AIPLA

American Intellectual Property Law Association

October 30, 2023

Ms. Suzanne Wilson
General Counsel and Associate Register of Copyrights
U.S. Copyright Office
101 Independence Avenue, SE
Washington, DC 20559-6000

Re: Comments Submitted Pursuant to Notice of Inquiry and Request for Comments, 88 Fed. Reg. 59942 (August 30, 2023)

Dear Ms. Wilson:

The American Intellectual Property Law Association ("AIPLA") is pleased to offer comments in response to the above-referenced U.S. Copyright Office Notice of Inquiry and Request for Comments ("RFC") related to the study of copyright law and policy issues raised by artificial intelligence ("AI") systems.

Founded in 1897, AIPLA is a national bar association of approximately 7,000 members who are engaged in private or corporate practice, in government service, and in the academic community. AIPLA members represent a wide and diverse spectrum of individuals, companies, and institutions involved directly or indirectly in the practice of patent, trademark, copyright, trade secret, and unfair competition, as well as other fields of law affecting intellectual property. Our members represent both owners and users of intellectual property. Our mission includes helping to establish and maintain fair and effective laws and policies that stimulate and reward authorship but that also balance the public's interest in healthy competition, reasonable costs, and basic fairness. Our members have a key interest in developing policies concerning AI that are clear, administrable, and resistant to misuse.

Copyright protection is essential for encouraging creativity by granting creators exclusive rights to their works and fostering innovation and investment in artistic and intellectual endeavors. The introduction of AI is the latest iteration of a long-standing debate, and ever-growing body of law, concerning the extent to which authorship and, by extension, copyright rights, are implicated by the use of tools to generate creative works. Conversely, AI also raises unique questions about the manner in which technology uses or facilitates the easy replication and distribution of copyrighted works. Striking the right balance between protecting the rights of creators and fostering technological innovation will be a challenging task for policymakers and the legal system.

We are compelled to note, as a fundamental value statement, that we believe human-generated creative works are irreplaceably possessed of inherent value. Art, music, dance, and storytelling are cultural universals found in all human societies throughout history. These expressive endeavors are foundational elements of human culture. They provide the medium for the inheritance of culture, and the shared exploration of the human condition. These forms of expression enrich our lives, define our humanity, and function as both a mirror reflecting who we are, and as a lens for perceiving and confronting the complexity of human existence.

Although AI is, in a sense, the latest continuation of an extant technological conundrum, recent rapid advances in AI threaten to replace or supplant human-authored creative works to a degree not previously seen. To that end, artistic communities, users, AI developers, legal practitioners, regulators, and legislators must join together to find common sense approaches to strike a balance between managing the impact of AI on the vibrancy of creative communities and the continued development of human-generated creative works without unduly stifling the innovation that its utilization could spark.

We are pleased to present our responses below, subject to a few prefatory remarks.

First, AI encompasses various types and subfields, each with its unique characteristics and applications. Our responses to these questions are predicated upon the current state of the art, with particular focus on the newest AIs recently made available to the general public, chiefly generative output systems. However, advancements in AI continue to evolve rapidly, and future AIs may have fundamentally different training, processing, or outputting aspects, about which our views may be different. This response thus should be understood in the context of the present technological landscape and is not necessarily transferrable to future developments.

Second, the RFC uses the term "AI-generated works" throughout, but this broad definition encompasses works whose genesis involves varying degrees of AI contribution. We suggest that this framework should be understood as encompassing three distinct categories of "works generated by or using AI." These are:

Category A: Works created by a human author with the assistance of an AI technology (AI as mere tool of human creativity);

Category B: Works created by a human in collaboration with AI, wherein the activity of the AI, if done by a human, would be considered sufficiently original and creative to merit copyright protection (AI as co-author with a human author); and

Category C: Works created by an AI that, if created by a human, would be considered sufficiently original and creative to merit copyright protection, but under circumstances in which there is no human creativity or originality sufficient to meet the legal standard for copyright protection (AI as sole author).

The issues presented in the RFC can differ among these categories of works generated by or using AI.

Third, AIPLA elected not to respond to all questions presented in the RFC. Although we, and our members, have viewpoints about these topics, we sought to generally confine our response to questions of copyright law and policy, and to allow others to provide responses to other types of questions, such as questions regarding the technological aspects of AI, and industry-specific concerns. However, our omission of responses to those questions should not be interpreted as an endorsement of any other views submitted, including separate submissions by organizations or individuals on their own behalf who are members of AIPLA, nor as a lack of interest in the subject matter.

Our responses are as follows.

QUESTION 1. As described above, generative AI systems have the ability to produce material that would be copyrightable if it were created by a human author. What are your views on the potential benefits and risks of this technology? How is the use of this technology currently affecting or likely to affect creators, copyright owners, technology developers, researchers, and the public?

Artificial intelligence is here, and it shows no signs of slowing down. Rather than ignoring its proliferation, the artistic community, users, AI developers, legal practitioners, regulators, and legislators must join together to find common sense ways to manage its impact on the creative community while keeping an open mind to encourage the innovation that its utilization could spark.

Artificial intelligence has the potential to allow human creators to maximize their efficiency, and enhance their creative outputs, by simplifying monotonous tasks and channeling human creativity. However, it can also be used to replace human authorship entirely, and independently generate content, or generate content using human-inputted parameters. We believe caution is merited due to the potential for AI to significantly supplant and impact the economic incentives for human creative endeavors.

We believe that copyright law has, so far, generally been applied correctly to AI, in that copyright protection is granted only to works created by humans that satisfy the threshold requirements for originality and creativity. However, this seemingly clear position gets fuzzy with the introduction of generative AI, as the line between human authorship and machine authorship can become muddled. Accordingly, we think it is important that the law have flexibility as to what constitutes "human creation." We believe existing copyright law is sufficiently flexible and adaptable for this purpose, and that courts should be afforded sufficient opportunity to apply existing copyright law to AI before intervention is warranted.

Although we believe existing copyright law is up to the task of addressing the issues raised by AI, to the extent that new legislation is considered, we believe that authors should retain the right (subject to existing defenses, such as fair use) to exclude their works from AI models, and that exclusion should be the status quo absent affirmative consent (i.e., an "opt-in" system). Though we do not believe it is necessary to legislate at this time, we remain open to, and interested in, considering any legislative or rulemaking proposals.

QUESTION 3. Please identify any papers or studies that you believe are relevant to this Notice. These may address, for example, the economic effects of generative AI on the creative industries or how different licensing regimes do or could operate to remunerate copyright owners and/or creators for the use of their works in training AI models. The Office requests that commenters provide a hyperlink to the identified papers.

Publication	HyperLink
Generative Artificial Intelligence and Copyright Law. Congressional Research Service. (Sep 29, 2023)	https://crsreports.congress.gov/product/pdf/LSB/LSB10922

Publication	HyperLink
Response to May 1, 2023 letter regarding AI Material (May 12, 2023)	https://www.copyright.gov/laws/hearings/Response-to-May-1-2023-letter-re-AI-Material.pdf
Letter to USPTO & USCO on National Commission on AI (Oct 27, 2022)	https://www.copyright.gov/laws/hearings/Letter-to-USPTO-USCO-on-National-Commission-on-AI-1.pdf
Fact Sheet: Biden-Harris Administration Secures Voluntary Commitments from Leading Artificial Intelligence Companies to Manage the Risks Posed by AI (July 21, 2023)	https://www.whitehouse.gov/briefing-room/statements-releases/2023/07/21/fact-sheet-biden-harris-administration-secures-voluntary-commitments-from-leading-artificial-intelligence-companies-to-manage-the-risks-posed-by-ai/
Generative AI, Productivity, the Labor Market, and Choice Behavior (September 22, 2023)	https://www.federalreserve.gov/newsevents/speech/cook20230922a.htm
Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence (March 16, 2023)	https://copyright.gov/ai/ai_policy_guidance.pdf
International Copyright Issues and Artificial Intelligence: Transcript from Online Webinar (July 26, 2023)	https://www.copyright.gov/events/international-ai-copyright-webinar/International-Copyright-Issues-and-Artificial-Intelligence.pdf
MEPs ready to negotiate first-ever rules for safe and transparent AI (June 14, 2023)	MEPs ready to negotiate first-ever rules for safe and transparent AI | News | European Parliament (europa.eu)
The EU AI Act: What businesses and their legal advisers need to know (July 03, 2023)	The EU AI Act: What businesses and their legal advisers need to know - The Global Legal Post
Amended EU AI Act strengthens copyright enforcement against Generative AI but IP issues remain. World Trademark Review. (June 6, 2023)	https://www.worldtrademarkreview.com/article/amended-eu-ai-act-strengthens-copyright-enforcement-against-generative-ai-ip-issues-remain
Regulating Generative AI: How Well Do LLMs Comply with the EU AI Act? (Aug 08, 2023)	https://towardsdatascience.com/regulating-generative-ai-e8b22525d71a
Principles for Copyright and Generative AI	https://cla.co.uk/about-us/copyright/principles-for-copyright-and-generative-ai/

Publication	HyperLink
The economic potential of generative AI: The next productivity frontier (June 14, 2023)	https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier?gclid=EAIaIQobChMIiLvMyt3KgQMVCDnUAR1wQggFEAAYASAAEgJ___D_BwE
Generative AI could raise global GDP by 7% (Apr 05, 2023)	https://www.goldmansachs.com/intelligence/pages/generative-ai-could-raise-global-gdp-by-7-percent.html
Sizing the prize: PwC's Global Artificial Intelligence Study: Exploiting the AI Revolution	https://www.pwc.com/gx/en/issues/data-and-analytics/publications/artificial-intelligence-study.html
The current legal cases against generative AI are just the beginning (Jan 27, 2023)	https://techcrunch.com/2023/01/27/the-current-legal-cases-against-generative-ai-are-just-the-beginning/
Marquette Intellectual Property Law Review: Artificial Intelligence, Copyright, and Copyright Infringement (Winter 2020)	https://scholarship.law.marquette.edu/cgi/viewcontent.cgi?article=1355&context=iplr
The U.S. Copyright Office said no to AI, but its explanation is unconvincing | GUEST COMMENTARY (May 31, 2023)	https://www.baltimoresun.com/opinion/op-ed/bs-ed-op-ai-copyright-office-20230531-7e2dthegbncx7i5y4zlp7qj4qu-story.html
2023.02.21 Zarya of the Dawn Letter by the Copyright Office	https://www.copyright.gov/docs/zarya-of-the-dawn.pdf
2023.09.05 Second Request for Reconsideration for Refusal to Register Théâtre D'opéra Spatial	https://www.copyright.gov/rulings-filings/review-board/docs/Theatre-Dopera-Spatial.pdf
2019 AIPPI World Congress – Adopted Resolution (Sep. 18, 2019) – Copyright in Artificially Generated Works	https://www.aippi.org/content/uploads/2022/11/Resolution_Copyright_in_artificially_generated_works_English.pdf

QUESTION 4. Are there any statutory or regulatory approaches that have been adopted or are under consideration in other countries that relate to copyright and AI that should be considered or avoided in the United States? How important a factor is international consistency in this area across borders?

Certainly, it is prudent and helpful to monitor the developments outside of the United States, particularly in the European Union and the United Kingdom. It is also helpful to participate in developing Voluntary Codes of Conduct relating to AI such as those that are being developed now in the United States, the European Union, and Canada.

However, we urge caution in the consideration of legislative and regulatory activity at this early stage, as it is not evident that statutory or regulatory intervention are merited. Existing legal doctrines have so far proven sufficiently adaptable to address AI, and the commercial challenges presented by AI may be adequately resolved in time by markets or the outcomes of currently pending litigation, discussed below. History has proved repeatedly that rushing to regulate new technology does not work, as the laws are necessarily premised on an early, and often incomplete, understanding of both the ultimate trajectory of the technology and the resulting markets for goods and services. The law is then outdated by the time it takes effect. We thus suggest giving markets and pending litigation time to play their respective roles in crystalizing what, if any, commercial and legal issues need to be addressed by new legislation.

Litigation is presently on-going in a number of U.S. Courts concerning the extent to which copyrighted material can be used for training purposes, for example *Reuters v. Ross Intelligence*. Tech companies including Meta Platforms, Stability AI and OpenAI are also facing lawsuits from copyright owners over the use of their work to train the companies' generative AI software. Use of existing works, whether in the public domain or held by copyright owners, is helpful to train AI platforms to obtain the benefits promised by AI. However, it is likely that the existing fair use framework will be able to adequately balance the rights of copyright owners with the rights of AI platforms who wish to use proprietary works. While international consistency of result is beneficial, it should be remembered that many other jurisdictions, particularly the European Union, do not have fair use provisions, which in the United States provides for a balance between the rights of copyright owners and users, and do have other forms of intellectual property, such as moral rights, that receive only limited recognition in the United States.

We do not believe that AI service providers should be compelled to disclose training data. The RFC makes specific reference to the EU Artificial Intelligence Act, amendment 399, art 28b, which would require providers of foundation AI models to "make publicly available a sufficiently detailed summary of the use of training data protected under copyright law." This proposed rule assumes that it will be easily possible to determine in all cases what training data is protected under copyright law. To the extent that legislation imposing such an obligation is considered, it should include broad exceptions for matters such as, but not limited to, information subject to non-disclosure obligations, trade secrets, and the use of one's own works as training inputs.

The nature and extent of the originality and skill that is necessary to imbue pure data with copyright protection may be best left to the courts at this time. If data monopolies develop or

access to works for training is inadequately addressed in the marketplace, it may be necessary to deal with that by legislation in the future.

QUESTION 5. Is new legislation warranted to address copyright or related issues with generative AI? If so, what should it entail? Specific proposals and legislative text are not necessary, but the Office welcomes any proposals or text for review.

As noted in prior responses, we do not think legislation is appropriate at this time.

See also the answer to questions 18 and 19 below.

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.

> **QUESTION 8.1. 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? 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?**
>
> **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?**
>
> **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?**
>
> **QUESTION 8.5. 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? 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?**

We encourage caution and deliberateness in addressing these issues.

The fair use analysis is contextual and flexible by design, and we advise against attempting to define, in advance, those circumstances under which the unauthorized use of copyrighted works to train AI models should and should not constitute fair use. Each use should be evaluated on its own facts and merits in the same manner that fair use is applied now. Additionally, we note that the fair use analysis may differ substantially for different uses within an AI model and should be independently evaluated for all such uses.

In *Google v. Oracle America*, the Supreme Court unambiguously limited its holding to the particular facts of that case, which are not implicated in the training and use of AI models. As such, we do not regard this case as meaningfully instructive.

The applicability of *Warhol* is also limited in that the sole issue on appeal was whether a work that conveys a different meaning or message from its source material can be "transformative" under the first fair use factor. The application of this holding will necessarily vary from case to case depending on the particular facts, and we again advise against attempting to determine, in advance of those facts, when and how to apply the ruling.

The source of funding currently does not play a meaningful role in the law of fair use, nor do we see any unique characteristics of AI, as currently constituted, which would compel a different result with respect to this technology. The applicability of the fair use doctrine and its constituent factors will necessarily vary from case to case depending on the particular facts, and we oppose attempts to determine in advance how it should apply.

The existing legal framework for the applicability of the fourth factor appears to be adequate and we advise against rewriting or revising this rule before courts have had a reasonable opportunity to apply it. We believe the fourth factor should look only to the work-in-suit, and not the effect on other works not at issue in a particular lawsuit. To the extent that broader classes of plaintiffs seek collective redress of harms suffered due to a single AI model, the Federal Rules of Civil Procedure provide the class action mechanism for this purpose.

We also note that the particular application of the fair use doctrine may apply differently to different actors. For example, the question of fair use may be applied in a different manner to the service provider offering an AI-based content generation service than to the ultimate end user of the outputs of such a system.

Finally, we also draw attention the decisions in *Sony Corp. of American v. Universal City Studios, Inc.*, 464 U.S. 417 (1984), and *Sega Enters. Ltd. V. Accolade, Inc.*, 977 F.2d 510 (9th Cir. 1992).

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)?

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?

QUESTION 9.2. If an "opt out" approach were adopted, how would that process work for a copyright owner who objected to the use of their works for training? Are there technical tools that might facilitate this process, such as a technical flag or metadata indicating that an automated service should not collect and store a work for AI training uses?

QUESTION 9.3. What legal, technical, or practical obstacles are there to establishing or using such a process? Given the volume of works used in training, is it feasible to get consent in advance from copyright owners?

QUESTION 9.4. If an objection is not honored, what remedies should be available? Are existing remedies for infringement appropriate or should there be a separate cause of action?

QUESTION 9.5. In cases where the human creator does not own the copyright—for example, because they have assigned it or because the work was made for hire—should they have a right to object to an AI model being trained on their work? If so, how would such a system work?

Starting with the defining purpose, "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," copyright laws provide incentive to authors by granting the author the right to control copying of the author's work for a period of time; the public benefits from the creative work being generally available to the public.

Copyright exclusivity is not unlimited; by statute, for example, registration is mandatory for enforcement and certain damages. While the copyright in a work persists, ensuring authors possess the right and ability to control and limit (subject to fair use and other defenses) the use of their works to train AI systems would arguably incentivize creativity, by creating demand for their work in any AI model.

We do not believe that the implementation of a compulsory licensing system is appropriate at this time. That said, we are also open to, and interested in, considering any other type of proposed legislation or rulemaking pertaining to simplified licensing of works. If such a system is implemented, we believe copyright owners should have a simple and clear means to opt out from use by AI, subject to existing exceptions, such as fair use.

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

QUESTION 10.1. Is direct voluntary licensing feasible in some or all creative sectors?

QUESTION 10.2. Is a voluntary collective licensing scheme a feasible or desirable approach? Are there existing collective management organizations that are well-suited to provide those licenses, and are there legal or other impediments that would prevent those organizations from performing this role? Should Congress consider statutory or other changes, such as an antitrust exception, to facilitate negotiation of collective licenses?

QUESTION 10.3. Should Congress consider establishing a compulsory licensing regime? If so, what should such a regime look like? What activities should the license cover, what works would be subject to the license, and would copyright owners have the ability to opt out? How should royalty rates and terms be set, allocated, reported and distributed?

QUESTION 10.4. Is an extended collective licensing scheme a feasible or desirable approach?

QUESTION 10.5. Should licensing regimes vary based on the type of work at issue?

The number of scenarios of potential licensing is astronomical and non-quantifiable. There is no one-size-fits-all license for the vast number and types of works available.

We do not believe that the implementation of a compulsory licensing system is appropriate at this time. That said, we are also open to, and interested in, considering any proposed legislation or rulemaking pertaining to such a system. If such a simplified licensing system is implemented, we believe copyright owners should have a simple and clear means to opt out from use by AI, subject to existing exceptions, such as fair use.

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? If so, what factors are relevant to that determination? For example, is selecting what material an AI model is trained on and/or providing an iterative series of text commands or prompts sufficient to claim authorship of the resulting output?

Because copyright rights arise as an incident of authorship, and because the law requires a human author, this question asks not only when a human should be regarded as "author," but when, and to what extent, the work in question is eligible for copyright protection at all. We believe the answer will vary depending on the nature and extent of AI-generated contribution and the particular facts of each instance of training, and using, an AI. We again note that legislation on this point is likely premature at this stage, and we advise caution and care in taking legislative action before existing legal disputes conclude and markets adapt, which may refine and narrow the issues for legislative intervention or obviate the need for intervention at all.

We also refer the Copyright Office to the AIPPI Group's Resolution answer to the AIPPI 2019 Study Question on copyright in AI Generated Works, and the law cited therein. These materials are included in our answer to Question No. 3.

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?

We do not believe revisions are required at this time.

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?

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?

Whether legal protection for AI-generated material is desirable as a policy matter will vary from case to case. The scope of what constitutes "AI-generated material" is broad, and the issues are complex and multifaceted. In some cases, protection may be desirable, and in others, it may not. We do believe that it is desirable as a policy matter to have as much clarity as is reasonably achievable regarding the scope and nature of the protection for AI-generated material, but the issues presented by AI are contextual and fact-specific, and not amenable to a single, comprehensive policy position.

At this time, it does not appear that legal protection for AI-generated outputs is critical to incentivizing the creation of AI technologies and systems; and the copyrightability of the AI system itself is sufficient. These systems were generated and commercialized in the absence of any clear authority providing legal protection to the outputs, and the absence of such protections does not appear to have diminished the public's interest in consuming AI, nor service-providers' interest in providing it. So far, market incentives do not appear to have turned on the legal status of AI outputs. However, we again observe that the AI economy is fluid and evolving rapidly, and this may change.

So far, it appears that existing copyright law is being applied correctly to AI-generated works, and we do not see any clear need for additional protections not provided by copyright. That said, we are open to, and interested in, considering any proposed legislation or rulemaking concerning such a regime. If there exists a desire to regulate in this area, we encourage careful consideration of the economic relationship between the inputs to, and outputs from, AI systems.

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?

Article I, Section 8, clause 8 of the US Constitution specifies the purpose of copyright: "to promote the progress of Science and the useful Arts." The basis for constitutional protection is intended to protect authors and publishers for a limited time from competing with copies or certain infringing derivatives. Obviously, the Framers were not aware of AI when the Constitution was drafted so it is a matter of judicial and administrative interpretation.

The Copyright Office has directly addressed the question of copyright in AI-generated material. The Office refused to register a claim for two-dimensional artwork described as "autonomously created by a computer algorithm running on a machine" and stated that that the work could not be registered because it was made "without any creative input or intervention from a human author," and that "statutory text, judicial precedent, and longstanding Copyright Office practice" all require human authorship as a condition of copyrightability. The Office's registration denial, as well as the supporting legal analysis, was affirmed in federal district court.

As noted in answer to Question No. 3, the *Zarya of the Dawn* provides an illustrative example of these principles in action.

There is flexibility and fluidity in the concept of "human authorship" such that artists who use AI may obtain copyright rights over their works, in whole or part, to the extent the human-contributed authorships is original and sufficiently creative.

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?

This is possible and the question can be, and for the moment should be, addressed through the application of the existing analytical framework for infringement of these rights, including

defenses against claims of infringement, such as fair use. Attempting to articulate in advance the circumstances under which these rights may be implicated is inadvisable, and the application of law to fact is best left to the factfinders, based on the evidentiary records before them.

We also note that the issues presented by AI infringement are analogous to human infringement. Conceptually, generative AIs use a similar process to that used by human authors. A human author learns her trade by absorbing a large amount of original source material, studying its structure and composition, and then practicing the skills used to produce them, and refining those skills based on feedback. The new original works produced by the artist inherently incorporate, on some level, the aesthetic sensibilities of the originals, sometimes unconsciously (e.g., "influenced by"), and sometimes consciously (e.g., homage). All of human endeavor is, to varying degrees, built upon what has come before, and the existing legal framework for infringement is capable of determining when an artist crosses the line to actionable infringement.

Likewise, an AI is trained in similar fashion, and the same legal framework we use to assess when a human author infringes may be applied to assess when an AI-generated work infringes.

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?

We believe substantial similarity may prove adequate in addressing claims of infringement based on generative AI outputs and that courts should be given an opportunity to apply this rule prior to formulating legislative interventions.

Substantial similarity examines the resemblance between forms and manifestation of expression, not mere ideas. A work generated using artificial intelligence can be evaluated under this standard in the same manner as other works because the analysis primarily considers the product as opposed to the manner of creation.

Unlawful appropriation of copyright has often been analyzed using two tests for substantial similarity: an extrinsic test to analyze the similarity of specific criteria between two works, and an intrinsic test to judge the similarity of the expressions based on the observations of a reasonable observer. *Sid & Marty Krofft TV Prods. v. McDonald's Corp.*, 562 F.2d 1157 (9th Cir. 1977). Copying also requires proof of access or actual copying, but proof of access or copying does not lower the threshold for proving substantial similarity. *Skidmore v. Led Zeppelin*, 952 F. 3d 1051, 1069 (9th Cir. 2020). Not all copying is prohibited, and it is a striking similarity between the protected elements of a work and the AI output that would give rise to infringement. See *Rentmeester v. Nike, Inc.*, 883 F. 3d 1111, 1117 (9th Cir. 2018).

The extrinsic test can be a question of law. It considers objective criteria like the type or category of work created, the subject matter, settings of the subject matter, and method of creation. Each of these elements can be answered simply by examining the model used in generation, the prompt, and the image or text output by the model. The intrinsic test is a factual inquiry. It remains unaffected by the method of creation because it examines the similarity of the expressions from the eye of the objective, reasonable observer. Here, the fact finder will

be comparing the AI output against the claimant's protected work to make an intrinsic judgment, not based on expert testimony or strict analysis of the objective elements of the works, but upon the similarity of the expressions wholistically.

Skidmore's clarification that access does not lower the threshold for finding substantial similarity marked the 9th Circuit's joining of a majority of circuits that reject the inverse ratio rule. This distinction helps answer the question posed here. Proof of copying and access poses difficult questions as to infringement by prompters who do not control the data upon which an AI model is trained. A majority denial of the inverse ratio rule separates substantial similarity from the murkier question of access/proof of copying. As such, the substantial similarity test, as it exists, is likely to be effective in infringement cases involving AI outputs.

The substantial similarity test need not be affected by a new method of creation. The extrinsic question of law may still be answered by examining the constituent parts of the AI output and the work allegedly infringed upon. The subject matter of each work, the methods of creation (prompt and AI generation against the method used by the claimant), the settings, and the categories or mediums of works created can all be compared through critical analysis of the AI output and the model's function. The intrinsic question regarding the similarity of expression remains one to be judged by a reasonable observer of the output and protected work.

Courts have "used the same term—'substantial similarity'—to describe both the degree of similarity relevant to proof of copying and the degree of similarity necessary to establish unlawful appropriation. The term means different things in those two contexts. To prove copying, the similarities between the two works need not be extensive, and they need not involve protected elements of the plaintiff's work. They just need to be similarities one would not expect to arise if the two works had been created independently. To prove unlawful appropriation, on the other hand, the similarities between the two works must be 'substantial' and they must involve protected elements of the plaintiff's work." *Rentmeester v. Nike, Inc.*, 883 F. 3d 1111, 1117 (9th Cir. 2018).

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?

The question essentially asks how infringement can be proven in the absence of direct evidence of copying. However, the absence of direct evidence of copying is not unique to AI-generated works, and there already exists a rich body of law for establishing copying under such circumstances.

For example, substantial similarity does not require a plaintiff to prove to a certainty that copying has occurred. In most cases, factfinders have no objective evidence of the process by which the challenged work was developed, and no direct evidence of copying, but rather must rely on circumstantial evidence and inference to find copying. *See, e.g., Murray Hill Publications, Inc. v. Twentieth Century Fox Film Corp.*, 361 F.3d 312, 316 (6th Cir. 2004) ("[W]here there is no direct evidence of copying, a plaintiff may establish 'an inference of copying by showing (1) access to the allegedly-infringed work by the defendant(s) and (2) a substantial similarity between the two works at issue.").

Accordingly, a plaintiff suing an AI operator would not necessarily need to conduct a forensic audit of the AI system to prove to a certainty that specific elements of an output were copied from specific elements of an input, but rather only access to the original and substantial similarity. Recently filed copyright cases may provide further guidance on this issue, and existing law has historically proven adaptable to situations where there is no clear evidence of copying. *See, e.g., Bright Tunes Music Corp. v. Harrisongs Music, Ltd.*, 420 F. Supp. 177 (S.D.N.Y. 1976) (finding "subconscious plagiarism" in the absence of any clear evidence defendant had previously encountered the allegedly infringed work).

There are other practical litigation considerations that could assist plaintiffs against AI operators. Most courts now require defendants put on notice of impending litigation to put a "litigation hold" on possibly relevant information. Any AI operator would be sophisticated and (a) be able to institute a hold and (b) be aware of this legal requirement. If an author believes her work has been "copied" by AI, she could promptly put the AI operator on notice to create an obligation on them to retain all records (including a list of all materials ingested with regard to the subject matter at issue). Courts have become more willing to enforce a lack of good record-keeping against defendants, by imposing presumptions of wrongdoing. Current discovery rules and tools are likely sufficient to plaintiffs to prove infringement in appropriate cases.

We further note that markets have historically stepped in to address or mitigate difficult issues in the use of creative works. For example, YouTube has implemented the Content ID system as a supplement to DMCA takedown requests. Markets are rapidly emerging and evolving for both AI services and secondary services that may address some of the legal issues raised in the RFC, such as tools that purport to determine whether a given image is AI-generated, and whether a given image has been included in AI training databases.

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?

Existing judicial doctrines concerning direct and secondary liability should be applied to make this determination on a case-by-case basis, and in light of the evidentiary record of each case. At this early stage, we believe it is inadvisable to attempt to anticipate and enumerate the circumstances under which various types of liability should attach before specific fact patterns emerge and specific evidentiary records are fully developed.

QUESTION 25.1. Do "open-source" AI models raise unique considerations with respect to infringement based on their outputs?

Based on how AIs are currently trained and constructed, we do not believe such models raise unique considerations concerning infringement based on their outputs. Open source is a software distribution model and should not impact the substantive legal analysis as to whether an AI model infringes.

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?

As with the other issues, the analysis is necessarily fact-specific, and the application of law to this circumstance will depend on what, specifically, is done with the copyright management information. We again believe existing law may be able to adequately address these issues in the context of AI. This is currently being tested in several pending cases that include claims under Section 1202(b). In at least one such case, *Doe v. GitHub, Inc.*, two of three asserted claims passed muster at the pleading stage. As with all CMI claims, plaintiffs must establish both that a defendant knew that CMI had been removed or altered and that this would likely result in copyright infringement. It should also be noted that, in *Doe v. Github*, common license requirements in open-source projects, such as passing along attribution, copyright notices, and licensing terms, were allegedly violated by source-code generating AI models, which were not programmed to treat those elements as legally essential, and which reproduced licensed code without them.

Another potentially illustrative case is *Stevens v. Corelogic*, wherein the defense won on summary judgment as the plaintiff could make an affirmative showing that the "defendant was aware of the probable future impact of its actions" in the software context.

QUESTION 30. What legal rights, if any, currently apply to AI-generated material that features the name or likeness, including vocal likeness, of a particular person?

Existing laws provide substantial protection against the unlawful use of the name, likeness, or voice of particular persons. Examples include state right of publicity laws that prohibit the use of a person's name, likeness, or voice in commercial endorsements or commercial products without permission. Other examples include Lanham Act §43(a) prohibiting false endorsements, state laws addressing "deepfakes" and "revenge porn," as well as collective bargaining agreements that address AI generated performances by actors and output affecting writers. Many such laws contain express exemptions protecting expressive works and free speech. In the absence of an express exemption, these laws should continue to be interpreted in a manner that does not conflict with or chill free speech and expression.

In addition, currently two states, Louisiana (La. Rev. Stat. §§470.1-470.6) and New York (New York Civil Rights Law §50(f)) provide very limited protection against the use of "digital replicas" of certain performers under limited circumstances, including in expressive works such as films. Both statutes appropriately contain numerous exceptions and limitations intended to respect free speech and expression.

QUESTION 31. Should Congress establish a new federal right, similar to state law rights of publicity, that would apply to AI-generated material? If so, should it preempt state laws or set a ceiling or floor for state law protections? What should be the contours of such a right?
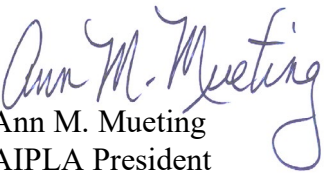
We are not yet convinced that the time has come to propose specific legislation concerning the regulation and use of AI, and new forms of intellectual property or other proprietary rights to AI-generated materials.

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?

Generally speaking, there is no copyright protection for a "style" *per se*. Accordingly, we do not believe the law should prohibit creations "in the style of" a given artist by an AI, either. To be clear, this does not mean that an AI-generated output in the style of a given artist can never infringe a copyright; rather, it means that infringement should be assessed using the existing framework for copyright infringement, under which the replication of a signature style is, standing alone, not infringement. Existing laws protecting against copyright infringement and passing off, as well as existing laws imposing liability for secondary liability where there is direct infringement, should be sufficient to protect against unlawful imitations without imposing overbroad protections on "style."

AIPLA appreciates the opportunity to provide feedback to the Copyright Office on the Notice and we look forward to responding to any questions you may have.

Respectfully submitted,

Ann M. Mueting
AIPLA President