Answers by Autumn Beverly, 10/30/2023

“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?”

My view is that the risks right now outweigh any benefits. And generative AI/ML are like a runaway train that needs to be managed ASAP to prevent more problems before we start looking at benefits. Otherwise the steep risk is to only obtain benefits at the expense of others, which would be the opposite of progress in any capacity.

I see the current negative effects on creators, copyright owners, and the public.

I am an animal/creature artist and I see how peoples art, photos, voices, identities, personalities, private info and more are being gobbled up by AI/ML companies and their users to be used in ways we didn’t and in most cases never would consent to even being a part of.

And I see companies using loopholes to do this (claiming research purposes on data collection, only for that data to then be provided to or down the line used in for-profit products) data laundering, and having closed datasets instead of being straightforward because they know it is legally dubious and that many would not agree to be part of this.

I have seen CEO’s involved with generative AI/ML misrepresent creativity and creators to suit their narrative.
For example dehumanizing artists as “tools” called our work “tedious” and us “elitists” “gatekeeping talent” for speaking up about being exploited, which has further stoked the flames of resentment towards artists by their follower/userbase.

Including new kinds of harassment their technology has opened the doors to.

Such as people fine tuning AI models on a specific artists entire body of work, deliberately in a targeted way. Even having contests to see who did it the best, such as what happened to Sam Yang after he expressed he did not want his art used in this way.

These AI enthusiasts have also come up with a number of uncreative slurs to refer to artists by.

I’ve seen undress apps that have been used on and with womens photos without their consent and other examples of non-consensual pornographic material. That has also been used to generate sexualized content of children.

I have seen minorities misrepresented by image generators by companies claiming to offer representation.

I also think of how people such as the Corridor Crew on YouTube are profiting off fine tuning models on copyrighted works (that they do not own the rights to). As they did with the animated film “Vampire Hunter D Bloodlust”. Even if it wasn’t done with malice, this should not be legal by existing copyright law.

I have seen artists names in search engines pull up more AI generated images using their name as prompts than their actual work. Making it harder to actually find them.

I’ve seen books on Amazon published using ChatGPT in authors names without their knowledge or consent.

I’ve seen peoples voices used without their consent to say things that could be damaging to their reputation and careers.

My grandmother was nearly scammed by someone using my cousins voice telling her he was arrested and needed bail money. It was not him, he was fine. She had no money to give and that is the only reason they didn’t get her money.

I have seen visual art hosting websites such as DeviantART and ArtStation become overflooded with generated images uploaded to their platforms so much so that it is now hard to find actual human artists under so much AI generated output.
Which is likely affecting artists' ability to be found for work opportunities. Rendering the use of these places nearly useless for finding work now, whereas prior to generative AI they were bustling hubs used by many for that purpose.
The websites themselves have taken no meaningful action against this issue, DeviantART even embraced it by implementing an AI model (DreamUp, using stable diffusion as a base) to their very website as a *feature*.The promise of profit outweighing even their target user base - artists, encouraging their exploitation even by places that were meant to be spaces designed for them, effectively pushing many of them out.
Other than adding filters for users to specifically select if they don’t want to see it, that are useless because users often do not tag things correctly.

It is also negatively affecting artists behaviors and mental health as well as their relationship with their work.
There was an uptick in artists talking about unaliving themselves in browser search results on the topic, because of generative AI in its current state many are fearing they have no possible way to make money and support themselves in a career they dedicated their entire lives skillbuilding and are in debt in schooling for nor any longer have control of their own works. We are also losing future artists before they even begin as this advent has influenced some to decide not to even think about pursuing art.

It's also affected even hobbyists, I've seen people say all of it has caused them to lose motivation to make art.

(Less human art is a negative impact on human culture.Of which art is a major facet to.)

Many artists have stopped posting online, moved to obscure websites or put their artwork behind paywalls, and/or are using adversarial measures to try protecting their work online such as Glaze. Which is all extra labor for the artist and the former two may limit who sees their art. And affects their ability to find work. But it's that important to them that many are actually weighing these options and risks and making these decisions right now. The ones who know and understand what is even happening anyway. Many still don't even know or understand what's happening to make any decisions at all.

And so much more, an entire book could be written about the risks and problems.

"2. Does the increasing use or distribution of AI-generated material raise any unique issues for your sector or industry as compared to other copyright stakeholders?"

I am a freelance visual artist. Right now artists have very poor union representation to negotiate not being exploited by this tech in the job market, we are also often so poorly paid that most of us cannot afford lawyers for good legal representation either. Which is further exasperated by image generators eliminating art related jobs and devaluing our work, So we are especially vulnerable to abuse of our work and rights to it, which is happening now at an unprecedented scale thanks to generative AI/ML.

"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"
Harry H. Jiang, Lauren Brown, Jessica Cheng, Mehtab Khan, Abhishek Gupta, Deja Workman, Alex Hanna, Jonathan Flowers, Timnit Gebru.
Link: https://dl.acm.org/doi/10.1145/3600211.3604681

"Glaze: Protecting Artists from Style Mimicry by Text-to-Image Models"
Shawn Shan, Jenna Cryan, Emily Wegner, Haitao Zheng, Rana Hanocka, Ben Y. Zhao
Link: https://arxiv.org/abs/2302.04222

"On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?"
Emily M. Bender, Timnit Gebru, Angela McMillan-Major, Shmargaret Shmitchell.

Link: https://dl.acm.org/doi/10.1145/3442188.3445922

“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?”

The Berne Convention should be considered as a way for everyone internationally to have their works protected. I think it’s very important that there is some consistency, however that should not mean we mimic excessive slack in some other countries. I like that China has regulation insisting AI generated outputs must be clearly labeled or identified as such.

“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. “

Yes. ML companies should have to comply with consent, credit, compensation and transparency regarding where they are getting all of their training data from and what material is in their datasets.
This is the only way to verify if they are being compliant with existing copyright law among other laws.

And to eliminate loopholes such as “for research purposes” being used currently by ML companies that are ultimately profiting off their products trained on copyrighted works and/or end users that may also use it for profit. The scope of exploitation with generative AI is so wide it could warrant having specific legislation to prevent such abuses of current loopholes that did not anticipate this technology.

Also, we desperately need ways for people with copyright but little means to officially register their works to have more protection. It should carry more weight than it currently does, and people without means to register everything are being taken unfair advantage of with too few ways to act against it. And/or make officially registering works significantly more affordable and accessible. Not having the money to register every creative work you make among other things should not mean you have less rights to your intellectual property.

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? 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)?

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?
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?
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. “

Anything that can be collected, is collected. Images, video, audio, text, biometric data, … ect. LAION 5B alone has over 5 billion images. It’s difficult to describe the scope of even just this.

Usually it’s scraped from various places across the web with bot web crawlers.
And the links are saved to a dataset.
But individual pieces can also be manually added.

Stock websites like Adobe Stock are beginning to have language in their user agreements that “allow” them to use stock contributors work in features like Firefly which is their generative AI model trained on contributors works that are hosted on the stock site, (along with “openly licensed content”, and public domain content (datasets closed we cannot verify)) often without explicitly telling stock contributors their work was being trained/fitted on AI/ML, and without explicitly asking their consent for this particular use that is fundamentally different and has many more implications that could be negative to the user than what stock images were previously known to be used for.
Despite the obvious deception these companies claim their models/systems are ethical and “safe for commercial use”.
Companies tend to boast moreso after a poor payment system that isn’t very transparent is implemented to “compensate” contributors for their work being used in generative AI/ML.

In the case of Adobe, it’s even more concerning since they do not seem to have the resources to curate and manage their stock website to abide by their own terms as works have been uploaded with artists names without their consent, and copyrighted materials are sometimes uploaded without the consent of the copyright owner.

Because of this it is possible copyrighted material was used to train Firefly without consent from the copyright owner.

Datasets are usually to be reused for retraining new, more advanced, or specialized models.
Companies claim they are not retaining the data, but only a database of links to the data.
Another abused loophole.

Though some individuals have said they have datasets on hard drives.

Corridor Crew showed they had computer folders for screenshots of “Vampire Hunter D Bloodlust” to fine tune an AI model on.

“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.
7.2. How are inferences gained from the training process stored or represented within an AI model?
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? 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? If so, is it economically feasible? In addition to retraining a model, are there other ways to “unlearn” inferences from training?
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? “

It’s different depending on the type of model.
But briefly, models are fed collected data which it is then able to copy in different ways.
Image generator models specifically are trained or “fitted’ on collected data, data which may be “temporarily” downloaded for this purpose for however long the company decides, images are associated with words, and through weights and balances, lossy compression, and usually - diffusion, are regenerated as output in somewhat randomized ways via an algorithm.

Inferences are numerical values that are set in the models parameters during training.
But this data is not currently identifiable back to the source training data.
In an analogy it’s almost like saving a lossy copy of a file in a different file format such as copying a .PNG but saving it as a .JPEG, except it doesn’t show you a picture to identify a resemblance to where it came from, and it can’t be converted to the original file type.

There currently is no known way for AI to “unlearn” inferences. Although it may be possible, it would also be hard to prove successful.

The surest and most thorough known way to remove inferences is to destroy the old model and retrain without the data you don’t want.

Retraining is a costly and environmentally unfriendly process, but the onus should not be on infringed copyright owners, who did not ask to be part of AI/ML models in the first place.
AI/ML companies should be more responsible.

Yes, and no.
It is possible to recognize many pieces easily from AI generated outputs that were “overfitted” due to many copies of the same image being in the training data, overfitting can result in near identical copies and close resemblances. But it can be difficult otherwise. Some images may have a few tells to a certain image. Although it should be noted that the training data is what gives a model its quality and without the fitting on training data it cannot produce good results because it is a reflection of such.

“Artificial Intelligence and Copyright 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.
8.1. In light of the Supreme Court’s recent decisions in Google v. Oracle America41 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?
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?
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?
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?
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? “

Research could be an exception as fair use but should actually be STRICTLY research, using research as a loophole to then build and distribute products built on then stolen copyrighted material to be used by individuals and companies alike for profit should fall against copyright law.

In Germany section 60d of “Text and Data Mining for Scientific Research Purposes” there is a rule where “non commercial research purposes” does not apply as an exception if research organizations are cooperating with a private enterprise which exerts a certain degree of influence.
I think this would be a good consideration for the U.S. too to prevent twisted conflicts of interest that can compromise copyright among other things

Entities supplying copyrighted material to another party for generative AI training should be held

responsible as an accomplice to a crime if it was found not to be used for research purposes and especially if that company had donated or invested funds into said entity or employees are shared between them.
Entities, individuals, companies etc. all should seek consent from copyright owners if they want to use their copyrighted materials for specific uses outside research.

The volume of data for training purposes is usually large, millions to billions.
But the volume of data should not affect fair use analysis.

Measuring market value of a copyrighted work used to train an AI model can be particular to the medium or individual work being infringed, as we are talking about many different kinds of mediums that are being infringed upon.
I think all the suggested examples should be considered for analysis and then some as applicable to the medium and work. But any one offense should be enough proof.

.

.

“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)?
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
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
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 feasable to get consent in advance from copyright owners?
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?
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? “

100% it should absolutely be a clear affirmative consent first basis. As copyright already functions this way.
Opt-out is okay to offer AFTER consent first.
Offering zero options for consent but providing an opt-out is ineffective and not a functional way to do this, this goes against current copyright laws functionality.
And puts the onus and unnecessary labor on the copyright owner.
Whom may not even be aware their work has been taken.or that opt-out even exists or may be recently deceased, or disabled, or otherwise not have access to the internet to opt-out.

Further this does not protect against models that have already been trained on a copyrighted work. As nobody knows how to untrain it, they cannot actually opt-out.
Opting-out is also tedious, and can be incredibly time consuming as a copyright owner may have many works made over a lifetime and each one can be scraped multiple times, it's hard to even search places like HaveIBeenTrained to opt-out successfully as you never really know if you found all of your work to opt it out entirely, and that's if you even know what all they took from you and doesn't account for things like other more private data that is harder to search for or one might not even think to check for. There are many different AI/ML models and only more will continue to be made, it is unrealistic to expect a copyright holder to have to search them all out and do this process for every model in existence and every work they made. I'd even argue that it would be physically impossible to maintain.
Consent first is the only way for copyright owners to have any control of their works. And the only way to make this functional and compliant with copyright at all.

Due to the unusually competitive nature and large scope of generative AI/ML which is difficult to predict, I think yes. Consent of the copyright owner should be required for most or all uses. To avoid loopholes and abuse.

I do think that again, because of the unusually competitive nature and large scope of generative AI/ML, which may be used in unpredictable ways that can negatively impact people, that a creator of a work who perhaps for their job has assigned copyright to say a company, should still have say whether their individual labor/contributions/and/or identity is used to train AI/ML.

"10. If copyright owners' consent is required to train generative AI models, how can or should licenses be obtained?
10.1. Is direct voluntary licensing feasible in some or all creative sectors?
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?
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?
10.4. Is an extended collective licensing scheme51 a feasible or desirable approach?
10.5. Should licensing regimes vary based on the type of work at issue? "

Licenses should be obtained via explicit clear fair legal contract.
They should describe exactly what the works will be used for, such as how long the dataset will exist and be used, what type of training, the type of model it will train, how the model will be distributed/used and so forth. And when if ever the dataset and/or model will be destroyed. Basically how long is the license for? And when will it expire?

“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)? “

Each step should require consent/licenses. And they must be clear, fair, and valid.
Otherwise it may be difficult for the copyright owner to know what their works are being used for and unauthorized use could slip through the cracks.
If the cost is high to obtain licensing that is not, should not, and has never been the copyright owners problem.

“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’s possible if obvious, though it might not always be obvious.
If an artists name is used as a prompt that can often account for the overall look of a generated image for one example. Overfitting is another example. Fine tuned/LoRA models are another example.
However this is not the limit of examples, some generated images just show obvious elements of existing artworks. Some are less obvious and harder to discern visually.
However, the existence of copyrighted works in datasets are hard evidence of their use in the models overall functioning. It cannot function without the training data, which is crucial to generate any output.

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

It would ensure copyright owners have control of their works, are less likely to be put out of work and replaced by AI/ML models copying and reproducing their works infinitely to compete against them in the same markets. It would ensure consent, credit, and compensation of the copyright owner. (We also need transparency)

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

Generative AI/ML models need to be held accountable for all kinds of infringement if committed, and people that cannot afford to register every work they have made over their lifetime should be better protected and still have control of their work. Again, not having the money or resources to register should not mean a person has less rights to their work.

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?
15.1. What level of specificity should be required?
15.2. To whom should disclosures be made?
15.3. What obligations, if any, should be placed on developers of AI systems that incorporate models from third parties?
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?”

Datasets should be 100% transparent and searchable in their entirety by the public.
Yes creators of datasets should also be obligated to be just as transparent.
Obligations are on everyone involved to ensure correct and appropriate licensing for all copyrighted works within a dataset. As well as verifying if any works collected are truly public domain if applicable.
This ensures copyright holders can verify if they have been infringed upon and allows trust between copyright holders, and consumers/users of the models.

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

Email, or physical letter with legal contract clearly describing intended use and asking the copyright owners clear consent. Or they destroy the model and retrain if they do not get consent from the copyright owner. All uses of the infringing model should be halted until they get clear consent via legal contract.

“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?”

Privacy laws.

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. With regards to generative AI/ML they are a customer, a client to a generative function. Not an author. They are not the "master mind" of an algorithmically generated output.

"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? "

I think it is already pretty clear.

"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?
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?"

No.
No.
No.
Legal or copyright protection for generative AI models could be harmful to copyright owners who are already suffering from a power imbalance with loopholes and such protecting these models from legal action despite mass exploitation.

"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.

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. All outputs are derivatives of works in the training data.

“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?”

Not alone. Some outputs can be recognizably similar but not all.
Hardest evidence is in training data. All outputs are a reproduction of what is in the training data.
And therefore derivative. The models cannot create outputs without the training data.
It is a crucial part of its function..

“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?”

It can be difficult unless there is a remarkable obvious likeness to the copied work.
Which is not always the case.
This is why datasets used in training MUST be transparent and public.
I do not think existing civil discovery rules are entirely 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?
25.1. Do “open-source” AI models raise unique considerations with respect to infringement based on their outputs?53 “

Everyone involved. There are many steps and each of those steps need to be responsible.
Otherwise we would have loopholes where one party escapes liability but may still profit from infringement (further encouraging reckless behavior/more infringement) by being able to just shift blame to another party.

“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? “

It should be treated the same way it is applied to everyone else.
Removal or altering of CMI is illegal without the authorization of the copyright owner.

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

The scope is inherently immense yet hard to get exact measures on with generative AI/ML models. Such cases where this technology is implicated in infringement should be taken as especially serious offenses.

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?
28.1. Who should be responsible for identifying a work as AI-generated?
28.2. Are there technical or practical barriers to labeling or identification requirements?
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?”

Yes. It should apply to all generated outputs. They should all be clearly labeled or publicly identified as AI/ML generated.

The people programming and distributing the AI model should be initially responsible for ensuring outputs are clearly identified as AI/ML generated. However this should also be made clear when posted online or applied anywhere by the end user, they should also be held responsible if misleading people or implying otherwise. Such as with news articles with images and/or text generated by AI/ML. Removal or obfuscation of labels or identifiers should be punishable by law.

Watermarks and/or clear obvious written or audible disclaimers paired with outputs should suffice.

“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? “

AI detection tools such as HIVE are notoriously inaccurate and give false positives.
They are easy to get around such as with image detection being fooled by flipping an image or adding filters/noise to images to not recognize an image as AI generated. Or mistaking authentic works made in popular art styles frequently reproduced by image generators as AI generated.
And has shown to cause harm and mistrust because people tend to put too much trust in the results of the AI detection, this has led to many students for example being falsely accused of using ChatGPT in school. Which can ruin their reputation and cost them their education unfairly with often no way or opportunity to prove their innocence.
Many visual artists have also been falsely accused, which can cause them to lose clients and networking.AI detectors are also too trusted in some art contests which can lead to unfair outcomes.

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?”

Not enough.

But some examples are Right to Publicity, Impersonation, Unfair Competition, Privacy, Defamation, Trade Secrets, etc. And state specific laws like the Biometric Privacy Act in Illinois.

“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 it should preempt state laws.
The contours should be thoroughly discussed and considered. As the scope of AI generated material and who, as well as how it impacts people is large.
But it is also moving fast, so we need to get the ball rolling on these types of protective actions now.

“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? “

This is an issue with how existing language used to describe art & art making prior to generative AI is now incorrectly being used to describe how AI models generate outputs.
A lot of what we call “style” when applied to humans describes the result of technique.

But these models are not using human techniques and processes to make their outputs nor merely just mimicking the “style” of images in the human sense of the word.
An AI doesn’t even “understand” what style exclusively is when asked to reproduce it.
They are copying/reproducing various elements USING the actual images from training data algorithmically, with lossy compression, and diffusion, etc.
As an analogy, if a human visual artist did this it would be more equivalent to a collage of other images, rather than merely a “style” from other images. If that makes sense.

So yes, I think there should be protections against this, specifically applying to mechanical/technological operations of reproduction.

“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?”

I think if it does apply it’s either not clear enough or enforced. Since people’s voices are being

exploited right now without much action being taken.
And may require legislative attention. Especially in cases where an actor is voicing a character they may not own the rights to in order to take legal action themselves. They must wait on the rights holder to act and they may not at all. Which leaves the voice actor in an unfair and vulnerable position where their voice can be abused and negatively impact their reputation and work.

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

Please please please make it easier for people to protect their works they actually created.
It is unreasonable to expect everyone to have the means to register every work they make throughout their lives under the current system. This puts so many in a vulnerable position to be exploited which has led to so much of the issues around generative AI and more.
I will reiterate that wealth, or accessibility, should not be a deciding factor for if you have less rights to your intellectual property.