Thorny AI ownership questions have Copyright Office seeking public input

reader comments

58 with

On Wednesday, the US Copyright Office began seeking public comment on issues surrounding generative AI systems and copyright. The public comment period, which starts on August 30, aims to explore the complex intersection of AI technology with copyright laws, and it closes on November 15. The comments could inform how the agency decides to grant copyrights in the future.

Outlined in a 24-page document published as a PDF file by the Federal Register, the “Notice of inquiry and request for comments” asks questions with far-reaching consequences for intellectual property in America.

Specifically, the Office is interested in four main areas: the use of copyrighted materials to train AI models and whether this constitutes infringement; the extent to which AI-generated content should or could be copyrighted, particularly when a human has exercised some degree of control over the AI model’s operations; how liability should be applied if AI-generated content infringes on existing copyrights; and the impact of AI mimicking voices or styles of human artists, which, while not strictly copyright issues, may engage state laws related to publicity rights and unfair competition.

Comments are sought to address these concerns because, in a relatively short period, we've seen the arrival of generative AI tools with surprising capabilities that can render images and video, imitate voices, answer questions, and compose text on demand. AI models have spawned hype and backlash, heavy corporate investment, and lawsuits. Those instances have put the Copyright Office on high alert as the frequency of AI-related copyright incidents increases:

> *Over the past several years, the Office has begun to receive applications to register works containing AI-generated material, some of which name AI systems as an author or co-author. At the same time, copyright owners have brought infringement claims against AI companies based on the training process for, and outputs derived from, generative AI systems.*
>
> *As concerns and uncertainties mount, Congress and the Copyright Office have been contacted by many stakeholders with diverse views. The Office has publicly announced a broad initiative earlier this year to explore these issues. This Notice is part of that initiative and builds on the Office's research, expertise, and prior work, as well as information that stakeholders have provided to the Office.*

Listed among the incidents noted in the Copyright Office document are stories we've reported on in the past, including the first registration of latent diffusion artwork to Kris Kashtanova, which was granted in September and conditionally revoked in February; an attempt to register an AI-created artwork solely to a machine by Stephen Thaler (who has attempted similar things with patents), which a judge declined this month; and a copyright lawsuit against OpenAI on behalf of Sarah Silverman regarding the company's use of her copyrighted works in ChatGPT training data. Additionally, artists sued Stability AI, Midjourney, and DeviantArt as a class action in January for alleged copyright violations.

Advertisement
In response to similar copyright concerns, several news organizations (including Ars Technica's parent company, Conde Nast) recently took steps to block OpenAI's web crawler, hoping it will prevent OpenAI from scraping data used to train future AI models. Most large language models, such as the ones that power ChatGPT, learned about the world by absorbing billions of documents found openly online—typically without consulting copyright holders.

The flexible arm of the law

Enlarge / The cover of "Zarya of the Dawn," a comic book created using Midjourney AI image synthesis in 2022 and subject to a copyright dispute.
Kris Kashtanova
While US copyright law is fixed, its interpretation of emerging technology is nuanced and can change depending on how regulators (and judges) feel about the topic, which depends partially on public sentiment.

As the interpretation stands, AI artwork may be copyrighted in the US as part of a larger, human-authored work, but not individually. And we've seen that copyright does not extend to

creative works without any human authorship. However, there are still unresolved questions about what degree of human involvement would trigger the threshold for copyright registration, similar to questions that arose upon the invention of the camera in the 19th century.

In 1884, the defendant in *Burrow-Giles Lithographic Co. v. Sarony* claimed that photographs could not be subject to copyright because a photo is "a reproduction on paper of the exact features of some natural object or of some person." They argued that a photo is the work of a machine and not a creative expression. But the court ruled that photos can be copyrighted because they are "representatives of original intellectual conceptions of [an] author."

The question today is: What action in using a machine learning tool counts as the "original intellectual conception of an author?" Is it the selection of the particular AI tool? The input of a prompt? The curation or editing of the resulting machine-generated output? While current US copyright guidelines seem to sidestep these questions, the Copyright Office is open to hearing the public's opinions.

Advertisement

Also, the Copyright Office would like to hear opinions about the unsettled issue over the inclusion of copyrighted material in training data sets, notably writing, "the Office is aware that there is disagreement about whether or when the use of copyrighted works to develop datasets for training AI models (in both generative and non-generative systems) is infringing."

But that's not all. The original 24-page request document is a fascinating look into the minds of current US copyright regulators, and it includes 34 multi-part questions of interest for the Copyright Office that would take another 24-page document to straighten out in detail.

The deadline for written comments is October 18, and they can be submitted online through a Regulations.gov form. Reply comments must be submitted to the Copyright Office by November 15.

https://arstechnica.com/?p=1964248

OpenAI disputes authors' claims that every ChatGPT response is a derivative work