R·I·T

Rochester Institute of Technology

Medical Illustration
College of Health Sciences & Technology
CBET 75-2129
153 Lomb Memorial Drive
Rochester, NY 14623-5608
585-475-2443

21 October, 2023

United States Copyright Office
Library of Congress

Re: Notice of inquiry: Artificial Intelligence and Copyright

To whom it may concern,

I am writing to share my thoughts on copyright policy issues raised by generative Artificial Intelligence (AI) systems. I have been a practicing Medical Illustrator for over 30 years. In my profession, I have created (and continue to create) a large number of copyrightable works. I have also been a Professor of Medical Illustration for 25 years. For most of that time, I have taught copyright law to my students to help them protect their own intellectual property rights and to avoid infringing the rights of others. Therefore, I have an interest in generative AI as a creator of intellectual property, as an educator of other creatives, and because of my academic interest in the law.

I am responding to the Copyright Office's Notice of inquiry and request for comments on the issues raised by generative AI. The Copyright Office is seeking input specifically on the following issues:

(1) the use of copyrighted works to train AI models;
(2) the copyrightability of material generated using AI systems;
(3) potential liability for infringing works generated using AI systems; and
(4) the treatment of generative AI outputs that imitate the identity or style of human artists.

Let me address these issues in order.

(1) the use of copyrighted works to train AI models
In my opinion, the "use" of copyrighted works to train generative AI systems ("scraping") is, by definition, "usage", and therefore requires permission or compensation to the owner of those works. Many corporations, including social media sites and content aggregators, seem to asknowledge this fact and have inserted "opt-out" clauses in their privacy settings. By accepting the privacy terms, users unknowingly grant these sites the right to use their content to train AI systems.

But not all of the material that is "scraped" by AI systems comes from social media sites. Much of it is obtained from public sites where the users have not granted permission for its use (either knowingly or unknowingly). In these cases, copyright owners must be compensated for the use of their work. In my opinion, this sort of usage is similar to reprographic use of copyrighted works. At one time, this referred specifically to the large-scale photocopying of works, but now can include electronic distribution of copyrighted works. Copyright owners are compensated with reprographic royalties when their works are photocopied or distributed electronically. Numerous organizations exist to collect these royalties from users and distribute them to copyright owners. They have relationships through the International

Federation for Reproductive Rights Organizations (IFRRO) to return reprographic royalties to foreign rightsholder and to collect royalties generated internationally for distribution to their U.S. clients. Royalties generated from the scraping of works for AI could be collected and distributed in a similar manner.

In the U.S., one such reprographic rights organization is the Copyright Clearance Center. However, the CCC should not be the only U.S. entity that collects and distributes such royalties. They do not currently recognize visual artists as rightsholders, even though visual artists often retain the rights to their artwork when it is published. Although the CCC returns millions of dollars in royalties each year to publishers and authors, in their 45-year existence they have not returned one penny of royalties to visual artists. Artists like myself have turned to other organizations, such as the Artists Rights Society (ARS), to collect reprographic royalties on our behalf. Currently ARS is unable to collect reprographic royalties generated in the U.S. because CCC refuses to acknowledge artists as rightsholders. Instead, all of ARS's revenue comes from overseas, from IFRRO-member organizations to *do* acknowledge artists as copyright owners. Over just the past three years, I've personally received thousands of dollars in reprographic royalties from France, Spain, Germany, and the UK (all through the Artists Rights Society), but never a penny from the CCC for use of my work in the United States.

I ask the Copyright Office to keep this gross inequity in mind as they contemplate a system to compensate rightsholders for the scraping of their work by AI systems. Any such system must recognize ALL creators as rightsholders, visual artists included.

Some might argue that the use of copyrighted works to train AI systems falls under the Fair Use provision of the law and, therefore, does not require payment of royalties to copyright owners. However, scraping of works by AI systems fails all four tests for fair use, and therefore cannot be deemed Fair Use of those works. The four factors to determine fair use are:

a. The purpose and character of use
b. The nature of the copyrighted work
c. The amount or substantiality of the portion used
d. The effect of the use on the potential market for or value of the work

a. The purpose and character of use. The scraping of copyrighted works to train generative AI systems is clearly a commercial use of copyrighted works, which tends to argue against Fair Use. Although the images produced by AI systems may be used for non-commercial purposes, the AI systems themselves are clearly meant to generate a profit (whether by charging fees or generating advertising revenue). Some of these systems may not be generating revenue yet, but it is only a matter of time before they are monetized. Therefore, the *scraping* of copyrighted works is commercial, even if the ultimate use of some of the generative AI content is not.

Second, the character of the use is not transformative. Once again, some of the content that is generated by these systems may be transformative, but the act of scraping and warehousing existing works is not, in and of itself, transformative. It is mere copying – for commercial purposes – and, therefore, fails this test for Fair Use.

b. The nature of the copyrighted work. Historically, the courts have given greater protection to unpublished works since it is at the author's discretion if their work is disseminated or not. In order for copyrighted works to be scanned by an AI system, they need to be accessible in some public medium. Therefore, we can assume that most works scanned by AI systems are published works. This tends to argue in favor of Fair Use.

On the other hand, “Use of a work that is commercially available specifically for the educational market is generally disfavored and is unlikely to be considered a fair use.”[1] This is certainly true of the work I create as a Medical Illustrator. The same is true for anyone who creates written or artistic work specifically for the educational market. Furthermore, “Courts are usually more protective of art, music, poetry, feature films, and other creative works…”[1]. Much of the concern about generative AI is that it trains itself on existing creative works in order to generate new works. This was one of the primary reasons for the recent strikes by SAG-AFTRA and the Writers Guild of America[2]. There was widespread concern about the use of generative AI to scrape existing creative works in order to automatically generate new content. The widespread scraping of creative works, as well as the use of works available for the educational market, argue against a Fair Use defense.

c. The amount or substantiality of the portion used. Presumably, AI systems will scrape works in their entirety. This includes scraping written works, music, and images. It is hard to conceive of a system that would “know” to scan only part of a text, sample just a portion of a song, or copy just one part of an image. The use of works in their entirety argues strongly against Fair Use.

d. The effect of the use on the potential market for or value of the work. The widespread use of generative AI represents a clear threat to the livelihood of individuals, like myself, who make creative content. My existing body of medical artwork has significant market value. Why would anyone need to license my existing work when they can generate it themselves (using an AI system which, ironically, trained itself using my images)?

One of the rights guaranteed by copyright law is the right to make derivatives of existing work. This further increases the potential value of my work. By creating new work after training themselves using existing work, AI systems, in effect, are creating derivatives of existing work, further diluting the market value of that work. As already mentioned, this was one of the main reasons for the Hollywood actors and writers strike. There was widespread concern that movie studios could simply scrape existing content to generate new scripts and entirely new movies. Generative AI can and will have a massive impact on the potential market for existing work, including the very work that it uses to train itself. This argues very strongly against Fair Use.

(2) the copyrightability of material generated using AI systems

It seems that the Copyright Office has already provided clear guidance on this issue. “In the Office's view, it is well-established that copyright can protect only material that is the product of human creativity. Most fundamentally, the term ‘author,’ which is used in both the Constitution and the Copyright Act, excludes non-humans.”[3]. I believe this is the correct interpretation of the law.

The whole point of copyright law is to promote creativity. It does this by affording protections to the individual works of authors. But ultimately it’s not so much about protecting the works themselves as it is about protecting the *act* of imagination and the *effort* of fixing creative ideas in tangible form. Generative AI can spit out infinite numbers of “works”, but these works involve no imagination and no creative effort.

Although I create all of my medical artwork on a computer, I still “draw” (using a stylus), I decide where every object is placed, and I carefully choose every color to create shading, texture, and realism. I make all of the same movements and aesthetic decisions as if I were working in charcoal or watercolor. Only the medium is different. The same is true of photography. Although they use a device to capture images, photographers must choose their subject, compose and frame the shot, and choose settings to create the desired mood or effect. In both of these examples, the computer and the camera are simply tools – no

different from a pencil or paintbrush – that enable artists to turn their ideas into tangible expressions. Generative AI does none of this.

The only human creativity in AI is coming up with a set of text prompts. I suppose these prompts could be subject to copyright if they are sufficiently unique and creative. But I fear, if this were allowed, the Copyright Office would be inundated with trivial submissions like "the Pope in a biker jacket" or "chimp with a glass of wine." I don't believe these sorts of simple prompts are worthy of protection.

(3) potential liability for infringing works generated using AI systems

Clearly there's a chance that works generated by an AI system might infringe existing works. Even a human can inadvertently create something that is substantially similar to an existing work. I suspect this will be even more common with AI systems because they are "trained" using existing works, although I don't have hard evidence to back that up.

If an AI-generated work is "substantially similar" to an existing work, then clearly it infringes the existing work. One important question, however, is who is "guilty" of the infringement? Is it the AI system that generated the work, the user of the infringing work, or do they share liability? If the infringing work is found to be substantially similar to one of the works that was used to train the AI system, then it seems that the AI system itself should bear more (or all) of the liability. It should be the responsibility of the AI system to ensure that it does not replicate any existing work too closely.

It is tempting to say that the user of an AI-generated work bears some responsibility to research the work prior to using it and to ensure that it is not similar to existing works. This may be possible with text or music, which are easily searchable, but it is very difficult to search for existing images (artwork, photographs, designs, etc.). The Copyright Office itself does not have a database that is searchable by image. One must know the name of the artist and/or title of the image in order to locate its registration. Although tools exist to search for images (e.g., TinEye or Google's reverse image search), they are far from perfect. As an example, I have two illustrations that have been widely infringed (mostly in other countries where I have little recourse to stop the infringement). I've used TinEye and Google reverse image search to locate these infringements. Each of these tools yields completely different search results, suggesting that neither of them is locating all occurrences of these images. This underscores the difficulty a user would have in determining if a newly-generated AI image matched any existing images. This supports my argument that the AI system itself bears more of the responsibility to ensure that it does not infringe existing images, especially those images that were used to train the system.

A more interesting question, in my opinion, is whether infringement of a work by an AI system can be considered "willful". Clearly a machine doesn't have a "will". But the creators of generative AI systems are aware that their machines are training themselves on existing works. They must know there's a risk that their systems will occasionally infringe existing works, however rare this might be. Likewise, the users of generative AI systems must be aware that those systems are trained using existing works. This has been all over the media for the past year. A reasonable person would know there's some risk of creating an infringing work. For these reasons, I feel strongly that infringement by AI-generated works must be considered willful.

The question of willful infringement is extremely important for those of us who create copyrightable works. When my work is infringed, I have little incentive to pursue the infringer if all I hope to collect is actual damages. The cost of hiring an attorney and filing in Federal court far exceeds what I could collect in damages alone. The only time it makes sense to pursue infringement (beyond simply sending a demand letter) is if there's a chance to collect statutory damages and attorneys fees. This depends on a finding of willful infringement. If all cases of AI infringement are determined to be "accidental", this makes it

nearly impossible for independent artists and creators to protect their copyrights. Copyright owners could try to use the new Copyright Claims Board to adjudicate their smaller claims, but this system is brand new as of 2022 and it's not clear how well it will work. Furthermore, respondents have the right to opt out of the CCB process, forcing claimants to file in Federal court[4]. This renders the CCB system virtually useless.

(4) the treatment of generative AI outputs that imitate the identity or style of human artists
I do not have strong opinions on this issue. I think it is regrettable that AI systems can generate art in the style of a particular artist and that these works might be passed off as original creations of that artist. If that happens, it constitutes fraud, but it is not the concern of the Copyright Office. Artists have always been able to produce work in the style of other artists; AI just does it faster.

Thank you for the opportunity to submit my comments on these issues.

James A. Perkins, MFA, CMI, FAMI
Board Certified Medical Illustrator
Fellow, Association of Medical Illustrators

Distinguished Professor and Graduate Director, Medical Illustration
Head of the Department of Medical Sciences Health & Management
College of Health Sciences & Technology
Rochester Institute of Technology
Rochester, NY

References

1. Crews, Kenneth D. (n.d.) "Fair Use". Columbia Copyright Advisory Services. Columbia University Libraries. https://copyright.columbia.edu/basics/fair-use.html

2. Wikipedia. (n.d.). "2023 Hollywood labor disputes".
https://en.wikipedia.org/wiki/2023_Hollywood_labor_disputes

3. Copyright Office (2023). "Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence." Federal Register document number 2023-05321.
https://www.federalregister.gov/documents/2023/03/16/2023-05321/copyright-registration-guidance-works-containing-material-generated-by-artificial-intelligence

4. Copyright Claims Board. "About the Copyright Claims Board."
https://ccb.gov/about/