COMMENTS IN RESPONSE TO THE COPYRIGHT OFFICE'S NOTICE OF INQUIRY ON ARTIFICIAL INTELLIGENCE AND COPYRIGHT

By Pamela Samuelson, Christopher Jon Sprigman, and Matthew Sag

October 30, 2023

Suzanne V. Wilson General Counsel and Associate Register of Copyrights U.S. Copyright Office 101 Independence Ave, S.E. Washington, D.C. 20559-6000

Re: Notice of inquiry ("NOI") and request for comments, Artificial Intelligence and Copyright, Docket No. 2023-6

Dear Associate Register Wilson:

We are U.S. legal academics who write and teach in copyright law. Pamela Samuelson is the Richard M. Sherman Distinguished Professor of Law at Berkeley Law, as well as a Professor of School Information, and Co-Director, Berkeley Center for Law & Technology. Christopher Jon Sprigman is the Murray and Kathleen Bring Professor of Law at NYU School of Law. Matthew Sag is a Professor of Law in Artificial Intelligence, Machine Learning, and Data Science at Emory University Law School.¹

For clarity and concision, we will address below the following issues: (A) Copyrightability of Generative AI outputs, (B) Claims of infringement for use of works to train models, (C) Claims of infringement for Generative AI outputs (D) Transparency & Recordkeeping, (E)

¹ We offer these comments in response to the NOI in our personal capacities only and our views are not necessarily shared by our institutions or any other affiliated entities. Christopher Sprigman is a member of the law firm Lex Lumina PLLC, which represents clients in relation to copyright and AI issues; however, Sprigman has no involvement in that representation. Many of our comments are adapted from our prior and forthcoming articles addressing copyright and Generative AI, including Matthew Sag, *Copyright Safety for Generative AI*, 61 HOUS. L. REV. (forthcoming 2023) (available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4438593); Matthew Sag, *Fairness and Fair Use in Generative AI*, FORDHAM L. REV. (forthcoming 2024); Pamela Samuelson, *Generative AI Meets Copyright*, SCIENCE (July 14, 2023); and Pamela Samuelson, *Fair Use Defenses in Disruptive Technology Cases*, 71 UCLA L. REV. (forthcoming).

Labeling and (F) Non-copyright rights and interests. As a result, we will address some of the NOI questions out of order and by cross-reference.

(A) COPYRIGHTABILITY

NOI Question 18: Under copyright law, are there circumstances when a human using a generative AI system should be considered the "author" of material produced by the system?

When AI models produce content with little or no human intervention beyond simple prompts, there is no copyright in those outputs, because they do not qualify as original works of authorship.

The authorship that makes a work copyrightable is not dependent on the objective features of the work, but rather on the person who created it and the process by which the author created the work. The Copyright Act reserves copyright for "original works of authorship."² As the Supreme Court explained in the 1884 case of *Burrow-Giles Lithographic Co. v. Sarony*, authorship entails "original intellectual conception[]."³

Even when a Generative AI system produces texts, images, music, or other categorically copyrightable type of works that are indistinguishable from human-authored works, it makes no sense to think of a machine learning program as the author. The fact that Generative AI can now make works that are good enough to pass as human-created is impressive, but it is also beside the point. An AI system can't produce works that reflect its own "original intellectual conception" because an AI system is incapable of having one. AI systems do not "think" or "create" as we understand those terms in the context of human mental processes. Rather, AI systems employ math to make predictions. GPT-4, for example, is a text

² 17 U.S.C. § 102(a). Also note that "Authorship" is also a Constitutional requirement by virtue of the wording of the IP Clause, which gives Congress the power "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. Article I, Section 8, Clause 8 (emphasis added).

³ In *Burrow-Giles Lithographic Co. v. Sarony*, the Supreme Court held that photographs were protected by copyright because they were "representatives of original intellectual conceptions of the author," defining authors as "he to whom anything owes its origin; originator; maker; one who completes a work of science or literature." Burrow-Giles Lithographic Co. v. Sarony 111 U.S. 53, 57–59 (1884). See also Feist Publ'ns, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340, 345 (1991) (holding that originality requires both independent creation and sufficient creativity). In *Burrow-Giles*, Sarony was acknowledged as the author of a photo of Oscar Wilde because Sarony posed the subject in front of the camera, arranged the setting, and controlled the lighting. In other words, Sarony <u>made subjective choices</u> reflecting his own aesthetic judgment such that the resulting photo reflected his original intellectual conception.

prediction model that responds to prompts with statistically well-informed guesses about what the next word should be, and the word after that, and so on.⁴ The GPT-4 model has no internal mental state, no thoughts, and no feelings it is trying to express.⁵ Thus, the notion of AI being recognized as an author is a doctrinal non-starter.⁶

The idea that an AI could or should be recognized as the author of a work is also problematic for at least two additional reasons. First, AI systems are not legal entities. If AIs were authors, this would pose a slew of second-order questions about who owned the works they authored. There are no simple answers to these questions. Second, there is no rationale for treating AIs as authors based on the need for incentives or rewards. AIs do the work they are programmed to do, without regard to incentives.

However, humans using AI as a *tool* of expression may claim authorship if the final form of the work reflects their "original intellectual conception" in sufficient detail. We agree with the Copyright Office that simple text prompting is unlikely to meet this standard and that such prompts are more akin to an instruction to an assistant to create a work.⁷ However, there is no reason in principle why prompts couldn't be detailed enough to meet the traditional threshold of authorship in some cases. This will depend on the circumstances. Sophisticated prompts that specify details of an image should be sufficient to meet the requirement that the work that results from and reflects a person's original conception of the expression. A person who instructs a Generative AI with enough detail, such that model output reflects that person's original conception of the work, should be regarded as the author of the resulting work. It may be that current AI technologies do not give users sufficient control over outputs such that complex prompts and outputs are connected closely enough to qualify the output as the user's original intellectual conception. But it may

⁴ Technically, the next token, which in many cases is less than an individual word.

⁵ See Emily M. Bender, et al, On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? FACCT '21: PROCEEDINGS OF THE 2021 ACM CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY 610 (https://doi.org/10.1145/3442188.3445922).

⁶ See, for example, Pamela Samuelson, *Allocating Ownership Rights in Computer-Generated Works* 47:4 U. PITT. L. REV. 1185 (1986); James Grimmelmann, *There's No Such Thing as a Computer-Authored Work— And It's a Good Thing, Too* 39:3 COLUM JL & ARTS 403 (2016).

⁷ See Copyright Office correspondence Re: Zarya of the Dawn (Registration # VAu001480196), dated February 21, 2023. p.9. ("As the Supreme Court has explained, the 'author' of a copyrighted work is the one 'who has actually formed the picture,' the one who acts as 'the inventive or master mind.' *Burrow-Giles*, 111 U.S. at 61. A person who provides text prompts to Midjourney does not 'actually form' the generated images and is not the 'master mind' behind them.") See also Thaler v. Perlmutter, No. CV 22-1564 (BAH), 2023 WL 5333236, at *3 (D.D.C. Aug. 18, 2023) (holding that the Register of Copyrights "did not err in denying the copyright registration application presented by plaintiff [in relation to an autonomously generated digital artifact]. United States copyright law protects only works of human creation.")

be that as AI tools develop and allow for more fine-grained control of outputs by the user, that link will be evident in at least some outputs.

Furthermore, refining a series of text prompts and choosing among different outputs should also be recognized as a way in which a human using Generative AI could meet the authorship standard. This seems not to have been the case on the specific facts of the registration applications the Copyright Office has considered to date, but potentially it may in some future cases. Authorship often involves engaging with a physical medium in an iterative exploratory fashion, contemplating alternatives, embracing some and rejecting others. Consider, for example, a painter who flings paint at a canvas and then decides whether to fling more paint, or decides to start again on a fresh canvas. The painter has only a loose idea of what the work will look like as it takes shape, but when the work is finished, it is surely a work of authorship within the contemplation of the statute.⁸ Or, consider photographic authorship. The photographer's control over timing, lighting, and framing are often considered acts of authorship resulting in the copyrightability of photographs.

NOI Question 19. Are any revisions to the Copyright Act necessary to clarify the human authorship requirement or to provide additional standards to determine when content including AI-generated material is subject to copyright protection?

The law in this area will continue to develop through Copyright Office registration decisions and federal court litigation. We do not see any benefit in additional legislative guidance at this time.

NOI Question 20. Is legal protection for AI-generated material desirable as a policy matter? Is legal protection for AI-generated material necessary to encourage development of generative AI technologies and systems? Does existing copyright protection for computer code that operates a generative AI system provide sufficient incentives?

Our view is that existing copyright protection for computer code, as well as the application of existing protection to some AI outputs—those in which AI is used as a tool to assist in human creativity—will provide adequate incentives for creativity in this area. We see no need for special copyright or sui generis rules for AI.

⁸ Dan L. Burk explores many permutations of this hypothetical in Dan L. Burk, *Thirty-Six Views of Copyright Authorship, by Jackson Pollock,* 58 HOUS. L. REV. 263 (2020).

NOI Question 20.1. If you believe protection is desirable, should it be a form of copyright or a separate sui generis right? If the latter, in what respects should protection for AI-generated material differ from copyright?

Granting copyright protection to AI-generated works would be undesirable for reasons stated above. At this time, we perceive no need for sui generis protection for generative AI outputs.

NOI Question 21. Does the Copyright Clause in the U.S. Constitution permit copyright protection for AI-generated material? Would such protection "promote the progress of science and useful arts"? If so, how?

The Supreme Court in *Feist v. Rural Publications* ruled that the Constitution imposes a limit on eligibility for copyright protection to "original" works of authorship, whose author imbued it with some creative expression. This suggests that extending copyright protection to artificially generated outputs where there is no human intervention that introduces original expression into that AI-generated output would be of questionable constitutionality.

(B) INFRINGEMENT

Before addressing specific questions from the NOI, we offer several important observations about AI in general, and Generative AI in particular, that should inform the Copyright Office's deliberations and that frame our responses.

First, Generative AI poses a multitude of questions that are outside the boundaries of copyright law and that copyright law is ill-equipped to handle. Generative AI will enhance productivity for many types of professionals, including office workers and artists. It will allow even those without specific training or talents or with physical limitations to create art and music. Like other productivity-enhancing tools, the overall impact of Generative AI on employment is hard to predict. Generative AI will reshape perceptions of human-made works, as photography once did. It will also be employed for antisocial uses and, according to some, it may even be a precursor to systems with agency and superior intelligence that present unknown threats to human welfare. Policy initiatives to regulate AI to address harms beyond copyright are underway at the national and international levels.⁹

⁹ On October 30, 2023, President Biden issued "a landmark Executive Order" focused on "seizing the promise and managing the risks of artificial intelligence (AI)." Whitehouse, Briefing Room, Fact Sheet, *President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence*, Oct. 30 2023, (https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence). According to Whitehouse, "[t]he Executive Order establishes new standards for AI safety and security, protects Americans' privacy, advances equity and civil rights, stands up for consumers and workers,

Second, Generative AI is a general-purpose technology. One remarkable feature of generative AI is the ability to create new digital artifacts that are often indistinguishable from copyrightable human expression. However, the current and future uses of foundation models are expected to be broader than this. A chatbot trained on copyrighted works might be used to interface with other computer systems in ways that have nothing to do with creating pseudo-expression, for example, booking airline tickets.¹⁰

Third, the potential copyright infringement issues raised by Generative AI are not unique to LLMs or text-to-image models. U.S. courts have addressed the legality of non-expressive uses of copyrighted works in the context of other copy-reliant technologies, including, software reverse engineering,¹¹ plagiarism detection software,¹² and the digitization of millions of library books to enable meta-analysis, text data mining, and search engine indexing.¹³ *Authors Guild, Inc. v. HathiTrust* is a particularly significant case in this regard because the district court in that case directly addressed the issue of text data mining.¹⁴

As one of us explains in a forthcoming law review article:

promotes innovation and competition, advances American leadership around the world, and more." *Id.*

¹⁰ We note that AI and generative AI can mean very different things to different people. See Katherine Lee, A. Feder Cooper, and James Grimmelmann, *Talkin' Bout AI Generation: Copyright and the Generative-AI Supply Chain* (July 27, 2023). (Available at SSRN: https://ssrn.com/abstract=4523551) ("[G]enerative AI" is a catch-all name for a massive ecosystem of loosely related technologies, including conversational text chatbots like ChatGPT, image generators like Midjourney and DALL·E, coding assistants like GitHub Copilot, and systems that compose music, create videos, and suggest molecules for new medical drugs. Generative-AI models have different technical architectures and are trained on different kinds and sources of data using different algorithms. Some take months and cost millions of dollars to train; others can be spun up in a weekend. These models are also made accessible to users in very different ways. Some are offered through paid online services; others are distributed open-source, such that anyone could download and modify them.")

¹¹ Sega Enters. v. Accolade, Inc., 977 F.2d 1510, 1514 (9th Cir. 1992); Sony Computer Ent. v. Connectix Corp., 203 F.3d 596, 608 (9th Cir. 2000).

¹² A.V. ex rel. Vanderhye v. iParadigms, LLC, 562 F.3d 630, 644–45 (4th Cir. 2009).

¹³ See Authors Guild, Inc. v. HathiTrust, 755 F.3d 87, 100–01 (2d Cir. 2014); Authors Guild v. Google, Inc., 804 F.3d 202, 225 (2d Cir. 2015).

¹⁴ Authors Guild, Inc. v. HathiTrust, 902 F. Supp. 2d 445, 460, n22 (SDNY 2012) ("The use to which the works in the HDL are put is transformative because the copies serve an entirely different purpose than the original works: ... The search capabilities of the HDL have already given rise to new methods of academic inquiry such as text mining. ... Mass digitization allows new areas of non-expressive computational and statistical research, often called 'text mining.")

Text data mining is an umbrella term referring to computational processes for applying structure to unstructured electronic texts and employing statistical methods to discover new information and reveal patterns in the processed data. In other words, text data mining refers to any process using computers that creates metadata derived from something that was not initially conceived of as data. The process of text data mining can be used to produce statistics and facts about copyrightable works, but it can also be used to render copyrighted text, sounds, and images into uncopyrightable abstractions. These abstractions are not same, or even substantially similar to, the original expression, but in combination they are interesting and useful for generating insights about the original expression.¹⁵

Machine learning based on copyrighted works is an application of text data mining, not a separate technological phenomenon. The copyright issues raised by text data mining are by and large the same as those raised by machine learning and Generative AI. After all, it is hard to explain "why deriving metadata through technical acts of copying and analyzing that metadata through logistic regression should be fair use, but analyzing that data by training a machine learning classifier to perform a different kind of logistic regression that produces a predictive model wouldn't be."¹⁶ This is particularly significant given that the Copyright Office itself has recognized the fair use status of TDM research.¹⁷

Fourth, to the best of our information, the process of training Generative AI models is generally preceded by massive amounts of web scraping that results in the creation of locally stored copies of millions or billions of copyrighted works. And yet, just like text data mining and other non-expressive uses, that copying typically does not implicate the copyright owner's interest in controlling the communication of their original expression to the public because the copying is simply the first step in an analytical process that typically yields abstract metadata that is then used to create new digital artifacts that are not substantially similar to any of the particular works in the training data. Generative AI begins as a random set of weights assigned to parameters in very large statistical models. These weights are adjusted through exposure to the training data in successive rounds of training. The training data influences the model, but except in rare cases of overfitting, it is not "ingested" into the model. That is, the model does not embody the training data, and the training dataset is a separate object from the model and the software that produces outputs from inputs to the

¹⁵ Sag, Copyright Safety for Generative AI, supra note 1.

¹⁶ Id.

¹⁷ See U.S. COPYRIGHT OFF., SECTION 1201 RULEMAKING: EIGHTH TRIENNIAL PROCEEDING, RECOMMENDATION OF THE REGISTER OF COPYRIGHTS, 121–24 (2021), <u>https://cdn.loc.gov/copyright/1201/2021/2021_Section_1201_Registers_Recommendation.pdf</u>. (In evaluating the proposed DMCA § 1201 exemption to circumvent technological protection measures on DVDs and eBooks for the purpose of conducting TDM, the Copyright Office said: "Balancing the four fair use factors, with the limitations discussed, the Register concludes that the proposed use is likely to be a fair use.")

model. Put another way, Generative AI models are generally not designed to copy training data; they are designed to learn from the data at an abstract and uncopyrightable level.

The generative AI training process extracts information from millions or billions of works and, in the process, disassembles or tokenizes their elements to construct a very different representation in the models.¹⁸ In his testimony at a Congressional hearing, computer scientist Christopher Callison-Burch explained that the goal of training is to enable foundation models to discern, among other things, "the structure, syntax, and semantics of language," including "grammar, sentence construction, and how words and phrases are related to each other" in order to facilitate the generation of "coherent and contextually appropriate text."¹⁹ The training process exposes it to "a vast array of factual information, which they internalize and use to generate relevant responses or content," including "knowledge about geography, history, science, and various other domains." Models are trained to detect "different perspectives, opinions, and ideas expressed in their training data," which enables them to "generate text that reflects diverse viewpoints." During training, generative AI systems "gain some capacity for common sense reasoning, which allows them to understand basic cause-and-effect relationships, infer missing information, and make simple deductions." Training enables image generative systems to "recognize various objects, patterns, and features, such as common colors and shapes, and how it is typically used," so that AI systems can generate images that are "visually consistent and contextually appropriate." AI models discern "underlying patterns, relationships, and structures from the data, which allow them to generate entirely novel sentences, images, and other content."

Understanding the process of training foundation models is relevant to the generative AI systems' fair use defenses because the scope of copyright protection does not extend to "statistical information" such as "word frequencies, syntactic patterns, and thematic markers."²⁰ Processing in-copyright works to extract "information about the original [work]" does not infringe because it does not "replicat[e] protected expression."²¹ The unprotectability of facts embodied in works as well as high-level plots or themes is essential to achieving copyright's overall purpose of promoting progress in the creation and

¹⁸ See, e.g., Molly Shaffer Van Houweling, Freedom to Extract in Copyright Law, (working paper).

¹⁹ Written Testimony of Christopher Callison-Burch at 9, House Jud. Comm., Subcom. on Courts, Intell. Prop., and the Internet, Hearing on Artificial Intelligence and Intellectual Property: Part I – Interoperability of AI and Copyright Law, May 17, 2023, https://judiciary.house.gov/sites/evo-subsites/republicans-judiciary.house.gov/files/evo-media-document/callison-burch-testimony-sm.pdf [Callison-Burch Testimony].

²⁰ Authors Guild, 804 F.3d at 209.

²¹ Id. at 220.

dissemination of new knowledge so that "each author [can] build[] on the work of others."²² The extractive purpose of training favors generative AI fair use defenses.²³

Yet, in limited situations, Generative AI models do copy the training data.²⁴ So unlike prior copy-reliant technologies that courts have held are fair use, it is impossible to say categorically that inputs and outputs of Generative AI will always be fair use. We note in addition that some have argued that the ability of Generative AI to produce artifacts that could pass for human expression and the potential scale of such production may have implications not seen in previous non-expressive use cases. The difficulty with such arguments is that the harm asserted does not flow from the communication of protected expression to any human audience.

Finally, our answers to the questions in the NOI would be quite different if we believed that Generative AI was simply the next version of Napster. We believe that many of the sincere calls for regulation of Generative AI and for various licensing schemes are premised on misconceptions of how the technology works, the relationship between Generative AI models and the training data, and the extent that copyright law allows authors to control non-expressive use of their works.

NOI Question 1. Views on the potential benefits and risks of Generative AI? How is the use of this technology currently affecting or likely to affect creators, copyright owners, technology developers, researchers, and the public?

We address this question in the introduction to this section.

NOI Question 4. Comments on statutory or regulatory approaches adopted/under consideration in other countries and the importance of international consistency.

Copyright law is harmonized internationally to a significant degree through the TRIPs Agreement and the Berne Convention. Article 9(2) of TRIPs embodies a universal commitment to maintaining a distinction between copyrightable original expression and uncopyrightable "ideas, procedures, methods of operation or mathematical concepts as such." The doctrinal and policy argument in favor of allowing copying for non-expressive uses stems directly from the idea/expression distinction which has been conventionally

²² Nash v. CBS, Inc., 899 F.2d 1537, 1540 (7th Cir. 1990).

²³ See, e.g., Molly Shaffer Van Houweling, *Freedom to Extract in Copyright Law* (working paper) (explaining various copyright doctrines as well as constitutional policies that support freedom to extract data and other unprotectable elements in protected works, including when some expression must be taken to enable these extractions).

²⁴ See Sag, *Copyright Safety, supra* note 1 for a summary of the computer science literature explaining when such "memorization" is more likely. See also our answer to NOI Question 22, *infra*.

understood as part of the common law of copyright since at least the 1880 Supreme Court case of *Baker v. Selden*, 101 U.S. 99 (1880).

In the United States, the need to maintain the idea/expression distinction plays a significant role in the application of the fair use doctrine. Other jurisdictions also recognize the centrality of the idea/expression distinction, but have chosen a different route and have arrived at similar outcomes through a more prescriptive legislative process.²⁵ These different approaches reflect differences in legal culture and political economy. We believe that all countries should embrace the logical implication of the idea/expression distinction and provide some flexibility for non-expressive uses of copyrighted works, but the adoption of different mechanisms to realize this objective is not in itself cause for concern. We often see other jurisdictions modeling prescriptive copyright exceptions and limitations on the results of American fair use case law. We note that the Israeli Ministry of Justice has opined that the use of in-copyright works as training data is fair use and the EU has adopted text- and data mining exceptions that are pertinent to the Generative AI copyright debate.²⁶

We address additional details of overseas jurisdictions laws at NOI Question 5.

NOI Question 5. Is new legislation warranted to address copyright or related issues with generative AI? If so, what should it entail?

As noted earlier, the copying that precedes training machine learning models does not generally implicate the copyright owner's interest in controlling the communication of their original expression to the public. Current caselaw suggests that such copying will generally be

²⁵ The U.K. enacted a limited exception for TDM in 2014, see Section 29A of the Copyright, Designs and Patents Act 1988, but has announced plans to go further, see UKIPO, 'Press Release: Artificial Intelligence and Copyright and Patents' June IP (28)2022)(https://www.gov.uk/government/news/artificial-intelligence-and-ip-copyright-and-patents). Article 30(4) of the Japanese Copyright Act permits non-expressive use of copyrighted works so long as the use does not "unreasonably prejudice the interests of the copyright owner in light of the nature or purpose of the work or the circumstances of its exploitation..." See Japan, Copyright Act (Act No. 48 of May 6, 1970, as amended up to January 1, 2022), Article 30(4), available at (https://wipolex.wipo.int/en/legislation/details/21342). In April 2019, the European Union adopted the Digital Single Market Directive ("DSM Directive") featuring two mandatory exceptions for text and data mining. Article 3 of the DSM Directive requires all members of the European Union to implement a broad copyright exception for TDM in the not-for-profit research sector. Article 4 of the DSM Directive contains a second mandatory exemption that is more inclusive, but narrower in scope. See, Directive 2019/790, O.J. 2019 (L 130/92). See generally, Pamela Samuelson, Text and Data Mining of In-Copyright Works: Is It Legal? 64:11 COMMUNICATIONS OF THE ACM 20 (2021)

²⁶ Ministry of Justice, State of Israel, Opinion: Uses of Copyrighted Materials for Machine Learning (Dec. 2022); Directive 2019/790, of the European Parliament and of the Council of 17 April 2019 on Copyright and Related Rights in the Digital Single Market and Amending Directives 96/9/EC and 2001/29/EC, 2019 O.J. (L 130) 92, 113-14 (Articles 3 and 4).

fair use, so long as it falls within the parameters of non-expressive use.²⁷ As such there is no pressing need for new legislation with respect to text or data mining, or training machine learning models, whether they are Generative AI, or otherwise.

On the other hand, business and research communities may benefit from an express copyright limitation that allowed for text data mining, including as part of process of training machine learning models. An express limitation could be modeled on Articles 3 and 4 of the European Union's CDSM directive but we caution that the policy balance struck there may not be a good fit for the needs of the U.S. technology sector or U.S. copyright interests.

For proposals related to recordkeeping, see answers to NOI Questions 15-17 in part (C).

For proposals related to copyrightability, see answers to NOI Questions 18-21 in part (A).

NOI Question 8. Under what circumstances would the unauthorized use of copyrighted works to train AI models constitute fair use? Please discuss any case law you believe relevant to this question.

(a) Discussion of relevant case law

United States courts have consistently held that technical acts of copying which do not communicate an author's original expression to a new audience are fair use. The case law indicates that even though these "non-expressive uses" involved significant amounts of copying, they did not interfere with the interest in original expression that copyright is designed to protect.²⁸ Each use involved copying as an intermediate step towards producing something that either did not contain the original expression of the underlying work or contained a trivial amount. Non-expressive uses (although not labeled as such) have consistently held to be fair use.

• In 1992, in *Sega Enters., Ltd. v. Accolade, Inc.*, and again in 2000 in *Sony Computer Entm't, Inc. v. Connectix Corp.*, the Ninth Circuit held that reverse engineering computer programs—a process that involves making several copies of the code to extract vital but uncopyrightable elements needed to make interoperable programs—was fair use.²⁹ In *Sega v Accolade*, the court referred to copying to extract uncopyrightable

²⁷ See, *infra*, answer to NOI Q8 for elaboration.

²⁸ The terminology of "non-expressive use" originates with Matthew Sag, *Copyright and Copy-Reliant Technology*, 103 NW. U. L. Rev. 1607, 1610, 1682 (2009).

²⁹ Sega Enters., Ltd. v. Accolade, Inc., 977 F.2d 1510 (9th Cir. 1992); Sony Computer Entm't, Inc. v. Connectix Corp., 203 F.3d 596 (9th Cir. 2000). See also, Atari Games Corp. v. Nintendo of Am., Inc., 975 F.2d 832, 842–43 (Fed. Cir. 1992) (observing that Atari's reverse engineering of Nintendo's 10NES program would have been a fair use of the program, except that Atari did not possess an authorized copy of the work).

elements as "a legitimate, essentially non-exploitative purpose."³⁰ In *Sony Computer Entertainment v. Connectix, Inc.*, the court expressly recognized that "the fair use doctrine preserves public access to the ideas and functional elements embedded in copyrighted computer software programs."³¹

- In *A.V. ex rel. V anderbye v. iParadigms, LLC*, the Fourth Circuit held that copying student papers into a reference database for comparison against new student papers was fair use.³²
- In *Authors Guild, Inc. v. HathiTrust*, in 2014 the Second Circuit held that making digital versions of printed library books for research purposes that included text data mining and machine learning was fair use.³³
- A differently constituted panel of the Second Circuit reached much the same conclusion in 2015 in *Authors Guild, Inc. v. Google, Inc.*, (the *Google Books* case).³⁴ In *Google Books*, the court addressed both the complete copying of millions of library books to make them searchable, and the display of small snippets of the books in search result menus. The complete copying is an example of non-expressive use; the snippet displays illustrate the application of a more traditional transformative use analysis.³⁵

When courts have declined to find fair use in cases that are *superficially* similar to those discussed above, it is invariably because the challenged use was *not* non-expressive and thus, on the facts presented, the potential expressive substitution effect was too significant.

• In Associated Press v. Meltwater U.S. Holdings, Inc., 931 F. Supp. 2d 537 (S.D.N.Y. 2013), the Southern District Court of New York held that fair use did justify the actions of a

³⁰ Sega Enters., Ltd. v. Accolade, Inc., 977 F.2d 1510, 1523 (9th Cir. 1992).

³¹ Sony Computer Entm't, Inc. v. Connectix Corp., 203 F.3d 596, 603 (9th Cir. 2000).

³² A.V. ex rel. Vanderhye v. iParadigms, LLC, 562 F.3d 630 (4th Cir. 2009).

³³ Authors Guild, Inc. v. HathiTrust, 755 F.3d 87 (2d Cir. 2014). Note that the court's reasoning relied on the non-expressive nature of the use. The court explained "the creation of a full-text searchable database is a quintessentially transformative use [because] the result of a word search is different in purpose, character, expression, meaning, and message from the page (and the book) from which it is drawn. Indeed, we can discern little or no resemblance between the original text and the results of the [HathiTrust Digital Library] full-text search." *Id.* at 97-98.

³⁴ Authors Guild, Inc. v. Google, Inc., 804 F.3d 202 (2d Cir. 2015).

³⁵ The court held that the display of three-line snippets to add context to book search results was transformative in purpose and that it was reasonable in proportion to that purpose. Those snippets allowed a user to verify that a book suggested by the search engine was in fact relevant to her interests. In addition, the snippets were so brief that they did not pose any risk of fulfilling the reader's demand for the original expression of the underlying manuscripts. Authors Guild, Inc. v. Google, Inc., 804 F.3d 202 (2d Cir. 2015).

media monitoring company, Meltwater. Meltwater scraped news articles on the web to provide its subscribers with excerpts and analytics. However, the lawsuit did not challenge Meltwater's use of copyrighted news articles to provide metadata and analytics to its subscribers, even though these services also necessitated copying. The court noted that this was "an entirely separate service" and implied that if it had been challenged, it would have been found to be transformative, and thus fair use.³⁶ Instead of attacking Meltwater's non-expressive use, the Associated Press focused on the length and significance of Meltwater's extracts provided to subscribers. The court agreed that Meltwater's extracts were too long and too close to the heart of the work;³⁷ it also held that Meltwater had failed to show that the amount of the extracts was reasonable in light of its stated purpose to operate like a search engine.³⁸

• In *Fox News Network, LLC v. TVEyes, Inc.,* 883 F. 3d 169 (2nd Cir. 2018) the Second Circuit held that a media monitoring service that copied and electronically searched television broadcasts went beyond the scope of fair use when it allowed users to save, watch, and share ten-minute long video clips of the copyrighted programs. In the court's view, those ten-minute video clips would, "likely provide TVEyes's users with all of the Fox programming that they seek and the entirety of the message conveyed by Fox to authorized viewers of the original."³⁹ In other words, the court was concerned that rather than primarily providing information about the content of particular news segments, the length of the video clips was such that they would substitute for those segments in their entirety. The district court in *TVEyes* held that copying for search alone was fair use, and Fox did not contest this ruling on appeal.

³⁹ *Id.* at 179.

³⁶ *Id.* at 557. The court said: "The display of that analysis—whether it be a graphic display of geographic distribution of coverage or tone or any other variable included by Meltwater—is *an entirely separate service*, however, from the publishing of excerpts from copyrighted articles. The fact that Meltwater also offers a number of analysis tools does not render its copying and redistribution of article excerpts transformative." (emphasis added) *Id.*

³⁷ *Id.* at 558.

³⁸ Id.

In addition to these cases, one could also broaden the frame from non-expressive use to consider other "computational use" cases such as *Perfect 10, Inc. v. Amazon.com, Inc.*,⁴⁰ *Kelly v. Arriba Soft Corp.*,⁴¹ *Field v. Google, Inc.*⁴²

(b) Under what circumstances would the unauthorized use of copyrighted works to train AI models constitute fair use?

In summary: if the outputs of an AI model are not substantially similar to the protectable expression embodied in specific works in the training data, then, subject to potential fact-specific arguments that might be raised under the fourth factor, the copying involved in assembling and pre-processing the training data is fair use.⁴³

To elaborate in more detail, we consider each of the fair use factors below:

Factor one: If we assume that in its ordinary and routine operation, a generative AI model does not copy, or produce copies, of the original expression in its training data, then AI training qualifies as a non-expressive use. This kind of use surely has a "purpose and character" that is favored under the first fair use factor. Deriving uncopyrightable abstractions and associations from the training data and then using that knowledge to confect new digital artifacts is not just transformative, it is highly transformative.⁴⁴

⁴⁰ 508 F.3d 1146 (9th Cir. 2007) (fair use for search engine developer to make and display thumbnailsized images of photographs posted on the Internet for purposes of indexing content to enable users to find relevant content)

⁴¹ 336 F.3d 811 (9th Cir. 2003) (accord)

⁴² 412 F. Supp.2d 1106 (D. Nev. 2006) (fair use for Google to make webcrawl copies of texts that Field posted on the Internet and cache their contents for the purpose of indexing contents). See generally, Samuelson, *Fair Use Defenses in Disruptive Technology Cases*, *supra* note 1.

⁴³ We are assuming that substantial similarity only arises if the training data is inadvertently memorized, directly or indirectly, by the model. Similarity that resulted from pure coincidence, or copying from a common source would not take the use outside the bounds of non-expressive use.

⁴⁴ A.V. v. iParadigms Liab. Co., 544 F. Supp. 2d 473, 482 (E.D. Va. 2008): "This Court finds the "purpose and character" of iParadigms' use of Plaintiffs' written works to be *highly transformative*. Plaintiffs originally created and produced their works for the purpose of education and creative expression. iParadigms, through Turnitin, uses the papers for an entirely different purpose, namely, to prevent plagiarism and protect the students' written works from plagiarism. iParadigms achieves this by archiving the students' works as digital code and makes no use of any work's particular expressive or creative content beyond the limited use of comparison with other works." AV Ex Rel. Vanderhye v. iParadigms, LLC, 562 F. 3d 630, 640 (4th Cir, 2009): "The district court, in our view, correctly determined that the archiving of plaintiffs' papers was transformative and favored a finding of "fair use." *iParadigms' use of these works was completely unrelated to expressive content* and was instead aimed at detecting and discouraging plagiarism." Authors Guild, Inc. v. HathiTrust, 755 F. 3d 87, 97 (2nd Cir. 2014): "... we conclude that the creation of a full-text searchable database is a *quintessentially*

Moreover, the non-expressive use of copyrighted works by generative AI use does not usurp the copyright owner's interest in communicating her original expression to the public because that expression is not communicated. Given such non-expressive uses are highly transformative and that commerciality "is to be weighed against the degree to which the use has a further purpose or different character,"⁴⁵ the commerciality of the use has no independent significance under the first factor in this context.⁴⁶

Factor two: In most fair use cases it does not make sense to treat the second factor in isolation as something that weighs either for or against the defendant.⁴⁷ The second fair use factor, which directs courts to consider "the nature of the copyrighted work," has not loomed large in fair use cases involving other non-expressive uses cases, nor should we expect it to in the context of Generative AI.⁴⁸

transformative use."); Authors Guild, Inc. v. Google, Inc., 804 F.3d 202, 216-7 (2d Cir. 2015): "We have no difficulty concluding that Google's making of a digital copy of Plaintiffs' books for the purpose of enabling a search for identification of books containing a term of interest to the searcher involves a *highly transformative* purpose, in the sense intended by Campbell." Authors Guild, Inc. v. Google, Inc., 804 F.3d 202, 217 (2d Cir. 2015): "... through the ngrams tool, Google allows readers to learn the frequency of usage of selected words in the aggregate corpus of published books in different historical periods. *We have no doubt that the purpose of this copying is the sort of transformative purpose described in Campbell* as strongly favoring satisfaction of the first factor."

⁴⁵ Andy Warhol Foundation Visual Arts v. Goldsmith, 143 S. Ct. 1258, 1276 (2023).

⁴⁶ Writing for the majority in *Andy Warhol Foundation Visual Arts v. Goldsmith*, Justice Sotomayor explained that "the fact that a use is commercial as opposed to nonprofit is an additional element of the first factor" but that the commercial nature of the use, although relevant, "is not dispositive." See Andy Warhol Foundation Visual Arts v. Goldsmith, 143 S. Ct. 1258, 1276 (2023). The majority in *Andy Warhol Foundation* reiterated the holding in *Campbell*, that commerciality "is to be weighed against the degree to which the use has a further purpose or different character. The more transformative the new work, the less will be the significance of other factors, like commercialism, that may weigh against a finding of fair use." *Id.* at 1276 (quotations omitted, citing Campbell 510 U.S. at 579, and at 580, 585.)

⁴⁷ We note that works like computer software may have special characteristics that inform a fair use analysis. Google LLC v. Oracle America, Inc., 141 S. Ct. 1183, 1198 (2021) ("... fair use can play an important role in determining the lawful scope of a computer program copyright It can help to distinguish among technologies. It can distinguish between expressive and functional features of computer code where those features are mixed. It can focus on the legitimate need to provide incentives to produce copyrighted material while examining the extent to which yet further protection creates unrelated or illegitimate harms in other markets or to the development of other products.")

⁴⁸ See e.g., Authors Guild, Inc. v. HathiTrust, 755 F. 3d 87, 98 (2nd Cir. 2014) (Holding that the second fair-use factor "may be of limited usefulness where, as here, the creative work is being used for a transformative purpose" and that "[a]ccordingly, our fair-use analysis hinges on the other three factors.")(internal citations and quotations omitted).

Factor three: If a use is non-expressive, then the third statutory factor which considers "the amount and substantiality of the portion used" will also favor finding of fair use. The ultimate question under the third fair use factor is whether the amount of copying was reasonable in relation to a purpose favored by fair use.⁴⁹ Although non-expressive uses typically involve making complete literal copies, such copying is reasonable as an intermediate technical step in an analytical process that does not lead to the communication of the underlying original expression to a new audience. Accordingly, courts in non-expressive use cases have found the third factor weighs in favor of the defendant.⁵⁰

We address the fourth fair use factor, the effect on the "potential market for or value of the copyrighted work" in our answer to NOI Question 8.5(a).

NOI Question 8.1(a). In light of the Supreme Court's recent decisions in *Google v. Oracle America* and *Andy Warhol Foundation v. Goldsmith*, how should the "purpose and character" of the use of copyrighted works to train an AI model be evaluated?

The Supreme Court's recent decision in *Andy Warhol Foundation for Visual Arts v. Goldsmith* ("AWF") emphasizes that the question of "whether an allegedly infringing use has a further

⁴⁹ Campbell v. Acuff-Rose, 510 U.S. 569, 586-87 (1994). ("[T]he extent of permissible copying varies with the purpose and character of the use.") In *Campbell*, the Court characterized the relevant questions as whether "the amount and substantiality of the portion used ... are reasonable in relation to the purpose of the copying," and noting that the answer to that question depends on "the degree to which the [copying work] may serve as a market substitute for the original or potentially licensed derivatives[.]" *Id.* at 586-588.

⁵⁰ In Sega v. Accolade, a reverse engineering case decided before Campbell, the Ninth Circuit simply said that although the third weighed against the defendant, "where the ultimate (as opposed to direct) use is as limited as it was here, the factor is of very little weight." Sega Enters., Ltd. v. Accolade, Inc., 977 F.2d 1510, 1526-27 (9th Cir. 1992). The court in Sony v. Connectix took the same approach. Sony Computer Entm't, Inc. v. Connectix Corp., 203 F.3d 596, 606 (9th Cir. 2000). Subsequent nonexpressive use cases have applied Campbell and held that the third factor favors the defendant. AV Ex Rel. Vanderhye v. iParadigms, LLC, 562 F. 3d 630, 642 (4th Cir, 2009); Authors Guild, Inc. v. HathiTrust, 755 F. 3d 87, 98 (2nd Cir. 2014) "In order to enable the full-text search function, the Libraries, as we have seen, created digital copies of all the books in their collections. Because it was reasonably necessary for the HDL to make use of the entirety of the works in order to enable the fulltext search function, we do not believe the copying was excessive." Authors Guild, Inc. v. Google, Inc., 804 F.3d 202, 221 (2d Cir. 2015) "Complete unchanged copying has repeatedly been found justified as fair use when the copying was reasonably appropriate to achieve the copier's transformative purpose and was done in such a manner that it did not offer a competing substitute for the original." Authors Guild, Inc. v. Google, Inc., 804 F.3d 202, 221-222 (2d Cir. 2015) "As with HathiTrust, not only is the copying of the totality of the original reasonably appropriate to Google's transformative purpose, it is literally necessary to achieve that purpose. ... While Google makes an unauthorized digital copy of the entire book, it does not reveal that digital copy to the public. The copy is made to enable the search functions to reveal limited, important information about the books."

purpose or different character ... is a matter of degree, and the degree of difference must be weighed against other considerations, like commercialism."⁵¹ *AWF* reaffirms the importance of transformative use and implicitly calls into question lower court rulings that had found uses to be transformative where there was no significant difference in purpose.⁵² Simply adding a layer of new expression or a new aesthetic over-the-top of someone else's expressive work and communicating both the old and new expression to the public in a commercial context, without further justification, is not fair use. Yet, the Court cited approvingly to the *Google Books* decision, saying it is important "whether the purpose of the use is distinct from the original, for instance, because the use comments on, criticizes, or provides otherwise unavailable information about the original, see, e.g., *Authors Guild*, 804 F.3d at 215–216."⁵³

AWF reinforces the importance of focusing on the particular use made by the defendant and the prospect that the use might result in competitive substitution for the plaintiff's expressive work. AWF thus articulates a limiting principle that was only implicit in Judge Leval's original formulation of transformative use and in the Supreme Court's adoption of transformative use in *Campbell*. By tying the availability of fair use defenses to the likelihood of expressive substitution, AWF helpfully clarifies the reason why transformative use has featured so prominently in the case law: the more transformative a use is, the less likely it is to substitute for the copyright owner's original expression. Using the author's work to reflect back on the original is an intrinsically different purpose; that difference in purpose makes expressive substitution less likely. In contrast, merely adding an overlay of new expression while leaving the original expression intact provides no such comfort. The majority in AWF rightly focuses our attention on how the defendant's use is likely to substitute for the author's original expression and makes that the measure of when the defendant's use is sufficiently transformative.

This focus on expressive substitution makes it clear why non-expressive uses are strongly favored under the first fair use factor. By definition, non-expressive uses pose no threat of

⁵¹ Andy Warhol Foundation Visual Arts v. Goldsmith, 143 S. Ct. 1258, 1273 (2023).

⁵² *Id.* at 1275 "Most copying has some further purpose, in the sense that copying is socially useful ex post. Many secondary works add something new. That alone does not render such uses fair. Rather, the first factor (which is just one factor in a larger analysis) asks 'whether *and to what extent*' the use at issue has a purpose or character different from the original. Campbell, 510 U.S. at 579 (emphasis added). The larger the difference, the more likely the first factor weighs in favor of fair use. The smaller the difference, the less likely." See also *Id.* at 1273 emphasizing that non-critical transformative use must be "sufficiently distinct" from the original and that the overlay of a new aesthetic was not sufficient by itself.

⁵³ *Id.* at 1284.

direct expressive substitution (in the language of transformative use, they are not just transformative, they are highly transformative.)

- In the *iParadigms* plagiarism detection software case, the trial court found that "the 'purpose and character' of iParadigms' use of Plaintiffs' written works to be *highly transformative*." ⁵⁴ The court explained: "Plaintiffs originally created and produced their works for the purpose of education and creative expression. iParadigms, through Turnitin, uses the papers for an entirely different purpose, namely, to prevent plagiarism and protect the students' written works from plagiarism. iParadigms achieves this by archiving the students' works as digital code and makes no use of any work's particular expressive or creative content beyond the limited use of comparison with other works."⁵⁵
- On appeal the Circuit court said that the "district court, in our view, correctly determined that the archiving of plaintiffs' papers was transformative and favored a finding of fair use. *iParadigms' use of these works was completely unrelated to expressive content* and was instead aimed at detecting and discouraging plagiarism."⁵⁶
- In *HathiTrust* the Second Circuit concluded "that the creation of a full-text searchable database is a *quintessentially transformative* use."⁵⁷
- In *Google Books* the Second Circuit explained: "We have no difficulty concluding that Google's making of a digital copy of Plaintiffs' books for the purpose of enabling a search for identification of books containing a term of interest to the searcher involves a *highly transformative* purpose, in the sense intended by *Campbell*."⁵⁸
- The same court also observed "... through the ngrams tool, Google allows readers to learn the frequency of usage of selected words in the aggregate corpus of published books in different historical periods. *We have no doubt that the purpose of this copying is the sort of transformative purpose described in Campbell* as strongly favoring satisfaction of the first factor."⁵⁹

⁵⁴ A.V. v. iParadigms Liab. Co., 544 F. Supp. 2d 473, 482 (E.D. Va. 2008).

⁵⁵ Id.

⁵⁶ AV Ex Rel. Vanderhye v. iParadigms, LLC, 562 F. 3d 630, 640 (4th Cir, 2009) (emphasis added).

⁵⁷ Authors Guild, Inc. v. HathiTrust, 755 F. 3d 87, 97 (2nd Cir. 2014) (emphasis added).

⁵⁸ Authors Guild, Inc. v. Google, Inc., 804 F.3d 202, 216-7 (2d Cir. 2015) (emphasis added).

⁵⁹ *Id.* at 217 (2d Cir. 2015) (emphasis added).

NOI Question 8.1(b) What is the relevant use to be analyzed? Do different stages of training, such as pre-training and fine-tuning, raise different considerations under the first fair use factor?

Different phases of the process of creating and using AI models have different copyright implications. Not so much because they raise different issues under any of the statutory fair use factors, but because many of them involve no reproduction of the training data in copies, creation of derivative works based on works in the training data, etc.

Adapting Lee, Cooper, and Grimmelmann's description of what they call "the generative-AI supply chain,"⁶⁰ and combining it with Sag's model of the lifecycle of text data mining research,⁶¹ it is useful to think of the lifecycle of generative AI proceeding as follows:

(1) **Creation** of works such as books, artwork, software, and other products of human creativity.

(2) **Conversion** of works and other information into digitally encoded files in standard formats. This is the beginning of the process of turning text into data.

(3) **Compilation** of individual items of data into "vast and carefully structured" training datasets, a process that "requires both extensive automation and thoughtful human decision-making."⁶² This step may involve reproduction, but it may also simply involve curating collections of internet links. The difference is significant because the latter is not by itself an action that reads on any of the exclusive rights in Section 106 of the Copyright Act.

(4) **Assembly** of training datasets (i.e., an act of reproduction, usually undertaken by the entity training the model). There may be techniques that allow this step to be skipped, but in practice, we understand that is considered broadly impractical to proceed without creating a semi-permanent local copy of the training data.⁶³

⁶⁰ See Katherine Lee, A. Feder Cooper, and James Grimmelmann, *Talkin' Bout AI Generation: Copyright and the Generative-AI Supply Chain* (July 27, 2023). (Available at SSRN: https://ssrn.com/abstract=4523551).

⁶¹ Matthew Sag, *The New Legal Landscape for Text Mining and Machine Learning*, 66 J. COPYRIGHT SOC'Y U.S.A. 291, 345-365 (2019) (https://ssrn.com/abstract=3331606).

⁶² Lee et. al. *supra* note 60.

⁶³ To avoid overfitting (and thus hopefully minimize the risk of copyright infringement and other analogous harms), it is important to deduplicate the training data. To address questions of bias and filter out toxic materials, the potential training data needs to be analyzed carefully before training begins. These actions may be undertaken at (3) compilation, but otherwise they are likely to be done at (4) assembly. We note that storing a semi-permanent local copy also makes sense if the developer

(5) **Model training to produce a "pre-trained" model** (i.e., a base model). This includes the choice of technical architecture and actually running a training algorithm to encode features of the training data in the model.

The process of training an AI model does not inherently involve further acts of reproduction in machine learning models that we are familiar with. In the context of LLMs, as Sag explains,⁶⁴ LLMs are trained to predict the next token in a sequence of tokens (where a token is a word or part of a word). At the beginning of training, the weights attached to each one of the billions of parameters in the model are assigned randomly.⁶⁵ The model is only exposed to brief snippets of text in fleeting fashion. The first time the model encounters a phrase like "the girl with the dark [blank]" it would be just as likely to fill in the blank with a word like "watermelon," "galaxy," "harmonica," "propeller" or a random punctuation mark.⁶⁶ However, over the course of training, the system updates the weights in the model, reinforcing the weights that improve the guess and downgrading those that don't. Those weights don't reflect any single source, and they are not the result of any single round of training.

(6) **Fine-tuning** to improve the performance of the base model or adapt it to a specific problem domain. As Lee et. al. note, "[t]his process, too, involves extensive choices — and it need not be carried out by the same entity that did the initial training."⁶⁷

(7) **Deployment**: i.e., embedding the model in, or connecting it to, some larger system such as a web-interface, a Discord server, or a smartphone app.

(8) **Generation**: i.e., the point at which the model is actually prompted with inputs and produces some new digital artifact.

anticipates the need to retrain the model from time to time. Continued access to the training data in its original form may also be necessary to evaluate the performance of the model, and to take additional steps to mitigate the potential for copyright infringement, or other undesirable outcomes.

⁶⁴ Sag, Copyright Safety, supra note 1.

⁶⁵ Not entirely randomly, but randomly drawn from specific distributions (like a normal or uniform distribution). The random seeding is important because it helps the model to explore a wide range of possible solutions and to avoid getting stuck in one area of the solution space.

⁶⁶ Actually, the model is predicting the next token, but the differences this would make to the example are minor.

⁶⁷ Lee, et. al. *supra* note 60.

(9) **Alignment and recalibration**: processes such as reinforcement learning through human feedback evaluate the digital artifacts generated by the model and evaluate them according to criteria to improve subjective performance.

The processes we have described may interact with each other and may not always be cleanly separated. For example, alignment and recalibration may lead to additional fine-tuning; generated artifacts may form the basis of additional training runs; and a sufficiently complex series of prompts may begin to resemble fine-tuning.

NOI Question 8.2. How should the analysis apply to entities that collect and distribute copyrighted material for training but may not themselves engage in the training?

The steps in the generative-AI supply chain set out above may be entirely vertically integrated, or they may be undertaken by a number of different actors.

The entities responsible for collections of training data would face no direct copyright liability to the extent that those resources are merely collections of links. However, they may be liable for downstream infringements (if any) under doctrines of contributory, vicarious, and inducement-based liability, assuming the relevant tests are satisfied. Any such determination would be fact-specific and potentially subject to the DMCA safe harbor for information location tools. We address this scenario for the sake of completeness. Our impression is that the most widely used sets of training data such as the Common Crawl and LAION-5B are actual archives and not merely collections of links.⁶⁸ The Common Crawl is a collection of copied webpages "stored on Amazon Web Services' Public Data Sets and on multiple academic cloud platforms across the world."⁶⁹ LAION-5B, is a dataset constructed from images and alt-text from the Common Crawl corpus.⁷⁰ In what we believe to be the more widely applicable scenario where training data is collected through acts of reproduction, the relevant entity may be directly liable for infringement.

In either case (direct or indirect liability) an entity that collected and distributed (or made available links for) copyrighted material for training but did not itself engage in the training would be entitled to claim fair use if appropriate steps were taken to ensure that the works were only used in processes that qualified as fair use. In *Google Books*, for example, Google provided scanned copies of the relevant library books to the participating libraries, subject to

⁶⁸ Common Crawl is a 501(c)(3) non-profit founded in 2007 with the mission of making "wholesale extraction, transformation and analysis of open web data accessible to researchers." See Common Crawl, Overview, (<u>https://commoncrawl.org/overview</u>).

⁶⁹ Id.

⁷⁰ See generally Romain Beaumont, *LAION-5B: A New Era of Large-Scale Multi-Modal Datasets, LAION* (Mar. 31, 2022), https://laion.ai/blog/laion-5b/.

contractual terms committing the libraries to use their digital copies "only in a manner consistent with the copyright law, and to take precautions to prevent dissemination of their digital copies to the public at large."⁷¹ The court held that the arrangements under which "each participant library has contracted with Google that Google will create for it a digital copy of each book the library submits to Google, so as to permit the library to use its digital copy in a non-infringing fair use manner" did not defeat Google's fair use claim.⁷²

NOI Question 8.3. The use of copyrighted materials in a training dataset or to train generative AI models may be done for noncommercial or research purposes. How should the fair use analysis apply if AI models or datasets are later adapted for use of a commercial nature? Does it make a difference if funding for these noncommercial or research uses is provided by for-profit developers of AI systems?

As discussed above, so long as training materials are used non-expressively, the use is highly transformative and thus the commercial nature of the user has less significance under the first factor. The eventual commercial nature of the model trainer, the fine-tuner, or the end-user may be relevant to fact specific arguments under the fourth factor.

NOI Question 8.4. What quantity of training materials do developers of generative AI models use for training? Does the volume of material used to train an AI model affect the fair use analysis? If so, how?

The volume of material used has no obvious bearing on the fair use analysis. The issue is whether the use is non-expressive. It is possible that the volume of material used may have relevance under particular fact-specific arguments under the fourth factor.

NOI Question 8.5(a). Under the fourth factor of the fair use analysis, how should the effect on the potential market for or value of a copyrighted work used to train an AI model be measured?

In most cases, if a use is non-expressive, the fourth statutory factor which considers the effect on the "potential market for or value of the copyrighted work" will favor a finding of fair use. The training dataset is generally not a commercial product; insofar as the use is non-expressive and is not marketed, it is questionable that it would have any effect on the market for or value of the works in the training dataset.

We recognize that there are several fact specific arguments that can be made under the fourth factor that may lead to a contrary result. However, in general terms, if a use is non-expressive, then it poses no direct threat of expressive substitution and should generally be considered harmless under the fourth factor. There may be a market effect in the broader

⁷¹ Authors Guild v. Google, Inc., 804 F. 3d 202, 229 (2nd Cir. 2015).

⁷² *Id.* at 228-29.

economic sense, but the "market" and "value" referred to in the fourth fair use factor are not simply any benefit the copyright owner might choose to nominate. A critical book review that quotes from a novel does not have an adverse market effect if it persuades people to buy different book instead;⁷³ a report from a plagiarism detection service might depress the market for helping students cheat on their homework, but that is hardly a cognizable injury under copyright law.⁷⁴ More broadly, copyright owners have no protectable interest in preventing criticism, parody,⁷⁵ or simply locking up unprotectable ideas and expression.⁷⁶ Nor can they simply claim, in circular fashion, that the right to charge for nonexpressive uses is a cognizable harm and that to avoid that harm they must be given the right to charge for non-expressive uses.⁷⁷

⁷³ Campbell v. Acuff-Rose, 510 U.S. 569, 591-592 (1994) "We do not, of course, suggest that a parody may not harm the market at all, but when a lethal parody, like a scathing theater review, kills demand for the original, it does not produce a harm cognizable under the Copyright Act. Because parody may quite legitimately aim at garroting the original, destroying it commercially as well as artistically, the role of the courts is to distinguish between biting criticism that merely suppresses demand and copyright infringement, which usurps it." (quotations and internal citations omitted).

⁷⁴ AV ex rel. Vanderhye v. iParadigms, LLC, 562 F.3d 630, 464 (4th Cir. 2009) ("Clearly no market substitute was created by iParadigms, whose archived student works do not supplant the plaintiffs' works in the 'paper mill' market so much as merely suppress demand for them, by keeping record of the fact that such works had been previously submitted In our view, then, any harm here is not of the kind protected against by copyright law.")

⁷⁵ Campbell v. Acuff-Rose, 510 U.S. 569, 577-79 (1994); NXIVM Corp. v. Ross Inst., 364 F.3d 471, 482 (2d Cir. 2004) ("[C]riticisms of a seminar or organization cannot substitute for the seminar or organization itself or hijack its market."); Bill Graham Archives v. Dorling Kindersley, Ltd., 448 F.3d 605 (2d Cir. 2006) ("A copyright holder cannot prevent others from entering fair use markets merely by developing or licensing a market for parody ... or other uses of its own creative work.") (internal quotations omitted).

⁷⁶ Sega Enters., Ltd. v. Accolade, Inc., 977 F.2d 1510 (9th Cir. 1992); Sony Computer Entm't, Inc. v. Connectix Corp., 203 F.3d 596 (9th Cir. 2000).

⁷⁷ Campbell v. Acuff-Rose Music, 510 U.S. 569, 591-92 (1994) (no cognizable market effect where parody or criticism depress demand for the original work); see also Sony Computer Entm't, Inc. v. Connectix Corp., 203 F.3d 596, 607 (9th Cir. 2000) (noting that a videogame manufacturer's desire to foreclose competition in complementary products was understandable, but that "copyright law ... does not confer such a monopoly."); Bill Graham Archives v. Dorling Kindersley, Ltd., 448 F.3d 605, 615 (2d Cir. 2006) ("[A] copyright holder cannot prevent others from entering fair use markets merely by developing or licensing a market for parody, news reporting, educational or other transformative uses of its own creative work.") (citations and quotations omitted). See also Authors Guild, Inc. v. HathiTrust, 755 F.3d 87, 100 (2d Cir. 2014) ("Lost licensing revenue counts under Factor Four only when the use serves as a substitute for the original and the full-text-search use does not.")(emphasis added); Authors Guild v. Google, Inc., 804 F.3d 202, 223 (2d Cir. 2015) (Framing the question as "whether the copy brings to the marketplace a competing substitute for the original, or its derivative,

Nonetheless, there is still scope for considerations of fairness in fair use that go beyond direct expressive substitution.⁷⁸ A defendant's failure to adopt adequate security measures, circumvention of paywalls, or disregard of robots.txt exclusions and similar mechanisms could each be framed in terms of an argument against fair use under the fourth factor. No doubt, other fact specific arguments will be raised by plaintiffs and addressed by courts in the coming years.

One issue under the fourth factor that has received significant attention relates to the practice of training LLMs on sites of known infringement or so-called shadow libraries like Library Genesis and Sci-Hub.⁷⁹ This issue awaits resolution by the courts.⁸⁰ One might argue that although copyright owners do not have a right to charge for fair uses as such, they do have a right to charge for access to their works. As such, it may be deemed harmful or unfair for commercial users to bypass the market for access to train their LLMs without a compelling reason. Such conduct arguably undermines the economic incentives that copyright is designed to create. Context matters, however. It would be unwise to elevate lawful access to a per se rule, even for commercial defendants. Moreover, prohibiting academic research on illegal text corpuses will generally not benefit copyright owners or further the interests copyright is designed to promote.

so as to deprive the rights holder of significant revenues because of the likelihood that potential purchasers may opt to acquire the copy in preference to the original.")(emphasis added).

⁷⁸ In *Sony v. Universal* the Supreme Court majority looked to considerations beyond expressive substitution and held that non-commercial time-shifting broadcast television by VCR-users was a fair use because the technology merely allowed users to do something they were already authorized to do, but with more convenience. Sony Corp. of America v. Universal City Studios, Inc., 464 US 417, 449 (1984) ("time-shifting merely enables a viewer to see such a work which he had been invited to witness in its entirety free of charge.") The majority may have also been influenced by the prospect that potential market failures may have resulted in a significant public benefit being otherwise be foregone. Note also that in *HathiTrust*, the Second Circuit held that providing print-disabled patrons with full digital access to books was not transformative, but that it was still fair use because the ordinary publishing market failed to provide adequately for the print-disabled. Authors Guild, Inc. v. HathiTrust, 755 F. 3d 87, 101-02 (2nd Cir. 2014).

⁷⁹ As argued in pending cases including, Tremblay et al v. OpenAI, Inc. et al, Docket No. 4:23-cv-03223 (Doc. 1 at 7)(N.D. Cal. Jun 28, 2023). For an overview of the earliest filed cases, see, Pamela Samuelson, *Generative AI Meets Copyright*, SCIENCE (July 14, 2023) (describing lawsuits).

⁸⁰ The Federal Circuit stated in *Atari v. Nintendo* that "[t]o invoke the fair use exception, an individual must possess an authorized copy of a literary work." Atari Games Corp. v. Nintendo of Am. Inc., 975 F.2d 832, 834 (Fed. Cir. 1992). However, this predated the Supreme Court's comment in *Campbell* that "If the use is otherwise fair, then no permission need be sought or granted. Thus, being denied permission to use a work does not weigh against a finding of fair use." Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569, 585 n.18 (1994).

On the other hand, as Michael Carroll argues, there are strong arguments to be made that copying from an infringing source may still be fair use.⁸¹ Carroll argues that "[t]reating an otherwise fair use as unfair because it was made from an infringing source would lead a court to deny the public access to the products of secondary uses that fair use is designed to encourage."⁸² He notes that significant doubt exists as to whether good faith is a consideration in fair use at all.⁸³ Judge Pierre Leval has also persuasively argued that using a good faith inquiry in fair use analysis "produces anomalies that conflict with the goals of copyright and adds to the confusion surrounding the doctrine."⁸⁴ Moreover, even if good faith is part of the broader fair use calculus, courts have found that knowing use of an infringing source is not bad faith when the user acts in the reasonable belief that their use is a fair use.⁸⁵ There is no recognized "fruit of the poisonous tree" doctrine in copyright law.⁸⁶

NOI Question 8.5(b) Should the inquiry be whether the outputs of the AI system incorporating the model compete with a particular copyrighted work, the body of works of the same author, or the market for that general class of works?

As noted above, if a use is non-expressive, then it poses no direct threat of expressive substitution and should generally be considered harmless under the fourth factor. Copyright protects original expression, it does not traditionally protect the copyright owner from the

⁸⁵ See, e.g., NXIVM Corp. v. Ross Inst., 364 F.3d 471, 478-79, 482 (2d Cir. 2004).

⁸¹ Michael Carroll, Copyright and the Progress of Science, 53 U.C. DAVIS L. REV. 893, 951-59 (2019).

⁸² *Id.* at 955.

⁸³ For an overview, see Simon J. Frankel & Matt Kellogg, *Bad Faith and Fair Use*, 60 J. COPYRIGHT SOC'Y U.S. 1 (2012).

⁸⁴ Pierre N. Leval, *Toward a Fair Use Standard*, 103 HARV. L. REV. 1105, 1126 (1990); see also Pierre N. Leval, *Campbell as Fair Use Blueprint?*, 90 WASH. L. REV. 597, 612-13 (2015) ("The public's access to important knowledge should not be barred because of bad behavior by the purveyor of the knowledge. A copier's bad faith has no logical bearing on the scope of the original author's copyright.").

⁸⁶ See, e.g., Kepner-Tregoe, Inc. v. Leadership Software, Inc., 12 F.3d 527, 538 (5th Cir. 1994) ("[u]nder copyright law, the district court could enjoin only those future versions of [defendant's program] that are substantially similar to [plaintiff's] Licensed Materials"); Liu v. Price Waterhouse LLP, No. 97 CV 3093, 2000 WL 1644585, at *2 (N.D. Ill. Oct. 30, 2000) (rejecting proposed jury instruction because it was based on a "fruit of the poisonous tree" theory that would allow recovery for the sale of defendant's future works even if they were not substantially similar to plaintiff's original, saying"[s]uch relief is not provided in the [Copyright] Act and would constitute an end-run around the Act's mandate that copyright owners may recover profits only after proving that the work in question is an infringement.... That defendants may have viewed or studied plaintiff's program is irrelevant if defendants' resulting work is not substantially similar to plaintiff's."); Real View, LLC v. 20-20 Techs., Inc., 811 F. Supp. 2d 553, 561 (D. Mass. 2011) (remittitur disallowing award of profits on noninfringing products despite illegal download). See also Mark A. Lemley, *The Fruit of the Poisonous Tree in IP Law*, 103 IOWA L. REV. 245, 248 (2017).

threat of competition with works that do not embody the copyright owner's original expression. The rights of the author are grounded in specific works, not bodies of works (with the exception of anthologies, but not is not what we understand the question to mean.)

NOI Question 9. Should copyright owners have to affirmatively consent (opt in) to the use of their works for training materials, or should they be provided with the means to object (opt out)?

See answer to NOI Question 9.1

NOI Question 9.1. Should consent of the copyright owner be required for all uses of copyrighted works to train AI models or only commercial uses?

Express copyright owner consent is not required if the use is a fair use. As explained above, the limited copying that takes place in most AI applications is likely to be a non-expressive use, and thus highly likely to amount to fair use. Although there is limited case law on point,⁸⁷ it is arguable that respect for opt-outs should be part of the fair use analysis as a general consideration. We leave it to others with greater specific knowledge to comment on how the robots.txt mechanism that has worked so well for Internet search can be adapted for generative AI.

NOI Question 12. Is it possible or feasible to identify the degree to which a particular work contributes to a particular output from a generative AI system? Please explain.

The state of the art may change, but for the moment it does not appear to be feasible to identify the degree to which a particular work contributes to a particular output from a generative AI system due to the size and complexity of the models. When a model is trained on a specific work, the model weights are impacted in subtle, complex, and usually non-linear ways. Like neurons in the brain, individual nodes in an AI model serve multiple functions. We are not aware of any attribution technique that has been shown to work with complex Generative AI models.

It is possible to compare the output of a Generative AI model to the training data and infer something about attribution, but any such inference would be simply an educated guess. This kind of attribution by inference of similarity could be useful for linking outputs to inputs in some contexts. It could also be employed prophylactically to reduce the chances of accidental generation of infringing outputs.

⁸⁷ Field v. Google Inc., 412 F. Supp. 2d 1106 (D. Nev. 2006).

NOI Question 10. If copyright owners' consent is required to train generative AI models, how can or should licenses be obtained?

To the extent that copyright owners' consent is required to train generative AI models, such licensing can occur through open licenses such as Creative Commons, or by direct voluntary agreements with rights holders or rights aggregators. Media reports indicate several examples of companies like Reuters and Shutterstock entering into licensing deals with AI developers, but the feasibility of such direct licensing depends on the nature of the works and the concentration of rights in the relevant market. In many instances, transaction costs are likely to be high.

It is important to note that legally-mandated collective licensing faces a difficulty in the AI context that differentiates collective licensing for AI from the use of the mechanism in any other field: it will be impossible under current technologies to calibrate payments made under a collective licensing arrangement to actual usage of individual authors' works.⁸⁸ ASCAP is able to efficiently distribute money because it makes no payments to significant numbers of rights holders whose works are infrequently performed. In contrast, without information about the relative importance of individual works, the total sum available under a legally-mandated collective licensing scheme would have to be divided equally between all potential claimants. Indeed, we suspect that a collective licensing system for AI is unlikely to result in significant payments to individual authors because the transaction costs involved in making individual distributions are likely to exceed the amount available for distribution. The more likely way a legally-mandated collective licensing scheme would work would be, effectively, as a tax.⁸⁹

NOI Question 13. What would be the economic impacts of a licensing requirement on the development and adoption of generative AI systems?

It is difficult to predict the economic impact of a licensing requirement on the development and adoption of generative AI systems without knowing more about the nature and scope of such a requirement, including whether it would be applied retrospectively to activities that may have been considered fair use in the past. Any prediction about the economic impact of a licensing requirement would also involve making an implicit prediction about future technological developments.

⁸⁸ See, e.g., Pamela Samuelson, *Fair Use Defenses in Disruptive Technology Cases*, 71 UCLA L. REV., Part III-B (forthcoming) (questioning the feasibility of administering such a regime). See also our answer to NOI Question 12.

⁸⁹ See, e.g., Martin Senftleben, *Generative AI and Author Remuneration* (July 20, 2023) (suggesting that authors are unlikely to be compensated for training data collective license, but might if remuneration levy focused on output services).

The imposition of a licensing requirement may lead to industry consolidation in that it would benefit incumbents who can use their models to generate synthetic data and train future models on a combination of synthetic data, public domain data, open-access data, and licensed data. It may lead to the development of generative AI moving to overseas jurisdictions with more favorable laws. Or, it may have only limited impacts given the possibility of training on synthetic data, public domain materials, open-licensed materials, and repositories controlled by content aggregators.⁹⁰

(C) INFRINGEMENT - OUTPUTS

NOI Question 22. Can AI-generated outputs implicate the exclusive rights of preexisting copyrighted works, such as the right of reproduction or the derivative work right? If so, in what circumstances?

(a) Reproduction

AI-generated outputs implicate the exclusive rights of preexisting copyrighted works when those outputs are substantially similar to specific copyrighted works in the training data. The test is whether the ordinary observer would regard a given output as substantially similar to the protectible expression of a given input.

Infringing outputs are rare, but when they occur the computer science literature suggests that it is because: (i) models were trained on many duplicates of the same work; (ii) images were paired with unique text descriptions in text-to-image models; and (iii) the ratio of the size of the model to the training data was relatively large.⁹¹ In addition, there is the "Snoopy Problem:" the more abstractly a copyrighted work is protected, the more likely the model is to "copy" it.⁹² Thus infringement is more likely with works that are protected at a more abstract level, such as copyrightable characters, or works which customarily entail very broad derivative rights, such as novels. Sag has shown that text-to-image models are prone to produce potentially infringing works when the same text descriptions are paired with relatively simple images that vary only slightly and thus they are especially likely to generate

⁹⁰ Although some may have thought training a generative AI model on open-access data alone is infeasible, a recent paper suggests otherwise. See Aaron Gokaslan, et. al., CommonCanvas: An Open Diffusion Model Trained with Creative-Commons Images, ARXIV, October 25, 2023 (https://arxiv.org/abs/2310.16825) ("Using entirely Creative-Commons images and our synthetic captioning approach, we achieve comparable qualitative performance to Stable Diffusion 2 (SD2base), as seen in CommonCanvas generations, while only requiring a small fraction (< 3%) of the amount of training data.")

⁹¹ See Sag, *Copyright Safety, supra* note 1.

⁹² Id.

images that would infringe on copyrightable characters.⁹³ Characters like Snoopy are protected at a somewhat abstract level and they appear often enough in the training data that the model learns the consistent traits and attributes that are associated with those names. Even though Generative AI models like Stable Diffusion and Midjourney are unlikely to reproduce a specific copyrighted work featuring Snoopy, both models have clearly learned how to draw Snoopy to a degree that could easily be seen as infringing.

(b) Derivative works

The argument that AI-generated outputs implicate the exclusive right to prepare "derivative works based on the copyrighted work" in the absence of substantial similarity is simply mistaken.⁹⁴ We acknowledge that Generative AI models and their outputs would not exist without the training data, but causation is not the test for derivative works under U.S. law.

As Pamela Samuelson summarizes in a forthcoming paper:95

... in the absence of substantial similarity in expression between this input and that output,⁹⁶ the derivative work claims in the generative AI cases are at best novel and given precedents, they seem quite weak. The Andersen complaint, for example, concedes that "[i]n general, none of the Stable Diffusion output images provided in response to a particular Text Prompt is likely to be a close match for any specific image in the training data."⁹⁷ Insofar as this is true, derivative work claims cannot be sustained. While it is true that outputs are, in some sense, "based upon" the works on which foundation models were trained, this has never sufficed to support derivative work claims; there must be substantial similarity in expressions to infringe that right.⁹⁸ Nor is it sufficient to claim the AI outputs are the result of

⁹³ Id.

⁹⁴ One of the disturbing implications of the theory that output of generative AI systems must be considered an infringing derivative of the training data is that it would make infringers of the millions of end users whose prompts lead to those outputs, not to mention third-party application developers who integrate their systems with tools like GPT-4.

⁹⁵ Samuelson, Fair Use Defenses, supra note 84, Part III-A-2.

⁹⁶ For a discussion of the origins and proper scope of copyright's derivative work right, see, e.g., Pamela Samuelson, *The Quest for a Sound Conception of Copyright's Derivative Work Right*, 101 GEO. L.J. 1505 (2013).

⁹⁷ Complaint, Andersen v. Stability AI, Ltd., Case No. 3:23-cv-00201 ¶ 93 (N.D. Cal. Jan. 13, 2023). Andersen also claims that Stable Diffusion's model is a derivative work of the training data. Id. ¶ 57. This too is an implausible claim. Getty claims that some outputs of Stable Diffusion closely resemble particular Getty images, Getty Images (U.S.), Inc. v. Stability AI., No. 1:23-cv-00135 ¶51 (D. Del. 2023), but does not provide specific examples of what it considers infringing derivatives.

⁹⁸ See, e.g., Skidmore v. Led Zepplin, 952 F.3d 1051, 1064 (9th Cir. 2020) (substantial similarity in expression copied from the plaintiff's work is necessary to establish infringement); Litchfield v.

recasting, transforming, and adapting their works under the open-ended clause in the statutory definition, 17 U.S.C. § 101, because that clause does not do away with the requirement that expression in the defendants' works must be substantially similar to infringe.⁹⁹ Andersen's complaint fails to show any examples of Stable Diffusion outputs that are substantially similar to expressions in Andersen's works. Her theory is that all outputs are infringing derivatives.¹⁰⁰ Unsurprisingly, a trial judge has announced his intent to dismiss Anderson's derivative work claims for lack of substantial similarity.¹⁰¹

In summary, under U.S. law making a derivative work means recasting a qualitatively and quantitatively significant amount of the primary work's original expression into a new form or a new version. If the model and its outputs are not, in some sense, a "copy" of a work in the training data, it is not a "derivative work based upon the copyrighted work" either.

NOI Question 23. Is the substantial similarity test adequate to address claims of infringement based on outputs from a generative AI system, or is some other standard appropriate or necessary?

Yes, see above. It would violate fundamental principles of copyright law to impose liability on developers of AI systems for outputs in the absence of substantially similar expression derived from expression in specific inputs.

NOI Question 24. How can copyright owners prove the element of copying (such as by demonstrating access to a copyrighted work) if the developer of the AI model does not maintain or make available records of what training material it used? Are existing civil discovery rules sufficient to address this situation?

We are not aware of any reason why existing civil discovery rules would be insufficient to address this situation.

Speilberg, 736 F.2d 1352, 1357 (9th Cir. 1984) ("frivolous" to claim infringement of the derivative work right in the absence of substantial similarity). See also Authors Guild v. Google, 804 F.3d at 207 (rejecting Guild's derivative work claims).

⁹⁹ See generally WILLIAM F. PATRY, PATRY ON COPYRIGHT §12:13 (2022) ("there must be substantial similarity in protectible expression between the parties" works" to infringe the derivative work right). See Samuelson, *Quest, supra* note 92, at 1525-27 (discussing the last clause of §101 definition).

¹⁰⁰ Complaint, Andersen v. Stability AI, Ltd., Case No. 3:23-cv-00201 ¶ 4 (N.D. Cal. Jan. 13, 2023).

¹⁰¹ Transcript of Proceedings, Andersen v. Stability AI, Inc., Case No. 23-cv-201 5 (July 19, 2023). The judge announced his intent to dismiss all but one of Andersen's claims, but to give the plaintiffs leave to amend. *Id.*

NOI Question 25. If AI-generated material is found to infringe a copyrighted work, who should be directly or secondarily liable—the developer of a generative AI model, the developer of the system incorporating that model, end users of the system, or other parties?

Although it is very difficult to control the output of an LLM or a text-to-image model by simple prompting, a user with detailed knowledge of a copyrighted work might be able to remake it, at least at a vague level of similarity. In the right hands, a Generative AI model can be used as a tool of copyright infringement, but the same is also true of a typewriter.

Makers and deployers of Generative AI tools should only be indirectly liable for infringing outputs that the tool-user did not knowingly provoke or that were highly foreseeable and could be easily guarded against. Whether and how the volitional act requirement applies to Generative AI is an interesting question. Courts may need to consider whether Generative AI systems are technologies that have substantial non-infringing uses and hence eligible for the *Sony* safe harbor that the Court reaffirmed in *MGM v. Grokster*.¹⁰²

NOI Question 26. If a generative AI system is trained on copyrighted works containing copyright management information, how does 17 U.S.C. 1202(b) apply to the treatment of that information in outputs of the system?

Several complaints against developers of Generative AI systems claim that the defendants removed and/or altered copyright management information (CMI) and distributed works whose CMI had been altered or removed in violation of § 1202.

Section 1202(b)(1) prohibits the intentional removal or alteration of CMI and § 1202(b)(3) outlaws distribution of copies of original works when the distributor knows that CMI has been removed or altered. Both subsections also require a nexus to copyright infringement, specifically, that the defendant knew or had "reasonable grounds to know, that [removal or alteration of CMI] will induce, enable, facilitate, or conceal an infringement."¹⁰³ Congress deliberately adopted a double scienter requirement in § 1202—a showing of intentional removal or alteration and either knowledge or reason to know of intentional or knowing removals or alterations of CMI will induce or facilitate infringement—to limit the scope of liability to this category of indirect infringement. The double scienter requirement in § 1202 (b) means that if copying training data is reliably fair use and the outputs are reliably non-infringing, there would be no infringement to be "induced, enabled, facilitated, or concealed" by removing the CMI and thus no liability under § 1202 (b) of the Copyright Act.

¹⁰² Metro-Goldwyn-Mayer Studios Inc. v. Grokster, Ltd., 545 US 913, 931-34 (2005)

¹⁰³ 17 U.S.C. § 1202(b).

The legislative history of § 1202 should be taken into account as courts and the Copyright Office interpret this provision in the context of Generative AI systems.¹⁰⁴ The goal of § 1202 was to guard against internet piracy.¹⁰⁵ Proponents of what became § 1202 feared that Internet pirates would strip out CMI from digital copies of copyrighted works and insert false CMI so that they could distribute copies outside of copyright owners' controls, either by monetizing the infringing copies or releasing the copies as if they were in the public domain. The Generative AI litigations pose novel questions about the application of § 1202 that seem quite different from the scenarios that § 1202 was intended to address.

We recognize that § 1202 claims are highly significant in the Generative AI cases because violations qualify plaintiffs for awards of statutory damages without any need to register copyright claims. The awardable damages range from a minimum of \$2500 to a maximum of \$25,000 per violation. 17 U.S.C. § 1202(c)(3)(B). With millions or billions of works alleged to violate § 1202 in the Generative AI cases, the potential statutory damages would be staggering in amount. We think that staggering awards of this sort should be reserved for § 1202 violations that create high risks of piracy of the sort contemplated when Congress enacted § 1202.

Courts have generally given § 1202 a relatively limited interpretation. In *Stevens v. Corelogic*, Inc.,¹⁰⁶ for example, the Ninth Circuit affirmed summary judgment against a photographer's claim that the automated process by which Corelogic compressed digital files for real estate service webpostings did not violate § 1202. The automated nature of the processing was one consideration weighing against liability, but the court also held that there must be a showing that "future infringement is likely,"¹⁰⁷ not merely that there was a general possibility of infringement.¹⁰⁸ Automated removal or alteration of CMI in the course of training foundation models may not result in the high risks of infringement for which § 1202 was adopted.

The text of § 1202 refers to removal or alteration must be from "copies" of protected works, suggesting that § 1202 violations must involve identical copies, not derivative works.¹⁰⁹ Most

¹⁰⁶ 899 F.3d 666 (9th Cir. 2018).

¹⁰⁷ *Id.* at 675.

¹⁰⁴ See, e.g., Pamela Samuelson, The U.S. Digital Agenda at WIPO, 37 Va. J. Int'l L. 369, 415-18 (1996).

¹⁰⁵ See S. Rep. 105-190 at 8, 11 n.18 (1998) (noting the overarching legislative purpose to "discourage piracy" on the Internet).

¹⁰⁸ Id. at 673. See also Kelly v. Arriba Soft Corp., 77 F. Supp.2d 1116, 1122 (C.D. Cal. 1999), rev'd other grounds, 336 F. 3d 811 (9th Cir. 2003) (removal of CMI as unintended side effect of automated process not grounds for § 1202 liability).

¹⁰⁹ See, e.g., Robert L. Stark Enter., Inc. v. Neptune Design Group, LLC, 2017 WL 1345195 (N.D. Ohio) (§ 1202 not violated by new architectural plans alleged to infringe that do not include CMI).

of the Generative AI output claim infringement of the derivative work right. While we explained earlier why we think most of the derivative work claims in the Generative AI cases are weak, even assuming some outputs do infringe, § 1202 claims may nonetheless be inappropriate.

(D) TRANSPARENCY & RECORDKEEPING

NOI Question 15. In order to allow copyright owners to determine whether their works have been used, should developers of AI models be required to collect, retain, and disclose records regarding the materials used to train their models? Should creators of training datasets have a similar obligation?

(a) Collection and retention

Proposals for recordkeeping and disclosure in relation to the use of copyrighted works are beset with conceptual and practical difficulties.

If an AI developer is given notice of an impending legal claim, it would be required to maintain relevant and appropriate records or risk sanctions for spoliation. Following the normal course of civil litigation, those records would be subject to discovery. Beyond that, we are not aware of any authority that entitles a copyright owner to know whether and how third parties make fair uses or engage in other non-infringing conduct with respect to their works.

The nature of the recordkeeping and transparency obligations contemplated in NOI Questions 15-17 is unclear. Good data science practices often include the retention of records and data sufficient to enable replication, but we note that norms relating to replication and transparency vary by discipline and according to context. We are not aware of any academic or industry consensus that could provide clarity in the face of a generally stated obligation to "collect, retain, and disclose records regarding the materials used to train [AI] models." The replication norms that we are aware of would not rise to the level of allowing copyright owners to determine whether their works had been used in a particular piece of research.

The scope of the recordkeeping and transparency obligations contemplated in NOI Questions 15-17 is unclear. Specifically, it is unclear whether such a requirement is contemplated with respect to works subject to open-source or Creative Commons licenses.

Many AI researchers would face significant practical obstacles in complying with an obligation to "collect, retain, and disclose records" sufficient to "to allow copyright owners to determine whether their works have been used" in the course of developing a machine learning or AI model.

Those practical obstacles include:

(1) AI researchers may have only very limited information about the works that comprise the training data. They may rely on training materials collected or curated by third parties. In many cases, the only information researchers will have about a work is that it was associated with a particular URL at a particular point in time.

(2) Any requirement to independently derive accurate information about the title, ownership, or chain of licensing, of works in the training data would substantially increase the cost of developing AI models for research and commercial purposes.

(3) If such an obligation were imposed, it is unclear what level of effort in tracing ownership, etc., would satisfy it.

(4) Such an obligation may conflict with data privacy laws to the extent that it requires AI developers to collect and store personally identifying information about individuals in multiple jurisdictions.

Moreover, the term "developers of AI models" encompasses a vast range of individuals and entities—from a graduate student working on a project to train a model to upscale photographs, to a multi-billion-dollar corporation creating sophisticated language models. A one-size-fits-all set of recordkeeping and disclosure requirements would raise costs for academic and commercial researchers alike, regardless of the merits of any fair use defense that may apply.

(b) Disclosure:

Technology companies and researchers may have a legitimate interest in not publicly disclosing general or specific details of the works included in, and excluded from, the training data of a particular model. The disclosure of such information could undermine trade secret rights; reveal personal or sensitive information; and/or provide a roadmap to circumvent measures designed to achieve various "AI Safety" goals.

NOI Question 15.1. What level of specificity should be required?

See above.

NOI Question 15.2. To whom should disclosures be made?

See above.

NOI Question 15.3. What obligations, if any, should be placed on developers of AI systems that incorporate models from third parties?

It seems infeasible to make developers of second-generation models (e.g. fine-tuned models) obtain records from the entity that trained the initial model.

NOI Question 15.4. What would be the cost or other impact of such a recordkeeping system for developers of AI models or systems, creators, consumers, or other relevant parties?

The cost and broader impact of any recordkeeping and disclosure requirements would be entirely contingent on the scope and nature of those requirements. Ultimately, the Copyright Office would need to seek guidance from technical experts with relevant industry expertise on the likely costs and downstream consequences of any specific record keeping and disclosure mandate under contemplation.

NOI Question 16. What obligations, if any, should there be to notify copyright owners that their works have been used to train an AI model?

As noted above, we are not aware of any authority that entitles a copyright owner to know whether and how third parties make fair uses or engage in other non-infringing conduct with respect to their works. If a developer has used a copyright work in a non-expressive way, as described above, there should be no obligation to notify copyright owners that their works have been used to train an AI model.

NOI Question 17. Outside of copyright law, are there existing U.S. laws that could require developers of AI models or systems to retain or disclose records about the materials they used for training?

As noted above, if an AI developer is given notice of an impending legal claim, it would be required to maintain relevant and appropriate records or risk sanctions for spoliation. Following the normal course of civil litigation, those records would be subject to discovery.

(E) LABELING OR IDENTIFICATION

NOI Question 28. Should the law require AI-generated material to be labeled or otherwise publicly identified as being generated by AI? If so, in what context should the requirement apply and how should it work?

We appreciate that there are broader discussions taking place about the implications of AI generated text and synthetic media, but these seem tangential to copyright law. AI-generated text and synthetic media raises significant questions in relation to consumer protection, personal privacy and dignity, and more broadly in relation to the health of our information ecosystem.

Any proposal for labeling or otherwise identifying AI-generated material should be carefully calibrated to the specific public interest objective the regulator has in mind. This calibration is important because different policy objectives necessarily entail different kinds of labeling and different thresholds for identification.

The line between AI-generated and human-generated may be difficult to draw when a person frames a prompt, reviews the output, and then edits the AI-generated text. The line will also be difficult to draw when a person uses AI-powered editing tools to manipulate a work that was initially human-authored, or adds significant human authorship to an image that was initially AI-generated. For example, photos taken on an iPhone in the "portrait mode" would be AI-generated according to some definitions—significant aesthetic features of the work are determined by a machine learning algorithm—but there is no obvious consumer interest in having all such images labeled or watermarked. Indeed, we imagine many iPhone users would object to such an interference.

In other contexts, labeling may be important because the public needs to know whether the content they are being presented with has been manipulated, or even entirely manufactured. If a news report features an image of the Pope in a white puffy jacket,¹¹⁰ the tools used to create the image are far less important than the fact that the image is fake. Accordingly, in certain contexts, any manipulation of the image or text should be disclosed.

(F) NON-COPYRIGHT RIGHTS AND INTERESTS

NOI Question 32. Are there or should there be protections against an AI system generating outputs that imitate the artistic style of a human creator (such as an AI system producing visual works "in the style of" a specific artist)? Who should be eligible for such protection? What form should it take?

Copyright protection for "style" in the sense of an artist's distinctive appearance or voice is unnecessary. If generative AI were used to re-created someone's distinctive appearance or voice, that person should have recourse under right of publicity. If current right of publicity laws are inadequate, they should be reformed.

If the term "style" is meant to suggest an exclusive entitlement defined by abstract features ascertained by looking for distinctive features across a series of works, then protection of style is contrary to foundational principles of U.S. Copyright law. Personality and reputation are emergent properties that develop across an artist's career, and these are protected by laws in relation to the right of publicity and trademark. Copyright, in contrast, protects the author's interests in relation to singular works.¹¹¹ Furthermore, any concept of style that can

¹¹⁰ Fact Check-Image of Pope Francis wearing oversized white puffer coat is AI-generated, REUTERS FACT CHECK, Mar. 29, 2023 12:14 PM (https://www.reuters.com/article/factcheck-pope-ai/fact-check-image-of-pope-francis-wearing-oversized-white-puffer-coat-is-ai-generated-idUSL1N36120G)

¹¹¹ The exclusive rights in Section 106 are framed in terms of "the copyrighted work."

only be identified by considering several works collectively is far too abstract to merit copyright protection consistent with the idea/expression distinction and Section 102(b).

Even if proposed copyright protection for "style" were focused on stylistic features of individual works, it is difficult to see how copyright protection for style or artistic technique could be reconciled with the idea/expression distinction and Section 102(b).

Finally, we note that copyright in style should not be confused with the protection afforded to copyrightable characters. A character is a person, animal, or even an inanimate object imbued with personality that is a recognizable expressive creation. A character is more than a recognizable image in a certain style, a character necessarily has a narrative element. There may be infringement claims that are described in terms of copyrightable style that could be more effectively argued in terms of copyrightable characters.

The tendency of users of text-to-image generators to invoke the names of living artists in prompts has caused considerable consternation. This practice may often be harmless, but it can occasionally result in the names of particular artists being publicly associated with works they did not author, to an extent that dwarfs their own substantial artistic contributions. For example, one of the most commonly invoked style prompts in early 2023 was Greg Rutkowski,¹¹² an artist who is well known for his richly detailed depictions of Dungeons & Dragons and similar worlds in a style that has been compared to the romantic English painter, William Turner.¹¹³ Rutkowski's name is primarily used as a shortcut to invoke high-quality digital art generally, or in relation to fantasy motifs, and not in an attempt to recreate his style specifically.¹¹⁴

Rutkowski's genuine works appear to be crowded out in Internet searches by tens of thousands of images produced "in the style of Rutkowski."¹¹⁵ This seems like harm that trademark law and right of publicity could address more easily than copyright law. Accordingly, copyright protection for "style" in the sense of an artist's name or reputation is probably unnecessary and reform efforts in this regard should be focused on trademark and right of publicity law.

¹¹² Melissa Heikkilä, "This artist is dominating AI-generated art. And he's not happy about it" MIT Tech. Rev. Sep 16, 2022 (available at https://www.technologyreview.com/2022/09/16/1059598/this-artist-is-dominating-ai-generatedart-and-hes-not-happy-about-it/) (noting that prompts in Midjourney and Stable Diffusion for the artist Greg Rutkowski were more popular than for Picasso and other more famous artists.)

¹¹³ Sag, *Copyright Safety*, *supra* note 1.

¹¹⁴ Id. (assessment based on a review of prompts including "Greg Rutkowski" located using a Google image search on April 20, 2023.)

¹¹⁵ Id.

We also note that this dilution-related harm could be avoided if the names of individual artists paired with images in the training data of models like Stable Diffusion were replaced with more general descriptions, or if platforms like Midjourney and DALL·E did not allow unmixed prompts in the names of individual living artists. However, these kinds of restrictions also have costs and may curtail some free expression. We encourage thoughtful self-regulation in this regard, but we suspect that any attempt to impose such rules would be unworkable in light of the First Amendment.

CONCLUSION

We commend the Copyright Office for its initiative in addressing the implications of artificial intelligence for copyright law and policy in this Notice of Inquiry and we are grateful for the opportunity to offer our assistance.

Sincerely,

Pamela Samuelson

Christopher Jon Sprigman

Matthew Sag (corresponding author)