UNIVERSITY of HAWAI'I® MĀNOA

William S. Richardson School of Law

October 30, 2023

United States Copyright Office
Library of Congress
101 Independence Ave. S.E.
Washington, D.C. 20559-6000

RE: Comments by University of Hawai'i William S. Richardson School of Law Intellectual Property Law Students in Response to U.S. Copyright Office Artificial Intelligence and Copyright Notice of Inquiry and Request for Comments, 88 Fed. Reg. 59,942 (Aug. 30, 2023).

To Whom It May Concern:

Thank you for the opportunity to provided comments in support of the U.S. Copyright Office's Artificial Intelligence Initiative. We are a group of fifteen law students at the William S. Richardson School of Law, University of Hawai'i at Mānoa in Honolulu. We are classmates in the Intellectual Property Law course taught by Professor Emile Loza de Siles and contribute these comments as our class project, given the great importance of and our strong interest in the topic. We have considered and here provide comments in response to two questions posed in the above-referenced Notice, as follows:

Question 4: Are there 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?

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?

Respectfully submitted:

Ikaika Agard, JD 2025
Makela Aluesi, MSW, JD 2026
Lia Sabina Maria Martinez Cook, B.S. Biochemistry, JD 2025
Thomas Davis, JD 2024
Maile Edwards, JD 2025
Christian Ferrer, B.S., M.A. Economics, M. Math Finance, JD 2024
Michellei Fisher, B.S. Biochemistry, M.S. Biomedical Sciences, JD 2025
Jenna Hanohano, B.A. US History, JD 2025
Jana Kakugawa, B.S. Biology, JD 2026
Tina Loos, B.A. Information & Computer Sciences and Political Science, JD 2025
Mariah Morgado, B.S. Criminal Justice, JD 2025
Constancio Paranal III, JD, LLM 2024
Sean Quinn, B.S. Physics, M.S. Mechanical Engineering, JD 2024
Rev. Josiah Sewell, M.S. Zoology, B.A. Journalism, B.S. Biology, JD 2024
David Shinsato, M.S. Chemistry, JD 2025

Question 4: Varied International Approaches to AI and Copyright Highlight the Importance of the International Harmonization of Laws.

The artificial intelligence (AI) revolution is here, and the market is expected to reach over $400 billion by 2027, with a predicted annual growth rate of more than 37% from 2023 to 2030.[1] The democratization of AI technology means there is now a new category of AI users who are not computer scientists but who rely on AI to perform a range of tasks, including responding to people via text or email, answering financial questions, prepping for job interviews, and summarizing complex or lengthy copy.[2] Earlier this year, ChatGPT, a generative AI that can produce essays, articles, poetry, and even code, set the record for fastest-growing user base ever, reaching 100 million users in just three months.[3] The rapid evolution of AI technologies is already shaping our future, and the number of AI generated works will inevitably increase with widespread adoption.

The rise of AI-generated content poses significant challenges to the traditional notions of authorship and ownership within the framework of copyright law. The U.S. Copyright Office has consistently emphasized the need for a human author with the capacity for intellectual, creative, or artistic labor in order to attach copyright protection.[4] Some countries like India and China have so far adopted a similar stance, centering copyright protection of AI-generated content around human creation.[5] Most EU countries have adopted a wait-and-see approach, observing the emerging legal frameworks and their resulting impacts on copyright.[6] Others, like

[1] Kathy Haan*, 24 Top AI Statistics and Trends in 2023*, FORBES (Apr. 25, 2023), https://www.forbes.com/advisor/business/ai-statistics/.

[2] *See id*.

[3] Krystal Hu, *ChatGPT Sets Record for Fastest-growing User Base – Analyst Note*, REUTERS (Feb. 2, 2023), https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/.

[4] *See Thaler v. Perlmutter*, No. CV 22-1564 (BAH), 2023 WL 5333236 (D.D.C. Aug. 18, 2023).

[5] *See* Sukanya Sarkar, *Exclusive: India Recognises AI as Co-Author of Copyrighted Artwork, ManaginIP* (Oct. 27, 2023, 12:01 PM), https://www.managingip.com/article/2a5czmpwixyj23wyqct1c/exclusive-india-recognises-ai-as-co-author-of-copyrighted-artwork; Haochen Sun, *Redesigning Copyright Protection in the Era of Artificial Intelligence*, 107 IOWA L. REV. 1213 (2022) (quoting Decision of the State Council on Amending the Regulations for the Implementation of the Copyright Law of the People's Republic of China (promulgated by Decree No. 359 of the State Council of China, Aug. 2, 2002) (amendment promulgated by Decree No. 633 of the State Council of China, (Jan. 30, 2013, effective Mar. 1, 2013), art. 2, https://wipolex-res.wipo.int/edocs/lexdocs/laws/en/cn/cn213en.pdf [https://perma.cc/4SLL-EG9A])).

[6] *See* European Union Intellectual Property Office, *FAQs on Copyright for Consumers* (Oct. 27, 2023 2:30 PM), https://euipo.europa.eu.

Japan, however, have refused copyright protection for AI-generated works altogether.[7] Because there is very little established copyright law related to AI-generated content, policies created today will have significant impacts on the future of copyright law as a whole and should be carefully considered to ensure both the rights of copyright holders and the encouragement of creative works.

The India Copyright Office has recognized copyright protection for at least one AI-assisted work, although the office's overall position is unclear. In 2020, the Office rejected an application listing an AI as the sole author but went on to copyright an AI technology as a co-author of a painting titled "Suryast" in 2021.[8] However, later that year the Office issued a withdrawal notice to the copyright owner that may imply an office oversight or foreshadow an upcoming change of policy.[9] The human co-author has since challenged the office's authority to withdraw registration after it had been granted since the act had no provision for revocation; thus, the current status of this copyright is uncertain.[10]

Similarly, in China, a 2019 decision by the Shenzen National District People's Court determined that an AI system can hold copyright for a finance news article generated by Dreamwriter, an AI technology offered by Tencent.[11] The Court acknowledged that Dreamwriter did not produce the article without the creative human input of Tencent team members.[12] A two-step test was used to determine the originality of the news article. Interestingly, the Court found that the content contained a degree of creativity as generated by the Dreamwriter AI, but also that Tencent's team input constituted enough "intellectual activity" to allow for copyright protection.[13] The ruling suggests that, although human authorship may not be an absolute requirement for copyright protection in China, as it is in the United States, there is a degree to which human intellect and input play a role in the analysis of copyright decisions.

In contrast, Japan generally does not allow copyright of AI-generated products or allow the human who enters prompts into the generative AI to become a copyright owner.[14] The copyright laws of Japan view the simple instructions given to AI as insufficient to constitute creative intent or contribution.[15] In these instances, there is no copyright of the AI-generated product, and such content is thus free for anyone to use. However, in some circumstances, Japan

---

[7] *See* Shinnosuke Fukuoka, Tomonobu Murata & Atsuki Mizuguchi, *Legal Issues in Generative AI under Japanese Law - Copyright,* LEXOLOGY (2023), https://www.lexology.com/library/detail.aspx?g=68d490a1-3021-4040-afdd-90ae8fa69337.

[8] *See id.*

[9] *See id.*

[10] *See id.*

[11] *See* Sun, *supra* note 5.

[12] *See id*.

[13] *See id*.

[14] *See* Fukuoka, Murata & Mizuguchi, *supra* note 7.

[15] *See id.*

does allow copyright on a case-by-case basis, dependent on how much human input played in the creative intent and contribution.[16]

Regarding the use of copyrighted works in training datasets, Japan provides an exception under Article 30-4 of their Copyright Act that allows the use of copyrighted work in AI training datasets without the copyright owner's permission, unless exceptional requirements are met.[17] This is because training datasets are not considered an "enjoyment of the ideas or sentiments expressed in the work."[18] Japan's AI copyright laws are deemed some of the more "relaxed" copyright acts in the world, but this approach is used largely due to the complexity in regulating the large quantities of data used for AI development.[19]

In South Korea, the Ministry of Science and Technology announced in June 2021 a "strategy to realize artificial intelligence trustworthy for everyone" that outlined strategies and goals regarding AI with a focus on technology, system, and ethics.[20] With regards to the technology focus, the Ministry sought to "create an environment to realize a reliable artificial intelligence" by creating a systematic process to ensure trust at each stage of AI development, support private sector actors, and develop source technologies for a trustworthy AI.[21] The system focus then sought to "lay the foundation for safe use of artificial intelligence" by ensuring the trustworthiness of AI learning data, promoting security for high-risk AI, assessing the influence of AI, and improving regulations to increase trust in AI.[22] Finally, the Ministry also sought to "spread AI ethics across society" by providing education programs on AI ethics, creating and distributing checklists for stakeholder groups, and operating a platform for ethic policies.[23] This Ministry created this strategic plan to ensure the development of an AI that can be trusted, while ensuring that the focus of these strategies is centered around people.[24]

Furthermore, in early 2023, legislation was introduced to the National Assembly of the Republic of Korea to enact an Act on Promotion of AI Industry and Framework for Establishing

---

[16] *See id.*

[17] *See id.*

[18] *See id*.

[19] *See* Andrew Ng, *Training Data Free-For-All Japan's AI Data Laws Explained*, BATCH (2023) https://www.deeplearning.ai/the-batch/japan-ai-data-laws-explained/.

[20] Ministry of Science & ICT, *MIST Announced Strategy to Realize Trustworthy Artificial Intelligence* (updated, but 2021), https://www.msit.go.kr/eng/bbs/view.do?sCode=eng&mId=4&mPid=2&pageIndex=&bbsSeqNo=42&nttSeqNo=509&searchOpt=ALL&searchTxt= (last visited Oct. 30, 2023).

[21] *Id*.

[22] *Id.*

[23] *Id*.

[24] *See id*.

Trustworthy AI, or "AI Act."[25] Although not the first bill relating to AI, this bill is focused and expands upon the previous framework announced by the Ministry of Science and Technology, as well as incorporates previously proposed legislation.[26] The South Korean AI Act is designed to support the AI and related technological industries, while maintaining a focus on protecting the human AI-users by setting up a system to ensure trustworthiness of the AI systems.[27]

In conclusion, the current global legal landscape of AI and copyright reveals a wide range of approaches to copyrightability with most centering human authorship as a core tenet in order to attach protection. However, until a universal framework is adopted, copyright holders will have to contend with each country's individual approach to AI-generated works in order to determine their own copyright protection and possible liability for infringement. International harmonization of copyright laws as they relate to AI-generated works enhances transparency of the copyright process and incentivizes creators to produce original works, ultimately benefiting the public.

~Comments in Response to Question 22 Begin on the Next Page.~

[25] Ji Eun Nam & Taeyoung Roh, *South Korea: Legislation on Artificial Intelligence to Make Significant Progress*, KIM & CHANG (2023), https://www.kimchang.com/en/insights/detail.kc?sch_section=4&idx=26935

[26] *See id*.

[27] *See id.*

Question 22: AI Technologies Potentially Infringe Copyright Owners' Exclusive Rights.

The ultimate goal of copyright is to expand public knowledge and understanding by giving potential creators exclusive control over copying and use of their works, providing financial incentive to create intellectually enriching works for public consumption.[28] The Supreme Court has said, "The sole interest of the United States and the primary object in conferring the monopoly lie in the general benefits derived by the public from the labors of authors."[29] Without copyright protection for AI-generated works, there is little incentive for potential creators to use this technology and create new works which may greatly benefit the public.

Here, we argue that AI technologies potentially infringe on the copyright owners exclusive right to reproduction and derivative works when copyrighted materials are used in training datasets. We next use AI-generated music as an instrument to explore potential infringement. We then propose that copyright holders should receive remuneration for use of copyrighted materials in AI training datasets, through an appropriate licensing framework and share of revenues. We argue that AI-generated works made using licensed or otherwise non-infringing training datasets should be copyrightable to encourage creation of useful works. Finally, we examine AI-generated works under the fair use defense.

The Use of Copyrighted Works in AI Training Datasets Potentially Infringe on the Exclusive Right to Reproduction and Derivative Works.

Two standard approaches to AI are generative adversarial networks ("GANs") and autoencoders.[30] GANs use two learning models: one which generates a random output based on the user's command (the "generator") and another that uses a pre-programed dataset to critique the generator's output (the "discriminator").[31]

Autoencoder uses a slightly different method, ingesting input data, compressing the data into coded segments, then reconstructing the input data through a decoding process.[32] The machine then compares the discrepancies between the input and output layers to extract only the input's most relevant elements.[33] The process relies on large training data sets to then be able to create unrecognizable, seemingly unique output.[34]

---

[28] *See*, *e.g.*, *Twentieth Century Music Corp. v. Aiken*, 422 U.S. 151, 156 (1975); *Sony Corp. of Am. v. Universal City Studios, Inc.*, 464 U.S. 417, 429 (1984).

[29] *Twentieth Century Music Corp.*, 422 U.S at 156.

[30] *See* Eric Sunray, *Sounds of Science: Copyright Infringement in AI Music Generator Outputs*, 29 CATH. U. J. L. & TECH. 185, 190 (2021).

[31] *Id*.

[32] *See id.* at 191.

[33] *See id*.

[34] *See id*.

Both of these AI methods rely on large training datasets and materials to develop their AI model and improve generated content. Reproduction of substantial portions of copyrighted works, even for non-expressive purposes, implicates the exclusive reproduction right and, absent an applicable exception such as fair use, is an act of copyright infringement.[35] The infringement doctrine contains two distinct aspects, proving literal copying and improper appropriation, which involves assessing the substantiality of copying that occurred from the protected expression.[36]

Accordingly, the specific content of the training material and the resulting generated work is an important consideration. The more the training datasets or resulting generated works are directly copied or substantially similar to the expression of a pre-existing copyrighted work, the easier it will be for the copyright holder to prove substantial similarity and thus, liability for direct copyright infringement.[37]

However, even if the generated code or work is meaningfully different from the pre-existing work, a copyright holder who can prove its work was used to train the AI generator could still argue that any model output is "based upon" its preexisting work and is therefore a derivative work.[38] Derivative works are based on one or more preexisting works and cover a wide spectrum of expressions, such as a translation, musical arrangement, abridgement, condensation, or any other form in which a work may be recast, transformed, or adapted.[39]

Professor Gervais has opined that the derivative work right may be infringed even if the derivative production would not qualify for protection as a work.[40] Derivative works require a modicum of creativity and modification of one or more preexisting works.[41] However, if one considers that, to infringe the derivative work right, a reproduction must be an original, creative work, then an AI technology that cannot produce originality cannot infringe this right.

In conclusion, AI models use machine-learning algorithms to analyze existing data to generate outputs; where those outputs are sufficiently similar to the copyrighted works used to train the model, copyright infringement may occur. Since AI models can quickly generate limitless versions of creative works, the lines between direct reproduction and derivative work will likely become increasingly blurred. Therefore, the general public needs clear guidance on how works generated by AI could potentially infringe on the exclusive rights held by copyright holders.

---

[35] *See* 17 U.S.C. §101.

[36] *See id*.

[37] *See id*.

[38] *See id*.

[39] *See id*.

[40] *See* Daniel J. Gervais, *AI Derivatives: The Application to the Derivative Work Right to Literary and Artistic Productions of AI Machines*, 53 SETON HALL L. REV. (2022), https://scholarship.law.vanderbilt.edu/faculty-publications/1263.

[41] *See* 17 U.S.C. § 101.

AI-generated Musical Work May Serve as an Instrument to Explore Potential Infringement.

AI-generated music is a prime example of the contradictory law and policy regarding generative AI and copyrightability. Examining these issues through the lens of musical composition helps us explore the current jurisdictional split in the treatment of AI generated content. To begin this inquiry, we must first examine how AI technologies generate music.

Once a model has been trained and effectively learns the nuances of individual musical elements, it rearranges those elements into a new sample.[42] This creates a "novel" sound that is woven together solely through sampling the audio of other's previous works.[43] Although an ordinary person may perceive a generated sound to novel, the sounds generated are merely reproduced through the sampling of many sound recordings.[44]

Understanding how AI creates its "novel" sounds demonstrates that AI generation is literal reproduction, even if not apparent on its face.[45] The question now becomes whether AI generated music is infringing upon the original artists copyright rights. If the audio used by AI is protected by copyright, then any product of infringement is therefore uncopyrightable, and AI users would be unable to obtain copyright for any material generated by AI. On the other hand, if the audio clips used by AI are so minimal that the usage of the audio is not an infringement of the copyrights of audio owners, then it may be that AI users could potentially claim rights of reproduction. The law ultimately finds itself in a jurisdictional split on this matter, leading to some copyright holders having more rights than others based purely on their locale.[46]

In *Bridgeport Music, Inc. v. Dimension Films*, the Sixth Circuit Court of Appeals held that, where a defendant does not dispute having digitally sampled a sound recording, there is no need for an improper appropriation or *de minimis* inquiry to find infringement.[47] The Court reasoned that a digital sampling is fixed in a medium and thus can be viewed akin to taking because the digital sampling is taken directly from the audio clip itself.[48] Based upon this logic, any sound reproduction, even if very small, used without a license or permission is considered an infringement of copyright and implicates the liability of AI training dataset users.[49]

However, the Ninth Circuit Court of Appeals concluded differently on this question. In *VMG Salsoul, LLC v. Ciccone*, the Court did not interpret copyright law to uniquely protect

[42] *See* Sunray, *supra* note 30, at 190.

[43] *See id*.

[44] *See id*.

[45] *See id*.

[46] *See id.* at 196.

[47] 410 F.3d 792, 797 (6th Cir. 2005).

[48] *See id.* at 798.

[49] *See id.*

sound recording in comparison to other types of copyrightable material.[50] Unlike as in *Bridgeport Music*, the Court here reasoned that a copier does not benefit from the original artists expression if the general public is unable to recognize the usage of a sound recording.[51] In this case, the Court chose to adhere to the principle of the de minimis rules used in other cases involving many types of copyrightable material.[52] Notably, this jurisdictional split divides two of the United States' top music producing cities, Nashville and Los Angeles.[53]

When discussing how the law ought to view generative AI and copyright, it is important to consider the purpose of these laws. Copyright law was established to protect the creative expression and ideas of society in a way that encourages people to be creative while also encouraging the use and development of such ideas. Looking at the jurisdictional split of the Sixth and Ninth Circuit Courts of Appeal, it seems apparent that each Court is trying to protect these two values that are promoted by copyright law. The Sixth Circuit is attempting to protect the interests of the creators of musical works while the Ninth Circuit is attempting to progress the development of music through AI.[54] As these copyright issues are still novel, the decisions made by courts today will have massive implications on copyright law in the future and therefore, we ought to be cautious in the decisions that are made.

Although the benefits of AI are significant, following the Ninth Circuit Court of Appeals leaves many copyright holders' works unprotected when used in AI training datasets and vulnerable to having their work used without any compensation for their original expression. As the world becomes more centered around technology and intangible property, it is important to create law that reflects the values of traditional copyright: to protect creators and encourage the creation of works for the public benefit. To do this, we should not strip copyright holders of their rights, but instead hold those who use and create generative AI accountable and transparent for using copyrighted works.

Copyright Holders Should Receive Remuneration for Works Used in AI Training Datasets or Materials.

Generative AI models typically use text and data mining (TDM) to collect vast amounts of digitized material for "ingestion" into the model[55] that implicates the right of reproduction and absent an applicable exception, constitutes copyright infringement. [56]

---

[50] 824 F.3d 871, 880-81 (9th Cir. 2016)

[51] *See VMG Salsoul,* 824 F.3d at 881; *Bridgeport*, 410 F.3d at 797.

[52] *See VMG Salsoul,* 824 F.3d at 881.

[53] *See* Sunray, *supra* note 30, at 204.

[54] *See id.*

[55] Jonathan Band, *Copyright Implications Of The Relationship Between Generative Artificial Intelligence and Text and Data Mining,* INFOJUSTICE (Oct. 27, 2023), https://infojustice.org/archives/45509.

[56] *See id*.

Although the benefits of AI technologies are apparent, it also seems reasonable to assume that they are not trained on data that were stolen or otherwise used without permission from the copyright holder. Regardless of whether ingestion is determined to be an infringement, there is a separate issue of whether authors of such ingested works should be remunerated for these types of uses. [57]

AI technologies that use exclusively licensed training datasets are on the rise to avoid these exact types of issues.[58] Getty Images, an American visual media company, has recently launched a new service, Generative AI by Getty Images, that has been trained using their own library of copyrighted images.[59] It promises users "full indemnification from commercial use" in order to avoid the intellectual property risks that have generally made businesses wary of using AI tools.[60]

Advocates for authors have suggested that when copyrighted works are used in AI training datasets or engaged in other activities that result in remuneration, the authors should be entitled to a share of the revenues generated by the AI.[61] Some private contracts now include text and data mining ("TDM") terms for copyrighted materials in their terms.[62] These terms expressly set a licensing fee for for-profit entities or at no cost for researchers and public organizations to use licensed content that is machine-readable and searchable.[63]

However, there are many more cases where an agreement over compensation or other terms cannot be reached regarding licensing of materials for AI ingestion. Asymmetry in bargaining power may lead to forced, rather than negotiated, terms, and without a consistent framework, the result may be an inefficient and fragmented set of solutions across the market.

We support a statutory license requirement, specifically for TDM or data ingestion use by generative AI owners to bulk-access copyrighted work.[64] Although such a licensing framework

---

[57] *See id.*

[58] *See* Emillia David, *Getty Made an AI Generator that Only Trained on Its Licensed Images,* VERGE (Sept. 25, 2023), https://www.theverge.com/2023/9/25/23884679/getty-ai-generative-image-platform-launch.

[59] Matt O'Brien, *Photo Giant Getty Took a Leading AI Image-maker to Court. Now It's also Embracing the Technology*, AP (Sept. 25, 2023), https://apnews.com/article/getty-images-artificial-intelligence-ai-image-generator-stable-diffusion-a98eeaaeb2bf13c5e8874ceb6a8ce196.

[60] *See id*.

[61] *See id*.

[62] *See id*.

[63] *See id*.

[64] *Accord* Christophe Geiger & Vincenzo Iaia. *The Forgotten Creator: Towards a Statutory Remuneration Right for Machine Learning of Generative AI*, COMP. L. & SEC. REV. – (forthcoming 2023), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4594873 (last visited Oct. 28, 2023)

limits a copyright owner's right to exclude,[65] this limitation is offset somewhat by affording that owner with a new right to remuneration.

This approach offers several advantages. First, it avoids the obvious practical challenges associated with mass ingestion, involving potentially thousands of individualized infringement claims of copyrighted works, coupled with the proportionate effort of tracking and enforcing violations. Second, this removes the risk of potentially fragmented or incomplete solutions that may develop solely from the self-regulation of the market players.[66] Third, it is a compromise to the tension between copyright holders and generative AI companies created by the new exploitation of creative works via TDM/ingestion. [67] Fourth, this approach ensures an attractive environment for the continued development of AI without marginalizing the role of copyright authors, particularly human authors. [68] Licensing fees should be collected by the US Copyright Office, and the aggregate "revenue" collected would be remunerated back to the copyright owners. [69]

The proposed licensing framework, coupled with other market solutions, such as the collective rights organizations bargaining on behalf of rights holders, blanket licensing for copyright owners' entire portfolio of work, and the potential use of technologies such as blockchain would help to flesh out this solution.[70]

In reality, implementation would require granular traceability of copyright licenses used in large training datasets. We propose that every copyrighted work embed some form of a "UCN," or Universal Copyright Number, for each copyrighted work analogous to the Universal Product Code ("UPC") used by the consumer goods industry. The UPC is what manufacturers of consumer goods use to identify, track, and monitor their goods through the supply chain.[71]

Blockchain could provide the technology that could help creators track their copyrighted work online, for example, by efficiently facilitating copyright enforcement, remuneration, and

---

[65] *Cf.* limited performance right described in 17 U.S.C. §110.

[66] *See id*.

[67] *See id*.

[68] *See id*.

[69] *See id*.

[70] Camille Brown, *Coded Copyright?: How Copyright Enforcement, Remuneration, and Verification Terms in Blockchain-Enhanced Contract Models for Online Art Sales Compare to Their Traditional Counterparts*, 31 S. CAL. INTERDISC. L.J. 617, 626 (2022).

[71] *Guide to UPC codes*, GS1 US, https://www.gs1us.org/upcs-barcodes-prefixes/guide-to-upcs#:~:text=For%20most%20retailers%2C%20having%20a,can%20help%20you%20track%20shipments (last visited Oct. 28, 2023).

authentication practices. [72] A blockchain ledger's tamper-resistant "transaction" data can be used to create an immutable record of origin and authorship for a work helping to authenticate its "consumption chain" into downstream TDM/ingestion activities of generative AI. [73]

In conclusion, copyright holders whose work is used by AI technologies should receive remuneration and the Copyright Office should consider an explicit licensing framework or shared revenue program to achieve remuneration.

AI-generated Works Created Using Licensed or Otherwise Non-Infringing Training Datasets Should be Copyrightable to Encourage Creation of Useful Works.

There is a strong argument that AI should be viewed as a tool rather than an author. AI-generated work often includes a system that processes large amounts of data, but where the creative output by the AI is directly related to the level of creativity in structuring the prompts. If two different users prompt an AI to create a "red butterfly" and the AI outputs the same butterfly, should the prompter garner any credit? The central question is whether or not there is sufficient creative expression in the prompts that may qualify that individual for copyright protection of the resulting works, assuming the training datasets and materials are appropriately licensed or otherwise non-infringing.

The U.S. Copyright Office has already recognized the copyrightability of the output produced by a breakthrough tool at the time, the camera.[74] In its analysis, the Court reasoned that a human creator conceived and designed the image, then used the camera as a tool to simply capture the creative expression.[75] In *Feist Publications, Inc. v. Rural Telephone Service Co.*, the U.S. Supreme Court established that copyright protection is not granted solely based on the amount of effort or labor invested in compiling data or information; instead, a work must be original and should display a minimal level of creativity, independent of the amount of labor involved.[76]

Simply because a human did not paint an AI-generated image does not necessarily preclude copyright protection if they sculpted the tool's output through one or more creative prompts.[77] Declining to extend copyright protections to AI-generated works disincentivizes creators and thus is a disservice to the expansion of public knowledge, the central purpose of copyright law.[78] Although AI-generated works are not currently eligible for copyright protection in the U.S., they should be subjected to the same legal analyses as other creative works to determine whether the content is derivative, infringes copyright, and able to be availed of the fair

[72] *See* Camille Brown, *supra*, at 626.

[73] *See id*.

[74] *See Burrow-Giles Lithographic Co. v. Sarony*, 111 U.S. 53, 60 (1884).

[75] *See id*.

[76] *See* 499 U.S. 340, 364 (1991) ("[C]opyright rewards originality, not effort.").

[77] *See Burrow-Giles Lithographic*, 111 U.S. at 60.

[78] *See Twentieth Century Music Corp. v. Aiken*, 422 U.S. 151, 156 (1975).

use defense.

AI-generated Works Under the Fair Use Defense.

As discussed, AI-generated content may be considered derivative works when it builds upon or transforms existing content and has at least a "modicum of creativity".[79] AI creates a disconnect between the artist's creative vision and the ultimate expression of that vision through a black-box instrumentality. We cannot see the brushstroke or the chisel strike, and apart from evaluating the generative prompt given by the user, we cannot evaluate the "modicum of creativity" that goes into the generative work.

Although *intent* is not relevant to copyright infringement liability, *purpose* is relevant, and it is the first factor in the fair use analysis.[80] Will the AI be able to successfully defend against an infringement claim by stating that its purpose for creating the work was for criticism, comment, teaching, or research? How about if the intent was to create a parody?

The proposed rules for AI should incorporate all four factors the Court considers under the fair use analysis, including the purpose and character of the use, whether the use is of a commercial nature or is for nonprofit educational purposes; the nature of the copyrighted work; the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and the effect of the use upon the potential market for or value of the copyrighted work.[81]

The controlling question regarding fair use is often whether the work is sufficiently "transformative" to avoid infringement as an unlawful derivative work or falls under another "fair use" exception such as limited educational use.[82] Additionally, because copyright is a commercial doctrine whose purpose is to encourage creativity by enabling authors to earn money from their creations, courts often consider the effect of the proposed fair use on the potential market value for the copyrighted work an important factor in a fair use assessment.[83]

These issues were central to the analysis in *Authors Guild v. Google, Inc.,* where a copyright infringement suit was brought by a professional advocacy organization for book authors and publishers, when Google began digitally scanning large quantities of books to reproduce in search results.[84] Google argued that their reproduction of small portions of books was for the express purpose of enabling a search for the identification of books containing a key term of interest to the searcher, thus was not a meaningful substitute for the original work.[85] The Court held that copying small portions of a book to use in search results was a highly

[79] *Feist,* 499 U.S. at 340.

[80] *See* 17 U.S.C. § 501.

[81] *See* 17 U.S.C. § 107.

[82] *Id*.

[83] *See id*.

[84] *See* 804 F.3d 202 (2d Cir. 2015).

[85] *See id* at 216.

transformative purpose that supported a finding of fair use.[86] The Court also reasoned that under the fourth fair use factor regarding potential market impact, the book snippets Google reproduced were unlikely to compete with the market for the full original works that the searcher would not need to piece together.[87]

AI training datasets also take copyrighted works and reproduce small portions of the content in their generated output. However, in contrast to *Google*, AI-generated works may have substantial similarities to the original expression that does not require a user to string multiple outputs together.[88] This increases the likelihood that they may be considered a meaningful substitute for the original work.[89] Further, AI-generated content will more than likely compete directly in the same markets as past and present works produced by the copyright holders whose creative works were used in training datasets.[90]

Direct market competition and dilution by AI generated works has already been demonstrated and is the central controversy in several ongoing copyright infringement claims by prominent authors like John Grisham, Jonathan Franzen, and Elin Hilderbrand.[91] Additionally, this year Amazon was so overwhelmed by self-publishing submissions that it enacted a limit to publishing "only" three books per day.[92]

In conclusion, the number of lawsuits surrounding issues of copyright and AI will only increase as AI technologies become more widespread. Policies created today will have a significant effect on the copyrightability and potential infringement of AI-generated works and should be carefully considered to both protect existing copyright holders and encourage production of creative works for the public good.

---

[86] *See id* at 218.

[87] *See id* at 225.

[88] *See id.*

[89] *See id.*

[90] *See id.*

[91] *See*, *e.g.*, Alexandra Alter et al., *Franzen, Grisham and Other Prominent Authors Sue OpenAI*, N.Y. TIMES (Sept. 20, 2023), https://www.nytimes.com/2023/09/20/books/authors-openai-lawsuit-chatgpt-copyright.html; Alex Reisner, *Revealed: The Authors Whose Pirated Books are Powering Generative AI*, ATLANTIC (Aug. 19, 2023), https://www.theatlantic.com/technology/archive/2023/08/books3-ai-meta-llama-pirated-books/675063/.

[92] Ella Creamer, *Amazon Restricts Authors from Self-Publishing More Than Three Books a Day After AI Concerns*, GUARDIAN (Sept. 20, 2023), https://www.theguardian.com/books/2023/sep/20/amazon-restricts-authors-from-self-publishing-more-than-three-books-a-day-after-ai-concerns.