Submitted by: Mykel G. Larson – 909 530 0396 – mykel@galapagosengineering.com

General Questions

The Office has several general questions about generative AI in addition to the specific topics listed below. Commenters are encouraged to raise any positions or views that are not elicited by the more detailed questions further below.

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?

<u>**Answer:**</u> **Under the current functioning of Stable Diffusion and Large Language Models (LLMs), the ones who benefit currently are the individuals and firms who have a financial stake in their success. Content creators are not being compensated for the use of content to make other individuals and firms copious amounts of money. It is simply being "stolen" under the assumption, "If it's out there on the Internet, it must be public good." This is an error in reasoning and judgment. I foresee a future where content creators will be forced to "lock up" content behind a paywall before it can be accessed (or scraped by an AI system. An informal labor union will emerge to fight the tendency – and reality – that there are individuals and firms who have zero respect for Intellectual Property Law or the people it was designed to protect.**

The greater risk systemically is creating a zero-trust environment where suspicion and doubt force innovation to go underground, just like it did before England implemented a patent system in around the early 1600s. It ends up being a latent National Security Risk when innovators live in an environment of theft.

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?

<u>**Answer:**</u> **Indeed, it does. It will and has affected the entire economy. Too often the mistake is made to hyper-classify because as soon as that happens, it is easy to find oneself in a minority position with no legal standing or political recourse to right what are obvious wrongs. Philosophizing too much runs this risk, too, as discussion about policy reflects hand-wringing, a sort of excuse to not to do anything because there's still "thinking" going on. Thinking is great, I like thinking. But if I don't do anything with those thoughts and fail to operationalize them, I might as well get a position generating fumes as opposed to something useful for all who participate in society and the economy, and I can get paid to ideate. GDP growth doesn't happen when people are just thinking *ad nauseum*. It happens when people innovate; they think, create, and do.**

Additionally, there has been chatter about a possible robot tax that would redistribute taxes collected for other under-funded institutions in the United States. But laws and codes often have holes in them that can be exploited for financial gain. If this is a real consideration,

make sure those holes are plugged so any attempts to do an end-run around taxation are quickly dispatched. A really strong and competent auditing team, who are unimpeachable, and who are also empowered to their job will go a long way towards preventing a robot tax from being circumvented.

Now on to each of the GAI systems: Images and Words.

Stable diffusion "hallucinates" in a mechanistic fashion, taking visual inputs from multiple sources scraped from the internet, and generatively produces an amalgamation that attempts to closely match user prompts as best as possible. For the most part, this is benign, but bad actors can use freely accessible content to spread disinformation or outright lies to a species that is heavily influenced by images. We have already seen a few fake images and videos of former President Donald J. Trump to manipulate his perception amongst slices of the population who are vulnerable to imagery, as imagery tends to reflect reality. Deepfakes can be used to shame people publicly, destroy their reputations, eroding whatever Social Capital they have and depended on to "make a living."

LLMs, on the other hands, have one fundamental problem that no one really talks about. LLMs have no judgment. LLMs in essence are large Dr. Frankensteins that stich words together in something that *prima facie* looks like a sentence, consumable by most with a basic education. But mechanistic, statistical construction of sentences is no better than having a room full of Monkeys randomly typing on typewriters to eventually produce *Hamlet.* LLMs just made the Monkeys generate words in a lot less time due to the mechanistic and discrete functioning of advanced computing systems.

Why is there no judgment? Meaning cannot be encoded digitally; it is extrapolated via the context in which the word is used when a human being consumes the words in a sentence. It is possible to assign valence to statements, but even that is a gambit. I am not arguing that relativism rules the day here. *The real danger is the construction of the illusion of human thought.* Human thought, which is based on experience, also includes? What? The senses and emotions related to sensory perception. LLMs have no way to be authentically relatable to the prompt engineer, which is typically a human.

Trying to do and-run around judgment generates the strong possibility that LLMs (and they have) will generate human-like sentences that are either factually wrong, a hallucination, or both. Why would anyone let a system be deployed that runs the risk of giving false or bad information?

The only people who benefit from the confusion are the owners of the LLMs and Stable Diffusion themselves. But the information externalities are real and should be taken seriously if the United States really believes in truth and judgment. There will be those who will attempt to use the shield of the 1st amendment to use LLMs and Stable Diffusion as tools to express their positions. There is comedy and then there is information, what we used to call news. Plus, there are issues of theft, as previously mentioned that add a

particular flavor of anarchy to the whole endeavor. Last time I checked the US frowns upon anarchy, but allows people to congregate to nudge a change in the functioning of how the State works. January 6th, 2021 was just a preview of what happens when anarchy becomes the norm. Turn that up to volume 11 with Generative AI (GAI) and you get a recipe for chaos that cannot be contained except through extreme forms of authoritarianism.

There is a reason why we don't just let anyone play around with nuclear material, or let people smoke at a gas station while pumping gas. Similar reasons exist for limiting the use and functionality of GAI.

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.

<u>Answer:</u> Please see answer #2.

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? [40]

<u>Answer:</u> Not that I am aware of, but I can imagine most are unsuitable, especially in countries with a high Perception of Corruption Index score. Someone must stand up for what is right. Theft is theft.

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

<u>Answer:</u> Critically important. If the United States is worried about bad actors in questionable regimes, these tools are perfect for increasing anxiety in international relations and diplomacy. See Answer #2 for further elucidation.

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.

<u>Answer:</u> The spirit of the law should be to prevent theft and foster collaboration between different types of content creators. Easier said than done.

<u>*Components:*</u>
Content creators have the right to ask a particular entity to:

1. Stop "using without explicit permission." We typically call this theft.
2. To offer up a rate card for what the creator values the use (or purchase, like art) of their creations (content).
3. If the firm or individual rejects those rates, and if the firm or individual has been found in violation – and a trial might be necessary - they are required to pay the

initial cost for licensing, plus a 10,000% penalty for violating copyrights. Losing money for reckless and criminal behavior should have a real financial consequence. If it happens repeatedly, each violation carries an exponential growth rate in penalties. (10,000% becomes 100,000,000%. Why? This acts as a disincentive to repeatedly steal. At a certain point, the fines will bankrupt a firm or individual and that is the point. There is a real price to be paid for breaking the law in this fashion.)

4. The content creator should have reasonable proof that the content is theirs. Copyrights that have been granted automatically meet this condition. If a copyright has not been applied for or granted, other supporting evidence is sufficient. The USPTO has models for determining ownership of prior and current art, and what constitutes a fundamental deviation from the original prior art. That's good, but not good enough. Right now, the threshold is a 15% difference, but that is actually pretty tough to measure objectively. It might require juries to come to a decision in these matters, as a singular authoritative figure, like a judge, cannot reasonably do this on their own. With a jury there is a blend of biases that create enough experiential bias and judgement to essentially arrive at "the wisdom of the crowd."

But receiving a copyright need present such a financial burden so that it cannot be granted and the enforcement protection that goes with a copyright. The cost should be as minimal as possible. If an economy is an exchange between two parties that "have" something the other wants, the only noble way that gets done is if A pays B or B simply gives it away to A as a gift. Assuming that content is available without paywalls on the internet is a flimsy argument for "fair use," and does not fall under the auspices that creators are giving content away simply because it is on the internet and easily accessible.

5. If it goes to court and the violator refuses to pay for past, current, or future usage, with evidence of prior communication that steps 1-3 were ignored or dismissed unilaterally, they are responsible for 100% of the plaintiff's legal fees. This is on top of paying whatever fines have been levied.

6. The normal rules of discovery and evidence in a trial shall remain. Defendants will have every opportunity to defend themselves under the assumption of innocence. Justice is mostly about repairing harm done to another party and judges can use their judgment to determine reasonable restitution for the first offense, given the parameters set in (3). Every violation thereafter, if found guilty, automatically triggers an exponential growth rate in penalties.

Training

If your comment applies only to a specific subset of AI technologies, please make that clear.

6. What kinds of copyright-protected training materials are used to train AI models, and how are those materials collected and curated?

Answer: Any content that requires compensation AND permission by the original creator. This is not hard to figure out. Books, articles, visual likenesses, original sonic compositions, which also include "voice acting," traditional media with hand-based tools (a piece of original art done in oils, etc.), or digital media created with desktop publishing tools designed for human use (Photoshop, Illustrator, 3D modeling software – any software a human uses in their creation, excluding GAI).

6.1. How or where do developers of AI models acquire the materials or datasets that their models are trained on? To what extent is training material first collected by third-party entities (such as academic researchers or private companies)?

Answer: They use automation software, known as "spiders" to seek out content that is easily accessible from the corpus of content available on the internet. They steal it, they do not borrow it. The extent doesn't matter. If it is easy to get via automated means, it does not matter to these firms or individuals where it came from, only that they can get it. Or humans are deployed to manually acquire data for training sets.

In creative markets like the film industry, clever and powerful pieces of software can replicate likeness of an actor in every way imaginable. All one must do is look at how motion capture was done for Gollum in the LoTR movies and realize that it's not hard to replicate a person's likeness digitally. This should be obvious.

6.2. To what extent are copyrighted works licensed from copyright owners for use as training materials? To your knowledge, what licensing models are currently being offered and used?

Answer: Look to websites and firms, like Adobe, who have each formulated their own rate card for access to images. Similar models exist for accessing research papers or individuals who sell subscriptions to their content. And, similar models exists for accessing 3D meshes for use in 3D modeling programs to produce games, film, and simulations.

Daneille DiMartino Booth is a proximate example. She sells a subscription model for accessing her insights on macroeconomic conditions. Newspapers and other Publications do the exact same thing. Think The Wall Street Journal or Bloomberg.

https://quillintelligence.com/

I won't reference the other websites because they are easily accessed.

6.3. To what extent is non-copyrighted material (such as public domain works) used for AI training? Alternatively, to what extent is training material created or commissioned by developers of AI models?

Answer: **Read this article to get a clearer picture what actually goes on "behind the scenes." Those with a conscience should be concerned not only at how training sets are built, but the conditions under which these laborers were exploited. Is this the kind of country we want to be? I hope not.**

https://time.com/6247678/openai-chatgpt-kenya-workers/

6.4. Are some or all training materials retained by developers of AI models after training is complete, and for what purpose(s)? Please describe any relevant storage and retention practices.

Answer: **Once it's stored on a hard drive or other types of digital archival media, there is absolutely nothing stopping them from using it again. When models must be retrained due to model drift, either the entire previous training sets are used in addition to new training data, or the previous training data isn't used; it can either sit on a hard drive or be purged. That kind of discretion is left up to each individual firm.**

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.

Answer: **For my particular project with Galapagos Engineering, Inc., we build our training data ourselves. However, it is partially constructed via a subscription to IEX Cloud, which we pay a monthly subscription fee to get access to real-time and historical securities information. We construct some other variables, but the actual historical and real-time data feeds we pay for. We don't steal content even if it's cheaper.**

Our models retrain the ML (Deep Learning) systems using data we have deemed that has an immediate self-life. But we are allowed to save this data forever based on the fees we pay to IEX Cloud and operate under their Terms of Service (ToS) and End User License Agreement(EULA). We save this data in the event we have the desire or need to go back in history to do analysis that impacts our and our customers' bottom line.

7.2. How are inferences gained from the training process stored or represented within an AI model?

Answer: **With many epochs of forward and backward propagation, nodes in a Neural Net are given weights, which act as scalars to data. If the model generalizes the trend without overfitting or underfitting, the weights are stored in the model. Then as new data is ingested into the Neural Net, those weights on each node constitute how well the model performs given new data it has not seen explicitly before. It's glorified pattern matching**

with brute forcing powered by CPUs, GPUs, lot of RAM, and lots of data storage. Statistical Inference is NOT the same as general human inference.

7.3. Is it possible for an AI model to "unlearn" inferences it gained from training on a particular piece of training material? If so, is it economically feasible? In addition to retraining a model, are there other ways to "unlearn" inferences from training?

Answer: Yes. You delete the entire model that is in production. Or you feed the Neural Net with a bunch of meaningless data and that will affect model performance in production, which is just about equivalent to unlearning.

Please note that while GAI has nodes in its NN, the NN architecture is a just a variant of NNs. GAI has to be built specifically to generate content as a amalgamation from its inputs.

7.4. Absent access to the underlying dataset, is it possible to identify whether an AI model was trained on a particular piece of training material?

Answer: Only through human inference and multiple people coming to the same conclusion.

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.

Answer: None. The laws regarding IP theft simply need to be enforced.

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?**

Answer: The entire lifecycle of creation must be audited from inputs used to outputs generated. It gets easy to determine the extent of unauthorized use of inputs to generate an output that generates a financial windfall. Ingesting entire works and calling it fair use is just an attempt to move the goal posts so it is easier to steal without consequences.

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?

Answer: That all depends on how these entities structured their own ToSs and EULAs. But there should be a lineage stored in a database along with transaction records to exculpate these entities for wrongdoing. If the original creator and content distributors have a legally binding agreement in place, this is a non-issue. However, if the purchaser of the content is found to violate the Terms and Condition under which a license was granted for said content, the content distributor should not be at fault, since the creator is being

compensated, by their own free will of entering into an agreement, which includes compensation by the Content Distributor (Repackager).

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]

<u>**Answer:**</u> **This is a form of credit. The problem is that there is no guarantee the endeavor will succeed, but that does not mean that fair use is subsumed by an anticipation of future profits. Actors get hired all the time to perform in a movie regardless of how it actually performs in the market. Some never make it to the market and are shelved, but the actor is paid regardless.**

Does it make a difference if funding for these noncommercial or research uses is provided by for-profit developers of AI systems?

<u>**Answer:**</u> **No. Pay for it or be forced to create your own data. Only with explicit permission that it to be used for noncommercial or research purposes would be allowed, with the right to audit its use when requested. The Grantor can certainly decide that a fee be paid to alleviate fears that the Grantee is not living up to its promises as described in the contractual agreement. This simplifies the market and avoids legal externalities that cost money. It is a common tactic for firms to outspend their adversaries in court, to win by attrition. This cannot be allowed.**

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?

<u>**Answer:**</u> **No. This question is about giving wiggle room to thieves. The aggregate constitution of the training data used does not matter. If copyrighted material was used without permission or compensation to the creator, it's an instant violation.**

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]

<u>**Answer:**</u> **Wrong question. Must use economic analysis to determine concentration of use in the market. Herfindahl-Hirschman Index (HHI) is useful for this purpose.**

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?

<u>**Answer:**</u> **Both. The attempt to delineate here just creates opportunities for theft, making the entire endeavor prone to being gamed by thinking too much. If it looks like a duck, quacks like a duck, it probably IS a duck.**

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)?

Answer: They need to give their explicit consent. The absence of an answer could be due to several factors. It is simple: pay for what you use.

9.1. Should consent of the copyright owner be required for all uses of copyrighted works to train AI models or only commercial uses? [47]

Answer: Yes. Don't give any wiggle room to be exploited or cheated.

9.2. If an "opt out" approach were adopted, how would that process work for a copyright owner who objected to the use of their works for training? Are there technical tools that might facilitate this process, such as a technical flag or metadata indicating that an automated service should not collect and store a work for AI training uses? [48]

Answer: Watermarks and other encryption schemes can slow down unauthorized use, but not prevent it completely. With anything digital, it is built up of an aggregation of 1's and 0's. Clever people can find ways to circumvent protective measures by surgically removing bit sequences to unlock content.

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?

Answer: See answer 9.2.

In legal environments, the question that needs answering is if there is evidence of tampering to avoid compensating the creator. It could be burdensome in terms of time and effort. But that is why the enforcement mechanisms need to be structured so that if caught, it hurts really bad and inspires a fundamental change in behavior. Plus, a certain bonus: companies who are caught trying to cover up their theft will pay a price not only financially, but reputationally, too.

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?

Answer: See answer #5.

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?

Answer: Whoever owns the copyright by either mechanism has the right to seek restitution.

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

Answer: Simple. They must start a dialogue and come to a mutually beneficial agreement that's legally binding.

10.1. Is direct voluntary licensing feasible in some or all creative sectors?

Answer: Sure, it is. The © symbol should automatically grant protection and start the process of negotiating the terms of a *quid pro quo* arrangement.

10.2. Is a voluntary collective licensing scheme a feasible or desirable approach? [49]

Answer: Only if the rate cards accompany such a collective licensing scheme. Prices must be known to the purchaser so they can evaluate if they wish to proceed.

Are there existing collective management organizations that are well-suited to provide those licenses, and are there legal or other impediments that would prevent those organizations from performing this role? Should Congress consider statutory or other changes, such as an antitrust exception, to facilitate negotiation of collective licenses?

Answer: Congress should first consider preventing theft and having proper enforcement before that question is addressed. Foundations matter first. Trying to find holes in the system invites bad behavior. This is not an anti-trust issue, as this has to do with theft, not market (pricing) power.

10.3. Should Congress consider establishing a compulsory licensing regime? [50]

Answer: Yes, they should. Prevent and punish theft. The United States in its early history was really upset about the Crown's propensity to acquire whatever it wanted, under the assumption that the colonies and everyone in them were there for plunder whenever the Crown felt like it. Let us not repeat stupid regimes. There is a point where the marginal cost of enforcement exceeds the marginal benefit of enforcement, but that is just a mathematical model designed to reduce cost. This is one area where the United States cannot screw up or else it runs the risk of turning into the evil regime the Colonies wished to escape.

If so, what should such a regime look like? What activities should the license cover, what works would be subject to the license, and would copyright owners have the ability to opt out? How should royalty rates and terms be set, allocated, reported and distributed?

Answer: The regime functions very much like a Republic: each individual sets their own terms and conditions for use of their creation and compensation is explicitly addressed in terms of price in dollars ($).

10.4. Is an extended collective licensing scheme [51] a feasible or desirable approach?

Unnecessary.

10.5. Should licensing regimes vary based on the type of work at issue?

Answer: Leave it up to the creators how they structure their use agreements. What's important are the enforcement mechanisms, funded properly to expose and rectify bad behavior.

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

Answer: The only practical issue that matters is did they get permission or did they pay for permission for use? Everything else is meaningless. A person employed at Company X is doing work for that Company to earn revenue. This even covers 1099 employees.

12. Is it possible or feasible to identify the degree to which a particular work contributes to a particular output from a generative AI system? Please explain.

Answer: With a jury there is a blend of biases that create enough experiential bias and judgement to essentially arrive at "the wisdom of the crowd." Similarity is not hard to ascertain. Some want to intentionally atomize this process so that it would be impossible to do so.

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

Answer: Finances internally matter before any statements or considerations on the possible effects on macroeconomic conditions. It would cost companies developing GAI systems more to develop, but that is a friction (cost) that is necessary. We do not let people steal patents without consequences, even if there is a positive effect on GDP or GDP per capita.

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

Answer: The only factor that matters is paying people for the work they have done. Failing to protect creators is not only a form of wage theft, but an invisible tax in the form of unrealized gains due to theft.

Transparency & Recordkeeping

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?

Answer: Yes to all of this.

15.1. What level of specificity should be required?

Answer: Time | Date | Source | Creator | Transaction logs – was there evidence of an exchange? Along with where the data currently resides.

15.2. To whom should disclosures be made?

Answer: Everyone. Why hide it?

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

Answer: They should make sure they're not implicitly or explicitly stealing. It's very easy to copy and paste code, change a few parameters, and call it "new." If the original creators allow use without monetary compensation, that is their choice to make. It's all very simple: follow the rules and stop trying to bend or break them.

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?

Answer: If spiders are being used, the transaction logs are already there unless someone intentionally turns off storing data lineages.

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

Answer: Notifying after the fact is like telling someone they should think about putting out a fire at their house after it's already burned down. Ask for permission. Communicating intent is not sufficient.

17. Outside of copyright law, are there existing U.S. laws that could require developers of AI models or systems to retain or disclose records about the materials they used for training?

Answer: I am not aware of any off the top of my head. But if companies must file their taxes, and if they get audited, they should have records to back up their filings. Trust but verify. Having poor records tends to indicate incompetence and criminal activity.

Generative AI Outputs

If your comment applies only to a particular subset of generative AI technologies, please make that clear.

Copyrightability

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?

Answer: No. None. An AI agent isn't just another entity that engages in production the way a human would. AI systems that use GAI are for one purpose ONLY: effort minimization, which economically and financially means minimize cost structures. It's the human that starts the inception process with GAI. This fact cannot be avoided.

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?

Answer: No. It's not protected. AI is not an entity with an SSN# like a human being. AI is a tool, nothing more. It is not its own independent agent more than a fork has a choice what food it goes into. The difference between AI and a fork that is AI can realize efficiency gains that a fork cannot. The only instance where a copyright could or should be granted is if an entity meets all the reporting requirements to ensure no content used to train GAI models has been stolen.

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?

Answer: No. GAI is all about theft. Wrong question, again.

20.1. If you believe protection is desirable, should it be a form of copyright or a separate sui generis right? If the latter, in what respects should protection for AI-generated material differ from copyright?

Answer: It doesn't need a normal copyright. It needs a label so it can be differentiated in the market properly. A copyright doesn't do that at all. Yes, questions will be raised by Big Tech and others, and there will be crocodile tears: "But I paid for all this hardware and software! It's mine. Except the data I used to create this service which I now charge money for." If people want to buy that service, fine, but the firms must follow the rules or pay the price. If they create their own content that is used as inputs, then there is no problem. Similarly, if they can meet the reporting requirements that no content has been stolen, a copyright can be granted, with a special additional mark, indicating it was built with the use of GAI.

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"? [52] If so, how?

Answer: No. But people will try to make this argument that stealing from a small few justifies progress. It does not. The only instance where it is allowed is IF the training data was either sourced internally or was properly paid for. In the banking world, suspicions

arise if $10,000 literally arrives *de novo* without any explanation or record where it came from. But, this is why some criminals like to sit on cash than have it in a bank, where the IRS can pull deposit records and levy taxes on financial windfalls, regardless where it came from.

Infringement

22. Can AI-generated outputs implicate the exclusive rights of preexisting copyrighted works, such as the right of reproduction or the derivative work right? If so, in what circumstances?

Answer: Absolutely. In all circumstances. Again, don't allow wiggle room.

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?

Answer: Yes. Adding additional mechanisms to test similarity only exacerbates the tendency to wiggle out of responsibility.

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?

Answer: Yes. Source and generated output are necessary and sufficient.

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?

Answer: The entity that owns the hardware and used software to generate the final output.

25.1. Do "open-source" AI models raise unique considerations with respect to infringement based on their outputs? [53]

Answer: No. They all require training sets to generate an output.

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?

(b)Removal or Alteration of Copyright Management Information.—No person shall, without the authority of the copyright owner or the law—

(1)intentionally remove or alter any copyright management information,

(2)distribute or import for distribution copyright management information knowing that the copyright management information has been removed or altered without authority of the copyright owner or the law, or

(3)distribute, import for distribution, or publicly perform works, copies of works, or phonorecords, knowing that copyright management information has been removed or altered without authority of the copyright owner or the law,

knowing, or, with respect to civil remedies under section 1203, having reasonable grounds to know, that it will induce, enable, facilitate, or conceal an infringement of any right under this title.

<u>Answer:</u> **It's people who use AI. The law would apply to these instances without special regard if it were AI or not.**

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

<u>Answer:</u> **I've covered all the pertinent bases.**

Labeling or Identification

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?

<u>Answer:</u> **Yes. Transparency is a must. The FDA requires food manufacturers to disclose their ingredients. This is no different in the case of AI, because content, when it is consumed by the conscious mind, cannot be reversed to return to prior the state of information one had previously. Information can cause harm. Go ask the NSA and CIA about this.**

28.1. Who should be responsible for identifying a work as AI-generated?

<u>Answer:</u> **The ones who used AI systems in the first place.**

28.2. Are there technical or practical barriers to labeling or identification requirements?

<u>Answer:</u> **No. If people or firms complain its because they do not wish to be held accountable if and when they are caught.**

28.3. If a notification or labeling requirement is adopted, what should be the consequences of the failure to label a particular work or the removal of a label?

<u>Answer:</u> **A series of fines and warnings, culminating in the termination of their entity enjoying their status in the eyes of the IRS, with no recourse to write off fines as an expense. This and in addition to having their product pulled off the market. They did not follow the rules, plain and simple.**

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?

<u>Answer:</u> **Previous attempts have been made in the education market to detect if submitted work by students was theirs, plagiarized, or used the assistance of LLMs. Automation will**

not and has not worked in accurate classification. Even trained professionals have a hard time determining if papers submitted are original works or constructed by means that violate the honor code. LLMs are not 100% fool proof, as LLMs have been caught generating fake references or making simple logical mistakes, even though the sentences are cogent.

Additional Questions About Issues Related to Copyright

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?

Answer: It's infringement if it is protected content, plain and simple. The owners of GAI systems have no rights, except the compulsion to stop acting badly and heal whatever harm was inflicted.

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?

Answer: No, this is unnecessary. We don't need a separate market. These products just require proper disclosure so that people know what they are buying. Then it is up to the purchaser to determine "if it is worth it."

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?

Answer: This question answered itself in the positive. The protection isn't absolute, however. Its protection manifests in the consequences when caught stealing.

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]

Answer: Federal Laws take priority over State Laws.

Does this issue require legislative attention in the context of generative AI?

Answer: No, it's already covered.

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

Answer: Here's a major one. The United States was partially founded on the principle that people have the right to shape their life as they see fit as long as it does not infringe or prevent others from enjoying those same rights. Trying to come up with escape clauses to escape responsibility only seeks to create a condition where people are incentivized to steal without getting caught or making the process so burdensome, it becomes economically and mentally taxing to seek relief. We are not a nation that prides itself on theft, but

cooperation, hope, faith, and trust. Agape love is what makes it all work and if laws are not in place to help bad actors choose a different path, we don't have a nation at all; just honor among thieves.