Artificial Intelligence and Copyright Comments

Date: Oct 16, 2023
Organization: Teacher Created Materials, Inc.
Role: Educational Publisher
Submitter:
Diana Kenney
VP, Business Contracts
dkenney@tcmpub.com

Numbers below refer to numbered questions in Request for Comments document.

1. Q: 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?
Comment: It will be very important to clearly define how human authors can use generative AI systems to produce copyrightable material. What human activity is being defined as the authorship? Modifying prompts? Writing original content that isn't produced, at all, by generative AI? Does this mean that authors should only use AI for research but cannot use any output? The use of this technology is going to help creators, researchers, and the public create content that will never need to be copyrighted—research, communications, contracts, etc. Copyright owners are at risk if the material that's used for training ends up in output that is then violating copyright (being sold, copyrighted, used for monetary gains).

9. Q: 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)?
Comment: Yes! Authors, Illustrators, and other developers (e.g., audio) are starting to demand contractual clauses from publishers restricting the use of their content for training AI. As a publisher, we need to be able to control where and how our content is used to protect our copyrights and our contractual obligations.

9.1 Q: Should consent of the copyright owner be required for all uses of copyrighted works to train AI models or only commercial uses?
Comment: Yes—for all.

9.2 Q: If an "opt out" approach were adopted, how would that process work for a copyright owner who objected to the use of their works for training? Are there technical tools that might facilitate this process, such as a technical flag or metadata indicating that an automated service should not collect and store a work for AI training uses?
Comment: Opt-out should not be used. It puts the responsibility on the creator, which is the opposite of our current system. If someone wants to use copyrighted material for any purpose, it should be their responsibility to seek permission/license.

9.4 Q: If an objection is not honored, what remedies should be available? Are existing remedies for infringement appropriate or should there be a separate cause of action?
Comment: This should be treated just like infringement.

9.5 Q: 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?
Comment: If a work was made-for-hire, then an entity owns copyright. Humans are still involved, and copyright is held/exists. Publishers in this scenario should have the right to object to their content being used for training AI.

10. Q: If copyright owners' consent is required to train generative AI models, how can or should licenses be obtained?
Comment: Consent should be handled how permissions and licensing are currently handled—directly through the publisher or through companies like Copyright Clearance Center or MOSAIQ.

10.3 Q: Should Congress consider establishing a compulsory licensing regime? If so, what should such a regime look like? What activities should the license cover, what works would be subject to the license, and would copyright owners have the ability to opt out? How should royalty rates and terms be set, allocated, reported and distributed?
Comment: This doesn't seem necessary. If it is, then it should be part of the copyright registration process (though, that's probably unconstitutional and an undue burden since you don't technically need to register to own a copyright).

12. Q: Is it possible or feasible to identify the degree to which a particular work contributes to a particular output from a generative AI system? Please explain.
Comment: This must be possible. There are applications that can analyze text for plagiarism. Wouldn't this be similar coding?

13. Q: What would be the economic impacts of a licensing requirement on the development and adoption of generative AI systems?
Comment: There is cost associated with producing content (as publishers). Why wouldn't these developers also have a cost associated with developing their product?

15. Q: 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?
Comment: yes; ISBN, title, copyright holder, publisher, date of publication (the usual metadata)

16. Q: What obligations, if any, should there be to notify copyright owners that their works have been used to train an AI model?
Comment: Notification should have been a request before use. Since this didn't happen, yes, they should notify copyright owners.

18. Q: Under copyright law, are there circumstances when a human using a generative AI system should be considered the "author" of material produced by the system? If so, what factors are relevant to that determination? For example, is selecting what material an AI model is trained on

and/or providing an iterative series of text commands or prompts sufficient to claim authorship of the resulting output?
Comment: It seems that this described scenario would be very difficult to document and validate.

19. Q: Are any revisions to the Copyright Act necessary to clarify the human authorship requirement or to provide additional standards to determine when content including AI-generated material is subject to copyright protection?
Comment: Based on question 18 (above), a clear definition of human authorship is needed.

20. Q: Is legal protection for AI-generated material desirable as a policy matter? Is legal protection for AI-generated material necessary to encourage development of generative AI technologies and systems? Does existing copyright protection for computer code that operates a generative AI system provide sufficient incentives?
Comment: AI-generated material should not be copyrightable. There should be transparency when AI-generated material is being used so that users/readers can assess whether it's a reliable source (e.g., disclaimer that content is AI-generated). The code being created in the application should be legally protected (e.g., copyright or patent) as long as the system was trained without infringing on copyrights.

23. Q: Is the substantial similarity test adequate to address claims of infringement based on outputs from a generative AI system, or is some other standard appropriate or necessary?
Comment: It depends on the efficacy of that test and the definition of "substantial."

24. Q: 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?
Comment: The existing situation makes defense of material nearly impossible. If they didn't keep track, the only way we can identify infringement is if generated content is matched with existing copyrighted material. That seems impossible to undertake as a publisher.

25. Q: If AI-generated material is found to infringe a copyrighted work, who should be directly or secondarily liable—the developer of a generative AI model, the developer of the system incorporating that model, end users of the system, or other parties?
Comment: Without any disclaimers on the output, the liability would be with the developers. If they add a disclaimer, the liability would pivot to the user.

28. Q: Should the law require AI-generated material to be labeled or otherwise publicly identified as being generated by AI? If so, in what context should the requirement apply and how should it work?
Comment: Yes! Citing sources is not a new concept.

28.1 Q: Who should be responsible for identifying a work as AI-generated?
Comment: The "creator."

28.2 Q: Are there technical or practical barriers to labeling or identification requirements?
Comment: This should be handled similar to how citing sources for text or licensing for photos is handled now. Every book contains a credits page. If AI generated content is used in publication, it should be cited.

32. Q: 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?
Comment: Current copyright law doesn't protect things like an artist's style. It only protects the finished work.

33. Q: With respect to sound recordings, how does section 114(b) of the Copyright Act relate to state law, such as state right of publicity laws? Does this issue require legislative attention in the context of generative AI?
Comment: We just need to make sure that sound recordings are used with permission.

Overarching comments:
Permissions with compensation need to be required. Teacher Created Materials, Inc., has authors, illustrators, and developers (e.g., audio) who are including clauses in contracts asking that their content not be used for training AI.

Liability for infringement should belong to the developers of the generative AI system. They need to build in checks and balances that prevent that from happening. Or, warnings should be included that the similarities exist and it can't be copyrighted. In those cases, a user moving forward with that content in ways that break copyright would be their own liability.