quality and importance." *Id.* at 587. To win on this factor, the alleged copier must not take the "heart" of the work. *Id.*

My prior opinion did not decide factor three but suggested that it leaned towards Ross. The opinion focused on Ross's claim that its output to an end user is a judicial opinion, not a West headnote, so it "communicates little sense of the original." 649 F. Supp. 3d at 485 (quoting *Authors Guild*, 804 F.3d at 223).

I stand by that reasoning, but now go a step further and decide factor three for Ross. There is no factual dispute: Ross's output to an end user does not include a West headnote. What matters is not "the amount and substantiality of the portion used *in making a copy*, but rather the amount and substantiality of *what is thereby made accessible* to a public for which it may serve as a competing substitute." *Authors Guild*, 804 F.3d at 222 (internal quotation marks omitted). Because Ross did not make West headnotes available to the public, Ross benefits from factor three.

In its briefing, Ross emphasized that the number of headnotes allegedly taken amounted to only a small percentage of the total headnotes owned by Westlaw. That argument is inapt. If taking 300 words out of President Ford's memoirs could count as taking the heart of the work, so too can taking several thousand headnotes out of Westlaw. *Campbell*, 510 U.S. at 587. The percentage of a total work copied is neither necessary nor sufficient to decide factor three. But Ross wins this factor anyway.

D. Factor four goes to Thomson Reuters

Factor four "is undoubtedly the single most important element of fair use." *Harper & Row*, 471 U.S. at 566. For this factor, I consider the "likely effect [of Ross's copying]

21

on the market for the original." *Campbell*, 510 U.S. at 590. I must consider not only current markets but also potential derivative ones "that creators of original works would in general develop or license others to develop." *Id.* at 592. I also consider any "public benefits the copying will likely produce." *Google*, 593 U.S. at 35. The original market is obvious: legal-research platforms. And at least one potential derivative market is also obvious: data to train legal AIs.

My prior opinion left this factor for the jury. I thought that "Ross's use might be transformative, creating a brand-new research platform that serves a different purpose than Westlaw." 694 F. Supp. 3d at 486. If that were true, then Ross would not be a market substitute for Westlaw. Plus, I worried whether there was a relevant, genuine issue of material fact about whether Thomson Reuters would use its data to train AI tools or sell its headnotes as training data. *Id.* And I thought a jury ought to sort out "whether the public's interest is better served by protecting a creator or a copier." *Id.*

In hindsight, those concerns are unpersuasive. Even taking all facts in favor of Ross, it meant to compete with Westlaw by developing a market substitute. D.I. 752-1 at 4. And it does not matter whether Thomson Reuters has used the data to train its own legal search tools; the effect on a *potential* market for AI training data is enough. Ross bears the burden of proof. It has not put forward enough facts to show that these markets do not exist and would not be affected.

Nor does a possible benefit to the public save Ross. Yes, there is a public interest in accessing the law. But legal opinions are freely available, and "the public's interest

in the subject matter" alone is not enough. *Harper & Row*, 471 U.S. at 569. The public has no right to Thomson Reuters's parsing of the law. Copyrights encourage people to develop things that help society, like good legal-research tools. Their builders earn the right to be paid accordingly. This case is distinguishable from *Google*, where the API was valuable "because users, including programmers, [were] just used to it." 593 U.S. at 38. There is nothing that Thomson Reuters created that Ross could not have created for itself or hired LegalEase to create for it without infringing Thomson Reuters's copyrights.

E. Balancing the factors, I reject Ross's fair-use defense

Factors one and four favor Thomson Reuters. Factors two and three favor Ross. Factor two matters less than the others, and factor four matters more. Weighing them all together, I grant summary judgment for Thomson Reuters on fair use.

* * * * *

I grant partial summary judgment to Thomson Reuters on direct copyright infringement for the headnotes in Appendix A. For those headnotes, the only remaining factual issue on liability is that some of those copyrights may have expired or been untimely created. This factual question underlying copyright validity is for the jury. I also grant summary judgment to Thomson Reuters against Ross's defenses of innocent infringement, copyright misuse, merger, *scenes à faire*, and fair use. I deny Ross's motions for summary judgment on direct copyright infringement and fair use. I revise all parts of my prior opinions that conflict with this one. I leave undisturbed the parts of my prior opinion not addressed in this one, such as my rulings on contributory liability, vicarious liability, and tortious interference with contract.

Copyright and Artificial Intelligence

Part 2: Copyrightability

COPYRIGHT AND ARTIFICIAL INTELLIGENCE
Part 2: Copyrightability

ABOUT THIS REPORT

This Report by the U.S. Copyright Office addresses the legal and policy issues related to artificial intelligence ("AI") and copyright, as outlined in the Office's August 2023 Notice of Inquiry ("NOI").

The Report will be published in several Parts, each one addressing a different topic. This Part addresses the copyrightability of works created using generative AI. The first Part, published in 2024, addresses the topic of digital replicas—the use of digital technology to realistically replicate an individual's voice or appearance. A subsequent part will turn to the training of AI models on copyrighted works, licensing considerations, and allocation of any liability. To learn more, visit www.copyright.gov/ai.

ABOUT THE U.S. COPYRIGHT OFFICE

The U.S. Copyright Office is the federal agency charged by statute with the administration of U.S. copyright law. The Register of Copyrights advises Congress, provides information and assistance to courts and executive branch agencies, and conducts studies on national and international issues relating to copyright, other matters arising under Title 17, and related matters. The Copyright Office is housed in the Library of Congress. Its mission is to promote "creativity and free expression by administering the nation's copyright laws and by providing impartial, expert advice on copyright law and policy for the benefit of all." For more information, visit www.copyright.gov.

PREFACE

In early 2023, the U.S. Copyright Office announced a broad initiative to explore the intersection of copyright and artificial intelligence.

In March of that year, the Office released a policy statement with registration guidance for works incorporating AI-generated content. Over the spring and summer, we hosted a series of online listening sessions, presented educational webinars, and engaged with numerous stakeholders to enhance our understanding of the technology and how it is used, the copyright implications, and the potential impact on businesses and individuals.

These activities culminated in an August 2023 Notice of Inquiry, formally seeking public input on the full range of copyright issues that had been raised. In response, we received more than 10,000 comments representing a broad range of perspectives, including from authors and composers, performers and artists, publishers and producers, lawyers and academics, technology companies, libraries, sports leagues, trade groups and public interest organizations, and even a class of middle school students. Comments came from all 50 states and from 67 countries. That valuable and extensive input, supplemented by additional Office research and information received from other agencies, forms the basis for the discussion and recommendations in this Report.

TABLE OF CONTENTS

EXECUTIVE SUMMARY

This second Part of the Copyright Office's Report on Copyright and Artificial Intelligence ("AI") addresses the copyrightability of outputs generated by AI systems. It analyzes the type and level of human contribution sufficient to bring these outputs within the scope of copyright protection in the United States.

Of the more than 10,000 comments the Office received in response to its Notice of Inquiry ("NOI"), approximately half addressed copyrightability. The vast majority of commenters agreed that existing law is adequate in this area and that material generated wholly by AI is not copyrightable.

Commenters differed, however, as to protection for generative AI outputs that involve some form of human contribution. They expressed divergent views on what types and amounts of contribution could constitute authorship under existing law. Many also stressed the desirability of greater clarity in this area, including with respect to the use of AI as a tool in the creative process.

As a matter of policy, some argued that extending protection to materials created by generative AI would encourage the creation of more works of authorship, furthering progress in culture and knowledge to the benefit of the public. The Office also heard concerns that an increased proliferation of AI-generated outputs would undermine incentives for humans to create.

While recognizing that copyrightability is determined on a case-by-case basis, in this Part the Office sets out the legal principles that govern the analysis and assesses their application to AI-generated content.

Section I identifies the copyrightability issues raised by AI technologies. It outlines the history of adapting copyright law to new technological developments and describes the Office's ongoing AI initiative.

Section II provides a brief background on the technologies involved. It then summarizes the existing legal framework, particularly the human authorship requirement, the idea/expression dichotomy, and the originality standard for copyright protection. After discussing the use of AI to assist authors in the process of creating works of authorship, it analyzes how the law may apply to various types of human contributions to AI-generated outputs: prompting, the inclusion of human-authored expressive inputs, and the modification or arrangement of AI-generated outputs.

Section III reports on the international landscape. It describes how other countries are approaching questions of copyrightability within their own legal systems.

Section IV addresses the policy implications of providing additional legal protection to AI-generated material and evaluates the arguments for and against legislative change.

Based on an analysis of copyright law and policy, informed by the many thoughtful comments in response to our NOI, the Office makes the following conclusions and recommendations:

- Questions of copyrightability and AI can be resolved pursuant to existing law, without the need for legislative change.

- The use of AI tools to assist rather than stand in for human creativity does not affect the availability of copyright protection for the output.

- Copyright protects the original expression in a work created by a human author, even if the work also includes AI-generated material.

- Copyright does not extend to purely AI-generated material, or material where there is insufficient human control over the expressive elements.

- Whether human contributions to AI-generated outputs are sufficient to constitute authorship must be analyzed on a case-by-case basis.

- Based on the functioning of current generally available technology, prompts do not alone provide sufficient control.

- Human authors are entitled to copyright in their works of authorship that are perceptible in AI-generated outputs, as well as the creative selection, coordination, or arrangement of material in the outputs, or creative modifications of the outputs.

- The case has not been made for additional copyright or *sui generis* protection for AI-generated content.

The Office will continue to monitor technological and legal developments to determine whether any of these conclusions should be revisited. It will also provide ongoing assistance to the public, including through additional registration guidance and an update to the *Compendium of U.S. Copyright Office Practices*.[1]

---

[1] U.S. COPYRIGHT OFFICE, COMPENDIUM OF U.S. COPYRIGHT OFFICE PRACTICES (3d ed. 2021) ("COMPENDIUM (THIRD)").

I.   INTRODUCTION

This Part of the Copyright Office's Report on Copyright and Artificial Intelligence addresses the use of AI systems to produce outputs that would be copyrightable if created by a human author.

The use of technology in the production of works of authorship is not new. Authors have used computer-assisted technology for decades to enhance, modify, and add to their creations—expanding their range of expression and advancing the goals of the copyright system. And today they are leveraging advancements in technology to push the boundaries of creativity in exciting ways. Neither the use of AI as an assistive tool nor the incorporation of AI-generated content into a larger copyrightable work affects the availability of copyright protection for the work as a whole. But the capabilities of the latest generative AI technologies[2] raise challenging questions about the nature and scope of human authorship.

These technologies now permit the creation of textual, visual, and sound outputs that resemble the creative works traditionally protected by copyright. Should these outputs also enjoy copyright protection? The answer will turn on the nature and extent of a human's contribution, and whether it qualifies as authorship of expressive elements contained in the output. Finally, to the extent that protection is not available under existing copyright principles, should the law be changed? If so, how?

A. Technology and Copyright

As stated in the legislative history of the 1976 Copyright Act, "[t]he history of copyright law has been one of gradual expansion in the types of works accorded protection."[3]

Over the years, copyright has proven flexible enough to respond to new technologies and mediums as they emerge. Protection has been extended to photographs, motion pictures, video games, and computer programs—to name just a few.[4] At the same time, courts have been called on to explore and analyze the nature of authorship. As authors have increasingly used

---

[2] "Generative AI" refers to "application[s] of AI used to generate outputs in the form of expressive material such as text, images, audio, or video." Artificial Intelligence Study: Notice of Inquiry, 88 Fed. Reg. 59942, 59948–49 (Aug. 30, 2023) ("NOI").

[3] H.R. REP. NO. 94-1496, at 51 (1976), *reprinted in* 1976 U.S.C.C.A.N. 5659, 5664.

[4] When Congress extended copyright protection to architecture, it explained that these types of works would be governed by "the general standards of originality applicable for all other copyrightable subject matter." H.R. REP. NO. 101-735, at 21 (1990), *reprinted in* 1990 U.S.C.C.A.N. 6935, 6952. Courts have also applied those standards to claims involving new technology in numerous cases. *See, e.g.*, *Meshwerks, Inc. v. Toyota Motor Sales U.S.A.*, 528 F.3d 1258, 1264–65 (10th Cir. 2008) (then-judge Neil Gorsuch stating "we do not doubt for an instant that the digital medium before us, like photography before it, can be employed to create vivid new expressions fully protectable in copyright"); *Stern Elecs., Inc. v. Kaufman*, 669 F.2d 852, 856–67 (2d Cir. 1982) (audiovisual work); *M. Kramer Mfg. Co. v. Andrews*, 783 F.2d 421, 436 (4th Cir. 1986) (video games); *Tandy Corp. v. Personal Micro Computs., Inc.*, 524 F. Supp. 171, 173 (N.D. Cal. 1981) (computer program and silicon chip).

technology in the process of creation, the relative roles of human and machine can be central to the analysis of copyrightability.

Given its role in registering claims to copyright,[5] the Copyright Office has considerable experience addressing technological developments related to the creation of works of authorship. As early as 1965, developments in computer technology began to raise "difficult questions of authorship," including whether material created using technology is "'written' by computers" or authored by human creators.[6] As then-Register of Copyrights Abraham Kaminstein observed, there is no one-size-fits-all answer:

> The crucial question appears to be whether the "work" is basically one of human authorship, with the computer merely being an assisting instrument, or whether the traditional elements of authorship in the work (literary, artistic, or musical expression or elements of selection, arrangement, etc.) were actually conceived and executed not by man but by a machine.[7]

Because the answer depends on the circumstances of a work's creation, Barbara Ringer (then-Chief of the Examining Division and future Register of Copyrights) noted that the Office could not "take the categorical position that registration will be denied merely because a computer may have been used in some manner in creating the work."[8]

The same analysis applies in the context of AI technology. For a work created using AI, like those created without it, a determination of copyrightability requires fact-specific consideration of the work and the circumstances of its creation. Where AI merely assists an author in the creative process, its use does not change the copyrightability of the output. At the other extreme, if content is entirely generated by AI, it cannot be protected by copyright.[9] Between these boundaries, various forms and combinations of human contributions can be involved in producing AI outputs.

While few bright-line rules are possible in assessing copyrightability, this Part of the Report seeks to shed more light on the relevant considerations.

---

[5] The Register of Copyrights is responsible for administering the copyright system, including examining claims for copyright registration. 17 U.S.C. §§ 410(a), 701(a). Although copyright vests automatically in an original work of authorship when fixed in a tangible medium, registration (or its refusal) provides a number of practical and legal benefits, including enabling U.S. copyright owners to enforce their exclusive rights in court. *See generally id.* §§ 106, 408(a), 410(c), 412, 411(a); U.S. Copyright Office, Circular 1: Copyright Basics (Sept. 2021), https://copyright.gov/circs/circ01.pdf.

[6] U.S. COPYRIGHT OFFICE, SIXTY-EIGHTH ANNUAL REPORT OF THE REGISTER OF COPYRIGHTS FOR THE FISCAL YEAR ENDING JUNE 30, 1965, at 5 (1966), https://www.copyright.gov/reports/annual/archive/ar-1965.pdf.

[7] *Id.*

[8] U.S. COPYRIGHT OFFICE, ANNUAL REPORT OF THE Examining Division, Copyright Office, for the Fiscal Year 1965, at 4 (1965), https://copyright.gov/reports/annual/archive/ar-examining1965.pdf.

[9] *See Thaler v. Perlmutter*, 687 F. Supp. 3d 140, 149–50 (D.D.C. 2023).

B. *The Copyright Office's AI Initiative*

In February 2022, the Copyright Office's Review Board issued a final decision affirming the refusal to register a work claimed to be generated with no human involvement.[10] A year later, the Office issued a registration for a comic book incorporating AI-generated material.[11]

In early 2023, the Office announced the launch of a broad AI Initiative and issued a statement of policy providing guidance on the registration of works incorporating AI-generated material (the "Guidance" or "AI Registration Guidance").[12] The Guidance reiterated the Office's longstanding position that human authorship is an essential requirement for copyright protection in the United States.[13] It explained that if a work contains more than a *de minimis* amount of AI-generated material, the applicant should disclose that information and provide a brief statement describing the human author's contribution.[14]

Since the Guidance was issued, the Office has registered hundreds of works that incorporate AI-generated material, with the registration covering the human author's contribution to the work.[15]

In August 2023, the Office issued a Notice of Inquiry seeking comments on a wide range of copyright law and policy issues arising from the development and use of generative AI.[16] The NOI asked five questions related to the copyrightability of material generated using AI systems:

> (1) Does the Copyright Clause in the U.S. Constitution permit copyright protection for AI-generated material?

---

[10] U.S. Copyright Office Review Board, *Decision Affirming Refusal of Registration of A Recent Entrance to Paradise* (Feb. 14, 2022), https://copyright.gov/rulings-filings/review-board/docs/a-recent-entrance-to-paradise.pdf.

[11] U.S. Copyright Office, *Cancellation Decision re: Zarya of the Dawn* (VAu001480196) at 5 (Feb. 21, 2023), https://www.copyright.gov/docs/zarya-of-the-dawn.pdf (explaining that registration covered the work's human-authored text as well as the human-authored selection, coordination, and arrangement of the work's written and visual elements, but not images generated by Midjourney that were not the product of human authorship).

[12] Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence, 88 Fed. Reg. 16190 (Mar. 16, 2023) ("AI Registration Guidance"). A copy of the guidance is available on the Office's website. U.S. COPYRIGHT OFFICE, COPYRIGHT REGISTRATION GUIDANCE: WORKS CONTAINING MATERIAL GENERATED BY ARTIFICIAL INTELLIGENCE (2023), https://copyright.gov/ai/ai_policy_guidance.pdf.

[13] AI Registration Guidance at 16191–92; *see also Thaler*, 687 F. Supp. 3d at 149–50.

[14] AI Registration Guidance at 16193; *see also* Registration Guidance for Works Containing AI-Generated Content Tr. (June 28, 2023), https://www.copyright.gov/events/ai-application-process/Registration-of-Works-with-AI-Transcript.pdf (webinar on registration of works incorporating AI-generated material).

[15] Registration records are searchable in the Office's public record, including by using keywords and filters to search the Copyright Public Record System. *Copyright Public Records System - Pilot*, U.S. COPYRIGHT OFFICE, https://publicrecords.copyright.gov/ (last visited Jan. 17, 2025).

[16] NOI.

(2) Under copyright law, are there circumstances when a human using a generative AI system should be considered the "author" of the material produced by the system?

(3) Is legal protection for AI-generated material desirable as a policy matter?

(4) If so, should it be a form of copyright or a separate *sui generis* right?

(5) Are any revisions to the Copyright Act necessary to clarify the human authorship requirement?[17]

Approximately fifty percent of the more than 10,000 comments received in response to the NOI addressed one or more of these questions. The Office refers to these comments throughout the discussion below.

---

[17] *Id.* at 59947–48.

II.   AUTHORSHIP AND ARTIFICIAL INTELLIGENCE

A. Technological Background

In the NOI, the Office defined an AI system as a "software product or service that substantially incorporates one or more AI models and is designed for use by an end-user."[18] As components to larger systems, AI models consist of computer code and numerical values (or "weights") designed to accomplish certain tasks, like generating text or images.[19]

Many of today's publicly available AI systems allow for the generation of an output from one or more inputs, such as text, images, audio, video, or a combination of mediums. A "prompt" is a common type of input, often in the form of text, that communicates the desired features of the output.[20] The AI system responds to these inputs by generating an output in the requested format (text, image, audio, video). Prompts typically describe a topic, theme, and/or subject that the user seeks to evoke, and may include the overall style, tone, and/or visual technique. Some are short and simple, such as a request for a "cartoon spaceship." Others are more detailed, requesting a litany of elements. Users may enter a prompt a single time or iteratively, refining it until the system generates an acceptable output.[21]

The practice of crafting prompts that are optimized to elicit a desired result is sometimes called "prompt engineering."[22] Prompts can also be automatically optimized by a generative AI

---

[18] NOI at 59948; *see also* James M. Inhofe National Defense Authorization Act for Fiscal Year 2023, Pub. L. 117–263, § 7223(4)(A), 136 Stat. 2395, 3669 (2022) (defining "artificial intelligence system" as "any data system, software, application, tool, or utility that operates in whole or in part using dynamic or static machine learning algorithms or other forms of artificial intelligence").

[19] NOI at 59948–49; *see* ZHANG ET AL., DIVE INTO DEEP LEARNING, ch. 1 (2023), https://d2l.ai/chapter_ introduction/index.html (ebook); GARETH JAMES ET AL., AN INTRODUCTION TO STATISTICAL LEARNING WITH APPLICATIONS IN PYTHON, at 404–05 (2023), https://www.statlearning.com/ (ebook) (explaining that the parameters of a neural network are sometimes referred to as "weights").

[20] *See, e.g.*, Leonardo Banh & Gero Strobel, *Generative Artificial Intelligence*, 33:63 ELEC. MKTS. 1, 3 (2023), https://doi.org/10.1007/s12525-023-00680-1 ("*Prompting* . . . enables end users using natural language to engage with and instruct [generative AI] application (e.g., LLMs) to create desired output such as text, images, or other types."); *Prompt*, GENLAW GLOSSARY, https://blog.genlaw.org/glossary.html#prompt ("Most generative-AI systems take [an] input (currently, this is often some text), which is then used to condition the output. This input is called the prompt.") (last visited Jan. 17, 2025); *Image Prompts*, MIDJOURNEY, https://docs.midjourney.com/docs/image-prompts ("You can use images as part of a prompt to influence a Job's composition, style, and colors.") (last visited Jan. 17, 2025); Sander Schulhoff et al., *The Prompt Report: A Systematic Survey of Prompting Techniques* at 5, ARXIV (Dec. 30, 2024), https://arxiv.org/abs/2406.06608 ("A prompt is an input to a Generative AI model, that is used to guide its output.").

[21] Other strategies are more complex, such as "prompt chaining" where a complex prompt is divided into a sequence of intermediate subtasks with a prompt for each step. Robert Clariso & Jordi Cabot, *Model-Driven Prompt Engineering*, IEEE *XPLORE*, 2023, at 48, DOI: 10.1109/MODELS58315.2023.00020.

[22] *See, e.g.*, *id.* at 47; Sander Schulhoff et al., *The Prompt Report: A Systematic Survey of Prompting Techniques* at 7, ARXIV (Dec. 30, 2024), https://arxiv.org/abs/2406.06608.

system that revises or expands them in order to improve the quality of outputs.[23] For example, ChatGPT "automatically generate[s] tailored, detailed prompts for [OpenAI's text-to-image model] DALL·E 3."[24]

As described below,[25] however, the output of current generative AI systems may include content that was not specified and exclude content that was. Although AI technology continues to advance, uncertainty around how a particular prompt or other input will influence the output may be inherent in complex AI systems built on models with billions of parameters.[26] Some observers describe AI as a "black box,"[27] and even expert researchers are limited in their ability to understand or predict the behavior of specific models.[28]

---

[23] *See, e.g.*, Siddhartha Datta et al., *Prompt Expansion for Adaptive Text-to-Image Generation* at 4, 14, ARXIV (Dec. 27, 2023), https://arxiv.org/abs/2312.16720 (describing a model that "takes a text prompt as input, . . . and outputs a set of *N* expanded text prompts that include specialized keywords (to improve image quality) and interesting additional details (to add diversity to the generated images)"); PROMPTPERFECT, https://promptperfect.jina.ai/ (last visited Jan. 17, 2025); PROMPTIST, https://foundr.ai/product/promptist (last visited Jan. 17, 2025).

[24] *DALL·E 3*, OPENAI, https://openai.com/index/dall-e-3/ (last visited Jan. 17, 2025).

[25] *See infra* notes 84–87 and pp. 24–25.

[26] *See, e.g.*, GARETH JAMES ET AL., AN INTRODUCTION TO STATISTICAL LEARNING WITH APPLICATIONS IN PYTHON, at 23–25 (2023), https://www.statlearning.com/ (ebook) (discussing the fundamental tradeoff between the flexibility and interpretability of statistical learning models, with neural networks as an example of highly flexible and difficult to interpret models); Christian Szegedy et al., *Intriguing properties of neural networks* at 1, ARXIV (Feb. 19, 2024), https://arxiv.org/abs/1312.6199 ("Neural networks achieve high performance because they can express arbitrary computation that consists of a modest number of massively parallel nonlinear steps. But as the resulting computation is automatically discovered[,] . . . it can be difficult to interpret and can have counter-intuitive properties."); Pantelis Linardatos et al., *Explainable AI: A Review of Machine Learning Interpretability Methods*, 23 ENTROPY 1, 1 (Dec. 25, 2020), https://dx.doi.org/10.3390/e23010018 (The "increasing complexity combined with the fact that vast amounts of data are used to train and develop such complex systems, while, in most cases, boost[ing] the systems' predictive power, inherently reduc[es] the[] ability to explain their inner workings and mechanisms. As a consequence, the rationale behind their decisions becomes quite hard to understand and, therefore, their predictions hard to interpret.").

[27] Steven Levy, *AI Is a Black Box. Anthropic Figured Out a Way to Look Inside*, WIRED (May 24, 2024), https://www.wired.com/story/anthropic-black-box-ai-research-neurons-features/ ("When I asked the researchers whether they were claiming to have *solved* the black box problem, their response was an instant and unanimous no."); Lou Blouin, *AI's mysterious 'black box' problem, explained*, UMDEARBORN.EDU NEWS (Mar. 6, 2023), https://umdearborn.edu/news/ais-mysterious-black-box-problem-explained. *See also infra* notes 84–87.

[28] *See, e.g.*, Trenton Bricken et al., *Towards Monosemanticity: Decomposing Language Models With Dictionary Learning*, TRANSFORMER CIRCUITS THREAD (Oct. 4, 2023), https://transformer-circuits.pub/2023/monosemantic-features/index.html ("Mechanistic interpretability seeks to understand neural networks by breaking them into components that are more easily understood than the whole. By understanding the function of each component, and how they interact, we hope to be able to reason about the behavior of the entire network."); Adly Templeton et al., *Scaling Monosemanticity: Extracting Interpretable Features from Claude 3 Sonnet*, TRANSFORMER CIRCUITS THREAD (May 21, 2024), https://transformer-circuits.pub/2024/scaling-monosemanticity/index.html ("Our work has many limitations. Some of these are superficial limitations relating to this work being early, but others are deeply fundamental challenges that require novel research to address.").

In addition, many popular AI systems are unpredictable in the sense that their outputs may vary from request to request, even with an identical prompt.[29] Some systems allow users to control this behavior and generate consistent results by setting a "seed" value, which is a number used to initialize the output generation process.[30] For example, Midjourney users can set a seed (*e.g.*, "123") and receive nearly identical images when repeating the same prompt.[31] Even these systems, however, are not always able to guarantee perfect consistency.[32]

B. *Legal Framework*

As the Office affirmed in the Guidance, copyright protection in the United States requires human authorship. This foundational principle is based on the Copyright Clause in the Constitution and the language of the Copyright Act as interpreted by the courts. The Copyright Clause grants Congress the authority to "secur[e] for limited times to authors . . . the exclusive right to their . . . writings."[33] As the Supreme Court has explained, "the author [of a copyrighted work] is . . . *the person* who translates an idea into a fixed, tangible expression entitled to copyright protection."[34]

No court has recognized copyright in material created by non-humans, and those that have spoken on this issue have rejected the possibility. In two well-known cases, the Ninth

---

[29] *See, e.g.*, *Reproducible Outputs*, OPENAI, https://platform.openai.com/docs/advanced-usage/reproducible-outputs (last visited Jan. 17, 2025); Shuyin Ouyang et al., *LLM is Like a Box of Chocolates: the Non-determinism of ChatGPT in Code Generation*, ARXIV (Oct. 17, 2024), https://arxiv.org/abs/2308.02828.

[30] *See, e.g.*, *Reproducible Outputs*, OPENAI, https://platform.openai.com/docs/advanced-usage/reproducible-outputs (last visited Jan. 17, 2025); *Seeds*, MIDJOURNEY, https://docs.midjourney.com/docs/seeds ("The Midjourney bot uses a seed number to create a field of visual noise, like television static, as a starting point to generate the initial image grids. Seed numbers are generated randomly for each image but can be specified with the --seed parameter. If you use the same seed number and prompt, you will get similar final images.") (last visited Jan. 17, 2025).

[31] *Seeds*, MIDJOURNEY, https://docs.midjourney.com/docs/seeds (last visited Jan. 17, 2025).

[32] *See* Alexander Schlögl et al., *Causes and Effects of Unanticipated Numerical Deviations in Neural Network Inference Framework*, *in* ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS 36 (A. Oh et al. eds., 2023), https://proceedings.neurips.cc/paper_files/paper/2023/hash/af076c3bdbf935b81d808e37c5ede463-Abstract-Conference.html; *Reproducible Outputs*, OPENAI, https://platform.openai.com/docs/advanced-usage/reproducible-outputs (explaining that users can obtain "mostly" deterministic outputs by setting the same seed value) (last visited Jan. 17, 2025); *Seeds*, MIDJOURNEY, https://docs.midjourney.com/docs/seeds ("Identical --seed values [for certain model versions] will produce *nearly* identical images.") (emphasis added) (last visited Jan. 17, 2025).

[33] U.S. Const. art. I, § 8, cl. 8.

[34] *Cmty. for Creative Non-Violence v. Reid* ("*CCNV*"), 490 U.S. 730, 737 (1989) (emphasis added).

Circuit held that text purportedly "authored by non-human spiritual beings"[35] and photographs that a monkey captured with a camera could not be protected by copyright.[36]

In 2023, the U.S. District Court for the District of Columbia became the first court to specifically address the copyrightability of AI-generated outputs.[37] The plaintiff challenged the Office's refusal to register an image that was described in his application as "autonomously created by a computer algorithm running on a machine."[38] Affirming the Office's refusal, the court stated that "copyright law protects only works of human creation," and that "human authorship is a bedrock requirement of copyright."[39] It found that "copyright has never stretched so far [as] . . . to protect works generated by new forms of technology operating absent any guiding human hand."[40] Because, by his own representation, the "plaintiff played no role in using the AI to generate the work," the court held that it did not meet the human authorship requirement.[41] The decision has been appealed.[42]

In most cases, however, humans will be involved in the creation process, and the work will be copyrightable to the extent that their contributions qualify as authorship. It is axiomatic that ideas or facts themselves are not protectible by copyright law,[43] and the Supreme Court has made clear that originality is required, not just time and effort. In *Feist Publications, Inc. v. Rural Telephone Service Co.*, the Court rejected the theory that "sweat of the brow" alone could be sufficient for copyright protection.[44] "To be sure," the Court further explained, "the requisite level of creativity is extremely low; even a slight amount will suffice. The vast majority of

---

[35] *Urantia Found. v. Kristen Maaherra*, 114 F.3d 955, 957–59 (9th Cir. 1997) (holding that "some element of human creativity must have occurred in order for the Book to be copyrightable" because "it is not creations of divine beings that the copyright laws were intended to protect"). While the compilation of the book was entitled to copyright, the alleged "divine messages" were not. *Id.*

[36] *Naruto v. Slater*, No. 15-cv-04324, 2016 U.S. Dist. LEXIS 11041, at *10 (N.D. Cal. Jan. 28, 2016) ("[Monkey] is not an 'author' within the meaning of the Copyright Act"), *aff'd*, 888 F.3d 418 (9th Cir. 2018) (finding that monkey cannot sue for copyright infringement).

[37] *Thaler*, 687 F. Supp. 3d 140. A second case challenging the Office's refusal to register an AI-generated output was recently filed. Compl., Allen v. Perlmutter, No. 1:24-cv-2665 (D. Colo. Sept. 26, 2024), Doc. No. 1.

[38] *Thaler*, 687 F. Supp. 3d at 142–43.

[39] *Id.* at 146.

[40] *Id.*

[41] *Id.* at 149–50.

[42] Notice of Appeal, Thaler v. Perlmutter, No. 23-5233 (D.C. Cir. Oct 18, 2023). Oral argument was heard on September 19, 2024.

[43] *See* 17 U.S.C. § 102(b); *Feist Publ'ns, Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340, 344-45 (1991) (explaining that "[t]he most fundamental axiom of copyright law is that 'no author may copyright his ideas or the facts he narrates'" (quoting *Harper & Row, Publishers, Inc. v. Nation Enters.*, 471 U.S. 539, 556 (1985))).

[44] 499 U.S. at 352–61.

works make the grade quite easily, as they possess some creative spark, 'no matter how crude, humble or obvious' it might be."[45]

More than a century ago, the Court analyzed the nature of authorship in a case involving the then-new technology of the camera. In *Burrow-Giles Lithographic Co. v. Sarony*, the Court considered a constitutional challenge to Congress's extension of copyright protection to photographs.[46] The defendant argued that photographs were not copyrightable because they lacked human authorship; instead, they were the product of a machine.[47]

The Court began its analysis by defining an "author" as "he to whom anything owes its origin; originator; maker; one who completes a work of science or literature."[48] It described copyright as "the exclusive right of a man to the production of his own genius or intellect."[49] Applying this framework, it identified numerous creative contributions made by the photographer, including "posing the [subject] in front of the camera, selecting and arranging the costume, draperies, and other various accessories," "arranging the subject so as to present graceful outlines," and "evoking the desired expression."[50] In sum, the use of a machine as a tool does not negate copyright protection, but the resulting work is copyrightable only if it contains sufficient human-authored expressive elements.

More recently, in cases involving more than one human contributor, courts have grappled with the nature of the contribution necessary to qualify as authorship. The Supreme Court provided additional guidance in the context of a commissioned sculpture. The parties in *Community for Creative Non-Violence v. Reid* ("*CCNV*") disputed who the author of the sculpture was: the nonprofit organization that conceived of it or the artist asked to make it. The Court concluded that the artist's contributions, which included sketching the design and executing his creative vision in a tangible medium of expression, made him an author.[51] In a remand to the trial court to determine whether the organization could be a joint author of the sculpture, the D.C. Circuit made clear that commissioning the sculpture and providing detailed suggestions and directions were insufficient, as such contributions constitute unprotectible ideas.[52]

The Third Circuit engaged in a similar analysis in *Andrien v. Southern Ocean County Chamber of Commerce*. *Andrien* involved an authorship claim by a plaintiff who had asked a

---

[45] *Id.* at 345.

[46] 111 U.S. 53, 55–57 (1884).

[47] *Id.* at 56, 59–60.

[48] *Id.* at 57–58 (internal quotation marks omitted).

[49] *Id.* at 58.

[50] *Id.* at 60 (internal quotation marks omitted).

[51] *CCNV*, 490 U.S. at 751–53.

[52] *Cmty. for Creative Non-Violence v. Reid*, 846 F.2d 1485, 1497 (D.C. Cir. 1988).

printer to rescale and print a collection of maps.[53] The plaintiff had "expressly directed the copy's preparation in specific detail," so that the "compilation needed only simple transcription to achieve final tangible form."[54] Because the printer "did not change the substance of [plaintiff's] original expression," the court held that the plaintiff was the author.[55] Applying *CCNV*, it stated that the author is the "party who actually creates the work, that is, the person who translates an idea into an expression that is embodied in a copy by himself or herself, or who authorizes another to embody the expression in a copy."[56]

Although an AI-generated output cannot be considered a joint work with respect to the user and AI system,[57] joint authorship provides a helpful analogy in assessing whether a party contributed sufficient expression to be considered an author.[58] To be a joint author, one must make a copyrightable contribution.[59] "A person who merely describes to an author what the

---

[53] 927 F.2d 132, 133 (3d Cir. 1991) (Under plaintiff's direction, the printer's work "included coordinating the scales, relettering the street names and adding designations for the diving sites as well as for local points of interest.").

[54] *Id.* at 135.

[55] *Id.* at 135–36. *Cf. S.O.S., Inc. v. Payday, Inc.*, 886 F.2d 1081, 1086–87 (9th Cir. 1989) (rejecting an authorship claim from a party who commissioned software noting that "[t]he supplier of an idea is no more an 'author' of a program than is the supplier of the disk on which the program is stored"); *M.G.B. Homes, Inc. v. Ameron Homes Inc.*, 903 F.2d 1486, 1493 (11th Cir. 1990) (providing sketches and ideas did not render client an "author" of the finished expression); *Geshwind v. Garrick*, 734 F. Supp. 644 (S.D.N.Y. 1990) (producer was not the author where he "wanted changes in details and aspects of the [animation clip] and even made suggestions," but did not materially constrain the animator's expression or otherwise influence how the animator executed the instructions), *vacated in part on other grounds*, 738 F. Supp. 792 (S.D.N.Y. 1990), *and aff'd*, 927 F.2d 594 (2d Cir. 1991); *Whelan Assocs. v. Jaslow Dental Lab'y*, 609 F. Supp. 1307, 1318–19 (E.D. Pa. 1985) ("general assistance and contributions to the fund of knowledge" do not make one "a creator of any original work"), *amended*, 609 F. Supp. 1325 (E.D. Pa. 1985), *aff'd*, 797 F.2d 1222 (3d Cir. 1986), *and cert. denied*, 479 U.S. 1031 (1987).

[56] *Andrien*, 927 F.2d at 134–35 ("When one authorizes embodiment, that process must be rote or mechanical transcription that does not require intellectual modification or highly technical enhancement.").

[57] A "joint work" is "a work prepared by two or more authors with the intention that their contributions be merged into inseparable or interdependent parts of a unitary whole." 17 U.S.C. § 101 (defining "joint work"). Because an AI system is not a human being, it cannot be considered an "author" in collaboration with a user. *See* Kernochan Center for Law, Media and the Arts ("Kernochan Center"), Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 6–9 (Oct. 30, 2023) ("Kernochan Center Initial Comments") (noting that machines are not "authors" within the meaning of the Copyright Act, nor are they capable of forming an intention to merge their output with the contributions from the user that interacts with these systems).

[58] *See* The Authors Guild, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 33 (Oct. 30, 2023) ("The Authors Guild Initial Comments") ("Areas of the law that will instruct courts in how to determine what is copyrightable in an AI-assisted human-created work or human-assisted AI-generated material include . . . joint work cases where the issue of whether a secondary creator contributed a sufficient amount to rise to the level of an author . . . .").

[59] *Brownstein v. Lindsay*, 742 F.3d 55, 64 (3d Cir. 2014) ("For two or more people to become co-authors, each author must contribute some non-trivial amount of creative, original, or intellectual expression to the work and both must intend that their contributions be combined."); *Ashton-Tate Corp. v. Ross*, 728 F. Supp. 597 (N.D. Cal. 1989) (finding that a contribution to a joint work must be protectable in itself and that only expressions of ideas, not ideas themselves, give rise to protected interest), *aff'd*, 916 F.2d 516, 521 (9th Cir. 1990).

commissioned work should do or look like is not a joint author for purposes of the Copyright Act."[60]

The following sections apply these legal principles in the context of generative AI systems. After describing uses of computer-assisted tools in the creation process, we discuss the following three kinds of human contribution to AI-generated outputs: (1) prompts that instruct an AI system to generate an output; (2) expressive inputs that can be perceived in AI-generated outputs; and (3) modifications or arrangements of AI-generated outputs.[61]

C. Assistive Uses of AI Systems

Many commenters expressed concern about continuing the longstanding and growing use of computer-assisted tools in the creation process.[62] They pointed to various tasks that have been performed in creative fields for years, some of which now incorporate recent developments in AI, such as "aging" or "de-aging" actors, identifying chord progressions, detecting errors in software code, and removing unwanted objects or crowds from a scene.[63]

---

[60] *Payday*, 886 F.2d at 1087; *see also Sullivan v. Flora, Inc*., 936 F.3d 562 (7th Cir. 2019) (upholding jury finding that plaintiff and defendant were not joint authors of illustrations because defendant merely offered suggestions on color, style, and text and rough outlines and sketches to guide the plaintiff's work, while the plaintiff used digital design software to create the illustrations, sometime incorporating defendant's suggestions and other times not); *BancTraining Video Sys. v. First American Corp.*, No. 91-cv-5340, 1992 U.S. App. LEXIS 3677, at *12 (6th Cir. 1992) ("Providing sketches, ideas or supervision over copyrightable material is not sufficient to make one a joint author.").

[61] Of course, many cases may involve a combination of two or more of these types of contributions. For example, a user could make creative modifications to an output generated using their own expressive input and multiple prompts.

[62] Commenters from the music industry noted that musicians and sound engineers have used such tools for many years, citing Autotune as one example. Songwriters of North America, et al., Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 6 (Oct. 30, 2023); *see also* Recording Academy, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 1 (Oct. 30, 2023) ("Recording Academy Initial Comments"). In the software industry, programmers and computer engineers use automated tools to modify software code, such as to perform refactoring and translate from one programming language into another. Apple Inc., Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 3–4 (Oct. 30, 2023) ("Apple Initial Comments").

[63] For example, commenters reported that musicians are beginning to use AI systems for developing beats or mixing a track. *See* Recording Academy Initial Comments at 3; *see also* Universal Music Group ("UMG"), Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 5–7 (Oct. 30, 2023) ("UMG Initial Comments"); Dina LaPolt, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 7 (Oct. 30, 2023) ("Dina LaPolt Initial Comments"). Motion picture companies use AI tools as part of their creative process, particularly in the context of visual effects and post-production. For example, these tools may be used for color correction, detail sharpening, or de-blurring. Motion Picture Association ("MPA"), Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 37–38 (Oct. 30, 2023) ("MPA Initial Comments"); *see also* Holton Lemaster, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry (Aug. 31, 2023) ("AI as a support tool for artists who choose to use them in their creation pipeline is fine. Crowd removal from photos, video stabilization tools, and ray tracing are all tools that really shine when enhanced by AI."). AI tools are frequently used in a process called rotoscoping, a time-consuming task that involves "altering individual frames within a single shot to align live-action and computer-generated images." MPA Initial Comments at 6, 37–38.

Commenters argued that these types of uses of AI should not affect the availability of copyright protection for the output.[64]

The Office agrees that there is an important distinction between using AI as a tool to assist in the creation of works and using AI as a stand-in for human creativity. While assistive uses that enhance human expression do not limit copyright protection, uses where an AI system makes expressive choices require further analysis. This distinction depends on how the system is being used, not on its inherent characteristics.[65]

Commenters also identified situations where creators have begun to experiment with using AI as a brainstorming tool. The Recording Academy, for instance, stated that "[m]any Academy members already use generative AI as a tool to assist them in creating new music," including through song ideation.[66] Another stakeholder noted that AI can be used to structure or create a preliminary outline for literary works.[67] In these cases, the user appears to be prompting a generative AI system and referencing, but not incorporating, the output in the development of her own work of authorship. Using AI in this way should not affect the copyrightability of the resulting human-authored work.[68]

D. Prompts

1. Commenters' Views

Many of the comments received in response to the NOI focused on the legal implications of creating outputs by providing prompts to an AI system. At the outset, as several

---

[64] *See, e.g.*, Intellectual Property Owners Association ("IPO"), Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 6–7 (Oct. 30, 2023) ("IPO Initial Comments") ("[I]t is desirable to provide copyright protection for works resulting from a human using an AI system as a tool of creativity and where that human activity satisfies the traditional requirements of human authorship. A lack of this protection would be detrimental to rights holders and creators alike.").

[65] One commenter urged the Office to adopt a distinction based on the type of AI platform a user employs. Scenario, Inc., Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 6 (Oct. 18, 2023) ("Scenario Initial Comments") (arguing that output generated by a multimodal generative AI platform should presumptively be deemed copyrightable, while output generated by a unimodal generative AI platform should presumptively be deemed uncopyrightable).

[66] Recording Academy Initial Comments at 10.

[67] Literary Works Listening Session Tr. at 31:18–23 (Apr. 19, 2023) (statement by Mary Rasenberger, The Authors Guild).

[68] Other examples of such uses provided by commenters include digital and copy editing and other uses that "are intended to assist, not displace, human creativity." Recording Academy Initial Comments at 3; Lori Wilde, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry (Oct. 24, 2023); IPO Initial Comments at 2; Authors Alliance, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 3 (Oct. 30, 2023) ("Authors Alliance Initial Comments").

commenters noted, prompts themselves, if sufficiently creative, may be copyrightable.[69] The copyright status of the output generated, however, is a separate question.[70]

Most commenters agreed that inputting simple prompts is insufficient to make a user the author of the AI-generated output.[71] Several described prompts as unprotectible ideas or instructions.[72] The American Society of Composers, Authors and Publishers ("ASCAP"), a performing rights organization, asserted that "[w]here a human's involvement is limited to the simple generation of minimal queries and prompts for an AI tool, the resulting material is not entitled to copyright protection."[73] The Brooklyn Law Incubator & Policy Clinic asserted that a simple, general prompt lacks "enough human creativity for the output to qualify for copyright protection."[74] Universal Music Group ("UMG") stated: "The prompting user is no more an

---

[69] *See* AI Registration Guidance at 16192 n.27; The Authors Guild Initial Comments at 32 n.39 (arguing that the creator of a prompt "has a copyright in the prompt assuming it has sufficient original expression"); American Association of Independent Music ("A2IM") and the Recording Industry Association of America, Inc. ("RIAA"), Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 34 (Oct. 30, 2023) ("A2IM-RIAA Joint Initial Comments"); Daniel Gervais, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 6 (Oct. 30, 2023) ("Daniel Gervais Initial Comments").

[70] *See generally* A2IM-RIAA Joint Initial Comments at 34 ("While the text of those prompts may be independently copyrightable if sufficiently expressive, that does not confer upon the author of the prompt any copyright in the output generated by the AI system."); Johan Brandstedt, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 29 (Oct. 29, 2023) ("Johan Brandstedt Initial Comments").

[71] Commenters used "simple" with varying degrees of specificity, generally referring to prompts that contain only generic descriptions or a short number of words. *See, e.g.*, Donaldson Callif Perez, LLP, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 2 (Oct. 30, 2023) ("Donaldson Callif Perez Initial Comments") ("[W]e agree that simple prompts by humans that result in a complex, creative work should not be granted copyright protection."); Dina LaPolt Initial Comments at 7 (stating that "a user inputting a simple generic prompt" should not be able to claim copyright protection); Edward Lee, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 11 (Oct. 30, 2023) (stating that "a simple one- or two-word prompt" is unlikely to satisfy the minimum standard for copyright protection in the output); Peer Music and Boomy, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 11 (Oct. 30, 2023) ("Peer Music-Boomy Joint Initial Comments") (finding it difficult to imagine how a single prompt that produces a complex output could provide a basis for claiming copyright protection in the output).

[72] *See, e.g.*, Adobe Inc., Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 5–6 (Oct. 30, 2023) ("Adobe Initial Comments") ("[A] prompt is not copyrightable because the prompt represents the idea."); Johan Brandstedt Initial Comments at 19 (stating that "prompts express *ideas*, image and text generators provide stored *expression*"); European Writers' Council ("EWC"), Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 16 (Oct. 30, 2023) ("EWC Initial Comments") (stating that "the person formulating the prompts [cannot] claim any rights with respect to the results on the basis of the prompts alone, because the mere formulation of the task and the choice between several results proposed by the AI system is not a creative or protectable act").

[73] ASCAP, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 48–49 (Oct. 30, 2023) ("ASCAP Initial Comments").

[74] Brooklyn Law Incubator & Policy Clinic ("BLIP"), Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 20 (Oct. 30, 2023) ("BLIP Initial Comments"); *see also* Qualcomm Incorporated, Reply Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 7 (Dec. 6, 2023) ("Qualcomm Reply Comments") (stating that output "based on a single, general prompt with *de minimis* creativity" lacks "requisite human expression").

author than someone who tells a musician friend to 'write me a pretty love song in a major key' and then falsely claims co-ownership."[75]

By contrast, other commenters disputed the notion that prompts merely "influence" the AI system and do not provide "specific instructions to create a particular expressive result."[76] For example, the Intellectual Property Owners Association stated that "[i]f a user prompts Midjourney to produce an image or series of images of a city scape under water, the user is going to get a city scape under water."[77]

Commenters' views on more detailed prompts, including those that are revised and repeated, varied. Some viewed highly detailed prompts as sufficient to make some AI-generated outputs copyrightable.[78] Professors Pamela Samuelson, Christopher Jon Sprigman, and Matthew Sag stated that "[s]ophisticated prompts that specify details of an image should be sufficient to meet the [human authorship] requirement," and that "[a] person who instructs a Generative AI with enough detail, such that model output reflects that person's original

---

[75] Letter from UMG, Summary of *Ex Parte* Meeting on Apr. 22, 2024, Regarding the Office's AI Study, to U.S. Copyright Office 11 (Dec. 3, 2024) (arguing that "users prompting [music generative AI companies] to generate audio files are not composing or writing anything, much less 'their own, original music'" and instead are "simply supply[ing] an uncopyrightable idea in a text prompt . . . and the software itself generates an audio track based on its own predictive algorithms").

[76] IPO Initial Comments at 5; Van Lindberg, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 41 (Oct. 30, 2023) ("Van Lindberg Initial Comments") ("Randomness is part of the generative process—but *the output of an AI model is not random*. A human using the AI system typically describes what should be generated and/or provides other inputs that are used to initialize and guide the generative process."); Ashley Greenwald, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 11 (Oct. 30, 2023) (arguing that interior designers initiate generative AI systems by "giv[ing] certain prompts and instructions," refining and modifying interim results, and "mak[ing] the *final determination* whether and how the output co-created with the help of generative AI tools should be utilized"); Christa Laser, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 5 (Oct. 30, 2023) ("Christa Laser Initial Comments") ("A few uses of generative AI employ random strings and undirected outcomes, but a more significant role of generative AI is to implement a human's extensive creativity, direction, and selection towards an outcome pre-dreamed in the human mind.").

[77] IPO Initial Comments at 5.

[78] *See* BLIP Initial Comments at 23 (stating that users "may provide very detailed and extensive prompts to an AI-system to ensure that its output is as close as possible to what they anticipated" and such outputs should be copyrightable if "they provided sufficient input and prompts to control the output of an AI system"); Van Lindberg Initial Comments at 42 (stating that "the more information that is given within the prompt, the more control is exerted over the output"); Law Office of Seth Polansky LLC, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 26 (Oct. 12, 2023) ("Seth Polansky Initial Comments") ("[A] human who closely guides the output of a generative AI system through curated training or by providing detailed prompts may be able to claim some form of 'joint authorship' with the machine."); Donaldson Callif Perez Initial Comments at 2 ("[I]f someone spends a significant amount of effort creating very specific and detailed prompts to create a complex work, perhaps there should be some copyright protection for that work.").

conception of the work, should be regarded as the author of the resulting work."[79] Another commenter asserted that, with detailed prompts, users "can achieve remarkable control over the expressive elements of the work, such as lighting, pose, style, expressions, and setting."[80]

In contrast, the Authors Guild argued that the unpredictability of the prompt-to-output generation process may make it "difficult to show that there was sufficient control and consequently a sufficient closeness between 'conception and execution.'"[81] Others agreed.[82] Adobe, for instance, stated that "[w]hen you submit a prompt (or idea), you then receive an output based solely on the AI's interpretation of that prompt," and the "AI's expression of [that] idea is not copyrightable."[83]

Several commenters described AI systems as black boxes,[84] meaning that not only do users in most cases not know what "will inform the [AI's] response" to prompts,[85] but that even developers of AI systems cannot generally predict outputs or explain why they include certain

---

[79] Pamela Samuelson et al., Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 3 (Oct. 30, 2023) ("Pamela Samuelson et al. Initial Comments"); *see also* MPA Initial Comments at 47 (predicting that prompts could become "much more detailed" as technology improves to the point where "the inputs themselves may provide the substantive content for the output" and concluding that "[a] rule that prompts would never satisfy the human authorship requirement neglects those likely possibilities").

[80] Christa Laser Initial Comments at 5. Several other commenters cautioned that while there may "be cases where the prompts are so directive and detailed" that the user could be entitled to copyright protection for the output, this is likely to be rare. The Authors Guild Initial Comments at 32; *see also* Daniel Gervais Initial Comments at 6 (describing as "exceptional" cases "in which a detailed prompt . . . could contain expressions of specific ideas that reflect human creative choices directly perceptible in the machine's output").

[81] The Authors Guild Initial Comments at 31.

[82] *See* Association of Medical Illustrators, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 8–9 (Oct. 30, 2023); Kernochan Center Initial Comments at 5–6.

[83] Adobe Initial Comments at 5–6; *see also* Johan Brandstedt Initial Comments at 14, 29; EWC Initial Comments at 16.

[84] *See, e.g.*, Professional Photographers of America, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 7 (Oct. 30, 2023); SeaQVN, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 63 (Sept. 13, 2023); IAC Inc. and Dotdash Media Inc., d/b/a Dotdash Meredith, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 3 (Oct. 30, 2023); Eric Bourdages, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 1 (Nov. 26, 2023); James Horvath, Reply Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 1 (Sep. 13, 2023); Cooper Reid, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry (Aug. 31, 2023).

[85] Kernochan Center Initial Comments at 5; *see also* Gabriel Moise, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry (Aug. 31, 2023); The Authors Guild Initial Comments at 31–32; Vikas Hassija et al., *Interpreting Black-Box Models: A Review on Explainable Artificial Intelligence*, 16 COGNITIVE COMPUTATION 45, 47 (2024), https://link.springer.com/article/10.1007/s12559-023-10179-8 (noting that "the internal workings of [a black-box] model are not easily accessible or interpretable" and that this "lack of transparency" makes it difficult "to understand the model's behavior").

elements and not others.[86] Some provided examples of prompts containing detailed descriptions of what the user had in mind, where the output omitted some elements requested and inserted others.[87]

Commenters also noted that prompts are often entered into an AI system in one medium (such as text) and the output is generated in a different medium (such as a visual image, video, or audio clip). Several commenters asserted that moving from one medium to another requires interpretation, and where AI provides that interpretation, the user's control over the execution of their idea is indirect.[88] UMG highlighted one popular text-to-music generator that cautions users, "[n]o matter how detailed[,] text prompts cannot fully define an actual piece of music."[89]

Some stressed that generative AI systems can produce a seemingly limitless number of variations in response to the same prompt, no matter how many times that prompt is used.[90] The Kernochan Center argued that "[e]xtending the scope of copyright protection in the written prompts to cover the multiplicity of potential outputs" that may be generated by an AI system "comes uncomfortably close to conferring a copyright in a method of generating images (or other works)," which would be prohibited under section 102(b).[91]

---

[86] *See* EWC Initial Comments at 9 ("In computer science, PROCESSING (computation) is consistently described as a black box; not even the operators of AI systems know exactly what happens during the learning process—and they do not control it."); *see also supra* Section II.A.

[87] *See* Kernochan Center Initial Comments at 8–9 & n.13 (noting that "even highly elaborated prompts will . . . yield multiple outputs (not all of them fully or accurately responsive to the prompts)" and providing examples). *See also* Tonio Inverness, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 5 (Sept. 12, 2023) (demonstrating labor that goes into refining prompts after the results of initial prompt were "not at all what [commenter] had in mind"); UMG Initial Comments at 76–77.

[88] Johan Brandstedt Initial Comments at 14, 19 (stating that "anything started in writing ought not to merit copyright claims over an image"); Kernochan Center Initial Comments at 8 (stating that a textual description "would need to evince an extremely high degree of precision" in order to claim copyright in a pictorial work produced through the use of those instructions); The Authors Guild Initial Comments at 31 n.36 (stating that converting a "text instruction to images created from training data makes the output unpredictable").

[89] Letter from UMG, Summary of *Ex Parte* Meeting on Apr. 22, 2024 Regarding the Office's AI Study, to U.S. Copyright Office 3 (Dec. 3, 2024) (internal citation omitted); *see also How do I make music with Udio?*, UDIO, https://www.udio.com/guide (last visited Jan. 17, 2025) (explaining that prompts cannot fully define an output because "the same text describes an infinite number of possible audio tracks").

[90] *See, e.g.*, Kernochan Center Initial Comments at 8–9; The Authors Guild Initial Comments at 32 n.39.

[91] Kernochan Center Initial Comments at 8–9; *see also* 17 U.S.C. § 102(b) (excluding from copyright protection "any idea, procedure, process, system, method of operation, concept, principle, or discovery, regardless of the form in which it is described, explained, illustrated, or embodied in such work").

A few commenters asserted that human-directed revisions to prompts may result in greater control over an output's expressive elements.[92] One technique entails submitting a prompt to the AI system, then revising the prompt, either by adding, removing, or replacing certain terms based on the initial output produced, to generate a new output. The user may revise and repeat upwards of hundreds of times.[93] Eventually the system may generate an output that meets the user's needs; if not, the user may decide to revise the prompt again or abandon the effort. Commenters noted that this process can require a significant amount of time and "demonstrable human effort."[94]

Some commenters advanced a theory of "authorship by adoption" (though few used that phrase).[95] They suggested that a user may exercise creative judgment when deciding to accept the output produced by a generative AI system. One suggested that a user who "repeatedly enters prompts until the output matches their desired expression" is no different than an "artist who continues to dab paint on the canvas until the image matches the painter's vision."[96] In contrast, the Authors Guild likened repetitive prompting to "spinning a roulette

---

[92] *See, e.g.*, Evangelical Christian Publishers Association ("ECPA"), Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 8 (Oct. 30, 2023) ("ECPA Initial Comments") ("If the issue is one of control and predictability, fine-tuning repeatedly until the final expression is satisfactory demonstrates the author's ultimate control of the final work, even if each iteration leading up to the final expression may be subject to unpredictability."); SCA Robotics, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 2 (Sept. 29, 2023) ("SCA Robotics Initial Comments") (stating that authorship should depend on factors such as "the human user's control of the artistic expression outputted by the platform," including "the extent of the human party's discretion over accepting and/or modifying the outputted work"); International Center for Law & Economics, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 18 (Oct. 30, 2023) ("AI systems remain tools that require human direction and judgment. As such, when a person provides the initial prompt or framing, makes choices regarding the iterative development of the AI output, and decides that the result is satisfactory for inclusion in a final work, they are fundamentally engaging in creative decision making that constitutes authorship under copyright law.").

[93] *See* IPO Initial Comments at 5 (noting that "[t]he same user might iterate on dozens, even hundreds, of prompts of greater complexity and specificity before achieving a desired result").

[94] Superframe, LLC, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry (Sept. 6, 2023); *see also* AI and Metaverse Task Force of the Trust over IP Foundation, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 4 (Oct. 30, 2023); Donaldson Callif Perez Initial Comments at 2.

[95] This theory would find authorship in the decision to adopt something unplanned or unexpected occurring in the course of creating a work. *See* Jane C. Ginsburg & Luke Ali Budiardjo, *Authors and Machines*, 34 BERKELEY TECH. L.J. 343 (2019). It can be traced to *Alfred Bell & Co. v. Catalda Fine Arts*, which assessed the originality of mezzotint engravings that were based on paintings in the public domain. 191 F.2d 99, 104 (2d Cir. 1951). The defendant argued that the engravings were mere copies of preexisting paintings, and therefore not protected by copyright. *Id.* In finding that the engraver's versions were sufficiently different, the court speculated that "[a] copyist's bad eyesight or defective musculature, or a shock caused by a clap of thunder, may yield sufficiently distinguishable variations." *Id.* at 105. "Having hit upon such a variation unintentionally," the court held that "the 'author' may adopt it as his and copyright it." *Id.*

[96] ECPA Initial Comments at 7.

wheel with infinite possibilities."[97] It argued that "when a user [metaphorically] spins the wheel dozens of times until they land on an output they like," such activity should not give the user a right to claim ownership of that output.[98]

Discussing the authorship by adoption theory, Professors Jane Ginsburg and Luke Ali Budiardjo concluded that, "[w]ere post-execution adoption to substitute for any authorial participation, even indirect or inadvertent, in giving physical form to a work, then, in addition to [naming] the 'wrong' author, copyright law would effectively vest adopters with rights in ideas."[99] Professor Daniel Gervais made a similar point with the following analogy: "If I walk into a gallery or shop that specializes in African savanna paintings or pictures because I am looking for a specific idea (say, an elephant at sunset, with trees in the distance), I may find a painting or picture that fits my idea," but "[t]hat in no way makes me an author."[100]

2. Analysis

The Office concludes that, given current generally available technology, prompts alone do not provide sufficient human control to make users of an AI system the authors of the output. Prompts essentially function as instructions that convey unprotectible ideas. While highly detailed prompts could contain the user's desired expressive elements, at present they do not control how the AI system processes them in generating the output.

Cases regarding joint authorship support this conclusion. These cases address the amount of control that is necessary to claim authorship. The provision of detailed directions, without influence over how those directions are executed, is insufficient.[101] As the Third Circuit explained, when a person hires someone to execute their expression, "that process must be rote or mechanical transcription that does not require intellectual modification or highly technical enhancement" for the delegating party to claim copyright authorship in the final work.[102] Although entering prompts into a generative AI system can be seen as similar to providing instructions to an artist commissioned to create a work, there are key differences. In a human-to-human collaboration, the hiring party is able to oversee, direct, and understand the contributions of a commissioned human artist. Depending on the nature of each party's contributions, the artist may be the sole author, or the outcome may be a joint work or work

---

[97] The Authors Guild Initial Comments at 31–32.

[98] *Id.*; *see also* Kernochan Center Initial Comments at 9 (asserting that "selection of a single output is not itself a creative act"); Daniel Gervais Initial Comments at 6–7; Johan Brandstedt Initial Comments at 29.

[99] Ginsburg & Budiardjo, *supra* note 95, at 370.

[100] Daniel Gervais Initial Comments at 7.

[101] *Payday*, 886 F.2d at 1087; *see, e.g.*, *CCNV*, 490 U.S. at 737 ("As a general rule, the author is the party who actually creates the work, that is, the person who translates an idea into a fixed, tangible expression entitled to copyright protection.").

[102] *Andrien*, 927 F.2d at 134–35.

made for hire.[103] In theory, AI systems could someday allow users to exert so much control over how their expression is reflected in an output that the system's contribution would become rote or mechanical.[104] The evidence as to the operation of today's AI systems indicates that this is not currently the case. Prompts do not appear to adequately determine the expressive elements produced, or control how the system translates them into an output.[105]

The gaps between prompts and resulting outputs demonstrate that the user lacks control over the conversion of their ideas into fixed expression, and the system is largely responsible for determining the expressive elements in the output. In other words, prompts may reflect a user's mental conception or idea, but they do not control the way that idea is expressed. This is even clearer in the case of generative AI systems that modify or rewrite prompts internally. That process recasts the human contribution—however detailed it may be—into a different form.

The following image, which the Office generated by entering a prompt into a popular commercially available AI system, illustrates this point:[106]

Prompt	*Output*
professional photo, bespectacled cat in a robe reading the Sunday newspaper and smoking a pipe, foggy, wet, stormy, 70mm, cinematic, highly detailed wood, cinematic lighting, intricate, sharp focus, medium shot, (centered image composition), (professionally color graded), ((bright soft diffused light)), volumetric fog, hdr 4k, 8k, realistic	

This prompt describes the subject matter of the desired output, the setting for the scene, the style of the image, and placement of the main subject. The resulting image reflects some of these instructions (*e.g.*, a bespectacled cat smoking a pipe), but not others (*e.g.*, a highly detailed wood environment). Where no instructions were provided, the AI system filled in the gaps.

---

[103] In contrast, AI systems cannot produce joint works or works made for hire because they are not "authors," they are not capable of forming an intention to merge their output with the user's contributions, and they cannot enter into binding contracts. *See* Kernochan Center Initial Comments at 7; Brief for Appellees, at 27, Thaler v. Perlmutter, No. 23-5233 (D.C. Cir. Mar. 6, 2024).

[104] This outcome would raise additional questions about the utility of AI in creative expression.

[105] *Cf. Geshwind*, 734 F. Supp. at 650–51 ("The fact that the agent, Geshwind, wanted changes in details and aspects of the portrait and even made suggestions, the compliance with which may or may not have improved the effect, does not make him the creator."); *M.G.B. Homes*, 903 F.2d at 1493; *Payday*, 886 F.2d at 1087.

[106] The Office used Google's generative AI chatbot Gemini to generate this image. GEMINI, https://gemini.google.com/ (last visited Jan. 17, 2025).

For instance, the prompt does not specify the cat's breed or coloring, size, pose, any attributes of its facial features or expression, or what clothes, if any, it should wear beneath the robe. Nothing in the prompt indicates that the newspaper should be held by an incongruous human hand.

The fact that identical prompts can generate multiple different outputs further indicates a lack of human control.[107] As one popular system explains on its website, "[n]o matter how detailed . . . the same text describes an infinite number of possible" outputs.[108] In these circumstances, the black box of the AI system is providing varying interpretations of the user's directions.

Repeatedly revising prompts does not change this analysis or provide a sufficient basis for claiming copyright in the output. First, the time, expense, or effort involved in creating a work by revising prompts is irrelevant, as copyright protects original authorship, not hard work or "sweat of the brow."[109] Second, inputting a revised prompt does not appear to be materially different in operation from inputting a single prompt. By revising and submitting prompts multiple times, the user is "re-rolling" the dice, causing the system to generate more outputs from which to select, but not altering the degree of control over the process.[110] No matter how many times a prompt is revised and resubmitted, the final output reflects the user's acceptance of the AI system's interpretation, rather than authorship of the expression it contains.

Some commenters drew analogies to a Jackson Pollock painting or to nature photography taken with a stationary camera, which may be eligible for copyright protection even if the author does not control where paint may hit the canvas or when a wild animal may step into the frame.[111] However, these works differ from AI-generated materials in that the human author is principally responsible for the execution of the idea and the determination of the expressive elements in the resulting work. Jackson Pollock's process of creation did not end with his vision of a work. He controlled the choice of colors, number of layers, depth of texture, placement of each addition to the overall composition—and used his own body movements to execute each of these choices. In the case of a nature photograph, any copyright protection is based primarily on the angle, location, speed, and exposure chosen by the photographer in

---

[107] *See supra* note 32. The Office re-ran the prompt above and received a much different image of a cat in a stormy setting.

[108] *How do I make music with Udio?*, UDIO, https://www.udio.com/guide (emphasis omitted) (last visited Jan. 17, 2025).

[109] *Feist*, 499 U.S. at 352.

[110] *See, e.g.*, Kernochan Center Initial Comments at 8 ("If each prompt newly rolls the dice, it is difficult to discern the dominance of will that 'direction' implies, and thus hard to classify it as meeting the requirement of an objective 'intent.'").

[111] *See, e.g.*, Tim Boucher, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 8 (Oct. 26, 2023); Christa Laser Initial Comments at 4; MPA Initial Comments at 47–50; Pamela Samuelson et al. Initial Comments at 4.

setting up the camera, and possibly post-production editing of the footage.[112] As one commenter explained, "some element of randomness does not eliminate authorship," but "the putative author must be able to constrain or channel the program's processing of the source material."[113] The issue is the degree of human control, rather than the predictability of the outcome.[114]

The Office also agrees that authorship by adoption does not in itself provide a basis for claiming copyright in AI-generated outputs. As commenters noted, providing instructions to a machine and selecting an output does not equate to authorship.[115] Selecting an AI-generated output among uncontrolled options is more analogous to curating a "living garden," than applying splattered paint.[116] As the Kernochan Center observed, "selection among the offered options" produced by such a system cannot be considered copyrightable authorship, because the "selection of a single output is not itself a creative act."[117]

There may come a time when prompts can sufficiently control expressive elements in AI-generated outputs to reflect human authorship. If further advances in technology provide users with increased control over those expressive elements, a different conclusion may be called for.[118] On the other hand, technological advancements that facilitate increased automation and optimization may bolster our current conclusions. For example, if generative

---

[112] Like other copyrighted works, nature photography must have a sufficient amount of creative expression to satisfy the originality standard.

[113] Kernochan Center Initial Comments at 5.

[114] *See* Digital Media Licensing Association ("DMLA"), Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 16 (Oct. 30, 2023) ("DMLA Initial Comments") (stating that "the foreseeability of the AI's results may bear on authorship" in cases "where there is a limited range of specific expressive output that is objectively foreseeable as a result of a human user's prompt"); Kernochan Center Initial Comments at 5; MPA Initial Comments at 45–46 (acknowledging that evaluating "the elements of predictability and control may be appropriate in certain cases"); International Trademark Association ("INTA"), Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 5 (Oct. 30, 2023) ("INTA Initial Comments") (acknowledging that if a program generated an image by simply populating "each pixel with a randomly-selected color, it seems obvious that the resulting work should not be considered a work of authorship"); The Authors Guild Initial Comments at 31.

[115] The Authors Guild Initial Comments at 31–32; Daniel Gervais Initial Comments at 6–7; Kernochan Center Initial Comments at 8.

[116] *Kelley v. Chicago Park Dist.*, 635 F.3d 290, 304 (7th Cir. 2011); *see also* COMPENDIUM (THIRD) § 306; *Thaler*, 687 F. Supp. 3d at 146 (holding that the "key" to copyright protection is "[h]uman involvement in, and ultimate creative control over, the work at issue").

[117] Kernochan Center Initial Comments at 9.

[118] *See* Authors Alliance Initial Comments at 19 ("[A]s both generative AI systems and the ways that creators use them change and evolve, the application of the human authorship requirement to content that is AI-generated or AI-assisted may also change. For example, if these tools developed in a way that would give creators more control over the outputs, works created with these tools could potentially be considered works of human authorship.").

AI systems integrate or further improve automated prompt optimization, users' control may be diminished.

E. *Expressive Inputs*

As discussed above, AI systems take inputs in the form of text, images, audio, video, or a combination of these mediums. Some systems—whether via tools, settings, or prompts—allow inputs to be substantially retained as part of the output. For example, one commenter noted that a human author may create an original illustration, input that work into an AI system, and instruct the system "to modify [the] color or layer portions of [the] existing image."[119] Another observed that an AI system may be used to modify or translate a copyrighted work,[120] such as uploading a story written in the first person and instructing the system to convert it to a third-person point of view.

These types of expressive inputs, while they may be seen as a form of prompts, are different from those that merely communicate desired outcomes. As commenters pointed out, where human-authored inputs are reflected in the output, they contribute more than just an intellectual conception. One explained that "a human author who inputs their own illustration or media file" into an AI system "may have a greater claim to authorship," because "there is a limited range of specific expressive output that is objectively foreseeable as a result of a human user's" contribution.[121] Another noted that when a user provides an input to an AI system such as "a traditional work created or designated by the user . . . the specified starting point constrains the 'autonomy' of the outputs" and thus may "present a more persuasive case of human intervention" than simply applying "prompts to an unknown starting point."[122]

---

[119] DMLA Initial Comments at 16.

[120] Pearson, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 7 (Oct. 30, 2023) ("Pearson Initial Comments").

[121] DMLA Initial Comments at 16; *see* Pearson Initial Comments at 7–8 (acknowledging that "copyright can only protect material that is the product of human activity" and stating that "further consideration should be given to whether a claim of authorship in output may exist where the input itself is a representation of the original intellectual conception of an author"); National Music Publishers' Association ("NMPA"), Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 30 (Oct. 30, 2023) ("NMPA Initial Comments") ("Creators that use AI to refine, recast, or modify, or to create new derivative works based on their preexisting works may also have legitimate claims of authorship over the resulting work in some circumstances.").

[122] Kernochan Center Initial Comments at 5–6; *see* MPA Initial Comments at 50 (noting that "material human creators provide to the AI tool" such as "inputs, like a drawing or photo" can be considered "intellectual and creative contributions that are inseparable from the ultimate work").

As an example, in the following work submitted to the Office for registration, the author had created a hand-drawn illustration and used it as an input, along with the prompt shown below.[123]

The AI system produced this output:

Prompt	*Input*	*Output*
"a young cyborg woman (((roses))) flowers coming out of her head, photorealism, cinematic lighting, hyper realism, 8k, hyper detailed."		

The drawing itself is a copyrightable work, and its expressive elements are clearly perceptible in the output, including the outline of the mask, the position of the nose, mouth, and cheekbones relative to the shape of the mask, the arrangement of the stems and rosebuds, and the shape and placement of the four leaves.

The applicant disclaimed "any non-human expression" appearing in the final work, such as the realistic, three-dimensional representation of the nose, lips, and rosebuds, as well as the lighting and shadows in the background. After reviewing the information provided in the application, the Office registered the work with an annotation stating: "Registration limited to unaltered human pictorial authorship that is clearly perceptible in the deposit and separable from the non-human expression that is excluded from the claim."[124]

---

[123] *Rose Enigma*, VAu001528922 (Mar. 21, 2023). More about the artist Kris Kashtanova's creation of this work is available on their website. *Portfolio: Rose Enigma*, KRIS KASHTANOVA, https://www.kris.art/portfolio-2/rose-enigma (last visited Jan. 17, 2025).

[124] *Rose Enigma*, VAu001528922 (Mar. 21, 2023). By contrast, the Office's Review Board upheld a refusal to register an image produced by an AI system with a human author's photograph as an input. U.S. Copyright Office Review Board, *Decision Affirming Refusal of Registration of Suryast* at 1 (Dec. 11, 2023), https://copyright.gov/rulings-filings/review-board/docs/SURYAST.pdf. The applicant disclosed that the image was generated by "RAGHAV Artificial Intelligence Painting App" ("RAGHAV"), which had been trained on Vincent van Gogh's *The Starry Night*—with an instruction to apply the style of *The Starry Night* to the photograph. *Id*. at 2. The Board found that the resulting image did not "contain sufficient human authorship necessary to sustain a claim to copyright" because the applicant "exerted insufficient creative control over RAGHAV's" generation of the output. *Id*. at 3, 7–8. Unlike *Rose Enigma*, the output did not clearly show the copyrightable work input by the applicant. *See id.* at 7–8.

As illustrated in this example, where a human inputs their own copyrightable work and that work is perceptible in the output, they will be the author of at least that portion of the output. Their own creative expression will be protected by copyright, with a scope analogous to that in a derivative work. Just as derivative work protection is limited to the material added by the later author,[125] copyright in this type of AI-generated output would cover the perceptible human expression. It may also cover the selection, coordination, and arrangement of the human-authored and AI-generated material, even though it would not extend to the AI-generated elements standing alone.

F. Modifying or Arranging AI-Generated Content

Generating content with AI is often an initial or intermediate step, and human authorship may be added in the final product. As explained in the AI Registration Guidance, "a human may select or arrange AI-generated material in a sufficiently creative way that 'the resulting work as a whole constitutes an original work of authorship.'"[126] A human may also "modify material originally generated by AI technology to such a degree that the modifications meet the standard for copyright protection."[127]

As several commenters noted, human authors should be able to claim copyright if they select, coordinate, and arrange AI-generated material in a creative way.[128] This would provide protection for the output as a whole (although not the AI-generated material alone).[129] A relatively common scenario in registration applications is the combination of human-authored text with AI-generated images. In one early case, for instance, the Office found that the selection and arrangement of AI-generated images with human-authored text in a comic book were protectable as a compilation. We explained:

---

[125] *See* H.R. REP. NO. 94-1476, at 57; S. REP. NO. 94-473, at 55 ("[C]opyright in a 'new version' covers only the material added by the later author, and has no effect one way or the other on the copyright or public domain status of the preexisting material.").

[126] AI Registration Guidance at 16192.

[127] *Id.* at 16192–93.

[128] *See, e.g.*, BLIP Initial Comments at 20; Center for Art Law, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 9 (Oct. 26, 2023); Cisco Systems, Inc., Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 7 (Oct. 30, 2023) ("Cisco Initial Comments"); IPO Initial Comments at 4–6; Peer Music-Boomy Joint Initial Comments at 12.

[129] *See Feist*, 499 U.S. at 348 (noting that copyright protection for a compilation "may extend only to those components of a work that are original to the author").

> [T]he Office finds that the compilation of these images and text throughout the Work contains sufficient creativity under *Feist* to be protected by copyright. Specifically, the Office finds the Work is the product of creative choices with respect to the selection of the images that make up the Work and the placement and arrangement of the images and text on each of the Work's pages. Copyright therefore protects [the applicant's] authorship of the overall selection, coordination, and arrangement of the text and visual elements that make up the Work.[130]

Multiple similar registrations have been made since then.[131]

A number of commenters also made the point that if a user edits, adapts, enhances, or modifies AI-generated output in a way that contributes new authorship, the output would be entitled to protection.[132] They argued that these modifications "should be assessed in the same way as . . . editorial or other changes to a pre-existing work."[133] Although such works would not technically qualify as "derivative works,"[134] derivative authorship provides a helpful analogy in identifying originality. Again, the copyright would extend to the material the human author contributed but would not extend to the underlying AI-generated content itself.[135]

---

[130] U.S. Copyright Office, *Cancellation Decision re: Zarya of the Dawn (VAu001480196)* at 5 (Feb. 21, 2023), https://www.copyright.gov/docs/zarya-of-the-dawn.pdf.

[131] *See supra* notes 15, 123.

[132] *See, e.g.*, Apple Initial Comments at 1; ASCAP Initial Comments at 49; The Authors Guild Initial Comments at 32; BLIP Initial Comments at 25; Cisco Initial Comments at 7; Graphic Artists Guild, Inc., Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 19 (Oct. 30, 2023) ("Graphic Artists Guild Initial Comments"); OpenAI, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 15 (Oct. 30, 2023).

[133] Kernochan Center Initial Comments at 6.

[134] A derivative work is "a work based upon one or more preexisting works." 17 U.S.C. § 101 (defining "derivative work"). Because entirely AI-generated outputs do not contain the human authorship required to be a "work of authorship," the modified versions cannot qualify under this definition. *See* H.R. REP. NO. 94-1476, at 57; S. REP. NO. 94-473, at 55 (noting that "the 'pre-existing work' must come within the general subject matter of copyright set forth in section 102, regardless of whether it is or was ever copyrighted").

[135] *See* H.R. REP. NO. 94-1476, at 57; S. REP. NO. 94-473, at 55.

Many popular AI platforms offer tools that encourage users to select, edit, and adapt AI-generated content in an iterative fashion. Midjourney, for instance, offers what it calls "Vary Region and Remix Prompting," which allow users to select and regenerate regions of an image with a modified prompt. In the "Getting Started" section of its website, Midjourney provides the following images to demonstrate how these tools work.[136]

| (1) Generate Candidate Images with Prompt: *meadow trail lithograph* | (2) Select and Upscale Image | (3) Use Freehand Editing Tool to Select Region | (4) Generate Candidate Images with Prompt: *meadow stream lithograph* | (5) Select and Upscale Image |

The image was further modified by repeating the editing process:

Other generative AI systems also offer tools that similarly allow users to exert control over the selection, arrangement, and content of the final output.[137]

---

[136] *Vary Region + Remix*, MIDJOURNEY, https://docs.midjourney.com/docs/vary-region-remix (last visited Jan. 17, 2025). Text descriptions below each image were added by the Office.

[137] OpenAI's ChatGPT, for instance, has a feature called "canvas," which provides an interactive interface for users to "collaborate" with the model while writing a document or code. Users can edit AI-generated text; highlight regions for the model to focus on; use built-in tools to request in-line suggestions, length adjustments, and changes to the reading level; and write instructions that detail particular edits to be made. *See Introducing Canvas*, OPENAI (Oct. 3, 2024), https://openai.com/index/introducing-canvas/.

Unlike prompts alone, these tools can enable the user to control the selection and placement of individual creative elements. Whether such modifications rise to the minimum standard of originality required under *Feist* will depend on a case-by-case determination.[138] In those cases where they do, the output should be copyrightable.

Similarly, the inclusion of elements of AI-generated content in a larger human-authored work does not affect the copyrightability of the larger human-authored work as a whole.[139] For example, a film that includes AI-generated special effects or background artwork is copyrightable, even if the AI effects and artwork separately are not.

---

[138] The selection, coordination, and arrangement of only two or three elements is not generally sufficient for copyright protection. *See* COMPENDIUM (THIRD) § 312.2 ("[T]he Office generally will not register a compilation containing only two or three elements, because the selection is necessarily *de minimis*." (citing H.R. REP. NO. 94-1476, at 122 (stating that a work does not qualify as a collective work "where relatively few separate elements have been brought together," as in the case of "a composition consisting of words and music, a work published with illustrations or front matter, or three one-act plays"))).

[139] *Cf.* AI Registration Guidance at 16192–93.

III. INTERNATIONAL APPROACHES

Other countries are also analyzing whether copyright protection should extend to works containing AI-generated material. Those that have addressed this issue so far have agreed that copyright requires human authorship.

The Korean Copyright Commission and the Ministry of Culture, Sports and Tourism issued *A Guide on Generative AI and Copyright* in 2023, in which it explained that "only a natural person can become an author"[140] and that "copyright registration for an AI output is impossible if a human did not contribute creatively to the expressive form."[141] The Korean guidance noted that "if a human had performed *additional work* on the AI output, such as modifying, or making additions or deletions, only the part that had undergone such change is copyrightable."[142] It also stated that an author can register a work as a compilation if he or she selected and rearranged the AI output creatively.[143]

In Japan, the Copyright Subdivision of the Cultural Council published a summary of its guidelines in May 2024 titled *General Understanding on AI and Copyright in Japan*.[144] The guidelines explained that the copyrightability of AI-generated content will be determined on a case-by-case basis, depending on the following factors: (1) the amount and content of the instructions and input prompts by the AI user; (2) the number of generation attempts; (3) the selection by the AI user from multiple output materials; and (4) any subsequent human additions and corrections to the AI-generated work.[145]

In the People's Republic of China, the Beijing Internet Court evaluated arguments in a copyright infringement case involving an AI-generated work in 2023, starting with the premise that human authorship was required for copyright protection.[146] It found that an image created

---

[140] Ministry of Culture, Sports and Tourism & Korea Copyright Comm'n, *A Guide on Generative AI and Copyright*, at 40 (2023), https://www.copyright.or.kr/eng/doc/etc_pdf/Guide_on_Generative_AI_and_Copyright.pdf.

[141] *Id.* at 41.

[142] *Id.*

[143] *Id.* It has been reported that a copyright registration was granted in December 2023 for an AI-generated film based on the "human editing of the AI[-]generated film and images." Edward Lee, *South Korea grants copyright to AI generated work, 'AI Suro's Wife' film as work edited by humans*, CHATGPT IS EATING THE WORLD (Jan. 8, 2024), https://chatgptiseatingtheworld.com/2024/01/08/south-korea-grants-copyright-to-ai-generated-work-ai-suros-wife-film-as-work-edited-by-humans/.

[144] Legal Subcommittee under the Copyright Subdivision of the Cultural Council, *General Understanding on AI and Copyright in Japan* (May 2024), https://www.bunka.go.jp/english/policy/copyright/pdf/94055801_01.pdf.

[145] *Id.* at 17.

[146] Li Mou Mou Su Liu Mou Mou Qin Hai Zuo Pin Shu Ming Quan, Xin Xi Wang Luo Chuan Bo Quan Jiu Fen An (李某某诉刘某某侵害作品署名权, 信息网络传播权纠纷案) [Li v. Liu, Dispute over Copyright Infringement of the Right of Attribution and Right of Information Network Distribution of Works], at 14 (Beijing Internet Ct. Nov. 27, 2023), https://english.bjinternetcourt.gov.cn/pdf/BeijingInternetCourtCivilJudgment112792023.pdf. Page numbers in this Report are based on the English translation released by the Beijing Internet Court online.

using Stable Diffusion was protected under China's copyright law,[147] and that the person who used AI to create the image was the author.[148] According to the court, the selection of over 150 prompts combined with subsequent adjustments and modifications demonstrated that the image was the result of the author's "intellectual achievements," reflecting his personalized expression.[149]

In the European Union, the majority of member states agreed, in response to a 2024 policy questionnaire on the relationship between generative AI and copyright, that current copyright principles adequately address the copyright eligibility of AI outputs and there is no need to provide new or additional protection.[150] Member states also shared the view that AI-generated content may be eligible for copyright "only if the *human input in* [*the*] *creative process was significant*."[151] This consensus extended to the understanding that purely AI-generated works cannot be protected by copyright, as only a natural person can be considered an author.[152] Based on similar reasoning, in 2024, a court in Czechia, also known as the Czech Republic, held that an AI tool cannot be the author of a copyrighted work.[153]

In the United Kingdom, a statute predating the development of generative AI technologies protects works "generated by computer in circumstances such that there is no human author of the work."[154] It designates the author as a "person by whom the arrangements

---

[147] *Id.* at 10–14; *see also* Copyright Law of the People's Republic of China (promulgated by the Standing Comm. Nat'l Cong., Feb. 26th, 2010, effective Apr. 1, 2010), art. 3.

[148] *See supra* note 146 at 14–15. While the ruling is not precedential under Chinese judicial practice, it may inform policies and practices about the copyrightability of AI-generated art under Chinese law. *Id.* at 11–12. China has recently considered statutory clarifications for when a work generated by AI is protected under copyright. A preliminary draft of China's proposed AI law states that when a work generated using AI meets the conditions under the Copyright Law then it can be protected under that law "based on the extent of the user's contribution to the final presentation of the content." Zhong Hua Ren Min Gong He Guo Ren Gong Zhi Neng Fa (Xue Zhe Jian Yi Gao) (中华人民共和国人工智能法 (学者建议稿)) [Law of the People's Republic of China on Artificial Intelligence (Scholar's Draft Proposal)], art. 36, Official WeChat account of the Digi. Rule of Law Inst. at East China Univ. of Political Sci. and L., *translated by* Center for Sec. and Emerging Tech., https://cset.georgetown.edu/wp-content/uploads/t0592 _china_ai_law_draft_EN.pdf.

[149] *See supra* note 146 at 11–12.

[150] Council of the European Union, *Policy questionnaire on the relationship between generative Artificial Intelligence and copyright and related rights – Revised Presidency summary of the Member States contributions*, at 16–18 (Dec. 20, 2024), https://data.consilium.europa.eu/doc/document/ST-16710-2024-REV-1/en/pdf.

[151] *Id.* at 16.

[152] *Id.* at 15.

[153] *See* Tomáš Ščerba & Jaroslav Fořt, *The first Czech case on generative AI*, TECH.'S LEGAL EDGE (Apr. 4, 2024), https://www.technologyslegaledge.com/2024/04/the-first-czech-case-on-generative-ai/; Alessandro Cerri, *Czech court finds that AI tool DALL-E cannot be the author of a copyright work*, THE IPKAT (Apr. 15, 2024), https://ipkitten.blogspot.com/2024/04/czech-court-finds-that-ai-tool-dall-e.html.

[154] Copyright, Designs and Patents Act 1988, c. X, I, §§ 178, 9(3) (UK), https://www.legislation.gov.uk /ukpga/1988/48/data.pdf. Protection lasts for fifty years from the date the work is made. *Id.*, c. I, § 12(7).

necessary for the creation of the work are undertaken."[155]　In 2021, the United Kingdom Intellectual Property Office ("UKIPO") sought public comment on whether to change this law, in light of advancements in generative AI.　Based on the lack of any case law applying this provision to AI,[156] the UKIPO concluded that "[i]t is unclear whether removing [protection for computer-generated works] would either promote or discourage innovation and the use of AI for the public good."[157]　It elected to leave the law in place but did not rule out future changes.[158] Since then, the UK government has initiated a new consultation on copyright and AI, including questions about prompts, computer-generated works, and outputs of AI models.[159]

Several other former and current commonwealth countries, such as Hong Kong,[160] India,[161] and New Zealand,[162] have enacted similar provisions, but there too it is unclear whether or how they will apply to AI-generated works.

In Canada, a 2021 review of the Copyright Act acknowledged a lack of clarity concerning the authorship of an AI-generated work.[163]　While the Standing Committee on Industry, Science and Technology, which led the review, recommended that legislation should

---

[155] *Id.*, c. I, § 9(3).

[156] UKIPO, *Consultation outcome of the Intell. Prop. Office on Artificial Intelligence and Intellectual Property: copyright and patents*, ¶ 22 (June 28, 2022), https://www.gov.uk/government/consultations/artificial-intelligence-and-ip-copyright-and-patents/outcome/artificial-intelligence-and-intellectual-property-copyright-and-patents-government-response-to-consultation#copyright-in-computer-generated-works.

[157] *Id.* ¶ 29.

[158] *Id.* ¶¶ 29–30.

[159] *See* UKIPO, *Open Consultation of the Intell. Prop. Office on Copyright and Artificial Intelligence* (Dec. 17, 2024), https://www.gov.uk/government/consultations/copyright-and-artificial-intelligence/copyright-and-artificial-intelligence#bcopyright-and-artificial-intelligence.

[160] Section 11(3) of Hong Kong's Copyright Ordinance states: "In the case of a literary, dramatic, musical or artistic work which is computer-generated, the author is taken to be the person by whom the arrangements necessary for the creation of the work are undertaken."　Copyright Ordinance, (2019) Cap. 528, § 11(3) (H.K.).

[161] Section 2(d)(vi) of India's Copyright Act defines author as "in relation to any literary, dramatic, musical or artistic work which is computer-generated, the person who causes the work to be created."　The Copyright Act, 1957, § 2(d)(vi).　Without citing that section, in 2020 the Indian Copyright Office registered the AI-generated work described in note 124, listing the AI tool as a co-author, but a year later issued a notice of withdrawal of the registration. Sukanya Sarkar, *Exclusive: Indian Copyright Office issues withdrawal notice to AI co-author*, MANAGINGIP (Dec. 13, 2021), https://www.managingip.com/article/2a5d0jj2zjo7fajsjwwlc/exclusive-indian-copyright-office-issues-withdrawal-notice-to-ai-co-author.

[162] Section 5(2)(a) of New Zealand's copyright law defines author as "in the case of a literary, dramatic, musical, or artistic work that is computer-generated, the person by whom the arrangements necessary for the creation of the work are undertaken."　Copyright Act 1994, s 5(2)(a).

[163] Innovation, Sci. and Econ. Dev. Canada ("ISED Canada"), *A Consultation on a Modern Copyright Framework for Artificial Intelligence and the Internet of Things*, at 12 (2021), https://ised-isde.canada.ca/site/strategic-policy-sector/sites/default/files/attachments/2022/ConsultationPaperAIEN.pdf.

provide greater clarity, the Canadian Parliament has not yet acted on the recommendation.[164] In 2023, Canada relaunched the consultation process, with a comment period that closed in January 2024.[165]

Similarly, in Australia, participants in 2024 consultations held by the Select Committee on Adopting Artificial Intelligence shared concerns over the lack of clarity in Australia's copyright laws regarding the "extent of copyright protection, if any, that is afforded to works created by humans with the assistance or augmentation of AI."[166] The Select Committee in its recommendations, however, did not specifically address this issue or suggest any action.

Although some level of consensus on the need for human authorship appears to be emerging, and most countries have so far continued to apply existing law, it is clear that views are still being formed. It remains to be seen how copyrightability standards will be interpreted and applied. The Office is closely monitoring developments abroad and evaluating how other countries' evolving approaches may ultimately overlap or differ from our own.

---

[164] *Id.* at 13.

[165] ISED Canada, *Consultation on Copyright in the Age of Generative Artificial Intelligence* (2021), https://ised-isde.canada.ca/site/strategic-policy-sector/en/marketplace-framework-policy/consultation-copyright-age-generative-artificial-intelligence.

[166] Select Committee on Adopting Artificial Intelligence, Parliament of Australia (Final Report, November 2024) ¶ 4.166, https://parlinfo.aph.gov.au/parlInfo/download/committees/reportsen/RB000470/toc_pdf/SelectCommitteeonAdoptingArtificialIntelligence(AI).pdf.

IV. THE ARGUMENTS FOR LEGAL CHANGE

A. Providing Incentives

Commenters generally stressed the value of incentives to produce new works of authorship.[167] They differed, however, in their interpretations of the Copyright Clause and their assessment of the impact of providing such incentives for AI-generated content.

Those supporting copyright protection for AI-generated material contended that it would encourage the creation of more works, furthering progress in culture and knowledge to the benefit of the public.[168] They took the position that the Copyright Clause should be read flexibly to encompass new technologies.[169] For instance, one commenter argued that this interpretation should "evolve with technological advancements" to ensure that "the spirit of this mandate continues to foster innovation and artistic expression in all its forms."[170]

Most commenters that opined on this issue, however, agreed with the Office's view that the Copyright Clause requires human authorship.[171] They supported the conclusion that AI-

---

[167] *See, e.g.*, A2IM-RIAA Joint Initial Comments at 4 n.11 (quoting *Thaler v. Perlmutter*, No. 22-cv-1564, 2023 WL 5333236, at *4 (D.D.C. Aug. 18, 2023)); Copyright Alliance, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 5 (Oct. 30, 2023) ("Copyright Alliance Initial Comments"); DMLA Initial Comments at 17–18; Graphic Artists Guild Initial Comments at 1; Internet Archive, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 10–11 (Oct. 30, 2023) ("Internet Archive Initial Comments").

[168] *See, e.g.*, Dallas Joder, Reply Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 3 (Nov. 30, 2023) ("Dallas Joder Reply Comments"); Peer Music-Boomy Joint Initial Comments at 14.

[169] For example, AI company BigBear.ai asserted that the Constitution "does not prohibit protection of AI-generated material," and that the availability of copyright protection "should not depend on the method through which [it] was generated." BigBear.ai Holdings, Inc., Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 25 (Oct. 18, 2023); *see also* Ryan Abbott, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 6 (Oct. 21, 2023) ("Ryan Abbott Initial Comments") ("The history and purpose of the Constitution and the Copyright Act both weigh in favor of the protection of AI-generated works because the public interest trumps any direct benefit to authors."); Peer Music-Boomy Joint Initial Comments at 15 ("[W]e do not believe that placing limitations on creators by limiting the sort of output we incentivize furthers the constitutional aims of copyright."); BLIP Initial Comments at 25 ("The Copyright Act should be amended to include a new section that provides protection for AI-generated material.").

[170] Dallas Joder Reply Comments at 3; *see also* Duane Valz, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 3 (Oct. 18, 2023) (While "the authors of the Constitution may not have imagined that entities other than natural persons would ever qualify as authors or inventors. . . .[, t]his doesn't mean that new types or persons or entities cannot be made eligible as authors or owners of copyrights if Congress sees fit to deem them such.").

[171] *See* A2IM-RIAA Joint Initial Comments at 34–35; The Authors Guild Initial Comments at 34; Anonymous AI Technical Writer, Reply Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 15 (Dec. 6, 2023) ("Anonymous AI Technical Writer Reply Comments"); Copyright Alliance Initial Comments at 96–97; DMLA Initial Comments at 17–18; Graphic Artists Guild Initial Comments at 20; David Newhoff, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 3 (Oct. 17, 2023) ("David Newhoff Initial Comments"); UMG Initial Comments at 81–82.

generated material can only be protected where there is sufficient human involvement or where AI is used as a tool to enhance human expression.[172]

These commenters emphasized that the Copyright Clause refers to promoting progress specifically by providing *authors* with legal and economic incentives.[173] They noted that AI systems, by contrast, are inanimate objects that "do not need an incentive to create."[174] As one commenter stated, "AIs do the work they are programmed to do, without regard to incentives."[175]

---

[172] *See* American Bar Association, Intellectual Property Law Section ("ABA-IPL"), Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 14 (Oct. 30, 2023) ("ABA-IPL Initial Comments"); American Intellectual Property Law Association ("AIPLA"), Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 11 (Oct. 30, 2023) ("AIPLA Initial Comments"); Johan Brandstedt Initial Comments at 30; ACT | The App Association ("App Association"), Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 6–7 (Oct. 30, 2023) ("App Association Initial Comments"); Entertainment Software Association, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 7 (Oct. 30, 2023); IPO Initial Comments at 7; Recording Academy Initial Comments at 11; Scenario Initial Comments at 16–17.

[173] *See* A2IM-RIAA Joint Initial Comments at 4 (quoting *Thaler v. Perlmutter*, No. 22-cv-1564, 2023 WL 5333236, at *4 (D.D.C. Aug. 18, 2023)), 35; ASCAP Initial Comments at 50; Authors Alliance Initial Comments at 18–19; DMLA Initial Comments at 17–18; Graphic Artists Guild Initial Comments at 1, 20; Daniel Gervais Initial Comments at 7; Fight for the Future, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 8 (Oct. 30, 2023) ("Fight for the Future Initial Comments"); Internet Archive Initial Comments at 10–11; Kernochan Center Initial Comments at 10–11; David Newhoff Initial Comments at 3; NMPA Initial Comments at 29–30; Seth Polansky Initial Comments at 29; Copyright Alliance Initial Comments at 5.

[174] Google LLC, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 12 (Oct. 30, 2023); *see also* Computer & Communications Industry Association ("CCIA"), Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 19 (Oct. 30, 2023) ("CCIA Initial Comments") ("Computers don't need incentives; only people do. And existing incentives—both legal, such as copyrights and patents, and non-legal, such as first-mover advantages and a desire to supply a commercial need—will suffice to ensure the development of generative AI technologies."); AIPLA Initial Comments at 11; NMPA Initial Comments at 29 ("As a policy matter, copyright law should never protect purely AI-generated content that does not represent human expression. Existing copyright law rightfully incentivizes human creativity by granting protection to the 'the fruits of intellectual labor' that 'are founded in the creative powers of the mind.'"); Xiyin Tang et al., Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 10–11 (Oct. 30, 2023) ("Xiyin Tang et al. Initial Comments") ("The artificial intelligence itself needs no incentive, as it is programmed to create, and needs only human prompting to generate works. The only other party that could need the incentive of copyright would be the users of AI systems. However, creation of works using AI technology requires substantially less time and effort than most human created works. Humans receive copyright protection for their works to balance against the cost of creating those works, and the risk in investing so much time and resources only for another party to copy the finished product. With AI-created works, 'both the fixed and variable costs of producing each copyrightable article are effectively zero, which allows producers to compete with imitators even absent legal protection.'" (citations omitted)).

[175] Pamela Samuelson et al. Initial Comments at 3. *See also* A2IM-RIAA Joint Initial Comments at 4 (quoting *Thaler v. Perlmutter*, No. 22-cv-1564, 2023 WL 5333236, at *4 (D.D.C. Aug. 18, 2023)); Association of American Publishers ("AAP"), Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 31–32 (Oct. 30, 2023) ("AAP Initial Comments"); CCIA Initial Comments at 19; Internet Archive Initial Comments at 10–11 ("The traditional policy foundations for extending copyright protection generally do not apply in the case of AI-generated material. There is no evidence that copyright law provides necessary incentives for the creation of AI-generated works, and regardless,

Several commenters asserted that there appear to be sufficient incentives for AI companies under existing law.[176] Some pointed out that the exponential growth of AI technologies—even in the absence of copyright protection—indicates that their developers do not need copyright incentives to produce these technologies.[177] "As machine learning practitioners," the AI company Hugging Face, stated: "[W]e find that very little to no innovation in generative AI is driven by the hope of obtaining copyright protection for model outputs. The incentives for innovation already exist without modifying copyright law."[178]

Finally, many expressed concern that providing legal protection to AI-generated content would discourage human authorship. Representatives of copyright owners maintained that the proliferation of legally protected AI-generated outputs would stifle creativity, leading to an overall decrease in human-authored works available to the public because humans will be disincentivized to create.[179] For example, the Copyright Alliance predicted that "[i]f . . .

---

the constitutional foundations of copyright make clear that its goal is to incentivize human authorship."). *But see* Dallas Joder Reply Comments at 4 (predicting that self-aware AI might someday "rationally respond to [intellectual property ("IP")] incentives just like humans," such that they should be "permitted to keep and profit from the fruits of their creativity").

[176] A2IM-RIAA Joint Initial Comments at 35; AAP Initial Comments at 31–32; AIPLA Initial Comments at 11 ("At this time, it does not appear that legal protection for AI-generated outputs is critical to incentivizing the creation of AI technologies and systems; and the copyrightability of the AI system itself is sufficient."); CCIA Initial Comments at 19; Copyright Alliance Initial Comments at 95–96. Commenters identified several incentives, separate from any potential legal protection in AI-generated outputs, that encourage the development of AI technologies. *See, e.g.*, R Street Institute, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 10 (Oct. 30, 2023) ("R Street Initial Comments") ("Existing copyright protection for computer code does offer some incentives for the development of generative AI technologies."); Xiyin Tang et al. Initial Comments at 10–11 ("There are already incentives for the creation and development of AI technology through patent and copyright protection in the machinery and software, so the developers of AI have been sufficiently incentivized to create and improve their programs."); CCIA Initial Comments at 19 (discussing perceived commercial need and first-mover advantage); Anonymous AI Technical Writer Reply Comments at 15 (discussing the availability of venture capital and stock-market funding for AI development); DMLA Initial Comments at 17 (discussing patents and trade secrets); UMG Initial Comments at 81 (discussing AI as a tool or service).

[177] AIPLA Initial Comments at 11 (noting that AI systems were "generated and commercialized in the absence of any clear authority providing legal protection to the outputs, and the absence of such protections does not appear to have diminished the public's interest in consuming AI, nor service-providers' interest in providing it"); The Authors Guild Initial Comments at 33; Copyright Alliance Initial Comments at 95–96; Graphic Artists Guild Initial Comments at 19–20.

[178] Hugging Face, Inc., Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 13 (Oct. 30, 2023).

[179] Take Creative Control, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 2 (Oct. 18, 2023); Software Freedom Conservancy, Reply Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 2 (Dec. 6, 2023); Timothy Allen, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry (Sept. 6, 2023) ("Not only does it prevent people from being able to claim any kind of ownership to their designs, it also creates a great degree of consumer confusion as to which pieces are real and which are not, and could have a chilling effect on further creative fields (many of which are already deeply suffering economically)[.]"); Anonymous Artist, Reply Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 1, 10 (Dec. 5, 2023); Letter from UMG, Summary of *Ex Parte* Meeting on Apr. 22, 2024 Regarding the Office's AI Study, to U.S. Copyright Office 6, 14 (Dec. 3, 2024).

policymakers give incentives to generate AI content, the sheer volume and speed with which AI material is generated could obliterate the markets for much human creation."[180] It further asserted that "[o]ur popular culture will be overtaken by low quality, AI-generated works because the cost of human creation would be deemed too burdensome in comparison to using AI."[181] The Authors Guild cautioned that if "AI-generated works were entitled to the same protection as human-created works," the producers of this material would have an "unfair leverage in the marketplace" which "would further incentivize the distribution of AI-generated content to the public, crowding and diluting the marketplace to the point that copyright incentives no longer function as intended."[182] It expressed particular concern that "[t]he creative middle class professions . . . will be drowned out and decimated," and that "our literary works and arts will suffer tremendously as a result."[183]

Some commenters sought to achieve the perceived value of incentives outside of the copyright system, proposing that AI-generated works could be protected instead through the establishment of new *sui generis* rights. They suggested that a "specialized right could be tailored to address the unique aspects of AI creations, including the balance between human input and AI processing," the term of protection, and the identity of rightsholders, among others.[184]

Of the commenters who addressed *sui generis* rights specifically, most opposed the idea. They saw *sui generis* rights as raising similar concerns about incentives and the impact on

---

[180] Copyright Alliance Initial Comments at 95; *see also* David Newhoff Initial Comments at 2–3 (explaining that vesting copyrights in corporate production of AI-generated material "pos[es] a threat to the careers of creative professionals" and that "[b]eyond posing a threat to the careers of creative professionals (and to the cultural value of creative work), at a certain point, the application of copyright law itself may become irrelevant and/or unconstitutional"); The Authors Guild Initial Comments at 34 ("Few human creators will be able to earn enough to sustain a profession and the human quality of work produced by professionals . . . will disappear."); Fight for the Future Initial Comments at 6. *But see* Donaldson Callif Perez Initial Comments at 2 ("Critics of artificial intelligence worry that the technology will eradicate jobs and be used to replace artists at the expense of human stories. Its proponents say that it is the way of the future and should be treated like just another tool in an artist's toolbox. The truth likely lies somewhere in the middle."); UMG, Reply Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 11 (Dec. 6, 2023); A2IM-RIAA Joint Initial Comments at 35.

[181] Copyright Alliance Initial Comments at 95.

[182] The Authors Guild Initial Comments at 34.

[183] *Id.*

[184] ImageRights International, Inc. ("ImageRights"), Reply Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 9 (Dec. 6, 2023) ("ImageRights Reply Comments"); *see also* Seth Polansky Initial Comments at 29 (suggesting shorter term for AI-generated material and clearer definition of who owns rights in outputs); Public Knowledge, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 19 (Oct. 30, 2023) (arguing that benefits of a *sui generis* right "may include faster and cheaper registration, and a lowered standard of documentation to illustrate which parts are attributable to AI, and (potentially) provenance of the work's AI components"); Rightsify Group LLC, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 10 (Oct. 30, 2023). A few advocated for *sui generis* protection specifically for AI model weights. *See* BLIP Initial Comments at 25–26; Van Lindberg Initial Comments at 5.

human authors.[185] Some also characterized past experience with *sui generis* regimes as problematic in various respects.[186]

In the Office's view, the case has not been made for additional protection for AI-generated material beyond that provided by existing law. As an initial matter, because copyright requires human authorship, copyright law cannot be the basis of protection for works that do not satisfy that requirement. As most commenters recognized, the incentives authorized by the Copyright Clause are to be provided to human authors as the means to promote progress. While Congress could instead consider establishing *sui generis* rights,[187] we do not find the policy arguments for additional protection to be persuasive.

To begin with, it is not clear that new incentives are needed. The developers of AI models and systems already enjoy meaningful incentives under existing law (as indicated by the rapid development and adoption of those models and systems). These incentives include patent, copyright, and trade-secret protection for the machinery and software, as well as potential funding and first-mover advantages. Moreover, we are not convinced that providing further incentives would promote progress. We share the concerns expressed about the impact of AI-generated material on human authors and the value that their creative expression provides to society. If a flood of easily and rapidly AI-generated content drowns out human-authored works in the marketplace, additional legal protection would undermine rather than advance the goals of the copyright system. The availability of vastly more works to choose from could actually make it harder to find inspiring or enlightening content. Indeed, AI

---

[185] *See, e.g.*, The Authors Guild Initial Comments at 33 (arguing that *sui generis* rights "will dilute the market for human-created works and . . . does not serve the goals of copyright or the needs of society"); EWC Initial Comments at 17; AAP Initial Comments at 31–32; ABA-IPL Initial Comments at 13–14; ASCAP Initial Comments at 49; Authors Alliance Initial Comments at 18–19; Kernochan Center Initial Comments at 10; NMPA Initial Comments at 29; App Association Initial Comments at 7; Pamela Samuelson et al. Initial Comments at 4; AIPLA Initial Comments at 11; R Street Initial Comments at 10.

[186] Consumer Technology Association, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 6 (Oct. 30, 2023) ("The history of *sui generis* approaches has been that as technology advances, they either quickly become obsolete (e.g., Semiconductor Chip Protection Act of 1984), or may raise uncertainties and impediments pertaining to copyright.").

[187] *See, e.g.*, the Semiconductor Chip Protection Act of 1984, establishing *sui generis* rights in mask works. H.R. REP. NO. 98-781, at 7–8 (1984), *reprinted in* 1984 U.S.C.C.A.N. 5750, 5756–57; Trademark Clarification Act of 1984, Pub. L. No. 98-620, § 301, 98 Stat. 3335, 3347 (1984); 17 U.S.C. §§ 901–14. *See also* the Vessel Hull Design Protection Act, establishing *sui generis* protection for original designs of vessel hulls. Digital Millenium Copyright Act, Pub. L. No. 105-304, Title V, § 502, 112 Stat. 2860, 2905 (1998), *amended by* the IP and Communications Omnibus Act of 1999, Pub. L. No. 106-113, § 5005, 113 Stat. 1536, 1501A–593 (1999); 17 U.S.C. §§ 1301–32. These rights differ from copyright in terms of eligibility, ownership rights, registration procedures, term, and remedies. It is difficult, however, to extrapolate from these examples, since experience with their use is limited and the context of today's widely used AI technologies is quite different.

training itself is reportedly reliant on human-generated content, with synthetic data leading to lower-quality results.[188]

There are already indications that AI-generated content has impacted some creators' ability to be compensated for their work.[189] Musicians and songwriters, for instance, have been impacted by the proliferation of AI-generated content on streaming services. UMG reported that "content oversupply," produced by an estimated 170 million AI-generated music tracks, currently threatens to dilute human creators' royalties.[190] AI-generated works have also threatened to reduce the pool of royalties available to human creators through the Mechanical Licensing Collective.[191]

If authors cannot make a living from their craft, they are likely to produce fewer works. And in our view, society would be poorer if the sparks of human creativity become fewer or dimmer.

B. *Empowering Creators with Disabilities*

A number of commenters asserted that extending protection to AI-generated works would empower more individuals with physical and cognitive disabilities to create.[192] The

---

[188] Kristian Hammond et al., *Degenerative AI: The Risks of Training Systems on their own Data*, NORTHWESTERN UNIV. CENTER FOR ADVANCING SAFETY OF MACHINE INTELL. (Sept. 6, 2024), https://casmi. northwestern.edu/news/articles/2024/degenerative-ai-the-risks-of-training-systems-on-their-own-data.html; Aatish Bhatia, *When A.I.'s Output Is a Threat to A.I. Itself*, N.Y. TIMES (Aug. 25, 2024), https://www. nytimes.com/interactive/2024/08/26/upshot/ai-synthetic-data.html.

[189] Researchers are beginning to seek to quantify the impacts of AI on artists' livelihoods. *See, e.g.*, International Confederation of Societies of Authors and Composers ("CISAC"), STUDY ON THE ECONOMIC IMPACT OF GENERATIVE AI IN THE MUSIC AND AUDIOVISUAL INDUSTRIES (Nov. 2024), https://www.cisac.org/services/reports-and-research/cisacpmp-strategy-ai-study; Gaétan de Rassenfosse et al., *Intellectual Property and Creative Machines*, NAT'L BUREAU OF ECON. RSCH. WORKING PAPERS, July 2024, Working Paper No. 32698, https://www.nber.org/papers/w32698.

[190] UMG Initial Comments at 13.

[191] Under a blanket license established in Section 115 of the Copyright Act, royalties for digital phonorecord deliveries of nondramatic musical works are paid into a pool for the mechanical licensing collective to divide and distribute to copyright owners. Although the Office has clarified that musical works that lack human authorship are not eligible for the blanket license, parties have attempted to obtain royalties for streams of AI-generated content. Letter from Suzanne V. Wilson, Gen. Couns. and Assoc. Register of Copyrights, U.S. Copyright Office, to Kris Ahrend, Chief Exec. Officer, The Mechanical Licensing Collective (Apr. 20, 2023), https://copyright.gov/ai/USCO-Guidance-Letter-to-The-MLC-Letter-on-AI-Created-Works.pdf. Such conduct has even been the basis of a criminal indictment for fraud. Press Release, U.S. Attorney's Office, Southern District of New York, North Carolina Musician Charged with Music Streaming Fraud Aided by Artificial Intelligence (Sept. 4, 2024), https://www.justice.gov/usao-sdny/pr/north-carolina-musician-charged-music-streaming-fraud-aided-artificial-intelligence.

[192] *See, e.g.*, BLIP Initial Comments at 24; ECPA Initial Comments at 8; Tom Yonge, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 33–36 (Sept. 18, 2023). Some commenters illustrated how generative AI has helped them create despite their disabilities. *See* Elisa Rae Shupe, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry Initial Comments at 1 (Oct. 27, 2023); Michael Summey, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 1 (Oct. 30, 2023).

specific applications they identified, however, involve the use of AI as a tool to assist in creating works, rather than to generate output without human authorship. The Brooklyn Law Incubator & Policy Clinic, for instance, cited functionalities like text-to-speech, visual art generative algorithms, and improving the written communication of those with cognitive disabilities.[193] Discussing creators with disabilities, another noted that "AI acts as a tool in the hands of an author," rather than a source of expressive content.[194]

The Office strongly supports the empowerment of all creators, including those with disabilities. We stress that to the extent these functionalities are used as tools to recast, transform, or adapt an author's expression, copyright protection would be available for the resulting work.[195] For example, the Office recently considered an application to register a sound recording by GRAMMY-winning country artist Randy Travis, who has limited speech function following a stroke.[196] The track was created based on the recording of a human voice, using "[a] special-purpose AI vocal model . . . as a tool . . . to help realize the sounds that Mr. Travis and the other members of the human creative team desired."[197] The result, which would have been infeasible without this technology, was a new track appearing to be sung in Travis's legendary voice. Because the sound recording used AI as a tool, not to generate expression, the Office registered the work.

The distinction between assistive uses and generative ones applies equally to creators with disabilities and other human authors. Copyright protection remains available where AI functions as an assistive tool that allows human authors to express their creativity.

C. Countering International Competition

A few commenters raised concerns about international competition. One organization warned that without copyright protection in the United States, "the scientific and creative communities will not be able to exploit the economic value of [AI-generated works]," which "may contribute to the U.S. lagging in the development of generative AI technologies and

---

[193] BLIP Initial Comments at 24.

[194] *See* ECPA Initial Comments at 8 (discussing artists who are not able to hold a paintbrush and stating that creators with disabilities are "wielding [AI] to create intended expression").

[195] Registration Guidance for Works Containing AI-Generated Content Tr. at 4–5 (June 28, 2023), https://www.copyright.gov/events/ai-application-process/Registration-of-Works-with-AI-Transcript.pdf.

[196] *Where That Came From*, SR0001018989 (May 29, 2024).

[197] Letter from Steven Englund to U.S. Copyright Office (Oct. 28, 2024). In correspondence with the Office, the applicant further explained that the model "was developed specifically for th[e] project under Mr. Travis'[s] supervision using a curated set of vocal tracks from prior recordings by Mr. Travis" and that "the creative team [used the tool] to translate a sonically-tailored recording of James Dupré singing the composition 'Where That Came From' into a vocal track in Mr. Travis'[s] distinctive voice, while preserving the original cadence, phrasing, articulation, dynamics and other musical characteristics of Mr. Dupré's human performance." *Id.*

systems."[198] Another commenter similarly stated that if the U.S. does not adopt copyright protection for AI-generated outputs, "the global locus of cultural [intellectual property] generation will . . . shift to other nations with more AI-friendly policy environments."[199] This commenter further argued that excluding AI-generated works from copyright protection would not actually serve artists' interests,[200] as American artists instead "will be swept away by a public domain flood of [low-cost] foreign AI content" with which they cannot compete.[201]

Regardless of what other countries conclude, however, the United States is bound by our own Constitution and copyright principles. We should not abandon or distort those principles simply because other countries may not share them. Rather, we should make a persuasive case that a human-centered approach is good policy and inherent to copyright.

In any event, as described above, it remains to be seen how other jurisdictions' copyright laws will address generative AI. Commenters' concerns assume a substantial disparity in legal protection for AI-generated material, but no such disparity has yet clearly emerged. As a group of law professors acknowledged, while generative AI is likely to have widespread impact on human creativity, its effects on employment are difficult to predict.[202]

D. Providing Greater Clarity

Some commenters stressed the benefits of clarity and certainty. They posited that creators would be better off with certainty that their works produced using AI would be protected and available to license or sell. One commenter said that otherwise, the "commercial viability of the works made using AI tools is undermined [and]
. . . [t]he adoption of these tools will also be impacted."[203] Some cautioned that, absent greater clarity, authors may question whether they own what they create using AI, whether they can license their content to other parties, whether they can register their works with the Office, and

---

[198] The Knot Worldwide Inc. ("TKWW"), Reply Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 2 (Dec. 6, 2023) ("TKWW Reply Comments").

[199] Dallas Joder Reply Comments at 2.

[200] *See id.*

[201] *Id.* This commenter further cautioned that American AI startups will expend more financial resources on IP litigation than competitors in other countries that offer more expansive legal protection but did not explain how the volume of litigation would hinge on the copyrightability of AI-generated works. *See id.*

[202] *See* Pamela Samuelson et al. Initial Comments at 5.

[203] Microsoft and Github, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 10 (Oct. 30, 2023); *see also* IPO Initial Comments at 6–7 ("[I]f works created by humans using AI tools are not protected, that creates uncertainty for companies. Uncertainty leads to difficulty planning, developing, and investing, which could undermine the encouragement and promotion of arts and sciences."); ABA-IPL Initial Comments at 13–14; App Association Initial Comments at 6; ECPA Initial Comments at 7–8; Van Lindberg Initial Comments at 46; MPA Initial Comments at 59; TKWW Reply Comments at 2; SCA Robotics Initial Comments at 1.

whether their registration certificates will be entitled to a presumption of validity in an infringement action.[204]

A number of commenters urged the enactment of legislation to articulate the scope of protection through guidelines or standards.[205] One suggested establishing a legal presumption that an AI system's owner is the author of any output that the system may generate.[206] Another contended that the law should clarify that an "insignificant use of an AI tool that is otherwise substantially created by a human" does not make that work ineligible for copyright protection.[207]

The Office understands the desire for clarity around the copyrightability of AI-generated material. We do not believe, however, that legislation is necessary at this point. Much of the concern expressed focused on the assistive use of AI tools, and this Report seeks to provide assurances that such uses do not undermine protection. As to determining the copyrightability of AI outputs, the courts will provide further guidance on the human authorship requirement as it applies to specific uses of AI (including in reviewing the Office's registration decisions). Meanwhile, the analysis in this Part of the Report can help to shed light on how existing principles and policies apply.

Even if Congress were to consider addressing this issue through legislation, greater clarity would be difficult to achieve. Because the copyrightability inquiry requires analysis of each work and the context of its creation, statutory language would be limited in its ability to provide brighter lines. Unless and until future developments raise new problems, the Office does not recommend a change in the law.

---

[204] *See* Sandra Aistars, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 10–11 (Oct. 30, 2023); Graphic Artists Guild Initial Comments at 2–3; Qualcomm Reply Comments at 6.

[205] BLIP Initial Comments at 22; CISAC, Comments Submitted in Response to U.S. Copyright Office's Aug. 30, 2023, Notice of Inquiry at 5–6 (Oct. 30, 2023); ImageRights Reply Comments at 8–9; INTA Initial Comments at 4–5; Seth Polansky Initial Comments at 27–28.

[206] Ryan Abbott Initial Comments at 18.

[207] ASCAP Initial Comments at 49.

V.   CONCLUSION

Based on the fundamental principles of copyright, the current state of fast-evolving technology, and the information received in response to the NOI, the Copyright Office concludes that existing legal doctrines are adequate and appropriate to resolve questions of copyrightability. Copyright law has long adapted to new technology and can enable case-by-case determinations as to whether AI-generated outputs reflect sufficient human contribution to warrant copyright protection. As described above, in many circumstances these outputs will be copyrightable in whole or in part—where AI is used as a tool, and where a human has been able to determine the expressive elements they contain. Prompts alone, however, at this stage are unlikely to satisfy those requirements. The Office continues to monitor technological and legal developments to evaluate any need for a different approach.

The Office will provide ongoing assistance to the public on the copyrightability issues related to generative AI, including by issuing additional registration guidance and updating the relevant sections of the *Compendium of U.S. Copyright Office Practices*. In doing so, we will rely on the comments received in response to the NOI, judicial developments, and other relevant input.

ACKNOWLEDGEMENTS

Shira Perlmutter
Register of Copyrights and Director
U.S. Copyright Office
January 29, 2025