Brooklyn Law Incubator & Policy Clinic

Professor Jonathan Askin
Founder/ Director

Nicholas Castro and Jimmy Nguyen
250 Joralemon Street
Brooklyn, NY 11201
(718) 780-0622
October 30, 2023

Response to the USCRO's Notice of Inquiry and Request for Comments re.: Copyright Law and AI Policy Issues [Docket No. 2023–6]

On behalf of the Brooklyn Law Incubator & Policy Clinic (BLIP) at Brooklyn Law School, we would like to submit our response to the USCRO's Notice of Inquiry (NOI) concerning artificial intelligence, Copyright law and policy issues. As passionate advocates for technological innovation and intellectual property rights, we commend the Office for its proactive stance in addressing critical issues surrounding generative AI and its impact on U.S. Copyright law and policy. You may find our response below, which contains answers to nearly all of the questions in the NOI.

General Questions:

1. 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 systems have many benefits. Firstly, material produced by these systems is a different and new type of creative expression that benefits the artistic landscape and society as a whole. Art, whether in the form of paintings, music, or literature, is often defined as whatever stirs emotions and stimulates thoughts in its beholder. Its role in society has been to promote the expression and communication of ideas and new perspectives, which in turn enriches society and fosters community. Understandably, what has been considered art has changed throughout the years, and technological advancement has often been the catalyst in the development and

recognition of new artistic works.

Copyright laws have come to recognize multiple novel types of artistic works over the past few centuries, evolving from written works and paintings to motion pictures and graphic designs. The kind of pushback AI-generated material has received is not new or peculiar to this novel technology. When photographic cameras were invented, many feared that photography would eventually replace the art of painting.[1] Painters thought that once people could easily access cameras, their job would become obsolete.[2] At the same time, critics argued that photographs were the result of mere tools and lacked any creativity and feeling and, thus, were not art.[3] Photographs, however, were eventually recognized for their artistic value and were granted copyright protection once Congress recognized that technological advancements had "made a change in the law necessary."[4] To recognize AI-generated works as original expressions of ideas and creativity would be in line with the protection and recognition of novel forms of art in the past.

AI-generated material has the same potential to benefit the artistic landscape and enrich society as any artistic media before it. It is able to evoke emotions and to stimulate thoughts and feelings, and it provides a glimpse into a world to which most people would otherwise have no access. Over the past few years, AI-generated works have permeated many spaces traditionally occupied exclusively by human-made artistic works. In many instances, the public and the artistic community have already come to recognize AI-generated materials as art. For example, in 2022, Colorado artist Jason M. Allen won the Colorado State Fair's annual art competition with an AI-generated artwork.[5] The judges stated that they awarded him first place based on the story and spirit of the work.[6] Yet, the same artwork that was praised as evocative and reminiscent of Renaissance art was denied copyright protection simply because it was AI-generated. Many other AI-generated creative works have suffered the same fate. Dr. Stephen L. Thaler's "A Recent Entrance to Paradise" and Kris Kashtanova's Zarya of the Dawn are among them. These examples make it abundantly clear that AI-generated creative works are already among us and are likely to become more prevalent. To continue to refuse to recognize that AI-generated works fall within the scope of the Copyright laws injures creators whose expressions are worthy of recognition and deprives both the creative community and the world of all that AI-generated material has to offer.

---

[1] *See* Jordan G. Teicher, *When Photography Wasn't Art*, JSTOR DAILY (Feb. 6, 2016), https://daily.jstor.org/when-photography-was-not-art/.
[2] *See id.*
[3] *See id.*
[4] S. COMM. ON PATENTS, TRADEMARKS, AND COPYRIGHT OF COMM. ON THE JUDICIARY, 86TH CONG., STUD. ON COPYRIGHT LAW 1–4 (Comm. Print 1960).
[5] Sarah Kuta, *Art Made With Artificial Intelligence Wins at State Fair*, SMITHSONIAN MAGAZINE (Sept. 6, 2022), https://www.smithsonianmag.com/smart-news/artificial-intelligence-art-wins-colorado-state-fair-180980703/.
[6] *Id.*

Secondly, some individuals benefit significantly from using generative AI as an assistive tool to produce creative works they could not create otherwise. For instance, individuals suffering from certain mental or physical health conditions, such as Bipolar Disorder, Borderline Personality Disorder, Down Syndrome, and many more, often face significant cognitive, emotional, and mental challenges. Due to their medical conditions, these individuals may require the assistance of AI in order to put their ideas into a tangible expression. In these instances, generative AI systems do not replace human creativity and effort, they simply complement it by functioning as a bridge between users' creative visions and their physical or cognitive limitations.

Finally, as with any new technology, generative AI systems present some risks. However, risk is not inherent to the technology itself but rather from its unregulated use. Copyright holders have expressed concerns that their copyrighted materials have been used to train AI systems without their authorization. As a result, generative AI systems may generate outputs that resemble existing works beyond what would be considered fair use. However, these concerns stem from a lack of clear guidelines on AI training and disclosure requirements on the processes through which AI generates creative works.

Some creators and artists have also voiced their fear that generative AI systems will replace them and make their art obsolete. These arguments are likely baseless if considered within the history of art and new artistic mediums – just like photographers did not replace painters, it is unlikely that AI-creative works will replace human-made ones. Even if, arguendo, a future where AI artists have replaced human artists could be conceived, this once again would be the result of allowing generative AI systems to produce art without regulation. The Copyright Office has the authority and resources to implement guidelines to ensure that human authorship is safeguarded while also allowing for the protection of AI-generated works. Although AI and humans are often pitted against each other in people's minds, the early implementation of regulations would ensure that both could flourish and co-exist without harming the other.

No one can realistically predict how art, writing, music, and other creative output will evolve over the next years, decades, and centuries. Nevertheless, we should know by now that when creative forces interact, they create a ripple effect that leads us into new creative vistas. When we allow artists to build new art on top of the art that preceded them, amazing, heretofore unimaginable creative works and forms emerge. There can be little doubt that by encouraging the mixing of human ingenuity with the enabling power of AI, an infinite array of new creations will emerge, which will ultimately enhance the human experience.

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.

1. What label should be applied to content produced by generative AI? – Epstein, Fang, Arechar, Rand (2023).

2. Achieving Balance in Generative AI: Open-Source Versus Proprietary Models – Dr. Ling Ge, General Manager and Chief European Officer, Tencent.

3. Why a Federal Right of Publicity Statute is Necessary – Kevin L. Vick, et al. (2011).

4. Thirty-Six Views of Copyright Authorship, By Jackson Pollock – Dan L. Burk (2020).

5. There's No Such Thing as a Computer-Authored Work - And It's a Good Thing, Too – James Grimmelmann (2016).

6. AI Chatbots, Health Privacy, and Challenges to HIPAA Compliance – Mason Marks & Claudia E. Haupt (2023)

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

Broad international consistency in the purpose and means of copyright law could enable greater commercial and artistic collaboration across the globe, deliver a simpler and more reliable expectation for rights holders in resolving disputes, and build further goodwill among member nations to an international agreement. We as law students hope to see America take a leading role in developing state of the art international agreements governing the emerging dynamics of generative AI and copyright. The following table has been constructed to help the USCRO stay informed regarding international approaches relating to copyright and AI:

Jurisdiction	**Author**	**Duration of Copyright in literary, dramatic, musical or artistic works**
Hong Kong	**Section 11(3) Copyright Ordinance.** In the case of a literary, dramatic, musical or artistic work which is computer-generated, the author is taken to be the person by whom the arrangements necessary for the creation of the work are undertaken.	**Section 17(6) Copyright Ordinance Duration of Copyright in literary, dramatic, musical or artistic works.** If the work is computer-generated the above

[7] For example, several jurisdictions have adopted copyright exceptions for text and data mining that could permit use of copyrighted material to train AI systems. Separately, the European Parliament passed its version of the Artificial Intelligence Act on June 14, 2023, which includes a requirement that providers of generative AI systems publish "a sufficiently detailed summary of the use of training data protected under copyright law." *See* Artificial Intelligence Act, amend. 399, art. 28b(4)(c), EUR. PARL. DOC. P9_TA (2023)0236 (2023), *https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_EN.html*.

		provisions do not apply and copyright expires at the end of the period of 50 years from the end of the calendar year in which the work was made.
India	**Section 2(d)(vi) Copyright Act 1957.** (d) "author" means, — (vi) in relation to any literary, dramatic, musical or artistic work which is computer-generated, the person who causes the work to be created.	No clear provision; Generally–life of author + 60 years (S22); Anonymous/Pseudonymous–60 years after first published.
Ireland	**Art. 21 Copyright and Related Rights Act 2000.** In this Act, "author" means the person who creates a work and includes: (f) in the case of a work which is computer-generated, the person by whom the arrangements necessary for the creation of the work are undertaken.	Generally–life of the author + 70 years (S24) **Art. 30 Duration of copyright in computer-generated works.** The copyright in a work which is computer-generated shall expire 70 years after the date on which the work is first lawfully made available to the public.
New Zealand	**Section 5(2)(a), Copyright Act 1994.** Meaning of author For the purposes of subsection (1), the person who creates a work shall be taken to be,— (a) in the case of a literary, dramatic, musical, or artistic work that is computer-generated, the person by whom the arrangements necessary for the creation of the work are undertaken.	Generally–life of the author + 50 years **Section 22(1), Copyright Act 1994.** If the work is computer-generated, copyright expires at the end of the period of 50 years from the end of the calendar year in which the work is made.
South Africa	**Section 2(h), Copyright Act 1978.** 'Author', in relation to- (h) a literary, dramatic, musical or artistic work or computer program which is computer-generated, means the person by whom the arrangements necessary for the creation of the work were undertaken.	Generally–life of the author + 50 years; if published, performed in public, offered for sale to the public, or broadcast, then 50 years from the end of the year in which the first of the said acts is done.

United Kingdom (UK)	**Copyright, Designs and Patents Act 1988, section 9(3).** Authorship of work— In the case of a literary, dramatic, musical or artistic work which is computer-generated, the author shall be taken to be the person by whom the arrangements necessary for the creation of the work are undertaken.	Generally–life of the author + 70 years **Section 12(7)** If the work is computer-generated the above provisions do not apply and copyright expires at the end of the period of 50 years from the end of the calendar year in which the work was made.

Training

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)?**

AI models can be trained on sets of data taken from incredibly large databases, usually run by third parties. These parties provide services to train machine learning models on datasets made from vast amounts of information collected from unwitting internet users. The United States has an insufficient legal framework for the internet that is allowing these databases to be comprised of billions of datapoints of user data.[8]

Publicly available information is irrevocably authorized for scraping by third parties unless otherwise specified through coding.[9] Publicly available information is information that can be publicly viewed on any website without the viewer having to create an account.[10] However, when users create accounts, they must agree to terms of use that almost always give platforms licenses to use whatever a user does on their platform.[11] Platforms rely on users clicking "I agree" without reading the terms of use. These agreements are overwhelming, and a

[8] Matthew Humphries, *Class-Action Lawsuit Accuses Oracle of Tracking 5 Billion People*, PCMAG (Aug. 23, 2022), https://www.pcmag.com/news/oracle-faces-class-action-lawsuit-over-tracking-5-billion-people.
[9] *See* hiQ Labs, Inc. v. LinkedIn Corp., 31 F.4th 1180, 1184–85 (9th Cir. 2022) (affirming a preliminary injunction granted to hiQ Labs in 2019 "forbidding LinkedIn from denying hiQ access to publicly available LinkedIn member profiles").
[10] *See id.* at 1199.
[11] For example, Google's Privacy and Terms says "We need your permission if your intellectual property rights restrict our use of your content. You provide Google with that permission through this license." Google, *Privacy & Terms*, https://policies.google.com/terms?hl=en-US#toc-permission (last visited Sept. 22, 2023).

consumer simply cannot use the internet without agreeing to terms of use of some browser, website, or other platform.[12]

Courts have recited a need for actual notice in these terms of use internet agreements, while, in actuality, requiring what many would understand to be inquiry notice.[13] Inquiry notice only requires that a user *should* know they agree to a website's terms of use. Nothing requires that a user knows what they are agreeing to, only that terms of use exists for a website.[14] Courts have been hesitant to find a lack of mutual assent between user and provider because of the far-reaching impact of finding lack-of-notice. Therefore, an incredible number of users give data away to websites without knowing that they are doing so. Frankly, they have no choice.

Because all this is legal, platforms collect data as first-party data brokers and then sell the data to third parties.[15] These third parties then use the data how they see fit, sometimes tailoring the data to fit regulations in the EU and California. Therefore, the data likely being used to train powerful commercial AI models necessitates large use of consumer data, obtained under an unfair (but not illegal) framework.

- **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?**

Unfortunately, many privacy notices and terms of use across the internet contain language that gives a website a license to use what would normally be considered a user's intellectual property. For example, Google's Privacy and Terms says "We need your permission if your intellectual property rights restrict our use of your content. You provide Google with that permission through this license."[16] Therefore, anyone using Google's popular platform to post or create gives Google license to use that post or creation.[17] This means that, if Google would like to use your copyrighted work to train an AI model, it could do so without asking your permission. All it takes is signing up to use the platform.

Another legal answer owners of AI models may levy is the *Sony* defense. This holds that one who merely supplies a device is not liable for the resulting infringements so long as the device is capable of substantial non-infringing uses.[18] Because displaying an AI model's work and processing in a readable format is so difficult (and even more difficult for a non-computer scientist to understand), it may be hard to prove the AI model is *not* capable of non-infringing

[12] Jason Cohen, *It Would Take 17 Hours to Read the Terms & Conditions of the 13 Most Popular Apps*, PCMAG (Dec. 4, 2020), https://www.pcmag.com/news/it-would-take-17-hours-to-read-the-terms-conditions-of-the-13-most-popular.
[13] Sw. Airlines Co. v. BoardFirst, L.L.C., No. 3:06-CV-0891-B, 2007 WL 4823761, at *5 (N.D. Tex. Sept. 12, 2007); Wu v. Uber Techs., Inc., No. 2022-05749, 2023 WL 6150415, at *1 (N.Y. App. Div. Sept. 21, 2023).
[14] *See id.*
[15] Charles Street, *How Data Brokers Steal & Sell Your Identity and How You Can Stop It*, CYBERGHOST, https://www.cyberghostvpn.com/en_US/privacyhub/data-brokers-put-a-price-tag-on-your-privacy-and-then-sell-it/#:~:text=First%2Dparty%20data%20broker%20companies,%2C%20online%20 activity%2C%20and%20more (last visited Sept. 22, 2023).
[16] Google, *supra* note 11.
[17] *See, e.g.*, Field v. Google Inc., 412 F. Supp. 2d 1106, 1115–16 (D. Nev. 2006).
[18] Sony Corp. of Am. v Universal City Studios, Inc., 464 U.S. 417, 442 (1984).

uses. This may allow creative lawyers to subvert legal remedies to infringement of copyrighted materials.

- **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?**

To our knowledge, there are no widely used AI models that advertise their models as pulling from exclusively non-copyrighted works. However, some are trying.[19] There are entities that are devoted to collecting a database of already-licensed works.[20] If an AI model were to pull from these databases, where users have intentionally created a library of works in the public domain, there would be little to no copyright issue.

- **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.**

Although training data can be deleted after training is completed in most cases, most entities keep the data.[21] This data must be kept according to current state and international law (e.g. CRPA, GDPR), but entities wish to retain the data for training more advanced models. For example, OpenAI's policy allows retention, and some models rely on retained data for their functionality.[22]

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.

A finding of fair use for AI training in certain models is highly likely when viewed in light of *Authors Guild v. Google, Inc.*[23] In *Author's Guild*, the Second Circuit addressed the issue of whether Google Books' use of complete digitized copies of books that resulted in "snippets" displayed to online users was fair use.[24] In determining that this was fair use, the court's analysis largely focused on the fact that despite Google's commercial motivations, their

[19] Andres Guadamuz, *Using Creative Commons Images to Train Artificial Intelligence*, TECHNOLLAMA (Mar. 13, 2019), https://www.technollama.co.uk/using-creative-commons-images-to-train-artificial-intelligence.
[20] *Share Your Work*, CREATIVE COMMONS, https://creativecommons.org/share-your-work/ (last visited Sept. 22, 2023).
[21] *See, e.g.*, *What Happens to the Training Data After Your Machine Learning Model Has Been Trained?*, STACKEXCHANGE, https://ai.stackexchange.com/questions/7739/what-happens-to-the-training-data-after-your-machine-learning-model-has-been-tra (last visited Oct. 20, 2023).
[22] "We retain certain data from your interactions with us, but we take steps to reduce the amount of personal information in our training datasets before they are used to improve our models. This data helps us better understand user needs and preferences, allowing our model to become more efficient over time." Michael Schade, *How Your Data Is Used to Improve Model Performance*, OPENAI, https://help.openai.com/en/articles/5722486-how-your-data-is-used-to-improve-model-performance (last visited Oct. 20, 2023).
[23] Authors Guild v. Google Inc., 804 F.3d 202, 214 (2015), *cert. denied,* 578 U.S. 941 (2016). Similarly, in *HathiTrust*, the Second Circuit found that creation of a database of digital works was fair use. Authors Guild v. HathiTrust, 755 F.3d 87, 105 (2014). Though this was a non-profit educational use, a court could find that AI models and systems with similar tasks and end-users could be fair use.
[24] *Authors Guild*, 804 F.3d at 214.

use of the digital copies was transformative, and that the small portions of the digital books that users could see were not a substitute for the original work.[25] Similarly, a court analyzing the use of copyrighted materials to train AI models would likely find the use transformative and weighing in favor of fair use. Even if an AI model's input contains data of an entire work, and the use is for commercial purposes, a court could find a transformative use if there is a justified purpose in "provid[ing] otherwise unavailable information about the originals"[26] because it serves the goal of copyright law.[27]

However, this analysis is qualified by several points mirrored in *Author's Guild.* First, the output displayed to users of Google Books was only "snippets" containing approximately "an eighth of a page" of text.[28] This implies that fair use in AI training would require that AI systems' outputs to end users be restricted to less than full reproductions of the work.[29] This requires AI models and systems to implement guardrails around user outputs, such as the inability for users to command a tool to make works in the style of a copyrighted work.[30] Second, the court notes that Google did not charge to use the search and snippet functions or display advertising on the search pages.[31] However, the court did not rely on the lack of direct profit from the service and still weighed Google's overall purpose as commercial. Given this, it is possible to assume that a direct profit from an AI system's output could weigh differently in a commercial analysis, towards a finding of infringement. Third, the court reasoned that the licensing market for the full version of the digitized books was more extensive, and unharmed by Google's use of the snippet function, and thus, there was not a licensing market harm under the fourth factor.[32]

Therefore, copyrighted works in training AI models may be transformative and fair use in certain circumstances; however, anticipated licensing regimes, the commerciality of the tool, and restrictions on end-user output will limit this. Additionally, due to the unprecedented volume of copyrighted works and unpredictable uses of the scraped data, we would strongly advocate for a licensing and consent regime as outlined below.[33]

---

[25] *Id.* at 207–08; *see also* H.R. REP. NO. 94-1476, at 66 (1976) (discussing that commerciality is not determinative in fair use, but should be weighed along with the other factors).
[26] *Authors Guild*, 804 F.3d at 215.
[27] The goal of copyright law is "promote the progress of science and the useful arts." U.S. CONST. art I, § 8, cl. 8.
[28] *Authors* Guild, 804 F.3d at 209.
[29] The Second Circuit notes that full "digitized versions" of books made available to the public would be infringing derivatives. *Authors Guild*, 804 F.3d at 225.
[30] We have already seen incidents of end users attempting to command AI tools to make an output in the style of a specific creator's work that is covered by Copyright. *See* Melissa Heikkilä, *This Artist is Dominating AI-Generated Art. And He's Not Happy About It*, MIT TECH. REV., (Sept 16, 2022) https://www.technologyreview.com/2022/09/16/1059598/this-artist-is-dominating-ai-generated-art-and-hes-not-happy-about-it/. Presumably, this type of activity would change a fair use analysis if the output were serving as replacement for the work. *See* Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith, 143 S. Ct. 1258, 1274 (2023) (discussing that if the original work is used by the second user to achieve a similar purpose as the original, a finding that the new work serves as a substitute is more likely).
[31] *Authors Guild*, 804 F.3d at 209.
[32] *Id.* at 226.
[33] *See* Question 10.

- **8.1. In light of the Supreme Court's recent decisions in *Google v. Oracle America*[34] and *Andy Warhol Foundation v. Goldsmith*,[35] 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?**

In *Goldsmith*, the Court clarified that under the first factor of "purpose and character," a court should look to whether the secondary use is transformative, which involves a secondary user's justification for using the original work.[36] In doing so, the Court noted that new expression, meaning, or message may be relevant, but is not dispositive.[37] This analysis will be particularly relevant in evaluating artistic uses of AI systems. In *Google v. Oracle*, the Court noted that software is different and the prior justifications for fair use in *Campbell*, which included commenting on the original copyrighted work or otherwise targeting the original work, did not as easily apply to a fair use analysis in software because of its functional use.[38]

In analyzing copyrighted works in training, AI models will have artistic uses and functional uses. While the use of copyrighted works in training is likely transformative, a fair use analysis will turn on the use of the AI models and systems. A finding of fair use will depend on the transformativeness of the output's use compared to the original's use.

The use to be analyzed in training AI models is the *input* of copyrighted works, *training* models, *and* the end use of the model. In *Authors Guild*, the court analyzed the digitization of complete books in connection with the display of "snippets" of text as a single use.[39] This analysis can be applied to AI in the same way. The analysis would be the input of the copyrighted work *and* how the trained model is used, as in *Authors Guild* the analyzed use was the input of books *and* the text snippets displayed to a user.[40]

- **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 analysis should apply similarly to entities that collect and distribute data for training. Here, the relevant use to be analyzed is the *collecting* of copyrighted materials *for the purpose* of training. A first factor analysis would mirror the analysis in AI model training. The scraped data is used for a commercial purpose, and it may be similarly transformative, but the data is being packaged and sold to an AI model or system.

In *Goldsmith*, the court cited *Campbell* and explained that neither a finding of commerciality nor whether the secondary output adds something new are dispositive factors in the analysis.[41] A transformative analysis under the first factor is a matter of degree, and the court

[34] Google LLC v. Oracle Am., Inc., 141 S. Ct. 1183 (2021).
[35] Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith, 42 143 S. Ct. 1258 (2023).
[36] *Goldsmith*, 143 S. Ct. at 1276–77.
[37] *Id.* at 82–83.
[38] *Google LLC*, 141 S. Ct. at 1202–05, 1208.
[39] *Authors Guild*, 804 F.3d at 216–17.
[40] *Id.*
[41] A*ndy Warhol Found.*, 143 S. Ct. at 1273, 1275; Campbell v. Acuff-Rose Music, Inc., 510 U.S. 569, 579 (1994).

would weigh the transformativeness of packaging copyrighted material for sale to AI models against the commercial purpose of for-profit data sales.

- **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.[42] How should the fair use analysis apply if AI models or datasets are later adapted for use of a commercial nature?[43] Does it make a difference if funding for these noncommercial or research uses is provided by for-profit developers of AI systems?**

As discussed, a fair use analysis in training will involve consideration of the model or dataset's purpose. If the purpose has changed from a nonprofit use to a commercial use, or vice versa, the fair use analysis would also change. Fair use is an affirmative defense applied to specific *uses*, and it does not consider the funding source for the use. It only considers the justification for the use. Therefore, if the purpose and character of the use have changed, a court would consider the change.

- **8.4. Does the volume of material used to train an AI model affect the fair use analysis? If so, how?**

The volume of training material would not affect a fair use analysis in assessing an AI model's use of copyrighted material to train. A fair use defense does not consider the number of works used, but under 17 U.S.C. § 107(3), the amount of a single work that is used is a factor courts should weigh. This would be particularly relevant in analyzing a model's user usage and outputs.

- **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?[44] 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?**

Courts note that the fourth factor considers "actual or potential market substitution."[45] Under a fourth factor analysis, a court should look to all relevant markets to determine harm. This is a case-by-case determination that cannot be answered in a single way. An increasingly expansive number of derivative markets for artistic expression and data precludes a bright line rule in the fourth factor. However, fair use is an affirmative defense asserted upon a finding of

---

[42] For example, the generative AI model, Stable Diffusion, was reportedly developed in part by researchers at the Ludwig Maximilian University of Munich but is used by the for-profit company Stability AI. *See* Kenrick Cai, *Startup Behind AI Image Generator Stable Diffusion Is in Talks to Raise at a Valuation Up to $1 Billion*, FORBES (Sept. 7, 2022), https://www.forbes.com/sites/kenrickcai/2022/09/07/stability-ai-funding-round-1-billion-valuation-stable-diffusion-text-to-image/?sh=31e11f5a24d6.

[43] 17 U.S.C. 107(1).

[44] *Id.* at 107(4).

[45] *Goldsmith*, 143 S. Ct. at n.12.

substantial similarity. Therefore, if there is substantial similarity, we would argue that any demonstrated market harm is relevant to a fair use analysis with AI training.

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?**

A direct voluntary licensing regime would be permissible in all or some creative sectors, however it is our view that this structure is insufficiently empowered to be the primary vehicle for licensing creative works. This is due to the nature of individual licensing agreements where both parties bear an administrative burden of negotiating the agreement and then monitoring compliance with the agreement. Inequitable bargaining is also commonplace in voluntary licensing regimes, where one party has (i) access to better alternatives (ii) more market share or (iii) more knowledge power.[46] There is also inherent inflexibility in direct voluntary licensing that is likely to exacerbate transactional cost with AI content. Any purpose not imagined or contemplated by the contractual language after the contract has been executed would require an addendum to the agreement before a new expressive use of the work is permissible. Generative AI is suitable to both multi-modal input and output & because of the scale that is enabled by this technology, it is foreseeable that licensees use will extend farther and wider than the format they've licensed the underlying content in.

- **10.3. Should Congress consider establishing a compulsory licensing regime?[47] 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?**

Yes, Congress should consider establishing a compulsory licensing regime that distributes mechanical licenses for AI-generated works. BLIP envisions the creation of a new class of licensing for AI-generated works, grounded in equitable financial incentives for all involved parties. The license would be based on the fundamental principle that while AI-generated works deserve some level of protection through copyright law, it is a lesser license that is not equivalent to works derived solely by human authorship. An additional fundamental principle is systemic flexibility that can adapt and grow along with changing circumstances in technology rather than establishing a first moving one-size fits all. Because the regime's policy

[46] Nancy S. Kim, *Bargaining Power and Background Law*, 12 VANDERBILT J. OF ENT. & TECH. L. 93, 95 (2009) (explaining that "'knowledge power' refers to the advantage that a superior understanding of the background law confers upon a contracting party").

[47] A compulsory or "statutory" license allows for certain uses of a copyrighted work "without the consent of the copyright owner provided that the person adhered to the provisions of the license, most notably paying a statutorily established royalty to the copyright owner." *Music Licensing Reform: Hearing Before the Subcomm. on Intell. Prop. of the S. Comm. on the Judiciary*, 109th Cong. (2005) (statement of Marybeth Peters, Register of Copyrights), http://copyright.gov/docs/regstat071205.html.

objective is to broaden access to copyrighting works containing or based on generative AI, there should not be a fixed statutory rate for licensing. The regime should be grounded in statutorily authorized penalties for violating its requirements and for willful non-payment of royalties.

A mechanical royalty licensing system, akin to the Mechanical Licensing Collective ("MLC") used in the Music Modernization Act ("MMA"), can achieve an innovative balance for many participants in the ecosystem. This kind of licensing system will continue to incentivize artists to use new technology in creating art, as they can retain their copyright. The system will also incentivize technical inventors to create Generative AI by offering them some financial interest in the use of their invention that does not interfere with an author's copyright. A mechanical royalty system will allow the author to retain sole authorship rights while also recognizing the contribution made by AI where such creative possibilities may not have existed previously. AI inventors can earn a fixed-rate royalty through a work for-hire arrangement when their technology is used. While sweat-of-the-brow doctrine has previously been rejected as a basis to earn a copyright, there is novel transformative use through generative AI. Our opinion is that works accomplished through generative AI deriving from an earlier in time licensed work used substantially in its creation justifies some financial interest for copyright holders in a derivative final work.

- **10.4. Is an extended collective licensing scheme[48] a feasible or desirable approach?**

Yes, an extended collective licensing scheme is socially desirable to protect the interests of stakeholders besides the immediate rights holders and derivative works creators. One of the most important classes of stakeholders is human users who have contributed underlying content to the training of generative AI models. As mentioned in earlier comments, this class of users provided valuable datasets and, in many instances, licensed works often unknowingly. These users can be paid out a portion of the fixed-rate royalty that creators of AI models or systems of models earn. The royalty would be managed by a third-party collective that has to vet and verify that its members can prove that they have a right to this royalty, i.e., that their content on the internet has been scraped and used for training AI. Royalties could be earned proportional to how often the underlying licensed or unlicensed work has been found in training datasets.

[48] "An Extended Collective Licensing scheme is one where a relevant licensing body, subject to certain safeguards, is authorised to license specified copyright works on behalf of all rights holders in its sector (including non-members), and not just members who have given specific permission for it to act." *Extended Collective Licensing (ECL) Scheme Definition*, LEXISNEXIS, (last visited Oct. 13, 2023), https://www.lexisnexis.co.uk/legal/glossary/extended-collective-licensing-ecl-scheme; *see also* Letter from Karyn A. Temple, Acting Register of Copyrights, U.S. Copyright Office, to Rep. Robert Goodlatte, Chair, and Rep. John Conyers, Ranking Member, H. Comm. on the Judiciary (Sept. 29, 2017), https://www.copyright.gov/policy/massdigitization/house-letter.pdf; Letter from Karyn A. Temple, Acting Register of Copyrights, U.S. Copyright Office, to Sen. Charles Grassley, Chair, and Sen. Dianne Feinstein, Ranking Member, S. Comm. on the Judiciary (Sept. 29, 2017), https://www.copyright.gov/policy/massdigitization/senate-letter.pdf.

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.

Yes, it is possible to identify the degree to which a particular work contributes to a particular output from a generative AI system. Maintaining an organized and labeled training dataset allows generative AI to attribute its output to its sources in the most efficient manner. One approach to identifying inputs and sources and the degree of their contribution to an output is by differentiable attribution. This approach has been used in natural language processing tasks such as machine translation, where the AI-model can determine which parts of the input text were most important in generating the output translation.

It must be noted however that the above example is a much more limited application of AI than a generative AI-model or system of models that produces new content based on the input data, training algorithms, and any additional parameters. Our current understanding is that while attribution is possible for generative-AI, it is a complex task that may produce unreliable results when combined with diffusion models. This is largely due to the algorithmic additions of noise to the training dataset, and subsequent random re-learning and removal processes of the noise when generating output. The well-documented issue of machine hallucinations can also add difficulty and confusion when assessing degrees of contribution from specific inputs to identifying outputs.

Transparency & Recordkeeping

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?

Under existing U.S. law, there are indeed statutes that could necessitate developers of AI models to maintain and reveal records pertaining to their training materials, albeit with varying focuses and applications. While the below laws do not explicitly address AI models, they each carry implications for entities that use AI systems.

1. Securities and Exchange Act (SEA) Rule 17a-4.

SEA Rule 17a-4, issued by the Securities and Exchange Commission, is a regulation which outlines requirements for records and data retention, indexing, and accessibility for financial institutions and broker-dealers.[49] The rule imposes strict preservation of specific records related to matters such as business operations and communications. Rule 17a-4 also requires that records of numerous types of transactions must be retained and indexed on inexpungeable media with immediate accessibility for two years, and with non-immediate access for at least six years.[50]

[49] 17 C.F.R. § 240.17a-4 (2023).
[50] 17 CFR §§ 240.17a-4(b)–(e) (2023).

In the event AI models are used in financial decision-making processes, such as trading algorithms or risk assessment systems, the records related to these AI models may fall under the purview of Rule 17a-4. This would include records detailing the development of AI algorithms, the sources of data used to train these model sets, the methodologies used for training and data gathering, and the various changes or updates the models may have undergone. Thus, under this Rule, developers and financial institutions would have to retain these records and have them readily available for audit by the SEC to ensure adequate compliance; this would also ensure transparency and integrity within financial markets.

2. Sarbanes–Oxley Act of 2002 (SOX).

Another piece of legislation that could require AI developers to disclose records pertaining to materials used for training is the Sarbanes–Oxley Act of 2002. SOX is a federally enacted statute that came about in response to corporate and accounting scandals in the early aughts.[51] It sets forth provisions to enhance corporate governance and financial reporting, focusing mainly on publicly traded companies.[52]

The sections of the Act cover responsibilities of a public company's board of directors, enhanced financial disclosures, and created the "Public Company Accounting Oversight Board" (PCAOB).[53] Section 302 of SOX, for instance, requires corporate officers to certify the accuracy and completeness of their company's financial statements.[54] Section 404 mandates the assessment and reporting on the effectiveness of a company's internal controls.[55] And lastly, Section 406 calls for the adoption of a code of ethics for senior corporate officers.

With consideration to AI systems in corporate governance and financial reporting, SOX would likely be the controlling federal statute. Auditors will look for compliance with Section 302 to ensure that financial statements created through AI systems are accurate. This is important in the event algorithms are tainted by "dirty data," thus contaminating the data set and decision-making processes.[56] Additionally, compliance under Section 404 will require ensuring that internal controls reports are properly assessed, including relevant information regarding internal controls related to AI systems used in financial processes. This Section will narrow in on companies that use AI systems in processes such as financial forecasting, auditing, or reporting. The internal controls report will require AI systems to be properly tested and to evaluate the effectiveness of their AI systems to ensure they are operating as intended as well as mitigating risks to financial reporting. Failing to do so will likely lead to an investigation and eventual order by the SEC. And lastly, Section 406 emphasizes the importance of ethical conduct and

[51] Sarbanes–Oxley Act of 2002, Pub. L. No. 107–204, 116 Stat. 745.
[52] *Id.*
[53] *Id.*
[54] 15 U.S.C. § 7241(a).
[55] 15 U.S.C. § 7262.
[56] Kathy Baxter, *Dirty Data or Biased Data?*, MEDIUM (Oct. 15, 2018), https://medium.com/salesforce-ux/dirty-data-or-biased-data-6d55db6b5dc6.

accountability for senior corporate officers. Compliance will mandate officers to consider whether their organization's AI usage aligns with ethical standards. Transparency is also a key aspect of ethical AI, especially with the disclosure of information such as data sources, algorithms, infrastructures, developers, and methodologies; moreover, transparency plays an important role in corporate accountability and integrity, especially when under scrutiny of SEC auditors and the general public.

3. Health Insurance Portability and Accountability Act of 1996 (HIPAA).

The Health Insurance Portability and Accountability Act[57] was passed by Congress in 1996. The uniform privacy standards within HIPAA were enacted with the goal of protecting the health information and medical records of individuals who receive any form of health care in the United States.[58] This protection includes allowing individuals to control how and where their information is used and limiting third-party knowledge.

Section 164.306 of HIPAA outlines the required security standards when protected health information is involved.[59] The listed requirements include the need for entities to protect against threats or disclosures which are not permitted, as well as the need to ensure confidentiality of health information which is created, received, maintained, or transmitted.[60]

With the emergence and growth of AI, it is critical to address its potential role in the health care field. If AI is implemented in health care recordation systems, there could be major risks involved. If the information is not protected against potential leaks, improper disclosures and breaches of confidentiality could lead to violations of HIPAA and resulting penalties.[61] Even if AI is not implemented in health care recordation, there is a chance of accidental disclosure if the information is not fully protected. Thus, under HIPAA, the use of AI could pose great risks if information is not protected adequately.

4. Food and Drug Administration (FDA).

While not a statute, this agency is relevant to our discussion on disclosure of information and materials used for AI systems. The Food and Drug Administration (FDA) is the federal agency responsible for protecting public health and well-being through the regulation and testing

[57] Health Insurance Portability and Accountability Act of 1996, Pub. L. No. 104-191 (1996).
[58] U.S. DEPT. OF HEALTH AND HUM. SERV., WHAT DOES THE HIPAA PRIVACY RULE DO? (2023).
[59] § 164.306.
[60] *Id.*
[61] Mason Marks & Claudia E. Haupt, *AI Chatbots, Health Privacy, and Challenges to HIPAA Compliance*, 330 JAMA 295, 309-310 (2023).

of the following non-exhaustive list: pharmaceuticals, cosmetics, medical devices, and food products.[62] FDA regulations are treated as binding federal law.[63]

Title 21 of the Code of Federal Regulations governs food and drugs within the United States for regulatory agencies such as the FDA, the Drug Enforcement Agency (DEA), and the Office of National Drug Control Policy (ONDCP).[64] Sections 21.70 and 21.71 discuss the disclosure of records in Privacy Act Records Systems.[65] Records concerning individuals in the Privacy Act Records System may not be disclosed except in instances of consent or a valid government interest such as carrying out a census or when the information is required for the performance of certain agency duties.[66] If information is requested and/or obtained under misleading or false circumstances, there will be violations under Title 5 of the United States, and penalties may follow.[67]

While it is not likely in the present time that AI will be used to maintain recordation, there are other ways it may play a role. Improper disclosure of records, such as medical device, pharmaceutical, or food information under false pretenses could lead to leaks across the public domain and potentially AI systems. It is thus crucial for agencies to enforce regulations to their fullest extent to ensure compliance and protection of consumer information and recordation. It may then be conceivable that FDA regulations could prevent AI from being implemented to hold or maintain any government information.

In recent news. the FDA has shed some insight on their view on AI. The possible risks of using AI, including biases and data protection risks, are acknowledged, however the agency seems to hold an overall favorable view.[68] The regulatory uses already exist across therapeutic areas, and it is likely that AI and machine learning will continue to be used in medical devices, as well as in the development of pharmaceuticals. [69]

5. Children's Online Privacy Protection Act (COPPA).

The Children's Online Privacy Protection Act was passed by Congress in 1998 to implement safeguards for children[70] from data collection on the internet.[71] It applies to

---

[62] *Food and Drug Administration (FDA)*, USAGOV, https://www.usa.gov/agencies/food-and-drug-administration (last visited Oct. 21, 2023).
[63] Food and Drug Admin., *FDA Laws, Regulations, and Guidance Documents*, https://www.fda.gov/media/133830/download#:~:text=Congress%20enacts%20statutes%2C%20and%20FDA,issuing%20more%20specific%20Agency%20regulations (last visited Oct. 21, 2023).
[64] 21 C.F.R. §§ 1–1402 (2023).
[65] 21 C.F.R. §§ 21.70-71 (2023).
[66] *Id.*
[67] *Id.* at § 21.71.
[68] Patrizia Cavazzoni, FDA Releases Two Discussion Papers to Spur Conversation about Artificial Intelligence and Machine Learning in Drug Development and Manufacturing, U.S. FOOD & DRUG ADMIN. (May 10, 2023), https://www.fda.gov/news-events/fda-voices/fda-releases-two-discussion-papers-spur-conversation-about-artificial-intelligence-and-machine.
[69] *Id.*
[70] A child is anyone that is "under the age of 13." 15 U.S.C. § 6501(1).
[71] 15 U.S.C. §§ 6501-6508.

operators[72] of websites that are either directed to children or have actual knowledge of collecting or maintaining personal information from children.[73] Under COPPA, operators are obligated to provide clear and conspicuous notice of data collection and are banned from collecting children's personal information, unless they obtain verifiable parental consent.

Obtaining verifiable consent means that the operator must make "any reasonable effort (taking into consideration available technology) to ensure that before personal information is collected from a child, a parent of the child: (1) [r]eceives notice of the operator's personal information collection, use, and disclosure practices; and (2) [a]uthorized any collection, use, and/or disclosure of the personal information."[74]

Furthermore, COPPA requires operators to provide public notice of their information practices on their web sites or online services.[75] One of the notice requirements involves operators providing,"[a] description of what information the operator collects from children,…;how the operator uses such information; and, the operator's disclosure practices for such information."[76] If a child provided their personal information to the operator, an operator is required to provide, at the parent's request, "a description of the specific types of personal information collected from the child by that operator."[77]

Additionally, a parent has a right to review personal information provided by their child to a web site or online service, so [a]n operator is obligated to provide a parent "[a] description of the specific types or categories of personal information collected from children by the operator, such as name, address, telephone number, email address, hobbies, and extracurricular activities."[78]

Regarding artificial intelligence, developers of AI models or systems that train their models on a child's personal information could be considered operators or agents of operators as defined under COPPA. As a result, developers of these AI models or systems using a child's personal information for training would be subject to the notice and disclosure requirements under COPPA mentioned above. Since COPPA was enacted to prevent unfair or deceptive practices related to the collection of data from children on the internet, it is highly likely that the collection of a child's personal information to train an AI model or system would be something COPPA would seek to have safeguards around.[79]

---

[72] An operator refers to "any person who operates a website located on the Internet or an online service and who collects or maintains personal information from or about the users of or visitors to such website or online service, or on whose behalf such information is collected or maintained, where such website or online service is operated for commercial purposes, including any person offering products or services for sale through that website or online service…" 15 U.S.C. § 6501(2).
[73] 16 C.F.R. § 312.2.
[74] 16 C.F.R. § 312.2.
[75] 16 C.F.R. §§312.4(d)(1)-(3).
[76] 16 C.F.R. § 312.4(d)(2).
[77] 15 U.S.C. § 6502(b)(1)(B)(i).
[78] 16 C.F.R. § 312.6(a)(1).
[79] 16 C.F.R. § 312.1.

6. California Consumer Privacy Act (CCPA) (as amended by the California Privacy Rights Act (CPRA)).

The California Privacy Rights Act amended the California Consumer Privacy Act and became fully effective on January 1, 2023 (collectively referred to herein as "the Act").[80] It is a state law that establishes privacy rights for consumers, who are California residents, like Europe's General Data Protection Regulation (GDPR).[81] The Act applies to for-profit businesses that collect personal information and meet a certain threshold.[82]

Under the Act, businesses that collect personal information are obligated to provide clear and conspicuous notice and disclosure of their data collection to consumers in a manner that is "reasonably accessible."[83] Additionally, numerous privacy rights are granted to consumers such as the "Right to Know What Personal Information is Being Collected."[84] As a result, consumers can request businesses that collect their personal information to disclose the following: "(1) The categories of personal information it has collected about that consumer. (2) The categories of sources from which the personal information is collected. (3) The business or commercial purpose for collecting, selling, or sharing personal information. (4) The categories of third parties to whom the business discloses personal information. (5) The specific pieces of personal information it has collected about that consumer."[85]

As a result, businesses that collect consumer personal information to develop AI models or systems could be subject to retaining or disclosing records about the materials used for training to comply with the Act's notice requirements. Compliance could involve mentioning such training materials in their privacy policies or responding to consumer privacy rights inquiries on what personal information is being collected, used, or shared.

Generative AI Outputs

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?

---

[80] CAL. CIV. CODE §§ 1790.100-1798.199.100

[81] "The GDPR sets out detailed requirements for companies and organizations on collecting, storing and managing personal data." *Data Protection Under GDPR*, EUROPA.EU, https://europa.eu/youreurope/business/dealing-with-customers/data-protection/data-protection-gdpr/index_en.htm (last visited Oct. 21, 2023).

[82] Business subject to the obligations under the CPRA must satisfy at least one of the following: annual gross revenue that exceeds twenty-five million dollars, collect personal information of at least one hundred thousand consumers, or generate more than half of their annual revenue from selling or sharing consumers' personal information. CAL. CIV. CODE § 1798.140(d).

[83] CAL. CIV. CODE § 1798.130.

[84] CAL. CIV. CODE §§ 1798.110; 1798.130.

[85] CAL. CIV. CODE §§ 1798.110(a)(1)-(5).

Yes, there are circumstances when a human using a generative AI system should be considered the "author" of material produced by the system. Before delving into what circumstances these may be, it is crucial to recognize that there exist many different types of generative AI systems. Some, like ChatGPT or Midjourney, can be classified as assistive tools that require human users to generate outputs. Put simply, without a prompt introduced by the user, these systems cannot autonomously generate outputs. Other, more advanced generative AI systems, such as Dr. Thaler's Device for the Autonomous Bootstrapping of Unified Sentience (DABUS), can autonomously generate output without the need for prompts or human intervention. In discussing when a human using a generative AI system should be considered the "author" of material produced by the system, the distinction between AI systems capable of autonomously creating and those requiring prompts is fundamental.

As far as systems like ChatGPT and Midjourney are concerned, copyright law already provides a clear answer as to when and how users should be considered authors of their outputs. The Copyright Office and the courts have long recognized that the author of a creative work is "he [or she] to whom anything owes its origin; originator; maker; one who completes a work of science or literature."[86] This definition does not change when a photographer relies on a tool, such as a camera of Photoshop, to assist in the creation of a photograph. Regardless of what tools are used to obtain the final product, photographers are considered authors because the USCRO recognizes that independent creativity exists in a photographer's artistic choices, like the angle and position of subject(s) in the photograph, lighting, and timing.

The authorship analysis should not change when we replace a tool with a more advanced one. Just as a photographer using a camera and Photoshop may have to show enough independent creativity in their final products, users of generative AI systems should be allowed to do the same. Recently, the USCRO attempted to create guidelines for applicants trying to register works created with the assistance of AI in an effort to clarify when these works could receive copyright. The guidelines state that simply providing a general prompt to a generative AI system ("write a poem about copyright law in the style of William Shakespeare") does not amount to showing enough human creativity for the output to qualify for copyright protection.[87] We agree. At the same time, however, the guidelines recognize that works created with the assistance of AI or containing AI-generated material may show sufficient human authorship to support a copyright claim.[88] This happens when a human "select[s] or arrange[s] AI-generated material in a sufficiently creative way that 'the resulting work as a whole constitutes an original work of authorship."[89]

Despite the clear language of the law and the somewhat clear guidelines recently released, the USCRO has nevertheless continued to force applicants to dissect their creative works to separate what is AI-generated and what is human-made. Any use of generative AI systems that is more invasive than Photoshop or Grammarly has been deemed more than *de*

[86] Burrow-Giles Lithographic Co. v. Sarony, 111 U.S. 53, 56 (1884).
[87] Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence, 88 Fed. Reg. 16190–92 (Mar. 16, 2023) (to be codified at 37 C.F.R. pt. 202).
[88] *Id.*
[89] *Id.*

minimis, and no amount of prompts or specificity of commands has been found satisfactory.[90] In fact, even if an applicant could succeed in demonstrating that they had materially changed and rearranged the AI output, the Office would still grant copyright only to "the human-authored aspects of the work," thus contradicting its own directive that the "resulting work *as a whole* [would] constitute an original work of authorship."[91]

This type of analysis, however, is inconsistent with practices that the Office has adopted for decades. In registering copyright to photographs, the Office does not ask photographers to dissect their works and remove all elements the photographer did not directly control or that are generated by tools. If a photographer uses Photoshop to automatically remove imperfections or adjust the lighting in a photograph, the photograph can still be copyrighted *as a whole* with the photographer listed as its author. Photographers are not forced to identify and remove all that was done by a machine in order to receive copyright in what's left. Creative work created with the assistance of AI should receive the same treatment.

Even assuming that AI-generated content cannot be copyrighted *per se* (meaning, outside of the boundaries of whatever creative work it is incorporated in), it is still inconsistent with existing copyright law to hold the work *as a whole* uncopyrightable. Photographs often contain elements that are uncopyrightable alongside ones that are the result of the photographer's creativity.[92] Yet, when reviewing copyright registrations for photographs, the Office does not require applicants to explicitly exclude those elements. It simply grants copyright if the photograph *as a whole* meets the requirements set out by the Copyright Act. The dissection of creative works, if any is needed, is done by the courts in the event of an infringement dispute. In those instances, if an alleged infringer only reproduced elements that are uncopyrightable, no infringement was committed. The overall work, however, remains nevertheless protected by copyright.

Other areas of the Copyright Law also support the view that creative works *as a whole* can receive copyright protection even if they contain AI-generated material. It has long been established that facts cannot be copyrighted, but compilations of facts can.[93] Compilation authors usually decide "which facts to include, in what order to place them, and how to arrange the collected data" in a way that "entail[s] a minimal degree of creativity."[94] In such cases, "even a directory that contains absolutely no protectible written expression . . . meets the constitutional minimum for copyright protection if it features an original selection or arrangement."[95] Assuming that AI-generated content is uncopyrightable, creative works containing AI-generated elements are thus no different than compilations containing facts. As such, they should be

---

[90] *See, e.g.*, Letter from Copyright Review Board to Tamara Pester, (Sept. 5, 2023), https://www.copyright.gov/rulings-filings/review-board/docs/Theatre-Dopera-Spatial.pdf

[91] Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence, 88 Fed. Reg. 16190, 16193–93 (Mar. 16, 2023) (to be codified at 37 C.F.R. pt. 202) (emphasis added).

[92] *See, e.g.*, Mannion v. Coors Brewing Co., 377 F. Supp. 2d 444, 454–55 (S.D.N.Y. 2005) (explaining that, while the "unusual angle and distinctive lighting" as well as the composition of a photograph were copyrightable elements, the subject's "face, torso, and hands" and the cloudy sky behind him were not).

[93] Feist Publications, Inc. v. Rural Tel. Serv. Co., Inc., 499 U.S. 340, 344 (1991); 17 U.S.C. § 103.

[94] *Id.* at 348.

[95] *Id.*

copyrightable as long as their authors can show that they modified, arranged, and selected the elements to include in a way that "entail[s] a *minimal degree of creativity*."[96] To ensure that copyright applicants are aware of what constitutes a "minimal degree of creativity" in relation to AI-generated elements, the Office could implement a threshold requirement, which we introduced in our response to Question 19 and 20.1.

The authorship analysis for the output of autonomous generative AI systems is different. In cases akin to that of Dr. Thaler and DABUS, the argument is not that a human user exercised enough creativity in relation to the AI output that he or she should be considered its author. Rather, the argument is that, despite the lack of human inputs and prompts, AI-generated works should still receive copyright protection. Whether this would be possible within the current copyright framework has already been discussed at length by Dr. Ryan Abbott.[97] Thus, we will address potential solutions for how the USCRO should protect these works in the event it finds Dr. Abbott's arguments unpersuasive in our answer to Question 20.1.

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, revisions to the Copyright Act are necessary to clarify the human authorship requirement in light of AI-generated content, and the Copyright Office should provide additional standards to determine when content that includes AI-generated material is subject to copyright protection. Simply revising the existing language of the Copyright Act without concerted action by the Office would not suffice to encompass all of the current uses of AI as a tool to create creative works. Thus, the required revisions and guidelines should take into account the distinction between autonomous AI systems and assistive AI systems delineated in our answer to Question 18.

To account for generative AI systems that can autonomously generate output without the need for prompts or human intervention, the Copyright Act should be revised to include a separate subsection. The new subsection should specifically address works autonomously generated by AI and outline requirements for AI developers and users. This suggestion is further described in our answer to Question 20.1.

As far as works created by human artists with the assistance of AI, we believe they already fall – or should fall – within the current authorship framework. Granting protection for AI-assisted works fits into existing case law that developed the human authorship requirement, as courts have held that authors seeking copyright protection need to show a human did enough with the work to qualify for registration.[98] Thus, rather than revising the Copyright Act, the

[96] *Id.* at 348 (emphasis added).
[97] *See generally* Ryan Abbott, *Artificial Intelligence and Intellectual Property: An Introduction*, *in* RESEARCH HANDBOOK ON INTELLECTUAL PROPERTY AND ARTIFICIAL INTELLIGENCE 1 (Ryan Abbott ed., 2022).
[98] *See* Thaler v. Perlmutter, No. CV 22-1564 (BAH), 2023 WL 5333236, at *3 (D.D.C. Aug. 18, 2023) ("Human involvement in, and ultimate creative control over, the work at issue was key to the conclusion that the new type of

Copyright Office should simply provide clearer guidelines and standards to explain when a human author has done enough work to qualify as the author of a work that includes AI-generated elements. In its most recent guidelines, the USCRO clarified that simply providing general prompts to a generative AI system does not amount to showing enough human creativity for the output to qualify for copyright protection.[99] AI users, however, may provide very detailed and extensive prompts to an AI-system to ensure that its output is as close as possible to what they anticipated. Once the AI generates output, authors may extensively alter it and rearrange it until the final product matches exactly what the author envisioned. This use of AI is starkly different from the one deemed uncopyrightable by the USSCO. As further described in Question 20.1, if a creator demonstrates that they provided sufficient input and prompts to control the output of an AI system, their work should be copyrightable. In its guidelines, the Office should clarify what kind and what amount of human input is enough to warrant copyright protection.

Additionally, the Copyright Office should alter its application process for AI-assisted or AI-autonomously-generated works. For AI-assisted works, a showing of sufficient human input should be required. The application should include a portion that requires authors to detail their creation process. This could be done by providing the Office with the entirety of the prompts submitted to the AI and a comparison between the AI-generated output and the final product submitted for copyright. For works autonomously generated by an AI system without human prompts, the application should require AI developers to describe and provide evidence of how the AI was trained and the process through which it autonomously generated its output. This framework would ensure honesty and transparency from AI users and developers and would provide enough information for the Office to determine the work's copyrightability.

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?

Legal protection for AI-generated material is desirable as a policy matter on several grounds. Firstly, granting legal protection to AI-generated materials can incentivize creators and organizations to invest in developing AI technologies. Copyright protection would enable companies to monetize their AI-generated content, which could be licensed or sold, thus creating revenue streams. This potential for financial gain would encourage investment in AI technology research and development and foster a competitive environment where companies are encouraged to create better and more innovative AI algorithms and applications to attract talent and users. Copyright protection would also make AI companies more attractive to investors.

---

work fell within the bounds of copyright."); Burrow-Giles, 111 U.S. at 58 (1884) (in which the Supreme Court held that photographs were copyrightable creations of authors, even if they are produced by a device, as it represents "original intellectual conceptions of the author"); Urantia Foundation v. Maaherra, 114 F.3d 955, 959 (9th Cir. 1997) (stating that "notwithstanding the . . . [book's] claimed non-human origin, the Papers in the form in which they were originally organized and compiled by the . . . [authors] were at least partially the product of human creativity. . . .Therefore, the Papers were works amenable to common law copyright protection.").

[99] Copyright Registration Guidance, *supra* note 91.

Investors are more likely to support companies that have legal protections for their intellectual property, as it provides a safeguard against competitors copying or replicating their works.

Secondly, granting legal protection to AI-generated materials discourages and actively prevents fraud in the copyright system. The Copyright Act clearly states that "knowingly mak[ing] a false representation of a material fact in the application for copyright registration" is a punishable offense.[100] However, while the possibility of being fined might be enough to stop some from attempting to register works generated by AI without disclosing such origin, it is likely not enough of a deterrent. Faced with the potentially high reward of obtaining copyright protection for their work and the relatively low probability of getting caught,[101] it is likely that applicants have already attempted to register AI-generated creative works as human-made and will continue to do so. By refusing to grant legal protection for AI-generated works, the current system ultimately favors applicants willing to lie about the origin of their works and punishes those who disclose AI's involvement in their creative process.

Thirdly, granting legal protection to AI-generated materials would remove a barrier to entry into the world of creative art for creators with mental and cognitive disabilities who rely on AI as an assistive tool in their creative process. Individuals suffering from certain mental or physical health conditions, such as Bipolar Disorder, Borderline Personality Disorder, Down Syndrome, and many more, often face significant cognitive, emotional, and mental challenges. For people facing these medical conditions, AI technology, particularly generative AI systems, can serve as a transformative tool. These systems have the capacity to assist individuals by providing a unique channel for creative expression. Instead of replacing human creativity, these AI systems act as valuable aids. They act as bridges between the users' creative visions and their physical or cognitive limitations.

In this context, generative AI systems can provide various types of assistance depending on the individual's specific needs. For example, text-to-speech applications can help those with speech impediments or communication difficulties to articulate their thoughts. Visual art generative algorithms can aid individuals with limited motor skills in creating digital artworks. In the realm of writing, AI systems like ChatGPT can assist individuals with cognitive disabilities in forming coherent sentences and expressing complex ideas. By facilitating and assisting in the process of creation, AI empowers people with disabilities and helps them express their unique perspectives, emotions, and imaginative concepts. This not only promotes their creative expression but also fosters a sense of independence and self-esteem, enabling them to participate more fully in society and the creative community.

- **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?**

[100] 17 U.S.C. § 506(e).
[101] Many studies have revealed that even tools intended to detect AI in texts and images often fail to do so. *See, e.g.*, DEBORA WEBER-WULFF, ET AL., TESTING OF DETECTION TOOLS FOR AI-GENERATED TEXT 28 (2023).

We believe that protection is desirable in the form of copyright. The Copyright Act should be amended to include a new section that provides protection for AI-generated material. This solution would give Congress and the Copyright Office enough flexibility to set specific limits and rules for creative works generated with or by AI without having to create a completely new *sui generis* right and legal framework. At the same time, a separate section would also be easier to amend and update as AI technologies continue to develop over time.

This new section should be further divided into two subsections to distinguish works created with the assistance of AI and works autonomously created by AI. The subsection on works created with the assistance of AI should clarify when these works can receive copyright protection. Specifically, when creators merely use AI as a tool to assist in their creative expression, standard copyright protection should be allowed if a certain threshold level of human input is met. For example, if an author gives general prompts to an AI system but then materially edits the AI output and adds her own creative contribution, she should be allowed to receive copyright protection for her final work. Similarly, if an author provides extensive and meticulous prompts to an AI system and then rearranges the outputs to combine them with his own creative expressions, he should be allowed to receive copyright protection. As we stated in our answer to Question 19, creators can demonstrate their human input by submitting evidence of the exact prompts they gave to an AI system and comparing the AI output and the work they are trying to protect. Recognition of AI-assisted creative works would level the copyright playing field for creators who cannot express their ideas without the help of an AI system. Additionally, it would recognize the reality that many creators are already using AI in some capacity to produce or supplement their work. Since more than a *de minimis* level of human creativity is involved in their creation, AI-assisted works should receive the same rights and protections as works created without AI assistance.

The subsection on works autonomously created by AI should instead mirror Section 178 of the United Kingdom's Copyright, Designs and Patents Act 1988. Section 178, which was implemented in the 1980s, grants copyright protection to computer-generated works (CGWs), which are defined as works "generated by computer in circumstances such that there is no human author of the work."[102] The Act further specifies that the legal author of CGWs is the person who made the "arrangements necessary for the creation of the work."[103]

AI systems capable of autonomously generating creative outputs without human-provided prompts are trained by human developers on carefully selected materials and data. In these circumstances, while the actual author of the creative output would be the AI, the legal author of its copyright should be the person who trained the AI or who selected the specific material and data it was trained on – in other words, the person who made the "arrangements necessary for the creation of the work."[104] As stated in our answer to Question 19, to receive copyright protection for works autonomously generated by AI, applicants should disclose how

[102] Copyright, Designs and Patents Act 1988, c. 48, § 178 (UK).
[103] *Id.* § 9 (UK).
[104] *Id.*

the system was trained and which data it had access to. To ensure transparency, applicants should also explain the process through which the AI generated the output in question. Just like CGWs, works autonomously created by AI should receive copyright of shorter duration compared to works created by human authors,[105] but they should receive some degree of protection, nevertheless. Protecting these works not only recognizes a new art form that has existed in the creative world for years and will continue to persist as systems further develop, but it also recognizes the high level of labor and creativity from the humans who train these autonomous systems, allowing them to have the same type of rights as any other creator.

Infringement

25. If AI-generated material is found to infringe a copyrighted work, who should be directly or secondarily liable—the developer of a generative AI model, the developer of the system incorporating that model, end users of the system, or other parties?

If AI-generated material is found to infringe a copyrighted work, end users of the system should be liable, and the developer of the system incorporating the model should be secondarily liable as a contributory or vicarious infringer. Copyright holders can sue parties who directly infringed, as well as a party who "helped the infringer or induced the infringement". Under 17 U.S.C. § 501(a), "Anyone who violates any of the exclusive rights of the copyright owner… is an infringer of the copyright…"[106] This includes preparing a derivative work based on a copyrighted work.[107]

Determining copyright authorship entails showing "a causal chain tracing the origin of fixed expression."[108] Therefore, to prove infringement, the owner only needs to show that she is the holder of a valid copyright, or in this instance, the creator of the copyrighted work that is being fed into the generative AI model, and that the defendant end-user copied the original expression from her work.[109] As Dan L. Burk notes, authorship of the final product that infringed upon the copyrighted work should be attributed to the end-user, and not the intermediary AI programmer, because the expressive components of the infringed copyrighted work merely "[passes] through the hands of the programmer" onto the end-user's "[110] canvas." [111] Access to the copyrighted material and substantial similarity between the protected and infringing works must be shown. Thus, if access is proven through a showing that the AI system was trained on the copyrighted work and substantial similarity is shown, an end-user may be for infringement, regardless of whether he knew that the system was unauthorizedly trained on protected works.

[105] CGWs are protected for 50 years from their creation rather than for the duration of the author's life plus seventy years from his or her death. *Id.* §12.
[106] 17 U.S.C. § 501(a).
[107] *See id.* at § 106.
[108] Dan L. Burk, *Thirty-Six Views of Copyright Authorship, by Jackson Pollock*, 58 HOUS. L. REV. 263, 263 (2020).
[109] 77 Am. Jur. Trials 449 *Copyright Infringement Litigation* § 48 (2000).
[110] Burk, *supra* note 108, at 321.
[111] *Id.*

James Grimmelmann highlights another instance where the end-user could be directly liable for infringing a copyrighted work through AI-generated material.[112] In his article, Grimmelman talks about six possible legal outcomes with allocating ownership of resulting work created by AI.[113] For our purposes, the legal outcome of "unlawful derivative work" is the most relevant; it occurs when the end-user modifies the original copyright creator's work enough to contribute the end-user copyright authorship.[114] Copyright status, however, depends on (i) whether the end-user has the original creator's permission; (ii) if the end-user makes a transformative fair use of the derivative work; or (iii) some other statutory authorization exists to allow the end-use to prepare their derivative work.[115] For the question at hand, setting aside (ii) and (iii), assuming that the end-user is unable to show clear and explicit consent from the owner of the original copyrighted work which the AI-generated model infringed upon, the end-user should be found directly liable for infringement.

Under the doctrines of contributory and vicarious infringement, copyright owners can also sue parties who did not directly infringe on protected works, but contributed to, or failed to prevent, someone else's infringement.[116] Developers of the system incorporating AI models may be liable under both doctrines. Under the doctrine of contributory infringement, the developer may be liable if they had knowledge of the infringing activity and "induced, caused, or materially contributed to the other party's infringement."[117] The developer must "know or have reason to know" of the direct infringement.[118] Because the developers are incorporating the AI model, they know or should know whether or not the model is unauthorizedly trained on protected works, and thus that users may directly infringe on those works through their output. Additionally, because the developers are making a system trained on copyrighted works available to public users, they materially contributed to the other party's infringement, as they made available "machinery… that provides the means to infringe."[119] However, under this doctrine, the end-user must be found to directly infringe in order for the developer to be liable for contributory infringement.[120]

Developers may also be liable under the doctrine of vicarious infringement. Defendants may be liable if they "have the right and ability to supervise the infringing activity" and "direct financial interest in the infringing activity."[121] Developers of systems incorporating AI models have the capability to ensure that the model is not trained on copyrighted materials by

---

[112] James Grimmelmann, *There's No Such Thing as a Computer-Authored Work - And It's a Good Thing, Too*, 39 COLUM. J. L. & ARTS 403, 410 (2016).
[113] *Id.*
[114] *Id*.
[115] *Id*.
[116] William S. Strong, Copyright Infringement: Direct, Contributory, and Vicarious Infringers § 9.3 (5th ed. 2020).
[117] *Id*. § 9.4.
[118] *Id*. § 9.4.1.
[119] *Id*. § 9.4.2.
[120] *Id*. § 9.5.2.
[121] *Id*. § 9.6.

incorporating technological features that reduce the likelihood of infringing output.[122] Developers also have direct financial interest in the end-user's activity and use of the system, particularly if they are dependent on increases in their user base.[123]

At its core, copyright law is constitutionally designed to reward, not to punish, and to incentivize and encourage creative works.[124] A key element underpinning the efficacy of copyright law is proximate cause determinations. Balancing both considerations in determining liability—direct or secondary—will reveal why end users or developers of the system incorporating that model, respectively, should be held to account for copyright infringement using AI-generated material.

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, any proposed federal legislation should have a framework for requiring publishers of AI-generated material to identify and label it prominently as AI generated content. This is a critical safeguard measure that is necessary to stem the tide of the scale of misinformation enabled by generative AI. Senator Michael Bennet (D-CO) wrote to technology CEO's in June of 2023 urging swift promulgation of labeling technology for AI-generated content. The European Union is leading the charge on several fronts regarding regulation of Artificial Intelligence. Notably, the commission has updated their voluntary anti-disinformation charter to urge signatories via their technologies to make it easier for people to distinguish facts from AI-generated fiction. The EU has also passed (as of 08/25/23) the Digital Services Act which grants the Commission among other powers the ability to regulate generated or manipulated content by very large online platforms and search engines. If entities do not employ "prominent markings" under DSA, they risk incurring multimillion-euro fines. In June 2023, EU Commission Vice President Vera Jourova was quoted stating that "EU regulations are aimed at protecting free speech, but when it comes to AI, "I don't see any right for the machines to have the freedom of speech."

Separately, the European Parliament is still deliberating over its Artificial Intelligence Act, which would come into effect no earlier than 2025 or 2026. The EU's AI Act will govern any product or service that uses an artificial intelligence system. The act will classify AI systems according to four levels of risk, from minimal to unacceptable. For now, general purpose generative AI systems resembling ChatGPT are subject to some of the requirements of high-risk

---

[122] *See* Brad Smith, *Microsoft announces new Copilot Copyright Commitment for customers*, MICROSOFT (Sept. 7, 2023), https://blogs.microsoft.com/on-the-issues/2023/09/07/copilot-copyright-commitment-ai-legal-concerns/ (stating, "We have incorporated filters and other technologies that are designed to reduce the likelihood that Copilots return infringing content").

[123] Strong, *supra* note 116, at § 9.6.2.

[124] Burk, *supra* note 108, at 323.

systems. This type of characterization is subject to chilling regulatory effects and may be too broad.

As law students we are fundamentally concerned about balancing the freedom of speech that human users enjoy, harnessing the innovative effect of this technology, and combating harmful disinformation. Prominently labeling AI-generated content is an effective and neutral tactic that gives human users a chance at verifying information while respecting their freedoms.

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

We believe that the responsibility for identifying or labeling a work as AI-generated depends on the context in which the material was produced. If the material was produced by an individual or groups of individuals in a non-commercial organization using either one generative AI or multiple open-source generative AI systems, the responsibility should remain with the individual to label the content as AI-Generated. Our concern is that this class of users would not be included in potential AI regulation. Individual AI-models enable sophisticated users to produce potentially harmful content at scale at dramatically reduced costs, and obligations should be imposed on individuals or small groups even operating in a typically non-commercial setting. This is a lower bar than the use of AI-systems, which serve to further amplify the ability of sophisticated users to direct sentiment with AI-generated material.

In situations unlike the above, such as commercial use, or by non-natural persons, we believe that the responsibility for identifying or labeling a work as AI-generated should lie with the cheapest cost avoider. The responsibility should be undertaken in these situations by the entity with the most resources and greatest capacity to label and identify AI-generated content as such. This will often be the same entity as the creator, developer, provider, broadly speaking the maintainer of the closed-source generative AI systems.

- **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?**

Congress could empower a newly created AI Agency through legislation and turn over discretion to experts within the agency. Consequences of the failure to label an AI-generated work or by removing labels improperly could incorporate some of the following ideas:

- Implementing regulations that mandate algorithmic “deboosting” of AI-generated material or mandate algorithmic disgorgement of the technology platform hosting the content after repeated or egregiously harmful violations. These actions have been pursued in the past in analogous actions by the FTC.
- Implementing regulations that call for the demonetization of AI-generated content including a strong joint commitment by platform operators and advertisers to avoid placing ads alongside AI-generated material that was either never labeled as such or where a label was improperly removed. Platform operators and advertisers can also make

a strong joint commitment to continually review advertisements under the scope of removing disinformation.

- Complete removal of the AI-generated content.
- Fines on a sliding scale similar to the European Union's progressive legislation. For instance:
    - Under the pending EU AI Act, breaches of a prohibited practice can amount to violations in ranges up to the higher of 40 million euros or 7% of a company's annual global revenue. The AI Act's fine provision awards the GDPR's fining range of up to 4% of a company's global revenue. Penalties for foundation model providers who breach the AI Act could amount to 10 million euros or 2% annual revenue, whichever is higher.

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?**

While not a standard-setting body, the actions of businesses in the free market seeking to self-regulate emerging technologies such as AI-Generated material should be considered influential and persuasive. The office should review the recent action in September of 2023 taken by TikTok to properly label and disclose AI-Generated material on its platform.[125] Additionally, TikTok's Community Guidelines Policy on Synthethic and Manipulated Media should also be considered as a persuasive tool for standard-setting bodies.[126]

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?

Individuals have a number of rights available to them for AI-generated material that features their name or likeness. Along with these rights are legal claims stemming mostly from tort law that litigants may use to enforce their rights.

1. Right of Publicity

This right refers to "the right to control the use of one's own name, picture, or likeness and to prevent another from using it for commercial benefit without one's consent."[127] The right of publicity is a state-law created right "whose infringement is a commercial tort of unfair competition."[128] Individuals whose name or likeness, vocal or otherwise, should be afforded the right of publicity if these features are used in AI-generated material. For a more in-depth

[125] *New Labels for Disclosing AI-Generated Content*, TIKTOK (Sept. 19, 2023), https://newsroom.tiktok.com/en-us/new-labels-for-disclosing-ai-generated-content.
[126] Synthetic and Manipulated Media, TIKTOK, https://www.tiktok.com/community-guidelines/en/integrity-authenticity/#3 (last visited Oct. 22, 2023).
[127] *Right of Publicity*, *Black's Law Dictionary* (11th ed. 2019).
[128] 1 J. Thomas McCarthy, *The Rights of Publicity and Privacy* § 1:3, at 1-2 (2d ed. 2000).

discussion on the right of privacy, please refer to Questions 31 and 32 below.

2. Right of Privacy

This right considers one's personal autonomy and the constitutional guarantees that safeguard certain "zones of privacy."[129] Included in this right is an individual's decision to hold their person and property free from unwarranted exposure or public scrutiny.[130] When an individual's privacy is intruded on, whether it be their name, likeness, personality, or private affairs among other factors, that individual could likely bring a tort claim to enforce their right of privacy.[131]

The below are two possible claims individuals may take to protect their above rights.

1. Defamation

Under the Second Restatement of Torts, defamation requires four factors: (i) there must have been a false and defamatory statement concerning another person; (ii) an unprivileged publication to a third-party; (iii) fault amounts to at least negligence from the publisher; and (iv) special harm was caused to the person of interest as a result of the defamatory statement.[132] Defamation includes libel or slander. Typically, public figures bring defamation suits, however, the Supreme Court has held private figures are entitled to protection through a showing of mere negligence from the defendant.[133] Private figures are ordinary individuals who do not public office and who have not sought out the public limelight; these figures did not invite attention or comment on their actions.

It is possible that if AI-programs are used to generate material that could damage a person's reputation, a defamation claim could reasonably be raised. AI-programs such as Bing's image creator, Microsoft's VALL-E, or other similarly designed "deepfake" software could be utilized to manipulate one's likeness to convincingly issue defamatory statements. Defamation should thus be a right that is affordable to both public and private figures in the event AI is used to harm the reputation of one's person, business, or organization.

2. Fraud

Fraud is the act of knowingly misrepresenting or concealing a material fact so as to induce another to act to their detriment.[134] Forgery is a crime of fraud and has been defined as a crime of fraudulently making a false document or altering a real document with the purpose of using it as though it were genuine.[135]

[129] *Right of Privacy*, *supra* note 127.
[130] *Id.*
[131] *Id.*
[132] Restatement (Second) of Torts § 558 (1977).
[133] *Gertz v. Welch*, 418 U.S. 323 (1974).
[134] Restatement (Second) of Torts § 525 (1977).
[135] 18 U.S.C. § 513.

In the case of AI, fraud can take the form of false impersonation through voice or appearance to deceive others, whether it be through financial fraud with voice imitation or forgery of signatures. Thus, there needs to be fraudulent detection measures not only to prevent or mitigate such risks, but to distinguish genuine and real representations from the disingenuous and deceitful ones. Otherwise, the public will hesitate to even believe real representations, leading to skepticism and distrust towards real people, public figures, or the government.

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, Congress should establish a new federal right, similar to state law rights of publicity, that would apply to AI-generated materials. The patchwork of state rights of publicity has resulted in a divergence of protection and application, leading to forum shopping and leaving national companies unsure of how their content will be treated from state to state.[136] To champion individuals' right to control their persona and establish uniformity across the nation, a federal right of publicity should be adopted. Congress has the power to do so through the Commerce Clause, as right of publicity affects interstate commerce activities like national advertisements of products or nationwide content and is infringed through interstate commerce channels like the Internet.[137] A federal right of publicity would need to establish the scope of what is to be protected, whether such a right would only attach to someone with pre-existing commercial value, and what constitutes infringement.[138] All people, celebrity or not, deserve the right to protect their image and persona from unwarranted use. A federal right of publicity should attach to every aspect of anyone's persona, regardless of preexisting commercial value. Mere use in commerce without their consent should constitute infringement of such a right. The federal right should preempt state rights. This would help to protect the rights of all people to their image while eliminating divergence of the right amongst the states.

However, considering the breadth of content on which AI systems can be trained, when it comes to infringement of the right of publicity through the use of AI systems, commercial use should not be a required element. Celebrities and those with commercial value have been combatting AI-generated deepfakes and music for the past few years.[139] However, sometimes the end-users do not produce these for commercial purposes, but for creativity or maliciousness. [140] Nonetheless, celebrities should be allowed to prevent this non-commercial, unauthorized use of

---

[136] Kevin L. Vick, et al., *Why a Federal Right of Publicity Statute is Necessary*, AM. BAR ASS'N (2011), https://www.americanbar.org/content/dam/aba/publications/communications_lawyer/august2011/why_federal_right_publicity_statute_is_necessary_comm_law_28_2.authcheckdam.pdf

[137] *Id.*

[138] *See* W. Woods Drinkwater, *Personality Beyond Borders: The Case for a Federal Right of Publicity*, 3 MISS. SPORTS L. REV. 115, 137-139 (2013).

[139] *See* Anumita Kaur, *Celebrities Warn Followers not to Be Duped by AI Deepfakes*, THE WASH. POST (Oct. 3, 2023), https://www.washingtonpost.com/entertainment/2023/10/03/tom-hanks-ai-ad-deepfake/.

[140] *See* Ian Sample, *What Are Deepfakes - And How Can You Spot Them?*, THE GUARDIAN (Jan. 13, 2020), https://www.theguardian.com/technology/2020/jan/13/what-are-deepfakes-and-how-can-you-spot-them.

their persona, as it opens the gates for other exploitation of their image. Additionally, AI-generated deepfakes and voices are also harmful to non-celebrity consumers,[141] whose images and voices may be used for "blackmail, bullying, defamation, harassment, identity theft, intimidation, and revenge porn."[142] If an individual cannot prove commercial use of their persona, we suggest there should be an exception if the individual can show a certain level of deception or confusion. This exception would be proactive in protecting all individuals from exploitation of their image or voice, commercial or not, as AI content continues to develop and grow.

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?

A federal right of publicity would serve to protect artists against AI-generated output that imitates their artistic style.[143] Any artist, no matter their degree of notoriety, should be eligible for such protection.[144] As an artist's style is intrinsically linked to their persona, a federal right of publicity should protect against unauthorized imitations, whether it be a painting style or an artist's musical voice.[145] As stated above in Question 31, a federal right of publicity should protect every aspect of anyone's persona, and in the context of imitation, this should include recognizable artistic style. If the end-user is profiting from imitating output, the artist should be able to recover.[146] However, if the imitation is merely for noncommercial creative purposes, but the artist still wants to protect their persona and artistic style, deception and confusion should be relevant factors in allowing the artist to succeed on such a claim.

---

[141] *See* Audrey Schomer, *Deepfakes Don't Just Harm Celebs. Consumers Are Also at Risk*, VARIETY (Oct. 12, 2023), https://variety.com/vip/deepfakes-dont-just-harm-celebrities-consumers-are-also-risk-1235752489/.

[142] Mekhail Mustak, et al., *Deepfakes: Deceptions, Mitigations, and Opportunities*, 154 J. OF BUS. RSCH 1, 1 (Jan. 2023), https://www.sciencedirect.com/science/article/abs/pii/S0148296322008335.

[143] Rachel Reed, *AI Created a Song Mimicking the Work of Drake and The Weeknd. What Does That Mean for Copyright Law?*, HARV. L. TODAY (May 2, 2023), https://hls.harvard.edu/today/ai-created-a-song-mimicking-the-work-of-drake-and-the-weeknd-what-does-that-mean-for-copyright-law/ (stating that when it comes to AI-generated output imitating artists, "...their best argument is not a copyright argument at all, but a right of publicity argument."

[144] *See* Beatrice Nolan, *Artists Say AI Image Generators Are Copying Their Style to Make Thousands of New Images – And It's Completely out of Their Control*, INSIDER (Oct. 17, 2022), https://www.businessinsider.com/ai-image-generators-artists-copying-style-thousands-images-2022-10 (quoting an artist, "People are pretending to be me... it seems unethical… We work for years on our portfolio... Now suddenly someone can produce tons of images with these generators and sign them with our name.").

[145] Waits v. Frito-Lay, Inc., 978 F.2d 1093, 1098 (9th Cir. 1992) ("[W]hen voice is a sufficient indicia of a celebrity's identity, the right of publicity protects against its imitation...").

[146] *See* Will Knight, *Algorithms Can Now Mimic Any Artist. Some Artists Hate It*, WIRED (Aug. 19, 2022), https://www.wired.com/story/artists-rage-against-machines-that-mimic-their-work/ ("[T]he idea of using these tools that feed on past work to create new works that make money feels wrong.").

The BLIP Clinic is deeply committed to furthering dialogue on emerging technologies and their impact on matters such as Copyright law and policy. We believe a well-informed forward-thinking society is crucial to navigate the evolving landscape of AI technology and its implications on the law. As such, we appreciate this opportunity to contribute to the USCRO's NOI and look forward to future engagement. We hope our response serves as a valuable resource in these ongoing discussions.

Sincerely,
/s/ Nicholas Castro
/s/ Jimmy Nguyen
/s/ Sara Sachs
/s/ Caroline Mattis
/s/ Juliette Adams
/s/ Austen Fisher
/s/ Sofia Vescovo
/s/ Michelle Agiviav

Brooklyn Law Incubator & Policy Clinic
Brooklyn Law School
250 Joralemon Street, Brooklyn New York, 11201
BLIP@brooklaw.edu
jonathan.askin@brooklaw.edu
nicholas.castro@brooklaw.edu
jimmy.nguyen@brooklaw.edu
sara.sachs@brooklaw.edu
sara.mattis@brooklaw.edu
juliette.adams@brooklaw.edu
austen.fisher@brooklaw.edu
sofia.vescovo@brooklaw.edu
michelle.agiviav@brooklaw.edu