

To Whom It May Concern,

I work in marketing; I do video editing, social media, graphic design and more, so I am one of the first who would be directly impacted by the negative changes AI will bring.

You are likely hearing from several developers of AI on this matter, some of whom may not be fully inclined to share the negative impacts of AI in good confidence in order to protect their product. You are also likely hearing from many in different fields, creative or otherwise, who are concerned about the impact this will have on their jobs and livelihoods and are therefore advocating for measures such as stricter copyright laws and clauses to protect intellectual properties from unauthorized AI data training. I will be among the latter because AI regulations, no matter how restrictive they may seem, are just too important to not implement.

I may not cover every specific section that you have outlined, but where I do offer my input, I will do to the best of my ability with the knowledge that I have in the hopes that again, it may convince you to strengthen existing copyright laws on a federal scale to reflect the current landscape being disrupted by AI.

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?

There are significant risks associated with generative AI technology. The most prominent of these is the justifiable concern that it will be used to replace human labor, thereby putting millions upon millions in the arts and creative centers out of work. Putting aside the fact that the quality of AI-generated works is limited at best, the prospect of quick, speedy content at the cost of people's livelihoods is not a desirable outcome, and one that many, myself included, are actively lobbying to prevent. Based on experts' predicted numbers, an estimated 45 million Americans could lose their jobs to AI by 2030, one quarter of the American workforce.¹ This is a smaller percentage than the rest of the world, but equally alarming. Over the next 10 years, on a global scale, AI could create 97 million jobs but destroy 375 million,¹ making it one of many cases where the negatives strongly outweigh the positives. These numbers would absolutely include marketers and creatives and would be among the first (but certainly not the last) to feel its negative consequences. We can see this issue playing out in the various lawsuits authors and artists have made against OpenAI and ChatGPT, among other open-source AI developers, as well as the WGA strike and the current SAG-AFTRA strike, with one of their key issues being limiting or prohibiting the use of AI in both writers' rooms and in a proposed system to pay a background actor for a day's work of being scanned for AI, only for that AI to be used without compensation to the individual, thereby destroying any authenticity and value a human creator's work contributes.

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?

The question does widely apply to marketing & creative fields, particularly the latter. In marketing, my current field, generative AI can better predict what target audiences are looking for via search engines and API, which can be greatly beneficial for those less focused on the technical side of the job, as it can otherwise be a time-consuming process. However, with generative AI being able to rapidly produce

¹ Zippia. "23+ Artificial Intelligence And Job Loss Statistics [2023]: How Job Automation Impacts the Workforce" Zippia.com. Jun. 11, 2023, <https://www.zippia.com/advice/ai-job-loss-statistics/>

content (albeit at varying quality and in many cases with much less engagement than human-made content), there is wide concern that companies, and to a great extent independent contractors, will view AI generated content as a cheap alternative to human labor, despite the fact it cannot be legally copyrighted, and as such threatens to cause a massive loss of job opportunities regardless of skill level.

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.

The following are prominent research articles I believe are relevant to this notice:

- [“Artificial Intelligence and Music: Open Questions of Copyright Law and Engineering Praxis”](#), by Bob L.T. Sturm, Maria Iglesias, Oded Ben-Tail, Marius Miron, and Emilia Gomez. This particular article may be relevant for the Copyright Office’s later questions pertaining to voice and likeness protections.
- [“Legal and human rights issues of AI: Gaps, challenges, and vulnerabilities”](#), by Rowena Rodrigues, published in the Journal of Responsible Technology, Volume 4, December 2020.

4. 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? How important a factor is international consistency in this area across borders?

There have been other statutory or regulatory approaches adopted or under consideration in other countries relating to copyright and AI. Among the examples shared, some should be considered in the United States, the others avoided:

- **Consider:** The Italian Data Protection Authority, in March of 2023, had officially (but temporarily) banned ChatGPT from use in Italy over data privacy concerns and accusations of its data collection practices. OpenAI complied with Italian regulators’ ban, blocking users in Italy from accessing the service for the duration of the ban until they had complied with the demands of the GPDP.
- **Consider:** The European Union (EU) has passed a draft of the A.I Act (voting on final legislation has not yet occurred and not expected to pass until the end of the year at the latest). The A.I Act, as it relates to the Copyright Office’s inquiries, includes requirements for AI developers to further disclose the data used to create their programs such as publishing summaries of copyright material used for training their systems and providing guardrails against the AI generating illegal content.² The legislation also includes a ban on many American companies providing API access for generative AI (most notably affected would be OpenAI and Google) without first complying with the provisions of the A.I Act, as well as provisions for third parties to sue national governments, such as the U.S, if any companies or developers operating in those countries were to provide AI services, among them open-source systems, without first complying with the entirety of the A.I Act, the most important provision being full transparency and disclosure of

² <https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence>

how and on what the AI systems are trained.³ The EU has been wrestling with AI legislation for the last couple of years, but the sudden emergence of ChatGPT has obviously forced governments worldwide to try and play catch-up with drafting AI regulations. As such, the EU's A.I Act would serve as a successful starting template for other governing bodies, including the Copyright Office and the United States Congress, to properly protect individual rights against AI and hold AI developers to greater accountability.

- **Consider:** While several governments worldwide are debating or fine-tuning AI legislation, China has successfully passed AI legislation of their own. While there are several pillars of AI regulation included in the legislation, the one most pertinent to the U.S Copyright Office, in relation to this inquiry, is the 2023 draft rules on generative AI, which determined that generative AI must not only be free of racial and gender bias, but also that training data for the AI does not violate IP rights, including training on copyrighted works.⁴ The updated terms for generative AI include measures such as labeling AI-generated content (Article 13) and placing the responsibility of pausing generation of (and removing) AI content that violates China's laws (Article 14).⁵ One caveat the Copyright Office should be aware of when implementing similar policies is Article 4, Section 1: "Uphold the Core Socialist Values",⁵ implying the responsibility of regulating AI should fall on, or align with, a single political system. It is entirely feasible for both the Copyright Office and the United States Congress to implement similar AI regulations that strengthen copyright laws and protect the rights of copyright holders, both their intellectual property and their personal data, while operating within its constitutional boundaries.
- **Consider:** Despite India previously pledging early in 2023 not to regulate AI in any way,⁶ the country appears to have reversed course with the Indian Digital Personal Data Protection (DPDP) Bill, which would prohibit generative AI platforms from processing the personal data of Indian citizens regardless of whether or not they are in the public domain, removing its previous exemption for search engines, and seem to have warned AI platforms they could face legal action if they violate the DPDP bill.⁷ Because the DPDP Bill's potential legal action against AI platforms was compared to the current legal battles generative AI faces in the U.S, it stands to reason that legislation in the United States would equally benefit from including language similar to the DPDP Bill in terms of accountability towards AI platforms. It is also important to note that the DPDP Bill emphasizes consent is key, something that applies to almost every legal and cultural issue we face today, including AI regulations.
- **Avoid:** Japanese copyright policymakers have recently ruled that in Japan, generative AI does not violate its copyright laws. It should be important to note here that while Japan has stronger enforcement of their copyright laws, those copyright laws in Japan have always been more lenient than the U.S, ergo adapting Japan's copyright policy worldwide would be disadvantageous. Furthermore, the language used that does open up legal action for copyright infringement puts too much of the onus on copyright holders and not enough on AI developers, mainly due to the fact that training data for AI models are not public,⁸ and therefore in a court of law, AI developers

³ <https://www.europarl.europa.eu/news/en/press-room/20230505IPR84904/ai-act-a-step-closer-to-the-first-rules-on-artificial-intelligence>

⁴ <https://carnegieendowment.org/2023/07/10/china-s-ai-regulations-and-how-they-get-made-pub-90117>

⁵ <https://www.chinalawtranslate.com/en/generative-ai-interim/>

⁶ <https://techcrunch.com/2023/04/05/india-opts-against-ai-regulation/>

⁷ [https://economictimes.indiatimes.com/tech/technology/personal-data-of-indians-in-public-domain-may-get-shielded-from-](https://economictimes.indiatimes.com/tech/technology/personal-data-of-indians-in-public-domain-may-get-shielded-from-ai/articleshow/101805843.cms?utm_source=contentofinterest&utm_medium=text&utm_campaign=cppst)

[ai/articleshow/101805843.cms?utm_source=contentofinterest&utm_medium=text&utm_campaign=cppst](https://economictimes.indiatimes.com/tech/technology/personal-data-of-indians-in-public-domain-may-get-shielded-from-ai/articleshow/101805843.cms?utm_source=contentofinterest&utm_medium=text&utm_campaign=cppst)

⁸ <https://restofworld.org/2023/japans-new-ai-rules-favor-copycats-over-artists/>

have an unfair advantage over copyright holders due to the lack of available information for discovery. Another important note is that Representative Takashi Kii of Japan acknowledged that despite Japan's current copyright law permitting generative AI, there would be a need for "new regulations to protect copyright holders",⁹ implying that their legislators are leaving the door open to the possibility of reversing course on its generative AI stance. When taking new legislation into consideration, the U.S Copyright Office should align less with the overall rigid legislation of Japan's copyright laws and more on advocates for Japan's copyright holders, such as the Japan Photographic Copyright Association, who argue for strong revisions of their constitutional amendments (specifically the 2018 amendment allowing for AI data scraping without the permission of copyright holders)¹⁰, as well as protecting copyright holders from generative AI scraping its data without consent, as stated in a joint statement from the aforementioned JPCA and the Japan Magazine Publishers Association and Japan Book Publishers Association.¹¹

In light of more governments working towards tightening and strongly enforcing AI regulation versus less, it is important for both the Copyright Office and the United States Congress to strongly consider international consistency in terms of AI and work towards adapting these regulations into U.S law.

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.

As previously stated, new legislation, as well as strengthening current legislation, is not only warranted in addressing copyright and generative AI issues but needed. In addition to adapting the regulations displayed in the previous question (namely China and the EU's AI regulations), there are other methods that could be uniquely applied to U.S Copyright law:

A major concern in every field affected by AI, not just creative fields, is mass job loss and layoffs, the idea being that an AI can replace a human worker. A great way to ensure this does not happen, or at least that the damage is limited, is to implement a clause stating that any company who fires, let's go of, or terminates an employee, provided they did not violate other company policies, will be held liable in wrongful termination lawsuits. Settlements for these lawsuits should be determined by the position of the individual prior to being laid off and the individual's living status. Should there not be a practical way to enforce such regulations, an alternative would be adhering to a set cap of how much labor AI is allowed to do in any given work environment (i.e., stating only 2-3% of job duties and tasks are permitted to be done with generative AI, with the remaining 97-98% requiring human labor only).

Another major concern is how works or products that are both human-made and AI-generated or AI-assisted would impact the human creator's monetary compensation, which in current theory would be significantly less for simply polishing up or editing something that is already generated by AI. The solutions to this issue would ideally be either strengthening copyright law to fully exclude AI-generated or AI-assisted content from being copyrighted or, if a copyrighted work is shared with AI, that the human in question receive a significant increase in monetary compensation for the loss or ordinary work. This method was recently implemented in writers' rooms as a result of the recently-resolved WGA strike as one of many guaranteed AI protections, giving the provision that individual writers would receive the

⁹ <https://finance.yahoo.com/news/ai-art-wars-japan-says-185350499.html>

¹⁰ <https://www.yomiuri.co.jp/editorial/20230613-OYT1T50164/>

¹¹ <https://www.japantimes.co.jp/news/2023/08/17/japan/crime-legal/japan-publisher-ai-copyright-concern/>

same amount of compensation regardless of whether or not they use AI to assist in their tasks, which may also reduce incentive for producers to attempt replacing humans with AI since they will still be essentially paying the same amount of money for those services. Implementing similar provisions in a variety of other job fields would not only remove incentive for employers to replace human employees with AI for 'cost-saving' measures but would also offset any other negative economic impacts that AI growth would inevitably bring.

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.

The only rational scenario of copyrighted works used to train AI constituting fair-use is if it was for non-commercial purposes. Unfortunately, given the hype of AI's economic value, it is almost impossible to separate training AI models from commercial use, increasing the need for such models to be regulated and requiring full disclosure of data used to train AI models.

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)?

Many resources implementing generative AI have provided a means for users to opt-out of having their work used to train AI models. Unfortunately, this is aligned with the notion of asking for forgiveness over permission, when realistically it must be the other way around. In order for AI developers to train on copyrighted works, copyright owners must be given the option to opt-in, rather than be asked to opt-out, before AI models can be allowed to train on this data, as consent is the most important aspect of the issue.

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 owners must be required for all uses regardless of whether or not they are for commercial purposes. We as humans must sign consent forms for sharing information with our healthcare providers, our employers, our web and communications providers, as well as any instance involving photo releases. Training AI models on copyrighted works *only* through consent of the copyright owners is no different.

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?

On a technical scale, it is very feasible to establish processes for requiring consent and cracking down on copyright infringement. YouTube, Facebook, TikTok, and many more platforms have automated systems to detect copyrighted material included, at which point it may either be removed or be allowed to stay up in return for ad revenue going directly to the copyright holders (for example, if I upload a video to YouTube that included a Michael Jackson song, YouTube's copyright detection system, Copyright ID, will find the song and where in the video the song is included, at which point they may decide to either block or remove the video altogether, or if they allow the video to remain up, I cannot monetize it, and any and all ad revenue generated goes directly to whomever owns the copyright to the song used). This is why the go-to solution for millions of content creators is royalty-free stock footage and music, which only requires either attribution or a paid membership plan that includes licensing rights, neither of which have a negative impact on copyright holders.

This should be no different for AI. There may be some difficulties in having AI developers get explicit permission from copyright owners or for providing compensation to copyright owners. This should not,

however, be an excuse under any circumstances to not implement such needed systems of accountability, especially when AI is poised to be as big an industry as social media.

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?

There are many existing remedies that may be applicable such as cease-and-desist letters, copyright lawsuits, and settlements for monetary damages. There will need to be solutions pertaining specifically to AI, which can begin with implementing a similar system mentioned in the previous question that flags copyrighted content being used to generate AI. It is debatable whether or not AI models can be *un-trained* on specific data, but in the event it cannot, there must at least be monetary compensation paid to copyright owners for previous training incidents as well as indefinite pauses on AI training on unauthorized copyrighted works.

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?

This particular question is tricky to answer, as there are numerous cases in sectors like the film industry or even freelance jobs involving work for hire in which the copyright belongs to an employer rather than whoever actually invests time and labor into the creative work in question. While the legal copyright holders can object to an AI model being trained on their work, the Berne Convention gives human creators a “moral right” to object to certain uses of work they contributed to, regardless of copyright law. The recently signed WGA contract is an example of this, allowing the Guild to raise objections if a script is being, or proposed to be, used to train AI. The easiest and most efficient fix for this issue would be to start including clauses in any form of contract or TOS that allows for a human creator who does not own the copyright to whatever they create raise objection to whether or not that work should be allowed to be used for training an AI model, much like the aforementioned WGA contract.

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

Licenses for copyright owners' consent may work similarly to how royalty-free stock footage resources provide licenses for their content, in which the AI developer establishes a centralized resource that gives users the option to use their content (albeit without being able to copyright that content, as the Office has regularly determined AI is ineligible for copyright protection), in exchange for attribution or a paid license, while the original copyright owners whose work would theoretically be used to train AI generated works are monetarily compensated at a fair rate. But the law must clearly state that any such resources are subject to the same terms as established data protection and privacy laws.

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

In order to illustrate some economic impact for licensing and copyrights, I will use the most recent YouTube transparency report as an example: in the second half of 2022 alone, their Content ID system flagged over 826 million possible copyright violations, and of those claims, YouTube paid \$9 billion in payouts to the copyright holders whose material was used in YouTube content.¹² This can give the Office an idea of what licensing and copyright holders' payouts could look like in relation to AI training on

¹² <https://www.digitalmusicnews.com/wp-content/uploads/2023/07/youtube-transparency-report-H2-2022.pdf>

copyrighted works. \$9 billion may seem like a large number even for a company like YouTube, but again, this does not constitute an excuse to not implement measures protecting copyright holders. In any case, since many pro-AI individuals tout the economic benefits of AI, this would not only provide incentive to contribute to what some believe to be a booming market, but also be more reason to properly compensate copyright owners monetarily since they would be generating the revenue for it.

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

Assuming you have not done so already, please take time to read through [the testimony of Karla Ortiz to the U.S Senate Judiciary Subcommittee on Intellectual Property](#), as she gives plenty of sensible ideas on how to approach legislation for AI in terms of training on copyrighted works and protecting copyright holders.

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?

Developers of AI models should absolutely be required by law to collect, retain, and disclose any and all records of materials used to train their models. This would give copyright owners a significant advantage in maintaining their rights as AI companies have thus far failed to properly disclose the full extent of their training data, and in some cases even deliberately attempt to keep that information hidden as OpenAI has done.¹³ This transparency would also be a significant advantage for copyright owners receiving monetary compensation for having their works used for training AI models, should eventual AI regulation permit it. Furthermore, many websites that center around user-submitted works contain metadata that reveals certain tools used, such as the type of camera or mobile device used for a photograph. Because some of this information is publicly available, it stands to reason that any and all AI-generated content should be publicly disclosed as well. These same legal obligations should also apply to any creator of training datasets.

15.1. What level of specificity should be required?

The specifics of collecting and disclosing training records should include, but certainly not be limited to, the copyright holder of any works used to train AI models, the specific details of the work(s) themselves, whether the intended use of AI training is commercial or non-commercial (which would also further determine whether or not it falls under fair use), and the number of copyrighted works used for AI training.

15.2. To whom should disclosures be made?

In relation to the specifics in the previous inquiry, the first and foremost disclosures made should be to the copyright owners, as they are the ones most directly affected by the use of copyrighted works for AI training. Many sites and platforms that have incorporated generative AI or allowed for AI-generated content to be submitted have provided the option for users to ‘opt-out’ of having their works used to train AI models, but having an ‘opt-in’ option, where users must agree beforehand to have their works train AI models rather than having to say they cannot, would be a significant step in the right direction for ensuring accountability. Disclosures should also be made to companies and platforms that currently utilize AI systems so that they are better informed and better protected if they find themselves facing legal action

¹³ <https://www.businessinsider.com/openais-latest-chatgpt-version-hides-training-on-copyrighted-material-2023-8>

over copyright infringement that results from AI. Most importantly, in the likely event more legal cases be presented involving copyright and AI, disclosures *must* be made to any and all legal, legislative, and judicial bodies in order to be as fair and impartial as possible when making legal rulings pertaining to AI and copyright.

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

Anything less than complete and transparent disclosure of all training data implies there is something in such data that AI developers do not want made public, which is indicative of being fully aware their training models are not in compliance with the law. Ergo, the developers of AI systems can and must have the highest of obligations to disclose all training data used, including the specifics referred to in Question 15.1.

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?

As demonstrated in the example provided in response to Question 13, there would most likely be a significant financial cost to implementing a recordkeeping system for developers of AI models or systems, creators, etc. However, that should not, under *any* circumstances, be used as an excuse to not implement such a vital and necessary recordkeeping system at all. In terms of impact, however, it would vastly increase transparency of AI development and additional security and privacy rights for human creators.

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

Notifying copyright owners of having their work used to train an AI model should absolutely be an obligation for AI developers, as well as strict penalties for failure to comply. That being said, if any proposed licensing regimes were to exist, notifying copyright owners would naturally be a part of that process anyway.

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?

Under current copyright law, there is no clause for copyrighting solely AI-generated material. The closest thing to it is partial AI-assisted copyright, which protects anything created by a human while excluding portions generated by an AI (as in the case involving ‘Zarya of the Dawn’ by Kristina Kashtanova, which copyrighted the text written by Kashtanova, but rejected copyright for the images generated through the Midjourney AI generator)¹⁴. Until or unless there is greater transparency on what, and how, AI generators and LLMs are trained on, and dependent on what that data reveals, no changes accommodating AI-generated material or partially AI-generated material should be made.

¹⁴ <https://www.copyright.gov/docs/zarya-of-the-dawn.pdf>

19. Are any revisions to the Copyright Act necessary to clarify the human authorship requirement or to provide additional standards to determine when content including AI-generated material is subject to copyright protection?

Any revisions to the Copyright Act should only further strengthen the human authorship requirement; again, until or unless there is transparency on what and how AI-generated material and LLMs are trained on and dependent on what the data reveals, no revisions to the Copyright Act should include anything protecting AI-generated material partially AI-generated material.

20. Is legal protection for AI-generated material desirable as a policy matter? Is legal protection for AI-generated material necessary to encourage development of generative AI technologies and systems? Does existing copyright protection for computer code that operates a generative AI system provide sufficient incentives?

Relating primarily to my career field, legal protection for AI-generated material is neither desirable as a policy matter nor necessary to encourage development of generative AI technologies and systems, as there are other fields that could benefit from such development such as in healthcare or cybersecurity, albeit within the same limits of data privacy concerns that are undoubtedly being raised. The Copyright Office should not consider any form of legal protection for AI generated content.

21. Does the Copyright Clause in the U.S. Constitution permit copyright protection for AI-generated material? Would such protection “promote the progress of science and useful arts”? If so, how?

Under the current language used to describe the Copyright Clause of the U.S Constitution, Congress may enact laws “to promote the progress of science and useful arts, by securing for limited times to authors and inventors the exclusive right to their respective writings and discoveries”, thereby encouraging creativity through protecting creators’ works via copyright. Were AI-generated material to fall under copyright protections, it would not promote the progress of science and useful arts due to the nature of business and tech executives using the AI-generated material as an excuse to implement ‘cost-cutting’ measures that would put millions out of work and destabilize the innovation only humans can bring.

When measuring the promotion of progress, an important factor is prioritizing, to borrow a rather famous quote, the needs of the many over the needs of the few. With this issue, the ‘many’ would refer to the 375 million Americans that could potentially lose their jobs and livelihoods,¹ among which those in creative fields would be one of the first, while the ‘few’ refers to the developers of AI who tout the ‘progress’ of generative AI by throwing out buzz phrases such as ‘increased productivity’ and citing the estimated 97 million jobs it would create while flimsily citing ‘fair use’ as a means to an end (97 million gained versus 375 million lost equals negative 278 million jobs actually ‘created’; $97 \text{ million} - 375 \text{ million} = -278 \text{ million}$). As such, not only would copyright protection for AI *not* promote the progress of science and useful arts, it would be economic suicide.

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?

Without clear and necessary transparency on AI training data (which they should be legally obligated to provide anyhow), proving copyright infringement can be a challenge for copyright owners. However, there are a number of existing privacy laws that would be applicable to this inquiry despite the current lack of much-needed AI regulation.

The Federal Trade Commission (FTC) Act includes clauses meant to “prevent unfair methods of competition and unfair or deceptive acts or practices in or affecting commerce”, as well as “seek monetary redress and other relief for conduct injurious to consumers”. Both the ‘unfair or deceptive acts’ and ‘monetary redress...for conduct injurious to consumers’ texts could easily be applicable in cases involving generative AI infringing copyrights, which would be greatly beneficial for protecting the rights of copyright owners until more concise AI regulation is passed. In addition, the FTC Act clause allowing the commission to “prescribe rules defining with specificity acts or practices that are unfair or deceptive, and establishing requirements designated to prevent such acts or practices” clearly and concisely opens the door for legislation requiring AI developers to fully comply with making all training data publicly available in order to better provide copyright owners with needed legal ground to prove the element of copying.¹⁵

The California Consumer Privacy Act of 2018 (CCPA) regulates the amount of personal information business are allowed to collect about consumers via giving those consumers more control over their personal information. This includes consumers’ rights to know about the personal information collected about them, how that information is used, the right to delete said information, the right to opt-out of the sharing and selling of personal information and, as of January 1, 2023, the right to limit the use and disclosure of sensitive personal information. While the CCPA may currently be limited to California residents or anyone collecting the data of California residents, many AI developers are still centered in Silicon Valley and are therefore legally obligated to comply with California law. While expanded legislation similar to the CCPA would be a welcome and extraordinarily beneficial change in all AI matters, the current text of the CCPA is designed in a way that it would easily cover the right to transparency about AI training data, which again would benefit copyright owners looking to prove the element of copying.¹⁶

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 AI-generated material is found to infringe a copyrighted work, both the developers of the generative AI model in question and the developer of the system incorporating that model should be directly liable due to the fact that having direct access to the codes and data the AI model is trained on makes them fully aware of whether or not copyrighted works were included in the AI model’s data (which again brings back the point of AI model developers not adequately disclosing how their models are trained and on what). The end users of the system using AI-generated material should be secondarily liable, as they may not intentionally set out to infringe on specific copyrights but nonetheless used a system that they presumably knew came with the risk of generating something based on copyrighted material.

25.1. Do “open-source” AI models raise unique considerations with respect to infringement based on their outputs?

The only significant difference between “open-source” AI models versus “closed-source” is its public availability, as well as, for open-source, the ability to take source code and modify it for different uses. The ethical considerations surrounding copyright infringement do not differ between open-source and closed-source AI models; both rely on training, or data scraping, from currently existing data, with or without the original human author’s consent.

¹⁵ <https://www.ftc.gov/legal-library/browse/statutes/federal-trade-commission-act>

¹⁶ <https://oag.ca.gov/privacy/ccpa>

26. If a generative AI system is trained on copyrighted works containing copyright management information, how does 17 U.S.C. 1202(b) apply to the treatment of that information in outputs of the system?

Under 17 U.S.C. 1202(b), it is illegal to intentionally remove or alter copyright management information, or to distribute any material knowing copyright information was removed or altered. The lack of transparency from AI developers on the data used to train their AI models gives a concerning perception that the nature of their training data would be defined as precisely the thing 17 U.S.C 1202(b) was implemented to prevent.

28. Should the law require AI-generated material to be labeled or otherwise publicly identified as being generated by AI? If so, in what context should the requirement apply and how should it work?

To my knowledge, many companies and websites to whom this question concerns, regardless of the law, already require their users to disclose whether or not anything they submit or share is generated by AI. Most users seem to respect these rules implemented without intervention, but any law implemented to solidify this stance would be a very welcome development. Enforcement, however, would require business owners, executives, and tech leaders to share liability in any case of noncompliance with this hypothetical law. The person submitting the material, though, would ultimately be responsible for disclosing whether it is AI generated, but as I will share in my response to the Copyright Office's next question (Question 29), there are a handful of tools developed that can assist with detecting AI-generated material, which could be used by site owners to ensure compliance with the law. Current law has the penalties for copyright violations at up to \$250,000 fines and no more than five years in prison. Prison time is a bit extreme for this particular issue, especially since copyright infringement is usually not done out of malice, so fines would likely be a more appropriate penalty, adjusted for the scope and scale of the infringement in question.

29. What tools exist or are in development to identify AI-generated material, including by standard-setting bodies? How accurate are these tools? What are their limitations?

There are a handful of tools that have been developed to either detect AI-generated material or protect human-made material *from* AI models:

- In the earliest days of ChatGPT's popularity, Princeton University computer science student Edward Tian developed GPTZero, which "quickly and efficiently" detects whether a human or ChatGPT wrote specific text, or how much was authored by one over the other. It has also recently implemented a sentence-highlighting model to differentiate mixed AI and human content. Like any tool, it has some margin of error and may not always get it right, but GPTZero operates with transparency as a core value and is consistently advancing their research to both be as accurate as possible and also to exclude ethnical bias from its data research, as one of the papers they studied for their research concluded AI detection algorithms may be biased towards ESL (English Second Language) writers.¹⁷
- Turnitin uses AI detection to identify AI-generated material through LLMs like ChatGPT. Turnitin primarily focuses on academic writing in order to combat plagiarism in students' work, but their model has shown the potential of using AI detection for determining anything AI-generated through LLMs.

¹⁷ <https://arxiv.org/pdf/2304.02819.pdf>

- The University of Chicago’s Glaze program, designed by computer science faculty and students in collaboration with a select number of human artists, essentially adds a digital ‘watermark’ to any work of art that forces AI learning tools to “incorrectly learn the unique features that define an artist’s style, thwarting subsequent efforts to generate artificial plagiarisms”,¹⁸ thus giving human artists a chance to protect their work from AI generators and LLMs. In 2020, the University of Chicago’s SAND (Security, Algorithms, Networking and Data) Lab had a similar program called Fawkes for adding a similar digital ‘watermark’ to personal photos in order to cloak them from facial recognition models, though Glaze is a step up in terms of fooling an AI model, thereby protecting artists’ work.
- Kudurru, an AI startup, is a network of websites that monitors AI models’ data scrapping and, after identifying scrapers, Kudurru not only blocks the scraper from downloading data, but it also provides an option to “poison” the AI developer’s scraping capabilities by feeding it an image of the user’s choice that will instead be fed to the scrapers. It is relatively new and in its early development stages, but in lieu of stricter data and copyright protections, it can be a powerful tool to help copyright owners to protect their work from AI developers’ data scraping methods.

30. What legal rights, if any, currently apply to AI-generated material that features the name or likeness, including vocal likeness, of a particular person?

As stated in Question 24, there are a number of privacy laws regarding people’s data and personal information. While some like the CCPA are state law rather than federal law, the recently introduced NO FAKES (Nurture Originals, Foster Art, and Keep Entertainment Safe) Act in the U.S Senate, which seeks to bar the unauthorized use of having the voice and likeness of actors, singers, etc. generated by AI and hold accountable those who violate the NO FAKES Act,¹⁹ is a significant step forward in further protecting legal rights for individuals’ name and likeness.

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?

As stated in the previous question, the NO FAKES Act is an excellent start to a much-needed process of Congress establishing federal rights applicable to AI-generated material. Although federal law pertaining to AI should not need to be modeled after state laws, some such as the CCPA serve as an efficient blueprint for regulating the use of AI in terms of copyright and individual likeness. Any and all federal laws considered by Congress should consider the need for stricter copyright protections, especially pertaining to the right of publicity that has frequently been violated through generative AI.

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?

There must absolutely be protections against AI systems generating outputs that imitate artistic styles of human creators, but those would be dependent on different circumstances. Like memes or social media filters, plenty of AI-generated images or text are used for non-commercial purposes and would fall under fair use of copyright, such as an artist using AI-generated images as references for their own work rather than generating an AI image and calling it their work. Where fair use would not apply is if AI-generated material imitating artistic styles of human creators is used for commercial purposes, i.e., selling

¹⁸ <https://technology.inquirer.net/127832/glaze-protects-artists-from-ai>

¹⁹ https://deadline.com/wp-content/uploads/2023/10/no_fakes_act_draft_text.pdf

bootlegged merchandise or other products (a prominent example is the frequent use of *Calvin & Hobbes* car decal, particularly the one of Calvin urinating on a custom image or text; though obviously not AI-generated, it illegally uses the *Calvin & Hobbes* trademark as such merchandise was not authorized by the official copyright holders of *Calvin & Hobbes*). Furthermore, in order to efficiently determine who is or is not eligible for copyright protection in regard to AI-generated material, both prompts used to generate AI material and the data used to train the AI model must be publicly available upon publishing the results of AI-generated material in order to efficiently and ethically determine whether the end users are intentionally violating copyrights.

33. With respect to sound recordings, how does section 114(b) of the Copyright Act relate to state law, such as state right of publicity laws? (54) Does this issue require legislative attention in the context of generative AI?

As determined by *Mider v. Ford Motor Co.*, a public figure's voice is distinctive enough to be considered part of their identity, or brand, so to speak, and is therefore legally protected from copyright infringement under Section 114(b) of the Copyright Act. It stands to reason, then, that additional language must be added to Section 114(b) to reiterate it is equally protected from generative AI, and to hold accountable both the developers of AI models and its end users for any infringement of such copyright protections.

Another case outside the court of law but within the same context of copyrights is the Recording Academy, which presents the Grammy Awards, recently rejected from consideration a song that used AI to simulate the voices of Drake and the Weeknd. Recording Academy CEO Harvey Mason Jr. gave a handful of reasons for this decision, but the relevant one for consideration by the Copyright Office is that the AI-generated vocals were not legally obtained or cleared by the respective artists or the labels they represent and was therefore ineligible despite the song's lyrics being written by a human creator.²⁰

34. Please identify any issues not mentioned above that the Copyright Office should consider in conducting this study.

Issues the Copyright Office should consider in this study is the fact that the use of AI in other fields, where it may be beneficial and have little to no risk of large-scale job loss, is not equal to the potential use of AI in creative and cultural fields. Even considering the Copyright Office's past legal rulings that AI-generated material cannot be copyrighted, there are currently no legal precedents for companies who may attempt to circumvent these rules. Regulatory procedures for holding such companies accountable should come *before* larger-scale AI regulations are established by Congress, not after.

I appreciate the Copyright Office's efforts in opening the discussion surrounding AI and copyright to the public forum and for taking the time to read through this letter, and I look forward to future action from all parties towards legislation that protects copyright holders and individual rights.

²⁰ <https://www.lexology.com/library/detail.aspx?g=f5446703-fb7c-4eb4-9d89-a53342689723#:~:text=In%20explaining%20why%20this%20song,that%2C%20it's%20not%20eligible.%E2%80%9D>