Responses to Copyright Office questions about AI

I am a technical writer in AI at one of the major tech companies at the forefront of AI, which I am leaving unnamed because I am speaking for myself. My work responsibilities include learning how AI works and training software engineers on that subject, both conceptually and in practice through most of the AI/ML pipeline. I am also a published author who depends on copyright for my livelihood.

Thank you for your time.

General Questions

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?

Generative AI, particularly generative text AI, is not capable of producing material of sufficient quality to be used for most purposes without training on an enormous dataset containing copyrighted material used without permission, e.g. Common Crawl (a copy of everything on the internet) or Books3 (a copy of all books on a notorious book-pirating website). Once a sufficiently high-quality model is trained, however, which requires a minimum investment of millions of dollars in computational resources, and then fine-tuned for a specific task, the statistical-probability-predicting program, comparable to an advanced form of autocomplete, is capable of flooding creative markets (and the Copyright Office's queue) with instantaneously generated, low-cost imitations of other artists' works. See for example:

- The Kim Jung Gi style-reproducing model created almost immediately after the artist's death without permission by fine-tuning on Kim Jung Gi's works. Note the explicitly derivative nature of these works.
- Theft of artists' works-in-progress, as happened during one artist's 12-hour livestream of a work-in-progress that was screenshot and instantly "completed" by another person using an AI model, who then claimed the artist was infringing on the AI-modified stolen work because the artist finished and posted hours after the AI-model image was completed.
- The Jane Friedman situation, where AI that was likely fine-tuned on Friedman's blog post was used to auto-generate books that were then published on Amazon under Friedman's name and linked to her profile, profiting from her reputation, with the proceeds going to the likely outside-of-US-jurisdiction scammer. Friedman had no recourse until she initiated a media barrage.
- Business Insider: "Eva Toorenent has been working as a freelance artist and illustrator since 2019. Late last year, she was shocked to discover that another artist had taken artwork she'd posted on Instagram and used it to "fine tune" the AI model Midjourney to produce AI art in her style. The other artist then sold the artwork to a gallery."

While it is possible to create and insert a filter to guard against "recitation," or regurgitation of some of the memorized training text (much of which is copyrighted), I am not aware of evaluations of how well this step works to actually prevent recitation. Also, only extremely

scrupulous model-training companies would create such a filter. Most models that are publicly available do not have this safeguard against repeating copyrighted text from the training corpus. I do not believe that such a model exists for sound or image generating models: see the nonconsensual use of various musicians' voices, explicitly taken from their copyrighted music, to produce covers of various other copyrighted songs.

The largest societal risk of generative AI is nonconsensual deepfake pornography of children and women (women constitute 99% of targets of deepfake pornography, which as of 2019 constituted 95% of deepfakes), which is likely to push them out of participating in education, professional, and political life. The phenomenon could be worsened if the Copyright Office allows the men who create these to copyright those images, giving them rights over them. Female students in NJ and Spain have already had deepfaked nude images generated by their male classmates and passed around. Further societal harms are raised in the Bender at al. paper "On the Dangers of Stochastic Parrots" (2021).

What is not mentioned in that paper is that because only extremely wealthy corporations (and well-bankrolled startups) can afford to create and train high-quality models, workers with six-figure incomes are ordered by executives with eight-figure incomes to take the works ("data") of low-compensated creators are being ingested at scale without permission and imitated in order to produce licensable software capable of generating billions in revenue for the benefit of those corporations meant to replace the creative output of those creators. Under such a regime, the few avenues by which artists and writers can make a small living are vanishing, and that money being rerouted to AI corporations (Business Insider).

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?

In addition to the issues above, generative AI is leading to an across-the-board devaluation of artistic skills. The contempt expressed by many wealthy, high-paid software engineers in the tech industry for artists' efforts is indescribable. They don't read books other than technical manuals and Brandon Sanderson so see no reason not to replace authors and artists with AI. They pirate music, movies, and books while self-justifying with the claim "information wants to be free!" and see no reason to respect artists' intellectual property and copyright. At the same time, they will condemn the theft of code and the leaking of proprietary information because that affects the corporations they work for and thus their own bottom line. These are some of the people who I believe will file for copyright on works derived via AI from real artists' pains and labors; the others are corporations.

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.

Jiang et al, "AI Art and Its Impact on Artists" (2023)

See also the WGA and SAG-AFRA 2023 labor agreements around AI for the main concerns of the effects of generative AI on the labor-industry front.

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?

It would be helpful to have international consistency, but the US is capable of leading by example, not least because the vast majority of leading AI companies (Microsoft, Google, OpenAI, Adobe, etc.) are headquartered in the US and governed by US law. It would be preferable to avoid the Japanese approach, which favors tech developers over the artists whose works are being used for training without consent.

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.

Because I am aware of the technological underpinnings of AI models, which despite the use of anthropomorphic language like "inspired" and "create" are in fact only learning probabilistic relationships between pixels and words in their training data and generating outputs based on those probabilities, I do not think their outputs should be ever be copyrightable. ("AI" itself is a misnomer; "machine learning," or ML, is more accurate.) Even in the case where an artist fine-tunes a base model on her own work, the model is simply learning probabilistic relationships between words and sentences and programmatically auto-completing text on the basis of those probabilities. There is no "authorship" involved.

That said, individual writers and artists will never wield the legal or lobbying power of billion-dollar technology companies. It may be more practicable to explicitly declare AI-generated work as derivative works, requiring the permission of artists whose works are used in the training data, and explicitly excluding licensable AI models' use of copyrighted works in training and fine-tuning data from "fair use," given its commercial use and threat to the originating artists/writers.

Note that current penalties and processes for copyright infringement are impractical for many writers/artists who are facing multibillion-dollar multinational corporations.

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?

The vast majority of training data for all corporations' base/foundational models is scraped from the internet, regardless of modality (text, sound, image, etc.) and without concern for copyright status. This is how and why an earlier Midjourney model kept producing watermarks over its

images: it had learned from its training data, which included unlicensed photos scraped from Getty Images and other stock photo providers, that photos had watermarks on them.

Common Crawl, a snapshot of all text on the internet, is the primary training data for the T5X family of text-generating models (open-source).

Stable Diffusion v2 is trained on the LAION-5B dataset, containing a large amount of copyrighted images (Jiang et al 2023, linked above).

OpenAI's GPT-3 model was trained on Books2, likely consisting of books pirated by LibGen (major pirate site).

Books3, containing books taken from book piracy site Bibliotik, was uploaded and made freely available by John Presser in order to allow other developers to compete with OpenAI. It has been used by Meta for their LlaMA model.

Many companies refuse to disclose what goes into their training data (a combination of the above) because they know the training data contains copyrighted material (it is not really possible for any sufficiently large-scale general-application training dataset to *not* contain copyrighted material) and that they are therefore at legal risk. I believe the bet they are making is that their technological advances will happen so fast, and generate so much profit for themselves, that they will be able to proclaim de facto "fair use" for what they've done (many software engineers are proudly declaring this already) or else extract enough profit that either the pending legal challenges will be of no real concern or (in the case of smaller companies) that they will exit via sale or merger before they have to pay any costs.

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)?

For **generative** base or foundational models, AI developers acquire data by any means possible, including unethical means (see the Meta and Bloomberg use of pirated books dataset Books3 or OpenAI's use of Books2), expecting that the fair use doctrine will eventually be stretched to cover their situation, since they have an overwhelming abundance of lawyers, and most creators do not. Anything on the internet can and will be scraped for AI training purposes. This includes third-party entities like pirate websites, YouTube, DeviantArt, stock image repositories, magazines that publish stories online, etc. in addition to the much smaller amount of text that is open-source and Creative Commons, including Wikipedia.

When **fine-tuning** these generative models, developers usually choose a high-quality subset of images or works, e.g. Kim Jung Gi's art, cited above, or Jane Friedman's blog posts. Copyrighted works can often be provided as a prompt to these models (or evoked via text prompt, e.g. "in the style of [artist name]" or "Taylor Swift singing Last Christmas") by the model user: see the case of Eva Toorenent, cited above, or the livestreaming artist whose work was stolen.

Because the question does not specify, I should add that there are other kinds of models, including:

- **Classification** models that predict what class an input belongs to (e.g., is this a photo of a chihuahua or a blueberry muffin? Does this retina scan contain a pathology?)
- **Regression** models that predict variable behavior (i.e. an advanced linear regression or logistic regression problem; what will housing prices in this zip code be in 2026?)

Datasets for these problems tend to be free of copyrighted data and are often created and released by government agencies and academic researchers under open licenses. See the UCI Machine Learning Repository, for one example, or Kaggle.

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?

Every major public model I am aware of is trained on copyrighted materials, with the possible exception of Adobe Firefly. Adobe Firefly (image generation model) claims to train only on materials it owns the license to, but I know no way to verify this claim.

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 works that are available on the internet are scraped and used for AI training data along with copyrighted works that are available on the internet. Some training material is created or commissioned by model developers, but due to the high expense, **in the case of generative models**, this relatively high-quality dataset is almost always used for fine-tuning a pre-existing, pre-trained base (foundational) model for a specific task.

(Research questions involving **classification** and **regression** problems, tending to focus on extremely specific tasks from the start, are more likely to commission data collection.)

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.

I believe this is often the case, though not always internally within a company. Data repositories like Huggingface, Common Crawl, and Github host data, including copyrighted and pirated works like the Books3 dataset, that many companies use. Companies may also retain data internally, whether for data audits anticipating EU regulations (DMA), for debugging issues that arise downstream (see the Sambasivan et al 2021 Data Cascades paper), or for reuse.

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.

Various datasets are ingested by the base (foundational) model during training (which these questions call "pre-training") over the course of, usually, several days or even months, at the size we are discussing for generative models, during which training the model learns to predict the next token in a series, in the case of Large Language Models, and also appears to learn probability distributions and interactions between words based on the training data. While I have not tested this, I suspect that a small model trained on 1,000 repetitions of one book would learn those relationships strongly enough to output mostly sentences (or verbatim excerpts of other lengths) from that book--because it is only learning probabilities through the frequency of words' appearing next to each other, and then outputting the most probable next token (word or subword). It is possible to get a generative model to output less probable next-tokens by increasing the **temperature** (a term referring to the model's degree of preference for the maximally-probable or not-the-maximally-probable next token).

In the case of **autoencoders** (e.g. Dall-E), an image is input into a neural network tasked with outputting the same image at the other end, but while compressing the image in the middle of the network. The model learns to take a 100px by 100px image (10,000 pixels, represented by a 10,000-element vector) and find a condensed representation of the relationships between the pixels so that, for example, one of the middle embedding vectors contains only 1000 pixels (a 10x compression), while still being able to output as much of the original 100px by 100px image on the other side--but for millions of images.

In a GAN (generative adversarial network) setup, a separate **discriminator** model provides signals on the autoencoder's output (often "detect whether this was produced by a generative AI model", with correct identification being a failure), which are backpropagated through the autoencoder to improve its performance. Once the autoencoder is succeeding by the discriminator's standards, the autoencoder has now learned an **embedding** for that image that can be imagined as a numerical representation within a large number of dimensions (that are hard to understand and explain). The discriminator can be discarded.

Interesting behavior emerges at scale. A "faceness" dimension might be governed by a handful of numbers, one controlling distance between the eyes, one controlling eye color, one controlling width of the nose at the nostrils, etc.

In **fine-tuning**, a trained base (foundational) model is further trained for a specific task by providing a high-quality specific dataset, often with example desired outputs. This can include copyrighted data and result in recognizably derivative outputs (see Kim Jung Gi, above). **The works from which the generated text derived may not always be acknowledged**. One can imagine an artist or writer (see Eva T. example above) fine-tuning a model on other artists' works without mentioning those artists' names, changing a few elements, if at all, and claiming ownership/authorship.

In **prompting**, a **user** can often input very long or large files (e.g. a whole book, at this point, or copyrighted images) and ask the model to output a text sequence of nearly arbitrary length or an image based on or using that input. Even in the unusual **hypothetical case where the developer has used only licensed and authorized data**, **the model can still take in copyrighted material and produce output based on that material when the user prompts the model to do so.**

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

Data is converted first to numerical representation of some kind, using one of many possible methods. This is called **encoding** and generally results in a vector (or equivalently column of data in a matrix; each element within the vector or matrix is called a **parameter**)**.** In the "hidden layers" (i.e. not the input or output layers) of the model, the encoding is generally turned into one or more **embedding vectors** that are ideally smaller than the original input vector, compressing the input information. At a very, very simple level, a dataset consisting of the following statements:

"I love dogs"

"Dogs love cats"

could be represented by the following vectors with a vocabulary size of 4:

I: [1, 0, 0, 0]

love: [0, 1, 0, 0]

dogs: [0, 0, 1, 0]

cats: [0, 0, 0, 1]

The sentences could be represented as:

"I love dogs" [1, 1, 1, 0]

"I love cats" [1, 1, 0, 1]

Given the word "I," and asked to predict the most probable next word based on the dataset, there is a 100% chance of the word "love" following "I," because in 100% of the instances where we use the word "I," the next word is "love." Given the word "love," there is a 50% chance the next word is "dogs" and a 50% chance that the next word is "cats." And so on.

For last-generation LLMs, a model often uses a pre-discovered (from another model) **static embedding** to represent words (e.g. the Word2Vec library) in an **embedding space**, some N-dimensional space that compresses certain values of semantic content in much the same way that an embedding vector can compress elements of "faceness". This is not the same thing as

"understanding the meaning of words." My understanding is that present-generation LLMs when encoding text incorporate not only **static embeddings** of words and subwords but also **positional matrices** that allow for contextual information and inter-word relationships to be embedded as well (turning a **static embedding** into a **contextual embedding**), creating a matrix representation of each document.

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?

While it is possible and common for model training to involve saved checkpoints so that interrupted training can be continued from a prior saved point, e.g. in the case of a power outage, it is not possible, as far as I know, for a base or foundational model to "untrain" on data, because what the model has learned in statistical probabilistic relationships between the "tokens" in its training set. A fine-tuned model can't relearn either; you would have to start the fine-tuning process over again from the base model. What certain companies are doing is suppressing outputs, not retraining models.

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?

In some cases it is obvious that material from an artist was included in the **training** data (try asking most image generators to create art in the style of Thomas Kinkade). In some cases it is obvious that material from an artist was included in the **fine-tuning** data (see Kim Jung Gi example above). In some cases it's not clear whether the data was included in training or fine-tuning, but it's obvious that it was included (having a generative audio AI model make Taylor Swift sing a Bing Crosby song). In other cases, the **model user**, rather than the developer, will **input copyrighted material** into the model, for example face-swapping a different person into a movie clip, or generating a graphic novel using copyrighted photos of an actress (as I suspect happened in *Zarya of the Dawn*).

In other cases it is not provable without access to the underlying dataset.

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

I do not think unauthorized use of copyrighted works would ever constitute fair use, but judges are human beings with high variability, legal might belongs with the technology corporations, and I suspect (and I believe corporations know) that, in practice, the legal outcome is a coin toss.

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?

The Google v. Oracle America case is not relevant to the generative AI case because, as the ruling notes, the code in question was an API, an application programming interface. APIs are explicitly meant for other programmers to use to call upon underlying programs. It would be like trying to copyright the words of a formal wedding invitation, e.g. "Mr and Mrs. X are pleased to invite you to celebrate the wedding of Y and Z on December 15 at two-o'-clock" when a recipient of said information decides to use near-identical wording for their own children's wedding.

Andy Warhol Foundation v. Goldsmith is relevant. The vast majority of generative models are extremely profitable commercial technology that would not exist without having ingested unauthorized copyrighted material, that compete with the originating creators--more accurately, outcompete the originating creators--in the same markets.

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?

You're asking how liable Library Genesis should be for pirating books, and how liable Huggingface and similar repositories should be for storing plaintext copies of Library Genesis. I think the Copyright Office is better suited to answering that question than I am.

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?

To the first question, fair use should not apply if the unauthorized copyrighted data **or the model** are made publicly available (vs. naming book titles and giving model benchmark scores in the paper) or adapted for commercial purposes. Morally, that is like a DNA analysis company taking volunteers' DNA samples for medical research, then selling that data to any buyer without re-acquiring volunteers' consent for that use. The research funding source should not be relevant as long as the for-profit developer is not in turn receiving permission to use the model trained on unauthorized copyrighted works or the "right" (access) to the unauthorized copyrighted works.

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

The question does not distinguish between training a base (foundational) model and fine-tuning a model, The first step requires far more data, in the case of generative text on the order of the entire internet. The second step requires far, far less data. The question also omits user prompting with inputs, which often include copyrighted data. The volume of material should not change the fair use analysis, but it should increase the number of copyright violations. I suspect that in many cases where artists' work is not in the base model's training data, fine-tuning and user prompting are far more damaging per work used without authorization to targeted individual artists, but when the artists' work is in the base model (DeviantArt artists, for example), the harms do not appear to be any less, and may be industry-wide as opposed to individual by individual.

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?

See Jiang et al 2023 (linked above) for a more thorough analysis. The answer is that, just as all factors of fair use matter in evaluation, all levels of impact of market and value matter, whether Eva T. and Jane Friedman are personally targeted by individual model users or whether many graphic artists are put out of business by Midjourney trained a model on their works and their peers' works, not least because the copyright infringement can occur at multiple levels, both from developers and other users.

Other important questions include:

- Does the model license compete with original artists/writers/musicians/creators' employment?

- Do the model outputs compete with original artists' work?

- Consider that a book takes years for a writer to finish (mine took about a decade of work each), while a model can almost instantaneously output a text sequence of the same length. The livestreamed artwork that was stolen and falsely claimed by a model user took 11 hours compared to the model's instantaneous "finishing" of the artwork from a screenshot. What prevents malicious actors from becoming "patent trolls" in the area of copyright and stealing and pre-empting artists' works by "finishing" them and registering them years in advance of completion, or overwhelming the Copyright Office with texts/stories/images of every possible subject in hopes of suing others for copyright infringement later, given that the later would take only a few hours and a few dollars, since each model generation output costs a few cents?

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)?

There is no technological means to remove a work from a training dataset once the model has been trained on it. Affirmative consent is the only possibility. Opting out after training is not actually possible.

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)

In research situations where the **model is not publicly released**, i.e. performance benchmarks are published, and where the **copyrighted data is not publicly released but documented**, academic use is reasonable. When those two cases do not hold, copyright use should be authorized before use. **Noncommercial use** outside of academic research is not practicably separable from **commercial use,** as the case of OpenAI, technically a nonprofit, demonstrates very well.

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 is not possible because it is not possible to "unlearn" the statistical probabilities of tokens within a work after model training. It also does not prevent model users from inputting copyrighted works.

It is also not feasible to put the onus of opting out of data scraping for AI on the artist, given how quickly images, music, video, and text are pirated. Requiring artists to make sure every pirated copy of their work has a flag opting it out of AI training is functionally equivalent to removing the ability to consent. My only-in-print works have been digitized and uploaded to pirate websites by others without my consent; I can't change how they were digitized or the flags on them.

I suspect industry advocates are trying to sell you some snake oil, given how many times this technical impossibility has been brought up in this question list.

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?

This question is posed in a way that seems to assume that "feasibility" is being judged by multibillion-dollar corporations and not by artists/writers/musicians. It is certainly not in the interests of AI corporations to have to retrain their very expensive models on licensed data when they used unlicensed data on the assumption that no one could stop them from doing so. The main objection is that it would cost the corporations a lot of time and money, and reduce model quality (and thus the profits from their business model that depends on unauthorized use of copyrighted material), to require artists to consent.

From the other perspective, I would ask: Is it feasible to become or continue to be an artist, writer, or musician when your works are being ingested and your style reproduced and profited off of by multibillion-dollar corporations and model users without your consent, and you have no recourse, and you can't get a job or make a living from your art any longer because your former employers are now licensing AI models?

From that perspective, the artist's perspective, I would argue that it is not feasible to **not** get consent in advance.

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?

Existing remedies for infringement don't work well for individual creators when the infringer is outside of the US in countries that do not honor copyright. Existing penalties for infringement *could* work if there were a separate and more straightforward path to judgment from what exists, which is an extremely onerous legal battle for poorly paid writers/artists against multibillion-dollar corporations. There should also be penalties for the copyright equivalent of patent trolling. I believe this would require the Copyright Office to explicitly declare certain uses to be infringing, certain types of model outputs to be derivative works, and no model output to be copyrightable, so that the gray areas become clearly one thing or another.

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?

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

In the same way that I had to obtain consent (and pay fees) to quote other writers for my third book: from the rightsholder. There could be a central clearinghouse for this purpose, like the one I used to find rightsholders, or the Copyright Office's records. I believe Authors Guild has proposed something along those lines.

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

Feasible for creators? Yes. Expensive and time-consuming for corporations? Yes. But they know how to license copyrighted works already--they have whole legal departments tasked with this. They simply don't want to do so for AI because they think they can get away without doing so.

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?

There are multiple music rights clearinghouses, I believe. Collective **creator** organizations like ASCAP, WGA, and Authors Guild should be able to come up with collective licenses, **but allowing rightsholders to opt in, rather than opt out**.

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?

I don't think Congress is necessarily capable of doing this at the moment without strongly favoring the multibillion-dollar corporations. Opt in is the only approach that allows copyright holders to exercise their rights, and it would require corporations to retrain their models on licensed data only, since data can't be removed from a model after training. Royalty rates for artistic works should be comparable to the opportunities and fees that are removed from the artist as a result of AI; possibly a floor of $10,000 per year for a 10-year license of an artist's 10-year body of work. The payment by corporations should be scaled to their revenue.

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

It's feasible but the corporations will protest. It's desirable for artists/writers/musicians as long as it's opt-in only, led by creator organization, and the rates are reasonable.

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

Code will probably be treated differently from writing/art/music/film.

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)?

Anyone making publicly available copyrighted materials or models trained on copyrighted materials should be responsible for securing the right to do so. End users who use models to infringe should be held liable for doing so.

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.

See answer to similar question above, re: Kinkade.

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

Artists who are losing their jobs and their ability to create art and making a living from that art would be able to survive and continue to make art. The only corporations who are currently able to compete in AI right now are all gigantic multi-billion-dollar multinational corporations or backed by those corporations and venture capitalists. If they were required to license copyrighted works, you would see more competition from smaller AI companies, who cannot at present compete.

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 to both. See also the Gebru et al. Datasheets for Datasets paper (2018) on the problems with poor data documentation overall.

15.1. What level of specificity should be required?

Metadata requirements for similar datasets in the realms of geospatial data, etc. are specified by ISO and tend to be fairly fine-grained. At minimum, **at the point of dataset creation**,

documentation should include titles of individual works, authors, and sources/source URLs for where the works were taken from.

15.2. To whom should disclosures be made?

The public.

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

Due diligence. They should be required to avoid, to the best of their knowledge, models trained on unauthorized copyrighted works.

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?

Model and data quality would increase (see Gebru et al, above). Similar metadata standards already exist for other forms of data, decided by ISO, and are in fact stricter in many cases. The impact would not be more than for other kinds of data used for non-ML purposes.

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

No obligation if there is an opt-in clearinghouse or some other way to opt-in.

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?

I am not aware.

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, the machine is creating probabilistic outputs through a statistical process that bears no relationship to authorship whatsoever. It analyzes, encodes, and embeds relationships between words that appear together or in proximity as learned through the text in its training set. Text

prompts get encoded and embedded as mathematical vectors that change the probabilities and influence the statistics of what word is output next, but there is no "creativity" involved at any point.

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?

Revisions may not be necessary but explicit exclusion of AI training from "fair use" and explicit refusal of copyrighting AI/ML outputs, e.g. through a circular, would be helpful.

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, it is not, and will lead to patent trolling and the Copyright Office being overwhelmed by registration applications for AI work. Generative AI systems and technologies are currently *highly* encouraged by billions of dollars of venture capital and stock-market funding being poured into development. They do not need extra protections. Existing copyright protections for code are sufficient.

Actual creators of artworks *are* being discouraged and *do* need protection.

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?

Stochastic outputs generated from patterns learned from other people's works used without authorization should not be copyrightable.

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?*

No and absolutely not.

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. Many examples have been given in prior answers.

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?

Because outputs are often stochastic/probabilistic and can be altered, and the same prompt may produce "All I Want for Christmas Is You" note for note on one run and *not* produce it on the next run, dataset auditing (and dataset transparency standards) for unauthorized use of copyrighted works is the only rigorous standard.

24. How can copyright owners prove the element of copying (such as by demonstrating access to a copyrighted work) if the developer of the AI model does not maintain or make available records of what training material it used? Are existing civil discovery rules sufficient to address this situation?

I am not familiar with to what extent civil discovery is able to force companies to admit what they trained on.

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?

The developer of a model that uses unauthorized copyrighted works in training or fine-tuning should be held primarily liable. End users who prompt models with unauthorized copyrighted works should be held primarily liable. Developers who knowingly incorporate infringing models should be held liable, but less so. The dataset provider, e.g. the sites hosting Books3, should be separately liable for hosting infringing content.

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

The apparent diffusion of legal responsibility makes open-source AI models more attractive for potentially harmful and illegal uses, e.g. the proliferation of deepfake AI models on Github that male students used to create pornography of their female classmates, and that male employees have used (BBC) to create and sell nonconsensual pornography of their female colleagues. I am not sure how judgments and penalties could be enforced against them, their contributors, and their users, but I am hopeful that someday they will be.

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?

I think that clause will not apply to outputs because there is no way to maintain the link between a specific work's metadata and the stochastic predicted output, but I think the section as a whole strongly applies to training data.

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

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, in all contexts and modalities, including at the beginning of AI-generated voice recordings and videos.

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

Both developers and users of models.

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

Not with good-faith actors, but it will be difficult to enforce against deliberate bad-faith actors.

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?

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?

Nothing is sufficiently accurate to be usable for text data from the highest quality models right now, because it can be passed multiple times through the model for revisions. I don't think image detectors work consistently well against the highest quality models either. However, human beings who have become familiar with AI-generated images and text seem to have been able to detect some attempts at passing off AI-generated work as human work, including for Magic the Gathering, Dungeons and Dragons, and a Clarkesworld magazine cover.

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?

Right of privacy, right of publicity, but these are not sufficient.

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, setting a floor. At minimum the right to not have one's face used in generative AI, with criminal penalties for nonconsensual pornography and child pornography generated with AI. Penalties should apply to model users, hosts of the rights-violating content, and I would argue model developers as well, because it is already technologically possible to block models' outputs if they contain sexual content/nudity/CSAI, and it is already technologically possible to perturb visual inputs so that a recognizable face is not "kept" by the model all the way to the output. The choice to *not* include these safeguards is just that: a choice.

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 are no protections at present but there should be; this could be done by explicitly stating that these are derivative works.

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?

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