My name is Christine Winckler; I write under the pseudonym Chris Winkle. I'm the editor in chief of Mythcreants, a small online publication for science fiction and fantasy writers. We write instructional material and offer fiction editing services. Previously, I worked as a web designer and developer.

2. Does the increasing use or distribution of AI-generated material raise any unique issues for your sector or industry as compared to other copyright stakeholders?

Writing advice online is a very competitive space. It is also considered a static or evergreen subject area. That means to produce content in this field, generative AI does not need to train on recent data, making the field particularly susceptible to disruption. However, not all writing advice is the same. In fact, it is an area where a great deal of innovation is happening, as people are starting to become aware that old models for writing and storytelling are inadequate.

In this space, Mythcreants focuses on high investment content. Our typical article 2,000-3,000 words long because we explore topics in great depth. In contrast, competing sites like Mythic Scribes and Writers Helping Writers typically publish articles between 1,000-1,500 words. Mythcreants innovates by creating new theories and terms to explain storytelling and help our audience engage readers. We also spend significant time locating real examples to help our audience learn.

However, it does not matter how high quality our content is if no one knows it exists. We depend on search engine listings to allow new readers to discover us. On a typical day, two thirds of our site traffic comes from organic searches. According to Statistica, Google has a 83-90% market share on online searches, so our ability to connect to potential readers is largely controlled by one company. Our competitors often succeed through the rapid production of low quality content designed to match search engine keywords. Generative AI allows these competitors to accelerate their quantity-over-quality strategy.

The acceleration of generative AI content farms with extremely low quality, inaccurate information has been noted by such respected publications as Futurism and The Verge. Notably, Futurism found CNET publishing inaccurate and plagiarized financial advice generated by AI without disclosing how the content was generated to users. Google did not penalize CNET in search rankings for this, which was seen as a green light for more websites to use this strategy even if misinformation and plagiarism resulted.

By crawling copyrighted material on the web and then generating content based on what it consumes, generative AI software takes the results of a business's investment in content and hands it to competitors who do not invest. This risks a tragedy of the commons scenario, or a race to the bottom, where businesses are under a great deal of competitive pressure to reap the benefit of communal knowledge without expending resources to contribute to it. As publications

like us go out of business or stop investing in content, this will lower the quality of content online.

This trend is likely to accelerate once generative AI inevitably begins training on its own output. As there currently is no reliable way for web crawlers to identify content generated by AI, training on AI generated content is inevitable as long as training data is gathered by indiscriminately scraping the web rather than licensing training content. A study published by Cornell University showed that generated AI output began to degrade after training on its content over five iterations.

However, the risks to our business from competing publications, even generative AI content farms, may pale in comparison to the risks that search engines such as Google pose by applying generative AI. A recent diary study on generative AI by user experience experts Nielsen Norman Group shows the strong business incentive entities such as Google have to use generative AI content while destroying the sources of training data that generative AI relies on.

In short, Nielsen Norman Group demonstrated that it is far more convenient for users to have information combined together and handed to them rather than visiting all of the websites that created that information. It is no surprise that Google's new generative search experience drastically reduces the presence of search listings in favor of generative AI answers to search queries. If Google wants to be competitive as a search engine, it has every reason to deprive websites such as Mythcreants of traffic. Many, if not a majority, of websites depend on search engine traffic to be viable, with up to 68% of all trackable traffic coming from search engines.

Nielsen Norman Group diary study showed users employed generative AI assistants for sophisticated tasks such as planning home school curricula. Producing such curricula for generative AI to train on would take any publication significant time and expertise. If those publications go out of business, the public will lose access to valuable knowledge.

Currently, the only way for sites like Mythcreants to preserve our investment in content is to erect walls around it, requiring users to login and agree to a terms of service to read content that was once publicly available. Large platforms such as Reddit and X are already placing restrictions on content access to prevent scraping for AI training.

A web where generative AI trains on copyrighted material without compensation is a web with fewer websites, lower quality content, inaccurate content, and more barriers to accessing content.

It is also a web with lower innovation in countless fields of human knowledge - technology is not the only sector where innovation occurs. We show our audience of fiction writers new ways of thinking about stories that help them learn and improve their skills. Without innovation, the same conventional wisdom that fails so many fiction writers today will still be failing them tomorrow.

Consumers who save time and effort by embracing generative AI will never know what knowledge or instruction would have been available if innovation was protected.

Industry: Fiction Writing

It is not unusual for a fiction writer to spend 20 years developing their skills. A writer may easily pay for editing, courses, and seminars that cost thousands of dollars. Mythcreants charges writers with a 100,000 word novel about $1000 for developmental editing or $2000 for copy editing. Some writers get three developmental edits and a copy edit before their novel is done. For most fiction writers, a single novel takes years of labor, and many writers will create multiple novels over decades without earning significant revenue. For a single person who hasn't made a cent selling fiction, this is an incredible investment.

The fiction writers we work with in editing are willing to put in this investment for two main reasons.

1. They are very emotionally involved in their work. They write stories that incorporate their grief after a loved one died, that embody their vision for a better world, or that express their lived experience and personal truth. For many of these writers, a generative AI that trains on their work without permission to reproduce pieces without credit is not simply denying them income. It is violating something sacred and personal. This is why so many fan writers at Archive of Our Own (AO3) were outraged that generative AI has been training off of their work, even though their work is free to read. To see some of this outrage, you can view the over 800 comments that were posted on this AO3 post before commenting was closed.
2. Fiction writers can become household names. Customers do not buy books because they were published by Penguin or Tor; they buy books because they were written by Stephan King or N.K. Jeminisin. But the ability of new writers to make their mark in this way will be greatly endangered if publishers can copy the style of famous writers for free and simply generate books based on market demand. Unlike individual writers, publishers have the funds for marketing and the connections for distribution. Why take a chance on a new writer when they can just copy prose that sells?

Fiction changes over time like any other field. Everyone benefits when a new generation of fiction writers with unique styles, voices, and viewpoints is discovered. But considering the grueling time and investment that is required to make it in this field, and the low likelihood of a liveable income already, generative AI could easily come to replace the next generation of fiction writers or dissuade them from putting in the investment necessary to hone their skills.

3. Please identify any papers or studies that you believe are relevant to this Notice. These may address, for example, the economic effects of generative AI on the creative industries or how different licensing regimes

do or could operate to remunerate copyright owners and/or creators for the use of their works in training AI models. The Office requests that commenters provide a hyperlink to the identified papers.

I used two scientific studies in writing these responses:

1. The Curse of Recursion: Training on Generated Data Makes Models Forget

2. Information Foraging with Generative AI: A Study of 3 Chatbots

5. Is new legislation warranted to address copyright or related issues with generative AI? If so, what should it entail? Specific proposals and legislative text are not necessary, but the Office welcomes any proposals or text for review.

New legislation is warranted for two reasons:

1. There is no federal right to publicity, allowing companies to intentionally mimic and sell a person's likeness and creative style. Creative professionals need a right to their personal brands and voice they have spent a lifetime investing in. New legislation should give all persons a right to their own image, voice, and personal artistic style. Legal or not, the theft of someone's personal image, voice, or style is violating.
2. Because fair use guidelines have been established by the court, their application can vary drastically, creating an uncertain business environment. Businesses need to know which behavior is legal and which is not. New legislation should clarify what constitutes fair use of copyrighted material.

8. Under what circumstances would the unauthorized use of copyrighted works to train AI models constitute fair use? Please discuss any case law you believe relevant to this question.

One factor in fair use is the impact on the market for the original work. Generative AI models allow the production of content at such a rate that its market impact will be enormous, if not astronomical. AI models that create materials with the potential compete with the original work, the author's body of works, or that general class of works should not constitute fair use for this reason.

8.5. Under the fourth factor of the fair use analysis, how should the effect on the potential market for or value of a copyrighted work used to train an AI model be measured? Should the inquiry be whether the outputs of the AI system incorporating the model compete with a particular copyrighted work, the body of works of the same author, or the market for that general class of works?

To protect innovation, fair use should only be applied when outputs of the AI system do not compete with the particular work, the author's body of works, or the general class of works. No one should have their work used, without permission, to create competing works that damage their livelihood. However, well-established artists, writers, and other creators may be less likely to incur significant market damage because of their high name recognition and high demand for their personal works.

However, new artists and writers must struggle to establish the name recognition that makes their livelihoods possible. It is these professionals who are most vulnerable to competition from mass production of derivative works. Flooding the market with AI outputs could prevent a new generation of creative professionals from becoming established and offering innovative ideas to the public.

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)?

Copyright owners should have to affirmatively consent, or opt in, for their works to be used in training materials, for the following reasons:
- Even the best opt out system imposes a burden on those who wish to opt out. This establishes a double standard by which the rights of those who are wealthy and well-organized are protected whereas the rights of the poor and vulnerable are not. The average person has little knowledge of the law and is unlikely to keep track of when and how they need to opt out. This doesn't mean they are agreeing to surrender their rights or that they do not deserve rights equal to everyone else.
- Opt out systems give companies a perverse incentive to make opting out difficult for copyright holders. Especially if there isn't a federal, standardized opt out system, it will be like canceling a subscription with a company today - likely employing tiny links and with several misleading buttons that abort the process. While it is possible to regulate opt out systems, this not only costs regulators time and money, but also requires regulators to play catch up every time a company finds a new way to impede their opt out system. For instance, Google technically allows websites to opt out of training data, but they have to also opt out of Google's search engine. Since this would render most websites obsolete, it is not an opt out system in any meaningful sense.

- An opt in system puts the burden of proof on companies gathering and using training data to prove their training data has been appropriately licensed. Without this, it will be easy for companies to claim that a copyright holder did not opt out, even if they did.
- An opt in system encourages companies to more carefully curate training data because they are paying for it. This will encourage them to train on high quality content containing less misinformation.

10.3. Should Congress consider establishing a compulsory licensing regime?

A compulsory licensing regime would not only be a violation of the moral rights of countless creators with works that are deeply personal, but also deny creators a fair market value for use of their works in training. AI companies should have to pay a fair market price to license content, and content creators should be rewarded for creating works that are in particularly high demand as training data. This will result in higher quality training data and high quality outputs.

15. 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?

Both developers of AI models and creators of training datasets should have an obligation to transparency.
- Without transparency, small copyright holders will have no practical way to protect their rights. While large copyright holders like Getty Image might spot their watermark on generated images, small copyright holders may not have knowledge of generated works that are employing the unique characteristics of their copyrighted art, writing, music, or other content. Without this knowledge, copyright isn't enforceable, but generated works will still negatively impact the market for similar content.
- Transparency also allows journalists and other third party watch dogs to better report on AI. Training data may determine to what extent generated content spreads misinformation or hateful content. Regardless of the fair use status of AI generated content, the public deserves to know how their content is being used, who is providing training data, and who is benefiting from that training data.

16. 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?

Authorship should be determined by the level of creative control a person uses when creating a work. A painter doesn't simply tell their brush to put a shadow under a tree, the painter determines the color and shape of the shadow. If the person leaves the creative control up to software, they are not the author of the work. However, if the person exercises control over the work at a detailed level, then generating content is not so different from using existing creative software like Photoshop to create an image. Using creative control as a standard allows professionals to benefit from the labor they invest in a work that uses AI generated components.

24. 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?

Civil discovery rules are not sufficient to address this situation. That's because discovery requires a lawsuit to already be filed. However, the strength of a copyright holder's case will determine which legal professionals are willing to file lawsuits on the copyright holder's behalf and whether the copyright holder decides to take the financial risk of a lawsuit. Many small creators do not file copyright lawsuits when someone violates their copyright because of the resources involved. Relying on civil discovery would allow those who violate the rights of small creators to do so with little accountability.

25. 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?

Any parties who gather training data, train a model on that data, or should have known about infringing data should have liability when the model infringes a copyrighted work. End users should not be liable unless they submit copyrighted material for the generative AI software to incorporate into its outputs. Otherwise, most end users do not have the means to discover if AI-generated material is infringing.

31. Should Congress establish a new federal right, similar to state law rights of publicity, that would apply to AI-generated material? If so, should it

preempt state laws or set a ceiling or floor for state law protections? What should be the contours of such a right?

Yes, a new federal right is needed to both protect the right to publicity and protect from defamation caused by AI-spread misinformation. This should be a floor for state law protections to allow policy innovation. States often create new policies that other states or the federal government later adopt. When change is rapid as it is with AI, this innovation is particularly valuable to the public.

This should ensure individuals have:
- The sole right to profit from their unique likeness, including face, body, voice, personality, or style, unless they license that ability to others.
- Protection from damage to their reputations caused by the reckless creation or dissemination of their likeness in generated works.
- Protection from fraud or deception through use of someone's likeness.

Coincidental likeness that doesn't generate additional revenue or inflict damage to real persons should not be penalized. Therefore the use of generated works should be protected unless it is demonstrated that those creating or disseminating the work should have known they were either profiting from the likeness to a real person, doing damage to a real person, or deceiving other persons through likeness to person.

32. 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?

These protections are absolutely essential to creators. Creators may spend a lifetime establishing a unique style that is not only deeply personal, but intrinsic to their public identity, and indispensable to their commercial value.

Thankfully, Getty Images has already demonstrated what a generative AI system that respects artists can look like. The Getty AI image generator does not have the data necessary to intentionally replicate an artist's style. This means training data has been scrubbed of any artist identifiers, so the system simply does not know which works are from which artists, and therefore cannot be specifically asked to mimic their works. This method should be mandatory for generative AI models.

There is no reason not to extend this protection to everyone, regardless of how well known they are. New creators should be able to build their reputation free of unauthorized copying of their style, and unknown creators will have further protection from the use of their style to engage in fraud or deception.

In cases where the user of a generative AI model submits works to be mimicked, the user should be liable for submitting copyrighted images. Generative AI software should be under an obligation to inform users of this liability. Normally, it is best to avoid any liability for users of software, but this would not be so different from standards today. If someone takes a copyrighted image and puts it into Photoshop to create a derivative work, they are liable for copyright violation, not Adobe.

There may also be causes where a user successfully uses text prompts to, for instance, coax generative AI software to mimic an artist's style even if the artist isn't named. There should be a federal law to protect creators against such mimicry, but in that case, it is reasonable to limit the protection to creators who have a reputation or notable work the violating party could profit from mimicking. This lowers the likelihood of liability due to coincidental likeness.