My name is Kristina Farnsworth. I am making these comments on my own behalf as a writer of creative works. I am seeking to publish my writing in the future and have concerns about how AI training will affect the creative writing industry.

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?**

I personally do not see the benefits of generative AI being able to produce material that could be created by a human author. I believe that it could be beneficial for AI to be used to archive creative works and analyze them for academic research purposes; but I do not believe there is any benefit to using AI to generate creative works. I believe this technology is more likely to harm the industry as well as creators by taking jobs from creators. I think this is especially true for freelance creators, such as artists. Instead of hiring artists, people are turning to AI generators for art for their creative projects. This is especially problematic when artists' works are being used to train AI generators without their consent. AI generated works that closely resemble an artist's or writer's style can also be damaging to that creator's reputation if the generated work is shared publicly.

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?**

I am a writer and I have a lot of concerns about the training of AI generators as well as their use. I do not want my writing to be used to train AI without my consent. When an AI is trained on a particular sample of writing, that AI is being trained to look for patterns to reproduce. These patterns can include writing style but it can also include characters' reactions and responses to different situations. The patterns that I use in my writing are unique to my writing and my stories. By reproducing these patterns, AI makes my writing less unique and less relevant. Why would anyone care about my writing if they can get the same product from an AI? It can also hurt my reputation as a writer if someone shares content that is similar to my writing style but contains content I would not normally produce (such as a different genre, explicit material, racial slurs, etc). Creators cannot control what their writing style is used for and that could be potentially damaging.

Also, not only are companies using writers' works to train their AI, once the AI generator is made public, there is nothing to stop someone from using large samples of an author's work as a prompt for an AI generator. Users are using authors' works without their permission to generate sequels and other stories. There are even instances of AI works being uploaded under the original author's name. This is very damaging to an author's reputation and livelihood.

According to the SEO.ai website "More and more, artificial intelligence (AI) is being used to automate mundane tasks such as writing articles, blog posts, and even books."

This attitude undervalues writers and undermines the amount of work, passion, and dedication they put into their work. AI companies seek to undermine individual creators by making them and their work seem mundane and unimportant. With this attitude, companies seek to profit off the idea that people are not needed because an AI can do the same "mundane" work at a lower cost.

Another quote from the SEO.ai website: (https://seo.ai/blog/ai-writing-articles)

> AI writing can provide benefits such as increased efficiency and productivity, improved consistency and accuracy, and the ability to generate large volumes of content quickly.

> Additionally, AI writing can help reduce costs associated with content creation and provide access to new types of content, such as personalized or dynamic content.

AI generators are using the works of actual writers, without their consent, in order to remove the need to pay writers to create works in the future.

3. **Are there any statutory or regulatory approaches that have been adopted or are under consideration in other countries that relate to copyright and AI that should be considered or avoided in the United States? [40] How important a factor is international consistency in this area across borders?**

I am not personally familiar with how this issue is being addressed in other countries, however I do believe there is value in international consistency. The internet is a global market. Internet users consume and share content from all over the world. I believe this should be taken into account when attempting to create legislature for this issue.

4. **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.**

I believe that current policies do not adequately address the issues surrounding the training and use of generative AI. I believe that the current fair use policies are not enough to adequately address this issue. It is important to also look at the ethical implications of generative AI. It is not enough to simply address the financial harm to creators. The use of a creator's works without their permission is unethical, especially if those works are being used to generate profit. I believe the rights of creators to share their works publicly should be protected, even if those creators are not sharing their works for profit. They have a right to not have their creative works stolen and reproduced by AI.

7. To the extent that it informs your views, please briefly describe your personal knowledge of the process by which AI models are trained. The Office is particularly interested in:

7.1. How are training materials used and/or reproduced when training an AI model? Please include your understanding of the nature and duration of any reproduction of works that

occur during the training process, as well as your views on the extent to which these activities implicate the exclusive rights of copyright owners.

AI are trained by providing input to the AI. The AI is evaluated based on the response to that input and then fed more input until the AI is able to generate a desired response. Many generative AI's provide users with "styles" to choose from. In order to be able to accurately reproduce these styles, the AI would need be fed large samples of that particular style, often from a single source. Because of this the AI's responses will often closely mimic the original input.

I believe that this cannot fall under the principles of fair use as large samples of a creator's work are often used without their consent. Also simply using a particular creator's style may hurt that creator's reputation. For example the AI art generator Wombo.ai gives the option to generate art using "Ghibli" style. This can be damaging to Studio Ghibli's reputation even if (perhaps especially if) the generated artworks do not closely match Studio Ghibli's art style. As stated previously, the copyright holder does not have any control over how their style is used. A user could use this generator to creator explicit content that does not match with the content Studio Ghibli normally creates. This could be damaging to their repuation.

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.

I do not believe most unauthorized use copyrighted works to train AI constitutes as fair use. Large samples of a creator's works are often used with the intent to reproduce their style and patterns of art or writing. This is often done in order for a company to sell services related to the use of their generative AI. Fair use would imply that the creator of the AI does not intend to profit from their software in the future or allow works generated by their AI to be shared publicly (as this sharing can be damaging to the copyright owners). Large samples of copyrighted material are often used to train AI and often the generated works are not transformative enough to be considered fair use. Output generated by AI is not meant to provide commentary on the original work and has no greater social benefit. Also many AI generators allow users to select the style of a popular author or artist without that author or artist's permission. These copyright owners are not compensated by the owner of the AI software.

I also believe that fair use does not apply even if a creator shares their creative works publicly for free. These creators still have a right to decide how their works can be used, shared, or redistributed. AI generators should not be able to profit off the works of artists who share their work on deviant art or public blogging spaces. Nor should AI generators be able to profit off the works of writers who share their writing on public blogging spaces or websites like AO3. Copywrite law should not only protect creators who are seeking to profit off their work now, but should also protect the potential of creators to profit off their work in the future, should they choose.

I believe creative works should be protected from copying and plagiarism, even if those works do not have an official copywrite on file.

8.1. In light of the Supreme Court's recent decisions in *Google* v. *Oracle America* [41] and *Andy Warhol Foundation* v. *Goldsmith*, [42] how should the "purpose and character" of the use of copyrighted works to train an AI model be evaluated? What is the relevant use to be analyzed? Do different stages of training, such as pre-training and fine-tuning, [43] raise different considerations under the first fair use factor?

I disagree with the supreme court's decision in Google vs Oracle. I believe the Google's use of the Oracle API's should not have constituted as fair use as Google's intention in using the API's was to profit off them. They should have used the proper channels to make an agreement for the use of the APIs.

Intention and character are important and our laws should reflect that. Laws exist to protect the rights of individuals. When a company uses the creative works of others without permission, their intent is to get something for nothing. They seek to profit off their AI software without having to jump through hoops to get what they want. They are seeking to take the easy way out in order to achieve profits for themselves without having to give up anything in return. This intention should not be rewarded. This is unethical and is not the type of business practice our laws should promote.

8.2. How should the analysis apply to entities that collect and distribute copyrighted material for training but may not themselves engage in the training?

These entities do not have the right to distribute the copyrighted materials of others and they certainly do not have the right to profit of the collection and distribution of copyrighted materials.

8.3. The use of copyrighted materials in a training dataset or to train generative AI models may be done for noncommercial or research purposes. [44] How should the fair use analysis apply if AI models or datasets are later adapted for use of a commercial nature? [45] Does it make a difference if funding for these noncommercial or research uses is provided by for-profit developers of AI systems?

I believe it could be considered fair use to gather or train materials for non-commercial research purposes (especially archival and academic purposes) however these datasets should not be allowed to be used later for commercial purposes.

8.4. What quantity of training materials do developers of generative AI models use for training? Does the volume of material used to train an AI model affect the fair use analysis? If so, how?

The volume of material used to train AI models should affect whether the AI training can be considered fair use. I believe that most AI training models use large samples from the same source and this should not constitute fair use as the AI would be too likely mimic the material from a particular source. It is also not fair use to explicitly allow users to select the style of a particular creator.

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? [46] 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?

I think all of the factors listed are important factors to consider. I belive that AI generators that produce creative works are hurting the creative industries as a whole as they are seeking to replace the need for individual creators. Software companies are using the works of creators without their permission in order to replace the need for those creators. Therefore the use of copyrighted and plagiarized materials is damaging to all creators in the industry and should not be considered fair use.

As stated previously, the use of copyrighted materials also has the potential to damage the reputations of individual creators and hurt their individual sales and profits.

The use of creative works without a creator's permission also limits that creators potential to make profits from their own work in the future, even if they are not currently selling their work. I believe fair use should not only evaluate whether the creator's marketability is being damaged now, but should also consider whether the creator's potential ability to market their skills is being hindered or limited.

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)?

Copyrighted materiels should not be used without the copyright owner's explicit consent. Creators should not have the responsibility of seeking out every AI generative software company in order to opt out of their AI training models. It should be the responsibility of the software company to seek out consent from copyright owners. This should include creators who have not filed for copyright but have published their work in some form online.

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 should be required for all AI training

9.4. If an objection is not honored, what remedies should be available? Are existing remedies for infringement appropriate or should there be a separate cause of action?

Creators should be able to easily appeal to have their work removed from a dataset or AI model without having to undergo the cost of taking legal action. Not all creatives necessarily want to sue for damages. Most would be happy to have their materials removed from the dataset and to have the AI unlearn their materials. Developers should be required to honor such appeals without legal action as this is costly and only hurts creators, especially those who make little to no profit of their current materials.

9.5. In cases where the human creator does not own the copyright—for example, because they have assigned it or because the work was made for hire—should they have a right to object to an AI model being trained on their work? If so, how would such a system work?

If it would affect their future marketability, I believe a creator should be able to object. For example if they created a piece of work as a commission, and that work was used to train an AI, that reduces the marketability of the creator to make future commissions. There should be a system for creators to appeal to have their work removed from the dataset. Software companies should be required to honor appeals without creators having to take legal actions.

13. What would be the economic impacts of a licensing requirement on the development and adoption of generative AI systems?

We should be less concerned with the impact licensing requirements would have on companies and corporations and we should be more concerned with securing and upholding the rights of individual creators. This is not a new concept. The music industry makes licensing agreements with songwriters. The movie industry makes licensing agreements with screenwriters. Why should the AI industry get to steal works from creators without being expected to make licensing agreements or pay the creators who contribute to their software? These other industries are not hurt because they have to pay their creators.

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?

Yes. An accurate list of all materials used in a dataset or training model should be publicly available and easy to find.

15.3. What obligations, if any, should be placed on developers of AI systems that incorporate models from third parties?

A list of all materials used in an AI dataset should be made available to the public. If a developer purchases a dataset they should make it clear to the public which dataset they are using and inform them of where they can find a list of materials.

15.4. What would be the cost or other impact of such a recordkeeping system for developers of AI models or systems, creators, consumers, or other relevant parties?

If an accurate list of materials us shared this should not negatively impact the company as they are not sharing their data simply a list of the materials used in the data.

16. What obligations, if any, should there be to notify copyright owners that their works have been used to train an AI model?

Companies should be required to notify the copyright holder if possible. They should be required to remove copyrighted material from their datasets and models upon request (without the copyright owner being required to sue).

It should be the responsibility of the company to ensure they are not using copyrighted materials. It should not be the responsibility of the copyright holder to ensure that their rights are not be ing violated.

18. 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?

An author could use AI software such as Grammarly to improve their own writing. They should be able to own the copyright to any changes they make based on suggestions from the software.

I do not believe users can expect to own the copywrite to works generated from simple prompts alone, even if the user spent a long time refining their prompts and adjusting filters on AI output.

Copyrighting the works generated by AI also incentivizes companies to replace actual writers and artists with AI software as the companies would be able to generate output and copyright that output for their own profit.

23. Is the substantial similarity test adequate to address claims of infringement based on outputs from a generative AI system, or is some other standard appropriate or necessary?

I believe this test is not adequate to determine if the output of an AI generator constitutes copyright infringement or fair use. While some outputs have been found to closely resemble their source material, many simply match the style or other patterns such as dialogue responses. However a copyright holder's reputation can be damaged by the output work, especially if that output work is shared online as having been produced by the original copyright holder.

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?

I do not believe there are currently adequate means for a copyright holder to prove that their work was used without permission. This is why developers should be required to share which materials were used to train their AI. It should not be the responsibility of the copyright holder to prove that their rights were violated. It should be the responsibility of the software company to prove that they did not use copyrighted materials without consent.

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?

If the copyrighted material was collected for the purpose of training an AI then the developer of the AI model, and the developer of the system incorporating the model should be directly responsible. If the copyrighted material was used by an end user as a prompt then the end user should be directly responsible, although the developer of the system should be secondarily responsible for not putting protections in place (ie many AI generators allow large samples of text to pasted in as a prompt, basically encouraging users to copy and paste copyrighted material). A terms of service that says "do not use copywrited materials" should not absolve developers from responsibility.

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?

As stated before, using a creator's style damages their reputation and therefore their marketability. An AI system should not be able to copy and reproduce such styles nor should they be able to advertise that they can reproduce a particular style (even if the model was not trained in that style).

Any creators should receive such protections.