Resource Guide for the Copyright Office Artificial Intelligence Initiative

James Gatto, Sheppard Mullin

Jgatto@sheppardmullin.com

SheppardMullin

CONTENTS

Copyright Office Artificial Intelligence Initiative Resource Guide

Using generative AI raises numerous copyright issues. In response, the US Copyright Office (USCO) has undertaken a new Artificial Intelligence Initiative ("AI Initiative"). This guide is a high level overview and a collection of materials relating to the AI Initiative.

New Artificial Intelligence Initiative

On March 16, 2023, the USCO launched the new AI Initiative to examine the copyright law and policy issues raised by artificial intelligence (AI), including the scope of copyright in works generated using AI tools and using copyrighted materials in AI training. According to the USCO: "This initiative is in direct response to the recent striking advances in generative AI technologies and their rapidly growing use by individuals and businesses." It is also a response to requests from Congress and the public.

Pursuant to this AI Initiative, the USCO will be holding public listening sessions:

- April 19, 2023 – Literary Works Listening Session
- May 2, 2023 – Visual Arts Listening Session
- May 17, 2023 – Audiovisual Works Listening Session
- May 31, 2023 – Music and Sound Recordings Listening Session

Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence

The USCO issued guidance covering copyrightability and registration issues raised by works produced by or with generative AI. We have provided a summary of this guidance here.

Some of the key issues from the guidance are:

- Copyright can protect only material that is the product of human creativity – the term "author" excludes non-humans
- The focus is on whether the AI contributions result from "mechanical reproduction" or instead of an author's "own original mental conception, to which [the author] gave visible form"
- If a work's traditional elements of authorship were produced by a machine, the Office will not register it
- Where a work containing AI-generated material also contains sufficient human authorship to support a copyright claim, copyright will only protect the human-authored aspects of the work, which are "independent of" and do not affect the copyright status of the AI-generated material itself
- This policy does not prohibit the user of technological tools in the creative process
- Applicants have a duty to disclose the inclusion of AI-generated content in a work submitted for registration and to briefly explain the human author's contributions to the work
- Individuals who use AI technology in creating a work may claim copyright protection for their own contributions to that work and describe the authorship contributed by a human
- Applicants should not list an AI technology or the company that provided it as an author or co-author
- AI-generated content that is more than *de minimis* should be explicitly excluded from the application
- Applicants must update pending applications and issued registrations to correct the record

Recent USCO Actions on Applications with Generative AI Materials

The USCO has acted in two high profile matters.

In the first application, Stephen Thaler attempted to obtain a copyright registration for "A Recent Entrance to Paradise," an AI-generated work that is the output of Thaler's AI system known as Creativity Machine. The copyright application listed the Creativity Machine as the author. This application was rejected by the Copyright Office based on its assertion that the author must be human. A copy of the Board Decision can be found here. Thaler filed a complaint in U.S. District Court seeking to overturn this refusal. That case remains pending.

In a second application, Kristina Kashtanova sought a copyright for the comic book "Zarya of the Dawn." The registration listed Kashtanova as the author and it was approved. Later, the Copyright Office learned that some materials were AI-generated. It reconsidered the registration after providing Kashtanova an opportunity to supply additional information. The result was that the copyright registration was maintained for the creative aspects provided by Kashtanova (the text and the selection, coordination, and arrangement of the work's written and visual elements) but canceled regarding the AI-generate images." A copy of the Registration Decision is here.

Past Activity

This is not the first foray by the USCO into the interplay between copyright and AI.

On October 26, 2021, the U.S. Copyright Office and the U.S. Patent and Trademark Office held a conference on machine learning and copyright law covering how existing copyright laws apply to the training of artificial intelligence and forward looking issues. Here is video of the conference.

On February 5, 2020, the USCO and the World Intellectual Property Organization (WIPO), held a symposium that dove into the use of AI to create original works and considered, the level of human input sufficient for the resulting work to be eligible for copyright protection and using copyright-protected works to train AI models. Here is video of the symposium.

Recent Litigations Involving Generative AI

Several lawsuits have been filed over use of copyright-protected materials. Some lawsuits of interest are:

- *Getty Images v. Stability AI* – Getty alleges misuse of over 12 million Getty photos to train the Stable Diffusion AI image-generation system. Here is a copy of the complaint

- *Anderson v. Stability AI* – putative class action lawsuit filed alleging that various AI companies of copyright infringement for scraping billions of images from the internet (including Ms. Anderson's art) to train AI models. This case also raises issues about whether copying an artist' style can be infringement. Here is a copy of the complaint

- *Doe v. GitHub* – another putative class action with various claims, including allegations of ongoing DMCA and open-source code license violations arising from GitHub's product called Copilot and the OpenAI product called Codex, which translate natural language requests into computer code. The tool is allegedly trained by scanning millions to billions of lines of open source computer code accessible on the Internet. While the code may be free to use, there are license conditions that allegedly are not complied with. Here is a copy of the complaint

Overview of Some Generative AI Issues

The following is an overview of some of the copyright issues that can arise with generative AI.

1. Collection of Training Materials

   a. Does the use of copyrighted materials to train an AI model constitute infringement? AI requires vast quantities of materials to train AI models. If these materials contain copyright-protected materials, is it copyright infringement to use these materials to train the model?

   b. Assuming there would be infringement, is fair use a defense? Many argue that using copyright-protected materials to train AI models constitutes fair use as it is transformative. Many cite *Authors Guild v. Google* to support this. Others distinguish the *Google* case arguing the model is trained with images and the system outputs new images, thus affecting the market for the images, which is another part of the fair use test that was not an issue in *Google*

2. License compliance

   a. If the content is licensed (e.g., a Creative Commons), what happens if the AI company does not comply with the license conditions? 3 of 6 Creative Commons licenses prohibit commercial use. Is using Creative Commons licensed content to train AI models a commercial use? Does that violate the license?

   b. If the content is open source code, are compliance obligations relevant and if so, how must AI tools comply. Many open source licenses have various compliance obligations but it is not always clear which obligations are relevant and how the AI tool can comply with these obligations

   c. If the content used to train the AI model includes name, image and likeness of celebrities, could that violate the right of publicity?

3. Protectability of Content

   a. Does a human prompt to an AI tool mean the human is an author of the output? The answer will be fact specific, but it is important to assess whether the input is just an "idea" for what the output should be or is sufficient human authorship

   b. Can AI-generated works be protected? According to the USCO guidance this requires looking at the work's traditional elements of authorship. If these elements were produced by a machine, the work lacks human authorship and the USCO will not register it. For example, when AI technology receives solely a prompt from a human and the AI produces complex written, visual, or musical works in response, this lacks human authorship

   c. Can works containing AI-generated material and materials with sufficient human authorship be protected? According to the USCO guidance, copyright will only protect the human-authored aspects of the work, which are "independent of" and do not affect the copyright status of the AI-generated material itself. AI-generated content that is more than *de minimis* should be explicitly excluded from the application. For example, this may apply when a human selects or arranges AI-generated material in a sufficiently creative way that the resulting work as a whole constitutes an original work of authorship or an artist modifies material originally generated by AI technology to such a degree that the modifications meet the standard for copyright protection

4. Liability for Infringement by Output of AI Tool

   a. If the content generated by AI tool infringes on copyright, who is liable? If the content generated by the AI tool infringes on copyright, who is liable? Is it the user who prompted the AI tool for the content or the AI tool developer who programmed the tool?

   b. If the content is based on Creative Commons or open source-licensed work, but does not comply with the license obligations, is that breach of contract? If so, who is liable? Some generative AI tools' terms of use state that users are responsible for infringement. Whether that will be sufficient to shift liability to users remains to be seen and likely will be fact specific

   c. Does breach of a license, terminate the right to use the content? With some open source and other licenses, the license terminates if the user breaches the license. This raise the question of whether this terminates the AI tool provider's right to use the materials covered by that license

5. What are the remedies for infringement?

   a. Besides the usual damage issues for copyright infringement that would apply to the output, can an AI tool be forced to cease using content to train the AI if use of that content is an infringement?

   b. Is algorithmic disgorgement appropriate? Some argue that if an AI tool improperly uses copyright-protected content resulting in a finding of infringement, algorithmic disgorgement should be imposed

   c. What is algorithmic disgorgement? Algorithmic disgorgement is a penalty that can be levied against companies that improperly use content to build algorithmic systems like AI and machine-learning models. This may require destroying various content, algorithms and/or models based on the improper use. In 2022, the FTC settled with a company for artificial intelligence/privacy violations imposing algorithmic disgorgement. This was not the first time the FTC applied this remedy

The foregoing is provided for information only and should not be construed as legal advice

SheppardMullin

Brussels | Century City | Chicago | Dallas | Houston | London | Los Angeles | New York | Orange County

San Diego (Downtown) | San Diego (Del Mar) | San Francisco | Seoul | Shanghai | Silicon Valley | Washington, D.C.

www.sheppardmullin.com