

October 18, 2023

U.S. Copyright Office
101 Independence Ave S.E.
Washington, DC 20559-6000

RE: [Artificial Intelligence and Copyright: Request For Comments](#)

Dear Members of the U.S. Copyright Office:

I am writing to you today as both the wife of an illustrator in the entertainment industry and as a community organizer. My husband's job, along with thousands of others in film, art, entertainment and so many other related creative industries are already dramatically affected by generative artificial intelligence (genAI).

As you well know, an AI doesn't work unless it has a dataset. Human artists, are the true creative core at the heart of genAI. In fact, no AI machine would be able to output **anything** if it were not filled with billions of images and copyrighted works which were collected from across the web, without the knowledge of consent from their creators. Data from these images and copyrighted works is what powers genAI. GenAI is then used to replace the human creators of those stolen images and copyrighted works in the workplace.

I hope this letter helps to address some of the questions that matter to me in your "Request For Comment."

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?*

GenAI technologies are directly competing against artists and creators of all kinds. Here is a link to a recent [FTC hearing](#). I draw your attention to the comments by Neil Clarke (ClarksWorld Magazine), Umair Kazi (Authors Guild) and Douglas Preston (Author). They did an incredible job explaining the issues of how this is affecting creators and copyright holders.

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)?*

There is ONLY one option that works - and that is an Opt-In system. This is because it's entirely impractical for any one individual to try to keep track of each and every one of these entities who are stealing the works of artists to be used as "data fuel" for the commercial use of others.

Most all of these AI companies have zero transparency into where they acquired the datasets for their AI models, so even if there was a way for an artist to check if their works were included in a dataset, they cannot do so. Additionally, the AI companies say that once an AI model has ingested an image, that the image cannot be removed. Opting-Out is entirely useless, because it's only Opted-Out from future trainings, and the stolen works are already part of the previous datasets being built upon.

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?*

Consent of the copyright owner should be required for ALL uses of copyrighted works to train AI models. There is no guarantee that outputs generated from any AI model will not be used for commercial purposes, and if you look at past practices that have already occurred - AI models that were purportedly to be used for "research" purposes, ended up being used for commercial purposes. So yes, consent of the copyright owner MUST be required for all uses of copyrighted works.

Question #9.3: *What legal, technical, or practical obstacles are there to establishing or using such a process? Given the volume of works used in training, is it feasible to get consent in advance from copyright owners?*

Given that these AI companies are valued in the billions of dollars, and that they will be displacing or degrading approximately 300 million jobs ([according to a report back in March of 2023 by Goldman-Sachs](#)), I believe they have the time, resources and responsibility to hire as many individuals as they can, paid at a living wage (in US dollars) to go out and acquire consent from each and every copyright holder whose works they wish to use.

Question #10: *If copyright owners' consent is required to train generative AI models, how can or should licenses be obtained?*

This can be patterned after currently existing licensing structures that already exist throughout many well established and regulated creative industries. It is entirely reasonable to expect that AI companies are treated with the same requirements and responsibilities that are commonplace and would be expected of any other individual who wanted to license a piece of music or work of art.

Question #11: *What legal, technical or practical issues might there be with respect to obtaining appropriate licenses for training? Who, if anyone, should be responsible for securing them (for example when the curator of a training dataset, the developer who trains an AI model, and the company employing that model in an AI system are different entities and may have different commercial or noncommercial roles)?*

This actually goes back to question 9.1 (above). This scenario used in your question, is exactly why it's imperative that consent of the copyright owner be required for ALL uses of copyrighted works to train AI models. To this question I would say, I believe the responsibility of acquiring consent should lie with the developer who trains an AI model. This is because, this is the part of the process where the data is ingested and it becomes much more difficult to remove something that should not be there.

While a curator may want to compile a number of works for and AI developer to use, if the AI company doesn't want to be held responsible for data that should not be incorporated, they should have to weed bad data or data that has yet to be authorized for use out of their datasets before it's put into the AI model.

The curator should have the responsibility to provide the AI developer with a list of how the works in the dataset were acquired (meaning the process used), where the works were acquired (meaning which websites and their corresponding URLs each item came from), the means by which the works were acquired (ie- web crawlers, etc.), the names and contact information of the original creators of the works (meaning the human creators who originated the works, not just a random person who shared an image they liked on a Facebook page).

This way, if a curator was acquiring data through predatory means, as has already been the case, the AI developer could simply refuse the materials from the curators. Bad actors who steal data indiscriminately, and through predatory means, would have no market for their ill gotten wares.

Additionally, the AI developer is the one who is using that data as fuel to create their product of an AI model, it's their model that will be damaged if they give it bad data, so they need to be extra careful not to damage their own product, since they are the manufacturer of that product.

It should also though, be incumbent upon the customer, the company employing that AI model to require proof from the developer of the AI model of consent for each and every work that has been included in the dataset.

I do think there needs to be some central place, and I don't know what that would be exactly, but there must be some central place where copyright holders can check which of their works have been used and where they have been used, so that they can track where their data is used, if they have given permission for it to be used, how that permission was acquired, and where they can file a grievance if something bad has occurred.

Question #14: *Please describe any other factors you believe are relevant with respect to potential copyright liability for training AI models.*

Here is one other example of a way that users of AI can also be in jeopardy and potentially liable for a copyright violation.

[AI Prompter Financially Liable] - Say an AI prompt person had an AI generate an image for them. And let's say the image ended up being nearly identical to a piece of copyrighted material. And let's say, the prompt person was unaware just how similar the image was to the original image and so they listed it for sale on a website. And let's say the copyright holder then went after AI prompter them for selling their work and they ended up having to pay the copyright holder. This would all have been because the AI company stole the data and put the AI prompter in jeopardy.

Question #27: Please describe any other issues that you believe policymakers should consider with respect to potential copyright liability based on AI-generated output.

Legislation regulating AI should require all existing generative AI algorithms that do not use public domain or consensually-obtained data to be algorithmically disgorged / destroyed, and companies should be required to start over with datasets that are fair, ethical and legal. Datasets used by AI products must have consent of the creators of the works ingested, credit those creators, compensate those creators, and be transparent about the contents of their datasets and how they were acquired.

Because of the previous mentioned predatory behaviors in the commercial sector, with respect to large foundational data acquisitions; trust from the public, especially the artistic and creative communities in the private sector is waning. Aggressive accountability standards for data acquisition must be immediately implemented by the government at all levels for every kind of generative AI, if we hope to truly safeguard people's rights (including intellectual property rights). We need to harness and regulate AI technology before it becomes too expansive to control.

I am not looking to hinder the progress of technological innovation. As I had stated at the beginning of my letter, my aim is only to protect those most vulnerable to generative AI's impact if it's left unregulated and unchecked. There are ethical ways genAI technologies can co-exist with human creatives and I want to thank the US Copyright office for letting artists and creatives weigh in on this matter. Thank you for your consideration.

Sincerely,
Deana Igelsrud