![Chamber of Progress logo]

October 30, 2023
U.S. Copyright Office | Notice of Inquiry - Artificial Intelligence and Copyright

Re: U.S. Copyright Office Notice of Inquiry - Artificial Intelligence and Copyright

Dear U.S. Copyright Office,

Chamber of Progress appreciates the opportunity to respond to your Notice of Inquiry on Artificial Intelligence and Copyright, 59942 FEDERAL REGISTER, VOL. 88, NO. 167. Our comments will focus on Questions 1, 5, and 8 of the Inquiry.

Chamber of Progress is a progressive tech industry group fighting for public policies that will build a fairer, more inclusive country in which all people benefit from technological leaps. Our partner companies include a diversity of social media, online marketplace, and other consumer-facing platforms, but our partner companies do not have a vote or veto over our positions.

Generative AI has revolutionized the way artists and researchers produce their work, become inspired, and create solutions to a wide-variety of issues facing the world. Chamber of Progress works to ensure that Generative AI enriches the American economy and expands the welfare of American society. Therefore, Chamber of Progress urges the U.S. Copyright Office to continue to abide by the well-equipped existing copyright laws to allow for Generative AI produced works to facilitate benefits to American society.

1. Question One: 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?

The discourse surrounding Generative AI primarily emphasizes its risks, such as data privacy issues, cybersecurity threats, algorithmic bias, deep fakes, factual inaccuracies (referred to as "hallucinations"), among others.[1][2] Nonetheless, it's essential to recognize that Generative AI offers substantial advantages, including improved decision-making, heightened human creativity, and increased resource efficiency.[3] Chamber of Progress advocates for the Copyright Office to acknowledge that enabling use of Generative AI can fundamentally enhance lives through fostering inspiration, boosting productivity, optimizing businesses, advancing research, and enriching the overall human experience.

A. Generative AI Beneficial Use-Cases are Pervasive

The capabilities of Generative AI can be a foundation for idea formulation and inspiration for artists and creators more generally.[4][5] From helping with writer's block, to exploring new styles, Generative AI tools push creators to explore unfamiliar territory.[6] Generative AI bridges the divide between smaller and underrepresented artists and consumers, enhancing accessibility and visibility, and offering artists a streamlined pathway to share their creativity with a broader audience like never before.[7]

Improving Content Moderation

---

[1] Sabrina Ortiz, "The 5 biggest risks of generative AI, according to an expert" Zdnet, (2023). Available at:
https://www.zdnet.com/article/the-5-biggest-risks-of-generative-ai-according-to-an-expert/

[2] Belle Lin, "AI Is Generating Security Risks Faster Than Companies Can Keep Up" Wall Street Journal, (2023). Available at:
https://www.wsj.com/articles/ai-is-generating-security-risks-faster-than-companies-can-keep-up-a2bdedd4

[3] Anujaa Singh, "9 Benefits of generative AI in enterprises" yellow.ai, (2023). Available at:
https://yellow.ai/blog/benefits-of-generative-ai/#:~:text=By%20automating%20tasks%20that%20previously,speeding%20up%20the%20design%20process.

[4] Ziv Epstein, et al., "Art and the science of generative AI" Science, (2023). Available at:
https://www.science.org/doi/10.1126/science.adh4451

[5] Tojin T. Eapen, et al., "How Generative AI Can Augment Human Creativity" Harvard Business Review, (2023). Available at:
https://hbr.org/2023/07/how-generative-ai-can-augment-human-creativity

[6] Kishor K., "How Generative AI is Revolutionizing Drawing and Art" Medium, (2023). Available at:
https://medium.com/@Nontechpreneur/how-generative-ai-is-revolutionizing-drawing-and-art-6c4e99e67ba9

[7] Sarah Shaffi "'It's the opposite of art': why illustrators are furious about AI.," The Guardian, (2023). Available at:
https://www.theguardian.com/artanddesign/2023/jan/23/its-the-opposite-of-art-why-illustrators-are-furious-about-ai

In the technology space, advancements in Generative AI have proven to bolster content moderation efforts. For example, human moderation decisions can be used to train and refine AI models used for automated content moderation for enhanced accuracy. Further, Generative AI plays a significant role in developing training content for AI and human moderators alike.[8]

Revolutionizing Medical Research

Advanced Generative AI tools also hold immense potential to revolutionize research, particularly in medicine, in extending healthcare access and accelerating diagnosis accuracy. AI algorithms, for example, excel at detecting early signs of cancer through more precise analysis of medical imaging than human observation. Additionally, AI techniques can efficiently manage extensive datasets and streamline labor-intensive aspects of drug development, enabling the rapid release of life-saving medications in response to emerging biological threats.[9] If medical researchers are denied access to the vast datasets from pharmaceutical and healthcare sectors for their AI tools, drug research and development will continue to be a cumbersome and time-consuming endeavor.

Benefiting Consumers

The burgeoning technology has vast benefits for consumers, notably in the realm of accessibility. In the United States, where residents speak over 350 languages, natural language processing AI can greatly enhance support for non-English speakers in crucial tasks such as communicating with healthcare professionals, meeting with their children's teachers, or participating in the democratic process, like voting.[10] What's more, consumers with physical disabilities can utilize AI "dressing rooms" for virtual clothing try-ons, deaf individuals can benefit from AI-assisted real-time sign language translation, and those who are visually

---

[8] Aditya Jain, "Impact Of Generative AI On Content Moderation," Avasant, (2023). Available at: https://avasant.com/report/impact-of-generative-ai-on-content-moderation.
[9] Matthew Huddle, et. al., "Generative AI Will Transform Health Care Sooner Than You Think" BCG, (2023). Available at:
https://www.bcg.com/publications/2023/how-generative-ai-is-transforming-health-care-sooner-than-expected
[10]USA Gov., "Official language of the United States" USA Gov, (2023). Available at:
https://www.usa.gov/official-language-of-us

impaired can use image assistance technology to identify objects real-time using their smartphone cameras.[11]

Enhancing Autonomous Vehicles

Generative AI also plays a crucial role in bolstering the safety of autonomous vehicles by producing extensive datasets and scenarios for honing autonomous systems. For example, Generative AI can be used to synthesize unlimited conditioned traffic and driving data to create immense and highly sophisticated simulations on which autonomous vehicles can train.[12] Consequently, the impending era of autonomous vehicle-based transportation stands to substantially diminish traffic accidents, enhance transportation alternatives for the elderly and individuals with disabilities, and foster equitable access to transit for underserved groups.[13]

Enhancing Educational Outcomes

As for students, Generative AI may be a crucial tool in the effort to reverse downward trends in education outcomes.[14] As highlighted by the U.S. Department of Education Office of Educational Technology, AI can "enable new forms of interaction," notably supporting students with disabilities, as well as "help educators address variability in student learning."[15] If educational AI tools are, for example, limited in drawing data from notable literary works due to copyright restrictions, students are hindered from better accessing and absorbing knowledge.

---

[11]Cristina Fonseca, "The role of AI in making CX more accessible and inclusive" Zendesk Blog, (2023). Available at: https://www.zendesk.com/blog/ai-cx-accessible/

[12] M. Xu et al., "Generative AI-empowered Simulation for Autonomous Driving in Vehicular Mixed Reality Metatverses," IEE Journal of Selected Topics in Signal Processing, (2023). Available at: https://ieeexplore.ieee.org/document/10177684.

[13] Chamber of Progress, "AUTONOMOUS VEHICLES: Leveraging Technology for Diverse Community Benefit," (2022). Available at: https://progresschamber.org/wp-content/uploads/2022/06/Report-AVs-Leveraging-Technology-for-Diverse-Community-Benefit.pdf.

[14] Sequoia Carrillo, "U.S. reading and math scores drop to lowest level in decades" NPR, (2023). Available at: https://www.npr.org/2023/06/21/1183445544/u-s-reading-and-math-scores-drop-to-lowest-level-in-decade

[15]Office of Educational Technology, "Artificial Intelligence" Office of Educational Technology, (2023). Available at: https://tech.ed.gov/ai/

2. Question Five: 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.

A. Any Changes to U.S. Copyright Laws Should Prioritize the Advancement of Generative AI and User Expression

"Copyright law is the only law that's already in existence that could bring generative AI systems to their knees,"[16] Pamela Samuelson.

Technology has continually reshaped our interaction with creative content. As these advancements create new opportunities for profitability, rights holders remain, understandably, vigilant.[17] Yet, the current wave of copyright litigation targeting Generative AI might suggest that it is the p*roviders* of Generative AI services that are at greater risk—given the emerging and impossibly complex legal landscape new providers must navigate.[18] Given this, new legislative frameworks to address the intersection of Generative AI and copyright are not needed.

The most prominent Generative AI services today necessitate exhaustive training, often involving the aggregation of publicly available content from the web.[19] This scraping exercise is pivotal to instruct large language models on the multifaceted ways humans interact and comprehend the world around them. Using this data, AI services generate outputs—whether text or images—that are comprehensible to humans. Consequently, today's rights holders contend that both the training data and the resultant outputs infringe upon their works, threatening the

---

[16] Stella Kotik, "Generative AI meets copyright law" Berkeley News, (2023). Available at: https://news.berkeley.edu/2023/05/16/generative-ai-meets-copyright-law

[17] *Andersen et al v. Stability AI Ltd. et al*, Docket No. 3:23-cv-00201 (N.D. Cal. Jan 13, 2023); *Walters v. OpenAI, L.L.C.,* Docket No. 1:23-cv-03122-MLB (N.D. Ga. Jul 14, 2023).

[18] *See* e.g., Act No. 2021-344, 2021 Ala. Acts & Vt. H. 410, An Act Relating to the Use and Oversight of Artificial Intelligence in State Government, Reg. Sess. (2023). Available at: https://www.ncsl.org/technology-and-communication/artificial-intelligence-2023-legislation

[19] Tom Davenport and Maryam Alavi, "How to Train Generative AI Using Your Company's Data" Harvard Business Review, (2023). Available at: https://hbr.org/2023/07/how-to-train-generative-ai-using-your-companys-data

continued existence of Generative AI, as its developers grapple with economically devastating lawsuits.[20]

Further, as Generative AI becomes more accessible to the broader public, the potential for its misuse correspondingly rises.[21] This triggers an additional concern for rights holders, especially when users deliberately craft prompts to generate infringing outputs, bypassing any safeguards implemented by the AI providers.[22] Consequently, user-induced infringement exposes Generative AI providers to claims of secondary liability, mirroring the challenges regularly faced by user-generated content services ('UGC services'), like YouTube.[23]

These liability concerns jeopardize the viability of Generative AI services. Therefore, if consumers and policymakers envision a vibrant future for Generative AI, it's imperative that the USCO emphasizes regulations that foster competition in this nascent field, rather than suppress it. Without such supportive measures, Generative AI risks becoming a fleeting phenomenon.

B. Any Proposed Legislation Must Uphold Existing Principles of Fair Use and Prevent Rights Holder Abuse

Generative AI services, exemplified by OpenAI's ChatGPT, Google's Bard, and Stability AI's Stable Diffusion, have elicited concerns from rights holders who fear their original works might be replicated en masse or even supplanted by such

---

[20] Kyle Wiggers, "The current legal cases against generative AI are just the beginning" Tech Crunch, (2023). Available at:
https://techcrunch.com/2023/01/27/the-current-legal-cases-against-generative-ai-are-just-the-beginning/

[21] Janna Anderson and Lee Rainie, "As AI Spreads, Experts Predict the Best and Worst Changes in Digital Life by 2035" Pew Research Center, (2023). Available at:
https://www.pewresearch.org/internet/wp-content/uploads/sites/9/2023/06/PI_2023.06.21_Best-Worst-Digital-Life_2035_FINAL.pdf

[22] Hungryminded, "Tricking ChatGPT: Do Anything Now Prompt Injection" Medium, (2023). Available at:
https://medium.com/seeds-for-the-future/tricking-chatgpt-do-anything-now-prompt-injection-a0f65c307f6b

[23] Brittany Rycoft, "4 Major Problems with User-Generated Content (and How to Fix Them)" GhostRetail, (2022). Available at:
https://www.ghostretail.com/post/user-generated-content-problems

technological breakthroughs.[24] While these concerns are legitimate, existing copyright doctrine is well equipped to address these concerns.

For example, fair use ensures that rights holders cannot monopolize creative ideas, creating an ecosystem conducive to continual innovation.[25] In practice, fair use requires a delicate and intricate analysis from the Courts. Its fact intensive, case-by-case nature lends it to stand-up to the latest technological innovations. Generative AI technology should not deviate from these principles as changes might not only impair its functionality but also stifle a whole new realm of expressive creations.

While some artists dub AI-text-to-image generators as "21st-century collage tools" that merge existing works into learning models, they miss an essential point.[26] Generative AI aligns more with human learning, where exposure to existing works shapes and influences fresh creations, rather than simply piecing together existing content.

In cases of direct copyright infringement, courts typically evaluate whether the defendant had access to the original work and if their creation is 'substantially similar' to it. In assessing substantial similarity, courts will assess both the quantitative and qualitative aspects of the copied content. They examine if the copying breaches a 'de minimis' limit and assess the significance of the copied sections in the context of the original work. An illustrative case is *Authors Guild v. Google Books*, weighing in favor of fair use where Google's snippet previews failed to reveal the "heart" of any of the protected works at issue.[27]

These established legal criteria are equally applicable to outputs produced by Generative AI technologies. When gauging access, courts might examine if the AI's training data incorporates the original work. User inputs can also be a source of original content. As for substantial similarity, the holistic aesthetics of the works are compared. Some courts may even consider whether a reasonable person can discern between the two.

---

[24] Gil Appel, et. al., "Generative AI Has an Intellectual Property Problem" Harvard Business Review, (2023). Available at:
https://hbr.org/2023/04/generative-ai-has-an-intellectual-property-problem
[25] Id.
[26] *Andersen et al v. Stability AI Ltd. et al,* Docket No. 3:23-cv-00201 (N.D. Cal. Jan 13, 2023).
[27] *Authors Guild v. Google, Inc.*, 578 U.S. 941, 136 S. Ct. 1658, 194 L. Ed. 2d 800 (2016).

Importantly, rights holders must substantiate their claims by pinpointing the exact infringed works. For instance, in *Andersen*, Judge William Orrick expressed a leaning towards dismissing most of the claims, not due to novel AI-specific concerns but because the plaintiffs failed to identify specific infringing works.[28] Broad allegations based solely on the potential of Generative AI to infringe cannot stand, echoing sentiments in the traditional copyright cases.

In Getty Images (US), Inc. v. Stability AI, Inc., Getty Images contends that Stable Diffusion violated their copyright by producing outputs that occasionally retain traces of Getty Images' watermarks.[29] The court is tasked with discerning substantial similarity. If the Plaintiffs can't prove similarity beyond the watermark traces, then copyright remedies shouldn't apply, irrespective of the Generative AI technology involved. And while the presence of watermarks on generated images might cause consumer confusion, this isn't a matter for Copyright examination.

Conversely, Generative AI services that use a limited training dataset, dominated by repetitive works of certain artists, are prone to generate substantially similar and potentially infringing outputs.[30] Just as artists exercise caution to avoid creating works that resemble others, there's no need for Copyright Law amendments; Generative AI providers and its users should naturally follow the same prudence.

In practice, fair use demands nuanced court analysis due to its detailed, case-specific nature, making it adaptable to the latest technological advances. Altering these principles for Generative AI might compromise its functionality and limit the emergence of innovative expressive works.

C. No Legislation Should Bring Artistic Style within the Scope of Copyright Protections

Though a growing concern among artists, policymakers should resist expanding copyright protections to cover artistic style. This issue is central to *Andersen et al*

---

[28] Blake Brittain, "US judge finds flaws in artists' lawsuit against AI companies," Reuters, (2023). Available at:
https://www.reuters.com/legal/litigation/us-judge-finds-flaws-artists-lawsuit-against-ai-compa nies-2023-07-19/

[29] *Getty Images (US), Inc. v. Stability AI, Inc.,* Docket No. 1:23-cv-00135 (D. Del. Feb 03, 2023).

[30] Mark Lemley & Bryan Casey, "Fair Learning" Texas Law Review, (2021). Available at:
https://texaslawreview.org/wp-content/uploads/2021/03/Lemley.Printer.pdf

v. Stability AI (2023), where AI tools like Stable Diffusion allegedly produce images that mimic the plaintiffs' distinctive artistic styles.[31]

Historically, courts are hesitant to offer broad copyright protections solely for artistic style. Rather, artistic style is viewed as a *component* of an artist's protected expression.[32] For instance, the Court in *Dave Grossman Designs v. Bortin* posited that while emulating Picasso's Cubist style is permissible, directly replicating his unique expression is not.[33]

The Bortin Court's stance is rational: Protecting artistic styles could curb the very essence of creativity.[34] Artists are regularly inspired by other artists and works of the past. For instance, the rising band, Greta Van Fleet, known for its Zeppelin-like sound, could not exist under a copyright regime that prohibits style imitation.[35] Artists harnessing Generative AI to do the same, perhaps just quicker and better, should enjoy similar freedoms under current copyright tenets.

However, this doesn't render rights holders powerless. While style could never be an artist's primary grievance, courts will still evaluate it alongside the defendant's overall expression. *Steinberg v. Columbia Pictures Industries, Inc.* exemplifies this, where the court acknowledged both the defendant's mimicry of Steinberg's style *and* the substantial similarity in their work.[36] The recent *Warhol* decision only further expands potential victories for rights holders under the substantial similarity doctrine.[37]

This ambiguity in the copyright landscape naturally discourages potential infringers, regardless of the tools and technology available to them.[38] For

---

[31] *Andersen et al v. Stability AI*, *supra* note 17.

[32] *Steinberg v. Columbia Pictures Indus.*, Inc., 663 F. Supp. 706, 3 U.S.P.Q.2d 1593 (S.D.N.Y. 1987), Court Opinion.

[33] *Dave Grossman Designs, Inc. v. Bortin*, 177 U.S.P.Q. 627 (N.D. Ill. 1973), Court Opinion.

[34] Brian Frye, " A Cento on Plagiarism" Social Epistemology Review and Reply Collective, (2022). Available at: https://social-epistemology.com/2022/07/28/a-cento-on-plagiarism-brian-l-frye/

[35] Sophie McVinnie, "Why does Greta Van Fleet's music sound so much like Led Zeppelin?" Guitar.com, (2021). Available at: https://guitar.com/features/opinion-analysis/why-does-greta-van-fleet-sound-like-led-zeppelin/

[36] *Steinberg,* style is one ingredient of "expression," at 712.

[37] *Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith*, 142 S. Ct. 1412, 212 L. Ed. 2d 402 (2022), Court Opinion.

[38] Matthew Sag, "Internet Safe Harbors and the Transformation of Copyright Law" Emory Law, (2017). Available at: https://scholarlycommons.law.emory.edu/cgi/viewcontent.cgi?article=1008&context=faculty-articles

example, while a user may be able to successfully coax an AI-image generator to produce substantially similar reproductions of another artist's works, the user is certain to face expensive repercussions under current copyright law should the user attempt to commercialize the generated works.

Similarly, without alterations to existing copyright regulations, AI service providers are already taking proactive measures to prevent potential copyright infringements, recognizing the gray areas around the fairness of AI outputs. For instance, Stability AI has modified its image generator, Stable Diffusion, to deny requests that mimic established artists' styles.[39] OpenAI made a parallel move with DALL-E 3, which now also refuses to generate images mirroring the style of any living artist and allows artists to opt-out of having their works included in any training sets.[40]

Concerns about AI serving as an artist's proxy are understandable. Yet, the annals of art reveal that imitations seldom grasp the distinct essence of an original artist. With a growing appreciation for authenticity, consumers will likely favor genuine art over AI-created pieces.[41] Given present copyright law already factors in market displacement for fair use assessments, there's no need for additional legislation aimed merely at propping up human artists at the expense of Generative AI advancements. Traditional artists will naturally adapt and differentiate their creations from AI-produced art, perhaps even through the assistance of Generative AI.[42]

The emergence of Generative AI suggests that rights holders may need to take up new proactive efforts, such as opting-out of the training sets that power

[39] Steve Dent, "Stable Diffusion update removes ability to copy artist styles or make NSFW works" Engadget, (2023). Available at:
https://www.engadget.com/stable-diffusion-version-2-update-artist-styles-nsfw-work-124513511.html
[40] OpenAI, "DALL·E 3" OpenAI, (2023) (DALL·E 3 is designed to decline requests that ask for an image in the style of a living artist. Creators can now also opt their images out from training of our future image generation models). Available at: https://openai.com/dall-e-3
[41] Kobe Millet, et al., "Defending humankind: Anthropocentric bias in the appreciation of AI art" Science Direct, (2023) (suggesting participants showed an overwhelming preference for artwork they thought was made by people). Available at:
https://www.sciencedirect.com/science/article/pii/S0747563223000584?via%3Dihub
[42] Rich Santon, "AI artist who won competition says art world is 'in denial' about the tech" PCGamer, (2023). Available at:
https://www.pcgamer.com/ai-artist-who-won-competition-says-art-world-is-in-denial-about-the-tech/

Generative AI services, to oversee their works. But should infringements arise, rights holders are well-equipped with a robust copyright framework, based on centuries of precedent, to zealously defend their rights. Generative AI neither diminishes the legal avenues available to rights holders nor calls for an immediate deviation from entrenched copyright norms.

D. Any Legislation Must Provide Incentives for AI Companies to Improve Their Models and Preserve User Expression

As Generative AI becomes more available to average consumers, the likelihood of its misuse inevitably increases. One concern for copyright pertains to users and their interactions with AI-image generator services. Considering the current capabilities of text-to-image AI generators, users might craft prompts that sidestep the protective measures set by AI providers, potentially leading to infringement.[43]

Analogous to user-generated content services (UGC services), the realm of bad things users can imagine and manifest online is limitless.[44] Similarly, providers of AI services will require a permissive regulatory environment that grants room for improvements. For this, AI providers need the liberty to incorporate human feedback and make adjustments without fearing overwhelming liability for what they know or learn in the process. For example, Generative AI providers should not be held liable for the infringement-driven prompts engineered and supplied by their users, nor for their efforts to learn about and guard against such abuses.

For instance, beyond the realm of copyright, immunities like Section 230[45] bolster First Amendment safeguards for UGC services by guaranteeing protection against liabilities arising from their content moderation efforts. Notably, the

---

[43]Pam Samuelson, "Generative AI meets copyright, " Science, Volume 381, Issue 6654 (2023), ("It is, however, possible for generative AI outputs to infringe copyrights. If the same input image (say, of Mickey Mouse) is present in many works on which the model was trained and its developer did not follow industry best practices by eliminating duplicates and using output filters to prevent infringements, user prompts could result in infringing outputs (although this user, not the developer of the generative AI system, may be the infringer."). Available at: https://www.science.org/doi/10.1126/science.adi0656?adobe_mc=MCMID%3D672063465477 01074570918360005019222071%7CMCORGID%3D242B6472541199F70A4C98A6%2540AdobeO rg%7CTS%3D1696466256

[44] Eric Goldman & Jess Miers, "Why Internet Companies Can't Stop Awful Content" Santa Clara University School of Law, (2020). Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3518970

[45] 47 U.S. Code § 230.

absence of a knowledge requirement in Section 230(c)(1)[46] empowers UGC services to refine their moderation methods, including their algorithms, without acquiring the legal awareness that could expose them to significant liability.[47] As a result, most UGC services go to great lengths to proactively clean-up awful content and provide a safe and trustworthy environment for their us ers.[48]

While Section 230 is not available for Copyright infringement claims, its immunity design has successfully bolstered and preserved the user-generated content ecosystem. Policymakers should look to Section 230 as a beacon when considering secondary liability for providers of Generative AI.

The counterpart to Section 230 for online copyright law is the 'safe harbor' provision provided to UGC services under Section 512 of the Digital Millennium Copyright Act (DMCA).[49] While this safe harbor is vital in warding off a myriad of secondary liability copyright claims for UGC services, its inherent design flaws, especially concerning the knowledge stipulation, have posed significant challenges for content creators.[50] Rather than encouraging UGC services to defend the availability of their creators' content, the design of Section 512 is infamous for promoting excessive content removal online.[51] Indeed, as some experts note, online copyright law seems to have carved a 'memory hole' in the digital content landscape.[52]

At the very least, Generative AI providers, which grapple with similar, if not more complex, challenges as UGC services, should have access to a safe harbor for

---

[46] 47 U.S. Code § 230(c)(1).

[47] Eric Goldman, "Why Section 230 Is Better Than the First Amendment" Notre Dame Law Review, (2019). Available at:
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3351323#:~:text=Eric%20Goldman,-Sant a%20Clara%20University&text=%C2%A7%20230%20(%E2%80%9CSection%20230%E2%80%9D,the% 20immunity%20in%20regulators'%20sights.

[48] Eric Goldman, "Content Moderation Remedies" Michigan Technology Law Review, (2023). Available at: https://repository.law.umich.edu/mtlr/vol28/iss1/2/

[49] 17 U.S.C. § 512 (2018).

[50]"Unintended Consequences: Sixteen Years under the DMCA," Electronic Frontier Foundation, (2014). Available at: https://www.eff.org/wp/unintended-consequences-under-dmca/archive.

[51] Daphne Keller, "EMPIRICAL EVIDENCE OF "OVER-REMOVAL" BY INTERNET COMPANIES UNDER INTERMEDIARY LIABILITY LAWS" The Center for Internet and Society, (2015). Available at: https://cyberlaw.stanford.edu/blog/2015/10/empirical-evidence-over-removal-internet-compa nies-under-intermediary-liability-laws

[52] Eric Goldman & Jessica Silbey, "Copyright's Memory Hole" BYU L. Rev. 929 (2020). Available at: https://digitalcommons.law.byu.edu/lawreview/vol2019/iss4/6/

secondary infringement claims.[53] However, as Generative AI booms, this moment offers the USCO and Congress a chance to leverage the insights gained from years of Section 512 litigation and introduce immunity-based enhancements to secure the future of Generative AI.[54]

For instance, one criterion for determining safe harbor eligibility could involve an evaluation of the size and diversity of the training dataset used for the model (e.g., whether it's sufficiently extensive to prevent substantially similar outputs and reasonably varied).[55] Further, given the inherently opaque nature of Generative AI models and the unpredictable behavior of human users, Congress may consider legislation that establishes a liability framework that shields Generative AI services from liability when users intentionally submit infringement-driven queries.[56] In the same vein, the providers of Generative AI services should not automatically obtain legal knowledge of infringement based on user input alone. Moreover, providers should not acquire legal knowledge based on what they learn from improving their data sets and safeguards to prevent copyright abuses.

A notice-and-takedown system akin to Section 512(c) may not align perfectly with the context of Generative AI, except in cases involving specific images present in the training datasets. If the USCO contemplates this approach for a Generative AI safe harbor, it should ensure that the responsibility to identify specific images in the training datasets remains with the copyright holders who can demonstrate legal ownership. Rights holders should not have the ability to assert that AI models might contain copies of their protected works solely because the models generally source data from the open Internet.[57] Likewise, Generative AI providers shouldn't be tasked with monitoring rights holders' collections of protected works

---

[53]Gil Appel, et. al., "Generative AI Has an Intellectual Property Problem" Harvard Business Review, (2023). Available at:
https://hbr.org/2023/04/generative-ai-has-an-intellectual-property-problem
[54] Joel Matteson, "Unfair Misuse: How Section 512 of the DMCA Allows Abuse of the Copyright Fair Use Doctrine and How to Fix It" Santa Clara High Technology Law Journal, (2018). Available at:
https://digitalcommons.law.scu.edu/cgi/viewcontent.cgi?article=1638&context=chtlj
[55] Pam Samuelson, "Generative AI meets copyright" Science, (2023). Available at:
https://www.science.org/doi/10.1126/science.adi0656
[56] Jeremy Goldman, "Do AI generators infringe? Three new lawsuits consider this mega question" FRANKFURT KURNIT KLEIN & SELZ PC, (2023). Available at:
https://ipandmedialaw.fkks.com/post/102i5io/do-ai-generators-infringe-three-new-lawsuits-consider-this-mega-question
[57] As was the case with the Andersen lawsuit. *Andersen v. Stability AI Ltd.*, *supra* note 17.

for exclusion.[58] The responsibility to safeguard specific protected works must squarely lie with the rights holder.

In practice, a safe harbor could enable rights holders to choose not to have their works used for training or to request the removal of their specific works once they realize they're part of the training data. This realization could arise from examining open-source training sets or identifying outputs closely resembling or containing elements of their specific protected works. However, the USCO should note that AI-image generators with larger training datasets are less prone to produce infringing outputs.[59] Therefore, providing rights holders the option to opt-out or withdraw from training sets might not be beneficial in the long run. Furthermore, this approach could limit the expressive capabilities of Generative AI services.

When determining secondary liability for Generative AI service providers, policymakers must recognize the distinct technological aspects that differentiate Generative AI from UGC services. These nuances present unique challenges in curbing copyright misuse. Policymakers must craft a balanced liability standard that protects rights holders while supporting the growth of Generative AI and its expressive potential.

3. Question Eight: 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.

The previous discussion honed in on the fair use implications of Generative AI outputs. Here, we'll delve deeper into how fair use doctrine interacts with the datasets AI models leverage during their training.

It's well understood that AI models have an insatiable appetite for data, scouring the vast expanses of the internet to obtain a rich tapestry of human context, enabling them to produce mesmerizing digital creations.[60] However, much of this online content, while publicly accessible, is protected by copyright law. This fact

---

[58] Policymakers should expressly denounce such a monitoring requirement as does 17 U.S.C. § 512(m).

[59] Gowthami Somepalli, "Diffusion Art or Digital Forgery? Investigating Data Replication in Diffusion Models," (2022). Available at: https://arxiv.org/pdf/2212.03860.pdf

[60] Sukhpal Singh Gill, et al., "AI for Next Generation Computing: Emerging Trends and Future Directions," (2022). Available at: https://arxiv.org/pdf/2203.04159.pdf

sits at the heart of ongoing copyright disputes targeting Generative AI platforms.[61]

A. Intermediate Copying Has Always Been Fair Use

One crucial differentiation in this debate is the intent behind copying: is it an intermediate step to enable new expression or merely an exploitation of existing copyrighted content?[62] Notably, even if an entire copyrighted work is copied during an intermediate step, it may still be considered fair use if the eventual output doesn't infringe on any rights.[63] This nuance isn't exclusive to Generative AI but has historical roots tracing back to the dawn of search engines. Consider *Field v. Google* where the Court concluded that Google's act of scraping websites to deliver search results was a form of fair use, emphasizing the transformative nature of caching.[64]

Similarly, the ruling in *Authors Guild v. Google Books* recognized that copying for improved search capabilities aligns with fair use principles.[65] Fast forward to Google *v. Oracle*, where the Supreme Court portrayed Google's replication of Oracle's Java APIs as non-expressive, posing a significant barrier for copyright plaintiffs.[66]

Drawing parallels with the precedent set by Google's caching mechanisms, Stability AI argues that their act of copying is a transitional measure intended to birth transformative creations. The *Andersen* Plaintiffs, on the other hand, advocate that the sheer act of mass copying through web scraping should be seen as infringement — a stance that current copyright rulings don't universally support.[67]

---

[61]*Andersen v. Stability AI Ltd.*, *supra* note 17., see also *Getty Images (US), Inc. v. Stability AI, Inc.,* *supra* note 29.

[62]Henderson, Peter and Li, Xuechen and Jurafsky, Dan and Hashimoto, Tatsunori and Lemley, Mark A. and Liang, Percy, Foundation Models and Fair Use (March 27, 2023). Stanford Law and Economics Olin Working Paper No. 584, Available at: https://ssrn.com/abstract=4404340 or http://dx.doi.org/10.2139/ssrn.4404340

[63] *Sega Enterprises Ltd. v. Accolade, Inc.*, 9th Cir. 1992; Sony Computer Entertainment v. Connectix Corp., 9th Cir. 2000.

[64] *Field v. Google Inc.,* 412 F. Supp. 2d 1106, 77 U.S.P.Q.2d 1738, 19 ILRD 355, 2006 ILRC 1037 (D. Nev. 2006), Court Opinion

[65] *Authors Guild v. Google, Inc.*, *supra* note 27.

[66]*Google LLC v. Oracle America, Inc.,* 141 S. Ct. 1183 (2021).

[67] *Andersen v. Stability AI Ltd.*, *supra* note 17.

Yet, a shadow of ambiguity persists. If copyright holders can pinpoint distinct infringements, Generative AI providers like Stability AI would be compelled to underscore their non-expressive and transformative intents, emphasizing that their compact replicas aren't exact surrogates for the original content.[68] These are nuanced, fact-driven arguments that might be ill-suited for quick judicial resolutions.

B. Training Sets Do Not Disrupt the Market for Current or Derivative Works

Additionally, judicial evaluations will also weigh in on how the defendant's usage might disrupt the potential market or value of the copyrighted or derivative works. The crux here is that any market disruption should stem from infringement, not mere competition.

Revisiting *Field v. Google*, the Court emphasized that Google's caching had no bearing on the market prospects of Field's creations.[69] Extrapolating this to the realm of Generative AI, a versatile model with a broad training set will likely fare better than models singularly focused on specific artists or genres. For instance, an AI service centered around Taylor Swift's artistry, fueled exclusively by her copyrighted material, might struggle to make a case for fair use.[70]

In addition to copyright law, there are other legal provisions addressing unauthorized access to protected materials such as trespass to chattels and the Computer Fraud and Abuse Act (CFAA). However, if data acquisition by Generative AI providers is blanketly labeled as infringement, we might witness the curtains fall on AI image generation. But, such a stance could also threaten the basic functionality of the Internet. And perhaps, for some rights holders, that's the endgame they're rooting for.

Conclusion

Generative AI has already changed the landscape of how art and research will be conducted. By encouraging artists to use these platforms through the Fair Use Doctrine will allow for greater economic innovation and an increase in welfare in American society.

---

[68] *Kelly v. Arriba Soft Corp.*, 280 F.3d 934 (9th Cir. 2002), finding fair use where the thumbnail images generated from copies could not substitute for full-size images and served a different purpose than the originals.
[69] *Field v. Google Inc.,* supra note 65.
[70] Will Knight, "Algorithms Can Now Mimic Any Artist. Some Artists Hate It" Wired, (2022). Available at: https://www.wired.com/story/artists-rage-against-machines-that-mimic-their-work/

Thank you for your leadership and consideration. We look forward to any discussion and are available to answer questions regarding our answers to your Inquiry.

Sincerely,

Jess Miers
Legal Advocacy Counsel
Chamber of Progress