



Writers Guild of America West and Writers Guild of America East Comment On USCO Notice of Inquiry on Copyright & Artificial Intelligence

Writers Guild of America West (“WGAW”) and Writers Guild of America East (“WGAE”) (jointly, “WGA”) are pleased to submit the following comments to the U.S. Copyright Office (USCO) in response to the Notice of Inquiry on Copyright and Artificial Intelligence. The WGAW is a labor union representing over 11,000 writers in the television, film, news, and streaming video industries. The WGAE is a labor union representing over 7,000 members in film, television, news, podcasts, and online media. WGA members recently concluded a 148-day strike in which artificial intelligence (AI) was a key bargaining subject. The new WGA agreement is one of the first collective bargaining agreements in the U.S. to restrict the use of AI in a creative industry. WGA has participated in numerous policy discussions around AI, including the USCO’s Copyright and Artificial Intelligence Audiovisual Works Listening Session on May 17, 2023.¹ Our comments will focus on the protections necessary to protect human artistry in the evolving AI landscape.

The current and potential implementations of AI technology threaten the craft and livelihood of our members. Film and television writers are still learning the full extent of this technology’s impacts, and the technology itself is still rapidly developing. Writers have established new protections related to AI through collective bargaining, but public policy solutions that prioritize human creativity and authorship are needed as well.

Training AI Models on Copyrighted Material is Not Fair Use

Question 8.1. In light of the Supreme Court’s recent decisions in Google v. Oracle America and Andy Warhol Foundation v. Goldsmith, how should the “purpose and character” of the use of copyrighted works to train an AI model be evaluated? What is the relevant use to be analyzed? Do different stages of training, such as pre-training and fine-tuning, raise different considerations under the first fair use factor?

Question 8.5. Under the fourth factor of the fair use analysis, how should the effect on the potential market for or value of a copyrighted work used to train an AI model be measured? Should the inquiry be whether the outputs of the AI system incorporating the model compete with a particular copyrighted work, the body of works of the same author, or the market for that general class of works?

WGA believes the use of copyrighted material to train AI models should not be protected by fair use. Under the first fair use factor, it is inadequate to claim that the purpose of AI training with copyrighted material is non-commercial because the training process was initially conducted by

¹ U.S. Copyright Office, *Copyright and Artificial Intelligence Audiovisual Works Listening Session*, (May 17, 2023), <https://www.copyright.gov/ai/transcripts/230517-Copyright-and-AI-Audiovisual-Works-Session.pdf>.

a non-commercial entity. The development of AI models is heavily intertwined with commercial use regardless of the entity that initiated the training. Many AI models have been developed by non-profit organizations, but are later easily adapted for commercial use. For example, OpenAI built its ChatGPT technology while operating exclusively as a non-profit, but the company is now licensing that same technology for commercial use under its for-profit arm, OpenAI, L.P.² In some cases, AI research and training are conducted by academic institutions but funded by commercial entities that plan to use that research to launch a commercial product. Stability AI, a for-profit company, funded the development of its AI product through a research group at the Ludwig Maximilian University of Munich.³ Once copyrighted material is ingested by an AI model, it is difficult to remove that material from the model when the use of the AI output becomes clearly commercial. Any *potential* use of copyrighted material by an AI model must be considered when assessing the first fair use factor, not just the initial use.

In addition, the resulting outputs from AI models are not sufficiently transformative of the original copyrighted work under the first of the four statutory fair use factors.⁴ AI models have been fed film and television scripts; some AI-supported products are explicitly intended to generate similar scripts.⁵ These AI-generated scripts mimic and regurgitate the work of human writers. Early studies have already found instances of text plagiarism in AI language models.⁶ An AI model does not have the capacity to create original work based on human experience—it can only exploit the existing work of human writers.

Furthermore, AI-generated scripts trained on copyrighted work can directly compete with human-authored scripts, affecting the commercial market for writers' work. Film and television writers invest significant time and labor to compete in the market for writing services. Writers often specialize in writing for a specific category of project or genre, and develop reputations for their writing styles. A studio interested in producing a project in a specific writer's style has an incentive to hire that writer. But AI models allow companies to generate work that draws from an individual writer's style when prompted to do so—without the writer's involvement. If fair use is applied to allow unauthorized ingestion of copyrighted works for AI training, writers will be forced to compete with AI-generated scripts trained on their work, without their authorization, and without fair compensation.

Application of the fourth statutory fair use factor must thus consider how the AI-generated outputs compete with the general class of works. Companies buy WGA members' work with the assumption that motion pictures—derivative works based on scripts writers create—are what bring in revenue. Companies may be able to circumvent the need for scripts with generative AI

² Devin Coldewey, *OpenAI shifts from nonprofit to 'capped-profit' to attract capital*, TechCrunch (Mar. 11, 2023), <https://techcrunch.com/2019/03/11/openai-shifts-from-nonprofit-to-capped-profit-to-attract-capital/>.

³ Andy Baio, *AI Data Laundering: How Academic and Nonprofit Researchers Shield Tech Companies from Accountability*, Waxy (Sep. 30, 2022), <https://waxy.org/2022/09/ai-data-laundering-how-academic-and-nonprofit-researchers-shield-tech-companies-from-accountability/>; Acknowledgement by the Computer Vision & Learning Group on GitHub of donation from Stability AI: <https://github.com/CompVis/stable-diffusion>.

⁴ See 17 U.S.C. § 107.

⁵ For example, NolanAI encourages users on the highest paid tier to “let Nolan generate the initial draft of your scene.” <https://www.nolanai.app/pricing>.

⁶ Jooyoung Lee, Thai Le, Jinghui Chen, and Dongwon Lee. *Do Language Models Plagiarize?*, In Proceedings of the ACM Web Conference 2023 (WWW'23), May 1–5, 2023, Austin, TX, USA. ACM, New York, NY, USA, 12 pages. <https://arxiv.org/pdf/2203.07618.pdf>.

models that produce fully formed films or television series. Writers may be forced to compete with both AI-generated scripts and AI-generated films and television series, both of which used writers' original work without their consent.

- The USCO should not consider fair use an adequate defense for feeding copyrighted material into an AI model without authorization.

The Use of Creative Work for AI Training Requires Affirmative Consent from Creators

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

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

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

Without a fair use defense, developers of AI models will be required to seek a license or other authorization from copyright owners to use copyrighted works to train AI models. Legislators and regulators like USCO should also require AI trainers to seek affirmative consent from the human creators of such works, whether or not those creators are the copyright owners.

The dominant production model in Hollywood is based on structures that have been in place for almost a century. In the 1930s, Hollywood writers organized for collective bargaining purposes as employees under the National Labor Relations Act.⁷ The corollary to this development is that most unionized writers, as employees, create works for hire, in which the copyright vests initially in their employer. Through collective bargaining, unionized writers have been able to negotiate important *contractual* rights in their work—relating, *inter alia*, to credits and attribution, compensation for reuse, and the right to produce derivative works—that exist alongside the copyright holder's statutory rights. Because of these contractual rights, and of the labor market considerations discussed in the previous section, writers are directly affected by the prospect of their work being fed into an AI model and they should be given a right to consent to such use, whether or not they are copyright holders.

A system of consent that involves creators and copyright owners should be opt-in rather than opt-out. There are numerous companies developing and training AI models; to be forced to continually track all of them down is an enormous administrative burden for individual artists. An opt-in system would allow creators and copyright owners to negotiate for fair-market valuation and licensing of the use of their individual works. WGA also has concerns about whether a framework of collective licensing would adequately compensate creators for the value of their work.

As detailed in the above section on fair use, the development of AI models will often lead to future commercial uses, even if the initial use is non-commercial. It is difficult to remove source

⁷ 29 U.S.C. § 151 *et seq.*

material in an AI model once it has been trained on such material, and thus it will be challenging or nearly impossible to integrate consent at later stages of development. An opt-in consent system should be required for all uses.

In sum, affirmative consent for use of copyrighted work in training AI models should:

- Include human creators even if they are not the copyright owners;
- Be an opt-in system rather than opt out; and
- Be required for all uses, not only instances where the initial use is commercial.

An Effective Consent and Compensation System Requires AI Models to Disclose Training Datasets

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

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

A system of affirmative consent will not function as intended unless the developers of AI models are transparent about what materials were used to train the model. Already, many writers have been unable to identify with certainty whether their copyrighted works have been fed into existing AI models. In some cases, it is unclear whether an AI model ingested a writer's work or another source that mentioned the work, or both. Any developer building an AI system should be required to disclose the material used to train the system, especially if that material is copyrighted. This disclosure should be made in a clear, publicly available, accessible, and searchable format that is legible to a layperson. The disclosure should include data that is obtained from third parties.

In addition, AI systems should preserve and disclose the prompting that leads to a specific output. This disclosure will be necessary to determine whether copyrighted material or a creator's written voice was appropriated to generate a specific output. As discussed above, prompts that include creators' names or their work are especially concerning for film and television writers who make their livelihoods from their skills, reputations, and distinctive style.

Transparency alone is not a sufficient protection for creative workers, but it is a necessary step. Preserving and disclosing AI datasets and prompts will provide the foundation for an affirmative consent model by requiring AI developers accurately to track the inputs to their models and giving creators the ability to understand where their work was used.

In sum, AI models should be required to disclose in a publicly accessible manner:

- Material, including copyrighted material or material sourced from third parties, that was used to train the model; and

- Prompts that led to a specific output, especially if those prompts included copyrighted material or the names of human creators.

AI-Generated Work Should Not Be Not Copyrightable

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

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

Fundamentally, AI models cannot create new works. Such models do not write from lived experience: only humans can experience emotion like joy, embarrassment, pain, and euphoria, and only humans can translate those experiences into human-to-human connection through stories. AI models have been fed human-authored works, including those of our members, and use that work to mimic human storytelling. The protection of human artistry through copyright is so valuable to American society that it is enshrined in the Constitution; granting copyright to AI-generated work is antithetical to the purpose of copyright law. WGA agrees with USCO’s view that “copyright can protect only material that is the product of human creativity.”⁸

Using technology that exists right now, AI models can be prompted to write a script “in the style of” one of our members or their work. This can be done without the writer’s knowledge, consent, or compensation, because these models were trained on writers’ work. As discussed above, writers earn their livelihoods from their distinct style and skills developed over time. AI-generated material that appropriates that style and skill can compete with human creators or damage their reputations, or both. Any work that was generated by an AI model using prompts containing artists’ names or their work threatens the livelihoods of human creators, even if the AI-generated work is later edited or arranged by a human.

On the issue of copyrightability of AI-generated work, WGA:

- Agrees with the U.S. Copyright Office’s interpretation that only human-authored material is eligible for copyright; and
- Has additional, independent objections to USCO granting copyright to material that was created by prompting “in the style of” a particular human creator or their work.

Conclusion

Copyright law exists to “promote the progress of . . . useful arts” by protecting the rights of “authors and inventors.”⁹ The advent of artificial intelligence presents new challenges to this

⁸ United States Copyright Office, *Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence*, (Mar. 16, 2023), https://copyright.gov/ai/ai_policy_guidance.pdf.

⁹ U.S. CONST. art I, § 8, cl. 8.

constitutional mandate, but the guiding principle of protecting human artistry must be preserved. Film and television writers are on the frontline in experiencing the impacts of this new technology. Their experiences can provide insight into the type of regulation that will be required to protect human artists. It is clear to writers that feeding their copyrighted work into AI models without their consent is theft—not fair use—and that any system of affirmative consent needs to include human creators, whether or not they are copyright owners. In keeping with the spirit of copyright law and the values of the Constitution, any policy adopted by USCO or other federal agency should prioritize human authorship and creativity.

Respectfully submitted,

/Laura Blum-Smith/

Laura Blum-Smith
Director of Research and Public Policy
Writers Guild of America West

/Erica Knox/

Erica Knox
Research and Policy Analyst
Writers Guild of America West

/Lowell Peterson/

Lowell Peterson
Executive Director
Writers Guild of America East