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?

- The risks of generative AI technologies outweigh the potential benefits by a large margin. It was created with no problems to solve and all it's doing is competing against creators which it uses the work of to exist, and uses that copyrighted work without any sort of compensation, consent, accreditation and sometimes knowledge of those same artists, whilst taking away job opportunities in these same related fields.

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?

- Yes, increasing the use in any sector will raise issues.

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.

- AI Art and its Impact on Artists
- How AI actually copies from its training material
- We Are All Raw Material for AI
- AI Lie: Machines Don't Learn Like Humans
- Ethics and AI Hardly Mix
- How much can artists make from generative AI?
- Popular "AI Hub" Discord Taken Down Following Copyright Complaints
- Microsoft's Github Copilot is Losing Huge Amounts of Money

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

- It is of upmost importance that a heavy worldwide AI policy is adopted, one that leaves no room for “Loopholes” or ways to go back to stage one.

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.

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?

- Materials used to train the AI vary on the kind of the AI, they may vary from pictures, artwork, videos, academic researches, books, audio, music, legal cases, medical cases, messages (private or public), amongst thousands of other kinds of materials. The collection of material for AI training is done through “Web-Crawlers”.

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)?

- They acquire the training material through “Web-Crawlers”, either “crawling” the “Common Crawl” or doing their own crawling through the internet, always with a third-party with the pretext of research, however as we’ve seen time and time again, that pretext is always a façade to train the models and profit off the work of others while avoiding licensing the material and compensating the original authors.

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?

- There’s no licensing acquired any current AI models. The only popular models with some sort of licensing were through Adobe Firefly, however the example mentioned prior, forced their contributors to join their models.
- **Adobe Firefly**: The whole selling point of Adobe Firefly was that it was a legal and ethically made AI image generator, trained with Adobe owned images. However, when it was announced that the images used for its training would be the work of the Adobe stock image contributors, with no way of removing the images from future training and models, the contributors voiced their dislike of the situation, as they were not warned about such before the announcement, basically being forced into this system they never agreed to with no prior warning. Adobe also did not disclose they would

also use the Adobe Creative Cloud to train their AI and that only became apparent after the release of Firefly, however, users of the Creative Cloud were given the option to “opt out” their works from future training, but in a silent manner, and still, many people have their works being used without their knowledge or consent, many not even knowing such option even exists, and some not even being able to say or do anything about such (E.g. death of user).

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?

- **Public Domain:** Public domain work is used for training. The extent of it is unknown, however it is known that it is heavily overshadowed by the billions of copyrighted work inside the database.
- **Created or Commissioned:** Little to none. There are no public numbers, nor any work or documentation that surfaced which proves AI companies created or commissioned material for AI training.

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.

- All training material is retained by AI developers, usually for the purpose of creating newer models and not have to “re-download” all the training material again.

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.

- AIs have 2 main methods of training, one is by having the AI make 1:1 copies of their training material and adding it to their database, and the second one is directly adding the training material inside the database. The material inside the database is actively used during the AI generation regardless of the training method used.

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

- Like said previously, it is gained by either having the AI make a 1:1 copy of the material or have the material directly included inside the database, any inferences gained are the AI copying from those training materials.

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?

- Retraining a model from scratch is the only feasible and reliable way for an AI to “unlearn” it’s training material. And yes, it is economically feasible, even if it wasn’t feasible they’d need to do it, the law shouldn’t care if an illegal action is economically feasible for the perpetrator or not, especially if the perpetrators are multi-billion dollar companies. And an add-on to this statement, AI technologies lose more money generating outputs rather they lose on training the models necessary for the generations.

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?

- No, without the access to a database, it is incredibly difficult to identify if their work was used for training the AI. That’s why **All** databases should be available for the public to see. As a way for the original authors to see if their work was used or not, and there’s not a single reason why a company would ever deny viewership of their database if all their material was acquired legally. And it does **not** create market competition, as it’d require for their competitors to also license the same image.

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.

- Under no circumstance AI should be considered fair use. AI generation inflicts against the third and forth base of fair use, its database not only includes the copyrighted material and distributes it illegally, it also is required to use entire materials to reach a conclusion, and by the use of that same unlicensed material, it negatively affects the market and hurts the original work(s) that were used for the output.

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?

- AI is not Fair Use. It is infringing at its base and core regardless of it either being noncommercial or commercial. It competes with the copyright owners it uses the works from, taking away their opportunity of acquiring work or severely lowering it, AI is not transformative in any way shape or form, it by all accounts cannot create or add anything new, because everything it generates or “adds” comes from other piece(s) that are owned by other authors. AI shouldn’t ever be considered fair use.

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?

- The responsibly to collect and distribute such materials should be handled entirely by the company behind the AI, if the company behind collecting and distributing said materials is proven guilty of disrespecting their responsibilities, they should be fully analyzed.

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

- Stable Diffusion used the pretext of research to train their AI merely as a façade to avoid licensing, this should not be considered fair use under any margin. Nothing about this question should even be considered under fair use.

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?

- Billions of materials are used to train various kinds of AI. An example being the "LAION-5B", an image dataset, a dataset with more than 5 billion images inside it, with a ludicrous amount of those images being copyrighted material which it didn't have any licensing for. Like said previously, this is not fair use, nor should it be considered fair use. These copyrighted materials are essential for AIs to function and should be properly licensed and its authors should be properly compensated for their work being used. An equivalent to this would be a factory stealing the materials necessary to reach a final product and such practice being acceptable by law.

8.5. Under the fourth factor of the fair use analysis, how should the effect on the potential market for or value of a copyrighted work used to train an AI model be measured? [46] Should the inquiry be whether the outputs of the AI system incorporating the model compete with a particular copyrighted work, the body of works of the same author, or the market for that general class of works?

- It is not fair use. It breaks the forth factor by using the original work(s) in a way that negatively the market and hurts the original work(s) that were used for the output, by taking away the chance of the original copyright owner from being contracted or having their work licensed. The inquiry should weather the outputs compete with the market as whole or general classes of works.

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)?

- The system should be "Opt In" where it's dictated by the copyright owners, where the original author of the works signs a non-lifetime (preferably, less than 5 years) contract to license for their work being used for the company and AI to use, while also being compensated with at least 1% of the original licensing fee for each generation and

accrediting them each time their work is used for a generation. This is possible, AI software keeps tracks of every work that is used for a generated output, however it is hidden from the user who generates it.

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]

- It should be required for all uses. And the authors of the copyrighted works should be compensated in full and credited for every single output that uses their material in the generation, regardless of how small or big the use may be.

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]

- "Opt Out"s have been the standard in AI generation. However, this system doesn't not work, as it has been violated many times at this point by AI companies, where they release new models where the "opted out" work are still in the database and are still actively used during generation, and is a backwards way of doing things, because how is someone to know if their work is being improperly used against their knowledge and consent. Because the companies behind the AI never had or acquired a license to use the images necessary for the training of the AI to begin with, this alone would be copyright infringement. This system only favors the AI companies, as they never get the proper license required to use the work, without the original author of said works being compensated, accredited, or the author even having knowledge of their work being used to train an AI and being inside a database.

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?

- It is feasible for AI companies to get a license and consent from copyright owners, the volume of currently used works does not matter, they were acquired illegally, and were never licensed. It should be redone from scratch, trained with the public domain, with an "Opt-in" system.

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?

- Retrain the AI model(s) without the unauthorized work in it, compensate the original copyright owner(s) in full for every single time their work was used in any generation, regardless of how big or small the use was, with the price increasing if the output was used commercially or alike.

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?

- Yes, the original authors should have every right to object their work being used for an AI model. For both uses, the original authors never agreed for their work to be used for AI training.
  Therefore, it should be a mutual, signed, agreement between the original creator and the current copyright owner, where the original creator must be contacted if the current copyright owner wants to give the work for AI training, and must get a signed a non-lifetime (preferably, less than 5 years) agreement from said author, and both of them must be compensated. With the process being the inverted if the original author wants to grant their work for an AI.

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

- Reaching out to copyright owners or waiting until they reach out with offers. Same way other sectors have been getting licenses. I don't see how AI would change proper material licensing.

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

- It is feasible and highly desirable for all creative sectors, especially in combination with the 9.5 question.

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

- No, this would be highly abusive and easy to exploit.

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

- No, this would be highly abusive if made and would just be another way to avoid properly licensing material, don't even consider this as an option.

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

- Such approach is possible, but would require a lot of effort to happen.

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

- No.

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)?

- There's no issues with respecting the proper acquiring of license. Those who should be responsible for securing the licenses should be both the curator of the training dataset, alongside the company behind or employing the model.

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.

- It is possible, and AI technologies already keep track of how much each individual work is included in each generated output. However, it is always hidden from the user/public.

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

- It'd cause a great positive impact to the copyright owners.

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

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?

- Yes, AI developers and companies should have to collect, retain and publicly disclose all the materials used for an AI generated output. This also applies for the trainers of the datasets.

15.1. What level of specificity should be required?

- Upmost specificity, everyone should know if something is AI generated.

15.2. To whom should disclosures be made?

- Everyone.

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

- Every obligation to properly license and compensate the original copyright owners of the materials included in such models.

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?

- Comparing to their annual earnings, little to none.

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

- Every single kind of obligation.

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?

- No comment, I'm not sure.

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?

- No, it is not possible to control AI, at its core, it is a slot machine that is slightly influenced by a prompt.

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?

- Yes, there should be a human authorship requirement, AI should not be considered human expression.

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?

- No, AI generated material should not be legally protected, it is made entirely by a machine, with almost no human input, which requires billions of other works to even function, most often than not copyrighted works the user of the AI do not own, under any circumstance it should be legally protected. Legal protection hasn't been at all thought of when these technologies were made, nor has it been relevant to them as it progresses, and legal protection for computer code cannot be applied to AI, it's a whole different medium.

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?

- AI should not be legally protected.

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?

- Generative AI does not promote progress of science, and much less promotes the useful arts. AI is limited to the scope of its material, it cannot surpass it, and relies on those same works to function.

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?

- Yes, if an author trains an AI with only their work, the AI would only be able to generate derivative works of the training material. Much like it normally does.

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?

- Substantial similarity should also be taken in consideration when reviewing a case of copyright infringement due to an AI generated output.

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?

- The only true reliable way to prove such is for the Databases to be fully visible to the public, there would be no problem in showing it to the public if all material was legally acquired or licensed. And for the latter part, all AI models contains records of the material used for every single generated output, yet, those are never shown to the public. Those records should also be made viewable to the public.

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?

- All of them.

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

- Yes, "Open-source" AI models allows for people to avoid proper licensing of material, and gives them the chance to train AIs with whatever material they want. "Open-source" AI models should not be legal in any way, shape or form, as it gives companies and people a way to avoid licensing and exploit copyright owners' works.

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?

- An AI would infringe on that law, as the copyright management information would either be fraudulent, altered or, possibly, removed.

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

- Do not legally protect or offer copyright to AI generated outputs.

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?

- Yes, regardless of context, all content generated, edited, assisted or otherwise influenced by AI, no matter how big or small the influence may be, should be clearly identified as such.

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

- The companies' behind the AI shall have full responsibility to have a clear and visible identifications for an AI generated output. With no option to remove the identifications included, as that would inherently defeat the purpose of said identification.
  Alongside that, all content generated should also be automatically sent and uploaded by the AI company owned site, as a method to prove a product was AI generated, in case of the illegal removal of said identification.

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

- No. The only reason it hasn't been done is because it'd affect the users of generative AI. A huge part of the user base are "grifters", con artists and various other deceptive kinds of people, those people thrive on deceiving and selling the AI generated output to others by never disclosing the sold product was generated with AI. With clear identifications on the products, this would decrease immensely. AI applications and technologies can easily incorporate the same technology behind "Stock Imagery", where an algorithm automatically adds a watermark on top of images.

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?

- If a company fails to provide the sufficient labels for identifying an AI generated work, they should be sentenced to pay a fine worth a sum of all the work used for the AI generated output with a 200% interest, where the user pays 50% of the fine and the company pays the remaining 150%, with additional payments being required depending on the damage it may cause, E.g. commercial use, deceiving of people and other related situations.

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?

- Tools for such identification exist, however, their accuracy on identifying AI generated material has proven to be incredibly flawed.

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?

- Currently, no legal rights or legislations are applied to AI generated content that attempts to copy someone's visual, vocal or other related kinds of likeness. Even though it is glorified identity theft, fraud or other related issues that intends to deceive people, or use someone's likeness for something the original people did not agree with or even know about.

31. Should Congress establish a new federal right, similar to state law rights of publicity, that would apply to AI-generated material? If so, should it preempt state laws or set a ceiling or floor for state law protections? What should be the contours of such a right?

- Yes, and no. Applying the existing rights of publicity to AI generated content would very likely be enough for such. However, if a new one was made, specifically for AI-generated material, it should be a countrywide or universal law.

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?

- If such happens, it must mean the training or prompting material used in the AI was the work of the human creator, which in such case would be copyright infringement if the work used in the AI was not licensed.

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?

- Yes, it does require legislative attention. AI generated sound or speech can be used to mimic someone's vocal likeness, to put them in situations they never agreed to or even know about. Legislative attention probably wouldn't affect that specific section of copyright, as AI-generated audio, while trained with voice samples, which are most likely taken from the recordings protected by such laws, do not modify the original recording in any way.

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

- Do not give the option for the general public to upload or give pictures into AI models. As that opens the opportunity for them to give copyrighted material they do not own, hurt someone's image or likeness without their knowledge or consent. At most they may curate the images already inside a database.
- Do not allow training with AI generated material generated before legislations passed, their whole existence is built on infringement and would case the created legislations to be exploitable. AI must have a requirement to only be trained on human created images.