INTELLECTUAL PROPERTY
MEETS ARTIFICIAL INTELLIGENCE

BY
DR. KIRTI GUPTA

Vice President, Cornerstone Research & Senior Advisor, CSIS. The views in this paper are those of the author alone and not necessarily of any of the professional affiliations.

TechREG CHRONICLE
SEPTEMBER 2024

INTELLECTUAL PROPERTY MEETS ARTIFICIAL INTELLIGENCE
By Dr. Kirti Gupta

As artificial intelligence ("AI") technologies advance, their integration into creative and inventive processes raises critical questions regarding the intellectual property ("IP") framework. This paper summarizes the evolving dynamics between AI and IP, focusing on the patentability of AI-assisted inventions, copyrightability of AI-generated works, and potential copyright infringement of content for training AI models and by AI generated outputs. Traditional IP laws, designed to protect human inventors and creators, face challenges as AI systems increasingly contribute to innovation and creativity. The paper examines recent legal rulings, such as those by the United States Patent and Trademark Office and the U.S. Copyright Office, highlighting the ongoing debates over AI as an inventor or creator. Furthermore, it discusses the implications of copyright infringement lawsuits and data licensing activity, emphasizing the need for clarity in IP rights and responsibilities. Through a comprehensive overview of these issues, this paper advocates for a harmonized approach that balances the promotion of innovation and the inevitable use of technology in inventive process of invention and creation, with the protection of original creators' rights, and identifying key issues to pay attention to, as the landscape of AI and IP evolves.

Scan to Stay Connected!

Scan here to subscribe to CPI's **FREE** daily newsletter.

Visit **www.competitionpolicyinternational.com** for access to these articles and more!

01
INTRODUCTION

The relationship between human creativity and AI has emerged as an area of inquiry as AI technologies continue to advance and increasingly infiltrate various creative and inventive processes. This evolving dynamic necessitates a reevaluation of our existing IP framework.

First, the IP system must adapt to address the challenges posed by AI-generated and AI-assisted inventions and creations. Traditionally, patent law has been designed to protect the rights of human inventors for inventions that are useful, novel, and non-obvious thereby incentivizing innovation.[2] However, with the increasing use of AI tools in the inventive process across critical technologies ranging from biopharmaceuticals to semiconductor chip design, the patentability of AI-assisted inventions is an evolving question.[3] Similarly, copyright law has been designed to protect the rights of human inventors and creators for incentivizing creative output.[4] The advent of AI-generated or AI-assisted creative works ranging from art and music to literature are raising questions regarding what constitutes as creative work that is eligible for copyright protections.[5]

Moreover, questions about potential copyright infringement of both the input and output of large AI models are currently being raised in legal disputes in courts.[6] AI models are trained on large and extensive datasets, often generated by crawling the Internet, which may potentially include copyrighted material. The legality of utilizing this data for training AI models without explicit permission has become a contentious issue within legal and creative-work communities,[7] raising the economic question of valuing different types of data and any potential licensing models.[8] Some argue that Gen-AI models pro-duce transformative works. Others posit that the use of data for training of Gen-AI models fall under the fair use doctrine.[9] Nevertheless, courts haven't yet established a clear precedent. At the same time, the U.S. Copyright Office is working on developing guidelines to help address some of these pressing issues and has received overwhelming interest and input from stakeholders, including the developers of AI models, content creators, and the general public.[10]

These issues highlight the complex interplay between fostering innovation, embracing the adoption of technology, and safeguarding patent and copyright protections. The purpose of this article is to provide an overview of the key issues at the intersection of AI and IP, and their potential implications for the inventive and creative process as AI tools are adopted more broadly. Section II provides an overview of the patentability of AI-generated or AI-assisted inventions. Section III summarizes issues related to copyrightability of AI-generated or AI-assisted creative works. Section IV discusses some of the key open questions and ongoing litigation concerning the use of training datasets in the training of GenAI models. Section V provides a conclusion.

02
AI AND PATENTS

New AI tools and applications are increasingly being used to assist in the process of creating inventions across industries. This has raised questions about the patentability of inventions generated by and assisted by AI.

---

2   35 U.S.C. § 101, § 102, and § 103 (United States Patent Act), defining what constitutes patentable subject matter. Merges, Robert P. & John F. Duffy, "Patents, Trade Secrets, and the New Economy," *Harvard Law Review,* vol. 113, no. 3, 2000, pp. 1322-1342.

3   See Kersten, Alex, "Assessing the Patent and Trademark Office's inventorship guidance on AI assisted inventions," *CSIS,* June 2024.

4   "Copyright Basics," *U.S. Government Publishing Office,* 2021.

5   Abbott, Ryan Benjamin and Rothman, Elizabeth, "Disrupting creativity: Copyright law in the age of Generative Artificial Intelligence," *Fla. L. Rev.*, Vol. 75, Issue 6, 2023.

6   Samuelson, Pamela, "Generative AI meets copyright," *Science,* 381.6654 (2023): 158-161.

7   Quang, Jenny, "Does training AI violate copyright law?," *Berkeley Tech. LJ* 36 (2021): 1407.

8   Benjamin, Misha, *et al.*, "Towards the standardization of data licenses," *AI for Social Good Workshop, ICLR*. 2019.

9   *Supra* notes 6 and 7.

10   See the Copyright Office's Notice of Inquiry, August 2023, at: Artificial Intelligence and Copyright.

One of the simpler questions –AI systems may invent but are they inventors? – has been asked and answered by the United States Patent and Trademark Office ("USPTO") and the United States Federal Circuit. In 2019, Dr. Thaler tested the limits of patent law by filing patent applications in more than a dozen countries for two inventions created by his AI machine, which he called Device for Autonomous Bootstrapping of Unified Sentience ("DABUS").[11] In the USPTO patent applications, Dr. Thaler wrote that "the invention [was] generated by artificial intelligence." The USPTO denied the applications on grounds that "a machine does not qualify as an inventor." Dr. Thaler challenged the USPTO's decision in the U.S. District Court which upheld the USPTO's decision, concluding that an "inventor" under the Patent Act must be an "individual," or a natural person. Dr. Thaler then appealed to the Federal Circuit, which ultimately found that there is no ambiguity in the Patent Act requirement that an inventor listed on a patent application be a human being.[12] Thus, the USPTO has made clear that AI cannot be an inventor.

However, given that inventorship is limited to natural persons under U.S. law, AI's growing use has raised questions around whether AI-assisted inventions should receive patents. There are industries where AI is an important tool for the inventive processes including design and discovery, especially in industries that are heavily reliant on patent protection. For example, AI is becoming an indispensable tool in the chemical, biological, and pharmaceutical industries to facilitate cheaper, quicker, and more effective discovery and development, such as by proposing, refining and even "inventing" new molecules and chemicals through iterative machine learning processes. Automated chip design software ("ADS") tools increasingly use AI-powered tools for the complex process of designing semiconductors. The degree to which those discoveries made with the help of AI tools are patentable may be the next frontier at the intersection of AI and patent law. At industry roundtables conducted by CSIS, there was broad consensus across large and small biotechnology and technology companies – for applications across drug discovery, code generation, and material sciences – that AI-enabled inventions should be entitled to patent protection.[13]

The U.S. government has a key role to play in clarifying the rules of invention and patentability with the growing use of AI in the inventive process. Accordingly, following President Joe Biden's Executive Order on AI in October 30, 2023, the USPTO put out a notice and called for public comment on proposed "Inventorship Guidance for AI-Assisted Inventions" in February 2024.[14] The USPTO's guidelines are an important step towards clarifying key principles for AI-assisted inventions — namely, that inventors and joint inventors must be natural persons, that AI-assisted inventions are not categorically unpatentable for improper inventorship, and that there must be "significant contribution" by a human inventor to conception based on joint inventorship law.[15]

Some areas requiring clarification will emerge from the application of these guidelines to real-world patent applications. For example, the USPTO guidance lists various examples that illustrate patentability in situations where a human has made a significant contribution to an invention. However, the USPTO guidance does not explore certain gray areas, such as what might constitute the minimal requirements for human contribution to satisfy inventorship, or the threshold of "significant contribution" by a human when judged against AI's contribution. In addition, the USPTO guidance includes a disclosure requirement for the use of AI in the inventive process, but not for the use of other tools, such as computers and algorithms. What constitutes AI and to what degree it may contribute to a patentable invention are not clearly defined yet.[16] These uncertainties leave the door open for the courts.

03
AI AND COPYRIGHT

Similar to patents, the question at the intersection of copyrights and AI is: when creative works are produced, in whole or in part, using AI, should they be protected by copyright?

On the extreme end of the spectrum of copyrightability, a work can be generated with no human creativity and thus is not protectable. For example, the Ninth Circuit made clear that a photograph produced by a camera triggered by a

---

11   See https://www.uspto.gov/sites/default/files/documents/ai-inventorship-memo.pdf.

12   *Thaler v. Vidal* 43 F.4th 1207 (Fed. Cir. 2022).

13   *Supra* note 3.

14   See, Federal Register :: Inventorship Guidance for AI-Assisted Inventions

15   *Supra* see Section IV.

16   *Supra* note 3.

monkey is not entitled to a copyright because the Copyright Act only recognizes human inventors.[17] At the other extreme, a work may be wholly created by a human and is protectable. The guidance on both is straightforward. The more difficult cases are those in between, and complex questions arise about where to draw the line between an AI-aided work that is and is not protectable by copyright law. The U.S. Copyright Office issued several rulings recently on the question of when works generated using AI technology are protected under U.S. copyright law and, so far, applicants have not been able to convince the U.S. Copyright Office that the AI-generated components of their works are protectable.

Stephen Thaler again filed a lawsuit against the U.S. Copyright Office in response to the denial for the registration of an artwork created by his AI system a computer scientist and creator of an AI system he dubbed the "Creativity Machine," on the grounds that it lacked human authorship. The U.S. District Court upheld the decision of the U.S. Copyright Office.[18]

On the question of whether AI-generated works contain sufficient human authorship to be copyrightable, the U.S. Copyright Office provided its first analysis in its ruling in *Kashtanova*. That ruling narrowly interpreted the human authorship requirement and refused the registration of AI-generated images in a graphic novel, finding that detailed text prompts did not sufficiently constitute human authorship.[19] Consistent with its decision in *Kashtanova*, the U.S. Copyright Office also refused to register a work titled "Théâtre D'opéra Spatial," a two dimensional artwork, whose copyright application described a detailed, iterative creation process that involved inputting numerous text prompts and hundreds of rounds of revisions in Midjourney, a generative AI tool for image creation. The Review Board found this was insufficient to constitute human authorship. More recently, the U.S. Copyright Office also rejected the registration of a two-dimensional computer-generated image titled "Suryast," created by inputting an original photograph into a style transfer tool called RAGHAV, to produce a highly stylized version of the original photograph. The office found the work non-registerable "because [it] is a derivative work that does not contain enough original human authorship to support a registration." It is worth noting that both the Indian Copyright Office and the Canadian Copyright Intellectual Property Office have registered "Suryast" and recognized RAGHAV AI painting app as its co-author along with its human creator, Sahni.

In March 2023, the U.S. Copyright Office did issue guidance on registration of works generated by AI.[20] Along the lines of the decision it made, the U.S. Copyright Office clarifies that while technological tools can be a part of the creative process, "what matters is the extent to which the human had creative control over the work's expression and 'actually formed' the traditional elements of authorship," which it will determine on a case-by-case basis.[21] While the U.S. Copyright Office has so far taken a narrower view of what is required to constitute sufficient human authorship in an AI-generated work, the law is still unclear, as no court has yet addressed the issue.

04
COPYRIGHT INFRINGEMENT

The issue that is receiving the most attention at the intersection of AI and IP is copyright infringement. Recently, the U.S. Copyright Office issued a Notice of Inquiry ("NOI") soliciting public comments regarding the collection and curation of sources for AI datasets, the methodologies employed in training AI models with these datasets, and the necessity for obtaining permission or providing compensation to copyright owners when their works are incorporated into this process.[22] The inquiry received over 10,000 comments from the public and stakeholders which the U.S. Copyright Office is in the process of evaluating.

In the meantime, over a dozen lawsuits are pending in various jurisdictions across the U.S. in which copyright holders

---

17 *Naruto v. Slater*, 2018 WL 1902414 (9th Cir. 2018). The famous case involved a selfie taken by a monkey in Naruto in Indonesia with the camera of the British nature photographer David Slater in 2011. The People for the Ethical Treatment of Animals ("PETA") commenced an action against Slater and his book publisher, claiming Naruto was the copyright owner of the selfie.

18 *Thaler vs. Pertmutter*, 2023 (U.S. District Court of Columbia), at: https://caselaw.findlaw.com/court/us-dis-crt-dis-col/114916944.html.

19 See https://copyright.gov/docs/zarya-of-the-dawn.pdf.

20 See https://copyright.gov/ai/ai_policy_guidance.pdf.

21 *Supra*.

22 Notice of Inquiry, 88 Fed. Reg. 59942 (U.S. Copyright Office Aug. 30, 2023), https://www.regulations.gov/document/COLC-2023-0006-0001.

are advancing multiple theories of infringement against AI platforms, specifically, the Gen-AI models. Plaintiffs in these actions generally contend that Gen-AI models infringe upon copyrights through impermissible inclusion of copyrighted materials in training data. Plaintiffs have also alleged that Gen-AI outputs do or are likely to infringe copyrighted materials.

Some argue that AI-generated works are transformative, thereby falling under fair use protections. However, courts have yet to establish clear precedents in this area. The application of fair use to the data used for training AI models is still being debated, especially with regard to the balance between innovation and copyright protection.

Several plaintiffs' cases have advanced direct and/or indirect infringement claims alleging that a Gen-AI model accessed and copied copyrighted material for the purpose of training the model.[23] Gen-AI models need substantial amounts of training data. For example, GPT-3 was trained on approximately 570 gigabytes of text data, derived from a diverse range of sources, including books, articles, and websites.[24] Some Gen-AI models employ techniques that "scrape" content from the internet, which may include potentially copyrighted content. However, the training that used by most Gen-AI models remains limited in public knowledge. Consequently, the viability of this infringement theory may vary depending on the facts and circumstances of each case.

There are some technical and economic arguments that are likely to surface in the ongoing copyright infringement disputes. For the question regarding the use and storage of datasets for training purposes, Gen-AI large language models need to break down text into smaller words and tokens for creating training datasets, and then correlate specific functions and linguistic data to tokens to probabilistically predict the next word given the previous words. Open AI has stated that the only thing stored in its model is the structure of the language itself, rather than the copyrightable expression of a given work.[25] An additional question would be about the incremental value of any specific input of data in terms of its contribution to

the training of a model, given the large size of the training datasets. Finally, another question would be whether a market for licensing of the training datasets would be viable in the cost-vs-benefit tradeoff of creating such a market. In the meantime, various licensing deals are emerging in the industry for the use of certain data by Gen-AI model developers. For example, Google signed a $60 million annualized licensing deal with Reddit to access Reddit Data APIs.[26] Open AI struck a deal with various media companies and outlets including the Financial Times, The Associated Press, Axel Springer, and News Corp (a media company that owns The Wall Street Journal, the New York Post, and The Daily Telegraph) for the use of current and archived articles.[27] Members of the newly formed Dataset Providers Alliance are looking to streamline the licensing process, ensuring fair compensation for rights holders and high-quality data for AI companies. It remains to be seen if collective licensing schemes for specific AI training datasets are viable or will gain traction.

> " *In the meantime, over a dozen lawsuits are pending in various jurisdictions across the U.S. in which copyright holders are advancing multiple theories of infringement against AI platforms, specifically, the Gen-AI models"*

Plaintiffs have also proposed theories of infringement that extend beyond the training phase of Gen-AI models. They contend that the operation of a given Gen-AI model constitutes an unauthorized derivative work, as it utilizes copyrighted materials in its outputs.[28] Additionally, plaintiffs maintain that the outputs generated by AI models can result

---

23   For example, *Getty Images (US), Inc. v. Stability AI, Inc.*, No. 1:2023cv01850 (S.D.N.Y. 2023), *Authors Guild et al vs. Open AI, Alter vs. Open AI*.

24   Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., ... & Amodei, D. (2020). *Language Models are Few-Shot Learners*. In *Advances in Neural Information Processing Systems*, 33, 1877-1901.

25   Comment from Open AI, Re: Notice of Inquiry and Request for Comment [Docket No. 2023-06] (Oct. 30, 2023), COLC-2023-0006-8906.

26   See "Reddit signs AI content deal ahead of IPO," *Bloomberg,* February 16th, 2024, available at: Reddit Is Said to Sign AI Content Licensing Deal Ahead of IPO - Bloomberg.

27   See "Open AI's News Corp deal licenses content from WSJ, New York Post, and more," *The Verge,* May 22, 2024, available at: OpenAI's News Corp deal licenses content from WSJ, New York Post, and more - The Verge.

28   See *Authors Guild v. OpenAI, Inc.*, No. 1:23-cv-08292 (S.D.N.Y. filed Sept. 19, 2023); *Andersen v. Stability AI Ltd.*, No. 3:23-cv-00201 (N.D. Cal. Oct. 30, 2023).

in substantially similar works, which may constitute copyright infringement.[29] It remains to be seen whether or not any of these claims would hold in the courts. At the very least, it is clear that courts require plaintiffs to sufficiently allege similarity of the output of the models to plaintiffs' works.[30] Courts have also requested evidence of sufficient economic injury, about a potential impact on the market for the copyrighted work.[31]

In the backdrop of the evolving litigation and licensing landscape, one of the issues for enterprises and customers of Gen-AI models is the indemnity protection against potential copyright infringement. While several AI providers indemnify enterprise and developers from copyright claims of their AI services, the scope this indemnity protection is often limited.[32] For example, the indemnification may only cover third-party infringement claims related to outputs, but not for claims that the training data and inputs were infringing.

The numerous ongoing copyright infringement lawsuits mark a critical juncture that is poised to shape the relationship between content creators and AI models. Until these are resolved, individuals and organizations engaging with AI platforms should be cognizant of the uncertainty and the risk of copyright liability. The forthcoming recommendations from the U.S. Copyright Office to Congress may provide more guidance towards the resolution of these pending issues.

"In the backdrop of the evolving litigation and licensing landscape, one of the issues for enterprises and customers of Gen-AI models is the indemnity protection against potential copyright infringement"

---

29   See *Doe 1 v. GitHub, Inc*., No. 4:22-cv-06823-JST (N.D. Cal. filed Nov. 3, 2022); *Getty Images, Inc. v. Stability AI, Inc.*, No. 1:23-cv-00135-JLH (D. Del. filed Feb. 3, 2023); *Concord Music Grp., Inc. v. Anthropic PBC*, No. 3:23-cv-01092 (M.D. Tenn. filed Oct. 18, 2023); *Andersen v. Stability AI Ltd.*, No. 3:23-cv-00201 (N.D. Cal. filed Oct. 30, 2023); *The N.Y. Times Co. v. Microsoft Corp*., No, 1:23-cv-11195, (S.D.N.Y. filed Dec. 27, 2023).

30   See, example, Silverman vs. Meta, No. 4-23-cv-03416-KAW.

31   Id.

32   See Regina Sam Penti, Georgina Jones Suzuki & Derek Mubiru, "Trouble Indemnity: IP Lawsuits In The Generative AI Boom," *Law360* (Jan. 3, 2024), https://www.law360.com/articles/1779936/trouble-indemnity-ip-lawsuits-in-the-generative-ai-boom.

05
CONCLUSION

In summary, the intersection of AI and IP presents a complex and evolving landscape for inventors, creators, and users of AI tools. As AI technologies increasingly permeate creative and inventive domains, significant questions arise regarding the patentability of AI-assisted inventions and the copyrightability of works generated with the help of AI. The ongoing legal disputes surrounding copyright infringement further underscore the need for a clear framework that delineates the rights and responsibilities of content creators, AI developers, and users.

The U.S. Copyright Office's efforts to develop guidelines and the recent public inquiries reflect an acknowledgment of the pressing challenges posed by AI in the realm of IP. However, the lack of established legal precedents, particularly concerning the classification of AI-generated outputs and the utilization of copyrighted materials for training purposes, leaves much ambiguity in the current legal framework. As litigations unfold, the outcomes may set critical precedents that will shape the future of both copyright and patent law.

Moreover, as various licensing agreements emerge within the industry, the question of fair compensation for content creators becomes increasingly pertinent. The potential for collective licensing schemes for AI training datasets may pave the way for a more equitable balance between fostering innovation and protecting the rights of original creators.

Ultimately, it is imperative for stakeholders — including legal practitioners, AI developers, and policymakers — to remain engaged in this discourse. By proactively addressing the implications of AI on intellectual property rights, we can work toward a more cohesive and adaptive legal framework that not only encourages technological advancement but also respects the fundamental rights of human creators. The ongoing evolution of this dialogue will be crucial in navigating the challenges and opportunities that lie ahead in the age of artificial intelligence. ■

"*The U.S. Copyright Office's efforts to develop guidelines and the recent public inquiries reflect an acknowledgment of the pressing challenges posed by AI in the realm of IP"*

CPI
SUBSCRIPTIONS

CPI reaches more than **35,000 readers** in over **150 countries** every day. Our online library houses over **23,000 papers**, articles and interviews.

Visit **competitionpolicyinternational.com** today to see our available plans and join CPI's global community of antitrust experts.

CPI COMPETITION POLICY® INTERNATIONAL